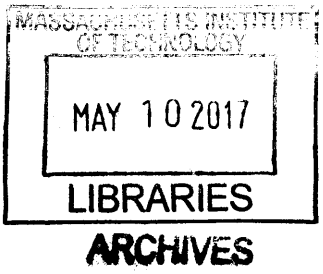
Discovery and Investigation of the Novel Overall Activity Allosteric Regulation of the *Bacillus subtilis* Class Ib Ribonucleotide Reductase

by

Mackenzie James Parker

B.S. Chemistry
Montana State University, 2007



Submitted to the Department of Chemistry
in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy in Biological Chemistry

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 2017

Signature of Author: Signature redacted
Department of Chemistry
November 4, 2016

Certified by: Signature redacted
JoAnne Stubbe
Novartis Professor of Chemistry and Professor of Biology
Thesis Supervisor

Accepted by: Signature redacted
Robert W. Field
Haslam and Dewey Professor of Chemistry
Chairman, Departmental Committee on Graduate Studies

This doctoral thesis has been examined by a Committee of the Department of Chemistry and Biology as follows:

Signature redacted

Barbara Imperiali
Class of 1922 Professor of Biology and Professor of Chemistry
Thesis Chair

Signature redacted

JoAnne Stubbe
Novartis Professor of Chemistry and Professor of Biology
Thesis Supervisor

Signature redacted

Elizabeth M. Nolan
Associate Professor of Chemistry

For my family

Acknowledgements

I first want to thank JoAnne for taking me into her lab and making me a better scientist. I am very grateful for this. I want to thank my thesis committee members Barbara and Liz for their support and intellectual discussions. I'm additionally indebted to Liz for allowing me to use some of her instruments over the past four years. I would not have been exposed to analytical ultra-centrifugation were it not for the Biophysical Instrumentation Facility, and I want to thank Barbara for establishing the facility and Debbie Pheasant for all her help and patience as I learned how to use the instrument. Debbie, we would be record-breaking gold medalists in window washing.

During the course of my graduate career, I've had the pleasure of working with several excellent collaborators. I want to thank Diana Downs and her students Lauren Palmer and Cecilia Martinez-Gomez for their tutelage, help, and for intellectual discussions we had when I was working on ThiC. I want to thank Amie Boal for all the intellectual discussions, support, and images she has provided, and I want thank her student Ailiena Maggiolo for her hard work in getting a crystal structure of dAMP-loaded NrdE. I want to thank Nozomi Ando and her students Will Thomas and Max Watkins for their support and intellectual contributions to this thesis. You guys put my mind at ease by confirming the strange behaviors I observed with *Bacillus subtilis* RNR using small angle X-ray scattering. Finally, I want to thank Cathy Drennan and her students Michael Funk, Christina Zimanyi, and Peter Goldman for intellectual discussions on the structure of RNR.

To my past lab mates Joey, Mimi, Rachael, Ellen, Quamrul, Yan, Ping, Jun, Henri, Danny, Arturo, Luke, Kathleen, Judy, and Thomas, thank you for teaching me, for commiserating with me, and helping me become a functional member of the lab. I am especially indebted to Ken and Crystal for their tutelage, inspiration, and support. I also want to thank Jeff Beyers for his support and help with the gloveboxes.

To my current lab mates Kanchana, Wankyu, Alex, Brandon, Denyce, Albert, Meng, Qinghui, Julia, Yifeng and Lisa, thank you for all your help and support over the past four years. Denyce, I particularly enjoyed our intellectual discussions about your work, and I'm excited to see where you go with it from here. Thank you also for reading through the rough drafts of my thesis chapters and providing thoughtful comments and edits. Brandon and Lisa, thank you for sharing your furry friends, Kingsley and Rafa, with the lab.

I had the honor of mentoring two very bright and talented students. Marica and Johannes: thank you for all your hard work and for being excellent pupils to teach. I wish you two well, and good luck in your future endeavors.

I had the pleasure of entering MIT with eight other peers in the Biological Chemistry division whom I have enjoyed working and interacting with during our time here. I especially want to thank Mike for his friendship and hosting our Thanksgiving potlucks every November, Alyssa and Jen for comradery and commiseration during TAing 5.07 and 5.08 our first year, and Stephanie, Michelle, Jen, and Wesley for their friendship and continued support over the years. Drew, you've

been my best friend out here in Boston, and I thank you for this, your support, and for all the dinners and MST3K movie nights.

One of the best distractions I had from the lab was book club. I want to thank Crystal and Jeff for originally organizing it, and to thank Laura, Jeremy, Sangyu, Megan, and Kanch for all the great discussions at the Muddy and for nominating some strange, yet delightful books. Making it through works like Anna Karinina, The Way We Live Now, War and Peace, Snow, and Orlando, to name a few -- that's quite an accomplishment.

I want to thank the PT cohort for their friendship and comradery. You guys got me out of the lab every now and then to have some fun and enjoy life and I really appreciate it. Melson, Ty, Jarrett, and Johnny "the PT," it was great living and getting to know you guys. Also, Johnny, thank you for giving me the opportunity to take care of the little Chupacabra. Walking him in the morning was a treat and made coming into work more bearable. Zach and Maria, congratulations again on your engagement. I'm sure I'll be hearing more from you guys in the future through Joe. In no particular order, Jenny, Thomas, Elyssa, Julie, and Alex--thank you for all the good times out here.

I had an excellent group of friends back home who have supported and encouraged me throughout my time here at MIT, and I am indebted to them all. I want to especially thank Chris and Lisa for all the phone calls and care packages, and Terra for keeping me grounded. Eric, Preston, Andy, Bryan, Krista, Tyson, and Chad, thank you all for your support and friendship. Finally, I'm extremely lucky to have Lee as my best friend since junior high school. Lee, thank you for all your support and encouragement over the years.

Over the past year, I've been especially fortunate to have met a remarkable family through my brother Joe. Mel, Alissa, Dave, Maryann, and the rest of the Belchertown family: thank you for your hospitality, encouragement, and support this year, and for making me feel like a part of the family. It was particularly fun to come out and join in making dozens upon dozens of pierogis (which are delicious) this past April. I look forward to continuing to build our relationship in the years to come.

Finally, and most importantly, I would not be where I am today without the love and support of my family. Dad and Judy: words cannot express my gratitude to you both. Your continued belief in me was motivation to finish this degree. Thank you for your love, support and encouragement. Joe, thank you for not only your support, encouragement, and love, but also for making my second stint in Boston a wonderful experience. Even though five years separates us, I feel like we got to know each other a lot better living together. Also, thank you for introducing and making me a part of the PT group. Alex, you helped me pull out of a dark time in the middle of my graduate school experience, and I wish I had been able to help you do the same. I hope you have found peace. Jesse and Kesley, thank you for your support. Believe me, school will end, so stick to your guns and finish those law degrees. To all of my grandparents, aunts, uncles, cousins, and step-relatives, thank you for you love, support, and encouragement. This thesis is dedicated to all of you.

Discovery and Investigation of the Novel Overall Activity Allosteric Regulation of the *Bacillus subtilis* Class Ib Ribonucleotide Reductase

by

Mackenzie James Parker

Submitted to the Department of Chemistry
On November 4, 2016 in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy in Biological Chemistry

ABSTRACT

Ribonucleotide reductases (RNRs) catalyze the reduction of nucleotides to 2′-deoxynucleotides in all organisms. Class Ib RNRs consist of two subunits: α houses the catalytic and allosteric effector binding sites, and β houses a catalytically essential dimanganic-tyrosyl radical ($Mn(III)_2$-Y•). The allosteric regulation of Ib RNR activity has only been studied with the *Salmonella enterica* enzyme, which exhibits substrate specificity allosteric regulation by ATP and 2′-deoxynucleoside 5′-triphosphates (dNTPs), but not overall activity regulation by ATP and dATP. However, the *S. enterica* enzyme is not a good general model for Ib RNRs because it is not essential under most growth conditions, including pathogenesis. Other bacteria pathogenic to humans utilize Ib RNRs as their sole source of dNTPs for DNA replication and repair. As RNR regulation plays a critical role in the high fidelity of these processes, the allosteric regulation of Ib RNRs used as the primary dNTP supplier for a bacterium should be distinct from the *S. enterica* enzyme and, therefore, could provide a potential target for therapeutic development.

Herein, the results of characterizing the allosteric regulation of the Ib RNR from the model organism *Bacillus subtilis* are presented. To facilitate these studies, we identified, cloned, and isolated the physiological reductant for RNR (thioredoxin/thioredoxin reductase/NADPH), thus allowing us to monitor activity spectrophotometrically. We discovered the effector dATP was a potent inhibitor of enzymatic activity at physiologically relevant concentrations, thereby demonstrating the first example of overall activity allosteric regulation in a class Ib system. In other RNRs, overall activity regulation is mediated by a domain called the ATP-cone. This domain is absent from the *B. subtilis* enzyme; therefore, the inhibition represents a new mechanism of overall activity regulation. Analytical ultracentrifugation studies suggest dATP inhibition may be mediated by formation of large protein complexes. Biophysical studies also led to the discovery of tightly bound dAMP associated with α that increases the susceptibility of RNR to dATP inhibition. The potential physiological importance of dAMP is supported by studies examining YmaB, the unique fourth member of the *B. subtilis* RNR operon, which revealed this enzyme can hydrolyze dATP into dAMP and pyrophosphate and, therefore, might insert dAMP into α.

Thesis Supervisor: JoAnne Stubbe
Title: Novartis Professor of Chemistry and Professor of Biology

Table of Contents

Dedication 5

Acknowledgements 7

Abstract 9

Table of Contents 11

List of Figures 21

List of Tables 27

Abbreviations 29

Chapter 1. Introduction to Ribonucleotide Reductases 33

1.1. Synopsis 34

1.1.1. Similarities and differences of class Ia and Ib RNRs. 35

1.1.2. Why should the allosteric regulation of class Ib RNRs be reinvestigated? 36

1.2. A general introduction to ribonucleotide reductases 39

1.2.1. The classes of RNRs. 39

1.2.2. The sources of reducing equivalents. 45

1.2.3. The mechanism of ribonucleotide reduction. 47

1.3. General introduction to class I RNR structure 50

1.3.1. Characteristics of class I RNR subunits. 50

1.3.2. Characteristics of class I RNR $\alpha_2\beta_2$ complexes. 55

1.3.3. The radical transfer pathway of the *E. coli* class Ia RNR. 61

1.4. The allosteric regulation of RNR activity 64

1.4.1. General scheme for allosteric substrate specificity and overall activity regulation of RNRs. 64

1.4.2. All RNRs have the same allostery for substrate specificity, but diversity in overall activity regulation. 66

1.4.3. Pyrimidine dNTP pools additionally regulated by deaminases of deoxy-cytidine analogs. 68

1.4.4. Structural basis for allosteric regulation of overall activity in class I RNRs. 69

1.4.5. Structural basis for allosteric regulation of substrate specificity in class I RNRs. 76
1.5. The *Bacillus subtilis* Class Ib RNR 85
1.5.1. Important residues in the class Ib RNR from *B. subtilis*. 85
1.5.2. Assembly of the $Mn(III)_2$-Y• cofactor in *B. subtilis* NrdF. 87
1.5.3. Influence of the lifestyles of *B. subtilis* on RNR regulation. 90
1.5.4. Some strains of *B. subtilis* encode a second, viral-associated class Ib RNR. 94
1.5.5. Nucleotide metabolism in *B. subtilis*. 96
1.5.6. Influence of *B. subtilis* nucleotide metabolism on RNR activity. 105
1.6. Nudix hydrolases 107
1.7. Chapter preview 111
1.8. References 117

Chapter 2. *Bacillus subtilis* Class Ib Ribonucleotide Reductase: High Activity and Dynamic Subunit Interactions 139
2.1. Introduction 140
2.2. Experimental 144
2.2.1. Materials 144
2.2.2. General Methodology 145
2.2.3. Buffers 146
2.2.4. Cloning of N-terminally His_6-tagged redoxins 147
2.2.5. General growth and expression procedure 147
2.2.6. Purification of N-terminally His_6-tagged RNR proteins, TrxA, TrxB, and YosR 148
2.2.7. Purification of His_6-YtnI 151
2.2.8. Determination of ε_{455} for FAD bound to His_6-TrxB 152
2.2.9. Redoxin activity assays 152
2.2.10. Reconstitution of $Fe(III)_2$-Y• His_6-NrdF 153
2.2.11. Reconstitution of $Mn(III)_2$-Y• His_6-NrdF 153
2.2.12. Purification of holo-$Me(III)_2$-Y• His_6-NrdF subsequent to cofactor assembly 154

2.2.13. Metal quantification 154
2.2.14. Y• quantitation by EPR spectroscopy 156
2.2.15. RNR radioactive activity assays 156
2.2.16. Quaternary structural analysis of 1:1 mixtures of α:β using biophysical methods 157
2.2.16A. Anion exchange chromatography 157
2.2.16B. Size exclusion chromatography 157
2.2.16C. Sedimentation velocity analytical ultracentrifugation (SV-AUC) 158
2.2.16D. Hydrodynamic modeling of the His_6-tagged *B. subtilis* RNR subunits 160
2.2.17. SDS-PAGE densitometry 161
2.3. Results 161
2.3.1. Characterization of as-isolated His_6-tagged *B. subtilis* RNR proteins. 161
2.3.2. Holo-$Me(III)_2$-Y• His_6-NrdF (Me = Fe or Mn) can be purified by anion exchange on a MonoQ column. 163
2.3.3. Purification and characterization of the *B. subtilis* redoxins. 164
2.3.4. The endogenous reducing system for *B. subtilis* RNR is TrxA, TrxB, and NADPH. 165
2.3.5. Activity assay optimization. 167
2.3.6. Reconstitution of holo-RNR from His_6-NrdE and reconstituted His_6-NrdFs. 169
2.3.7. Estimation of the α/β interaction strength. 170
2.3.8. Analysis of the quaternary structure of the *B. subtilis* class Ib RNR by multiple methods reveals a dynamic system. 172
2.3.8A. Analysis of quaternary structure by SEC. 172
2.3.8B. Analysis of quaternary structure by SV-AUC. 175
2.4. Discussion 179
2.5. References 184

Chapter 3. Characterization of the Allosteric Regulation of the Class Ib RNR from *Bacillus subtilis* and its Links to Quaternary Structure 193
3.1. Introduction 194

3.2. Experimental 199
3.2.1. Materials and methods 199
3.2.2. Expression and purification of SUMO protease 199
3.2.3. Cloning the *B. subtilis* RNR genes into pE-SUMO 201
3.2.4. Expression and purification of the His_6-Smt3-tagged *B. subtilis* RNR proteins 202
3.2.5. Removal of the His_6-Smt3 tag to produce tagless *B. subtilis* RNR subunits 203
3.2.6. Determination of NrdI FMN loading and ε_{449} for protein-bound flavin 204
3.2.7. Reconstitution and purification of holo-$Mn(III)_2$-Y• tagless NrdF 205
3.2.8. Measuring RNR activity using the spectrophotometric assay 205
3.2.9. Correlating NADPH consumption with [5-^{3}H]-dCDP production 205
3.2.10. Measuring steady-state kinetics of *B. subtilis* RNR with different substrate/effector pairs 205
3.2.11. Characterization of *B. subtilis* RNR subunit quaternary structure(s) in the presence of nucleotides by SV-AUC 206
3.2.11A. Sample preparation 206
3.2.11B. Instrument setup and experiment execution 207
3.2.11C. Data analysis 208
3.2.11D. Hydrodynamic modeling of the tagless *B. subtilis* RNR subunits 209
3.3. Results 209
3.3.1. Fusion of a cleavable His_6-Smt3 tag allows for relatively facile production of tagless *B. subtilis* RNR proteins. 209
3.3.2. Reconstitution of tagless $Mn(III)_2$-Y• NrdF is similar to the His_6-tagged counterpart. 212
3.3.3. NADPH oxidation and [5-^{3}H]-dCDP formation by Mn-loaded *B. subtilis* RNR are correlated. 213
3.3.4. The catalytic efficiency of *B. subtilis* RNR is highest when bound with dNTP effectors. 214
3.3.5. Cooperative effects on steady-state activity are observed with the purine substrates. 218
3.3.6. Substrate reduction is best stimulated by the expected prime effector. 219

3.3.7. Inhibition of *B. subtilis* RNR by dATP is an inherent property of the enzyme. 221
3.3.8. Maximum GDP reduction rates require the simultaneous presence of TTP and ATP. 222
3.3.9. *B. subtilis* RNR exhibits product inhibition by dADP that is not reversed by ATP. 225
3.3.10. The quaternary structure of the *B. subtilis* RNR subunits is not significantly affected by tag removal. 228
3.3.11. General considerations for SV-AUC analysis of *B. subtilis* RNR in the presence of nucleotides. 229
3.3.12. The quaternary structure of NrdF is unaffected by nucleotides. 230
3.3.13. The binding of dATP stabilizes NrdE dimers. 232
3.3.14. NrdE dimer is not stabilized by TTP. 235
3.3.15. GDP binding causes NrdE dimers to dissociate. 235
3.3.16. NrdE dimerization is not induced by the presence of saturating amounts of both substrate and effector. 236
3.3.17. Large NrdE/NrdF complexes form in the presence of dATP. 237
3.4. Discussion 238
3.4.1. Steady-state kinetic parameters are consistent with estimated *in vivo* concentrations of nucleotides in *B. subtilis*. 239
3.4.2. The role of substrate competition in specificity regulation of class I RNRs. 241
3.4.3. UDP is efficiently utilized by *B. subtilis* RNR. 242
3.4.4. The *S.* Typhimurium and *B. subtilis* Ib RNRs differ in response to ATP versus dATP. 243
3.4.5. The ability of ATP to augment TTP-stimulated GDP reduction is similar to eukaryotic class Ia RNRs. 244
3.4.6. Cooperative effects on RNR activity by purine substrates are very unusual. 245
3.4.7. Quaternary structural alterations are still the mechanism by which dATP inhibits the *B. subtilis* RNR. 246

3.5. References 248

Chapter 4: Discovery of a Tightly Bound Equivalent of dAMP Associated with *Bacillus subtilis* NrdE that Modulates Inhibition of Ribonucleotide Reductase by dATP 257

4.1. Introduction 258

4.2. Experimental 262

4.2.1. Materials and Methods 262

4.2.2. Identification of a tightly bound equivalent of dAMP associated with NrdE 262

4.2.2A. Quantitation of dAMP bound to NrdE 262

4.2.2B. Large-scale isolation of dAMP 263

4.2.2C. Determination of the nucleoside phosphorylation state by poly-ethyleneimine (PEI) cellulose thin layer chromatography (TLC) 263

4.2.2D. Characterization of the NrdE-associated nucleotide by ^{1}H-NMR 264

4.2.2E. Separation of (deoxy)nucleoside mono-, di-, and triphosphates by anion exchange chromatography 265

4.2.2F. Nucleotide isolation from NrdE using perchloric acid precipitation and assessment for non-specific phosphate hydrolysis. 265

4.2.3. Efforts to load or remove dAMP from as-isolated NrdE 266

4.2.4. Separation of holo- from apo-NrdE by MonoQ anion exchange chromatography 266

4.2.5. Assessment of the effects of dAMP loading on NrdE quaternary structure 267

4.2.5A. SEC 267

4.2.5B. SV-AUC 267

4.2.6 Assessing the effects of dAMP loading on RNR activity 268

4.2.7 Assessing the effects of dAMP loading on dATP binding to NrdE 268

4.2.8. Site-directed mutagenesis, expression, and purification of tagless holo-NrdE-$C_{382}S$ 270

4.2.9. Determination of the effects of dAMP loading on $\alpha_2\beta_2$ complex formation 270
4.2.9A. Assessing the strength of the interaction between tagless apo- or holo-NrdE and $Fe(III)_2$-Y• His_6-NrdF using pull-down assays 270
4.2.9B. Assessing the quaternary structure of 1:1 mixtures of holo-NrdE-$C_{382}S$ and $Mn(III)_2$-Y• β_2 in the absence and presence of nucleotides by SV-AUC 271
4.3. Results 271
4.3.1. Identification of dAMP as the major compound associated with heterologously-expressed as-isolated NrdE. 271
4.3.2. The presence of dAMP is not a result of non-specific hydrolysis of dATP during protein denaturation. 278
4.3.3. dAMP is tightly bound and cannot be installed or removed from as-isolated NrdE *in vitro* by non-denaturing methods. 280
4.3.4. Apo- and holo-NrdE can be separated by anion exchange chromatography. 282
4.3.5. Bound dAMP appears to stabilize NrdE tertiary structure. 285
4.3.6. NrdE exhibits a higher propensity to oligomerize with dAMP bound. 285
4.3.7. Under non-inhibitory conditions, RNR activity is not affected by the presence of dAMP. 287
4.3.8. Apo- and holo-NrdE exhibit differences in dATP inhibition. 289
4.3.9. Binding experiments suggest apo-NrdE can bind two equivalents of dATP/α whereas holo-NrdE can only bind one. 290
4.3.10. dAMP loading does not influence the ability of NrdE and NrdF to form a qualitatively "tight" complex. 293
4.3.11. Purification and characterization of NrdE-$C_{382}S$. 294
4.3.12. Biophysical analysis of holo-NrdE:NrdF complexes indicate dAMP does not influence overall RNR quaternary structure. 295
4.3.13. SV-AUC analysis of a 1:1 mixture of holo-NrdE-$C_{382}S$:$Mn(III)_2$-Y• NrdF in the presence of GDP and TTP reveals $\alpha_2\beta_2$ and higher order complexes. 296
4.3.14. A tightly bound dAMP associated with NrdE appears to be unique to the *B. subtilis* system. 297

4.4. Discussion 298
4.4.1. A role for tightly bound nucleotides in RNRs is unclear. 299
4.4.2. A preliminary crystal structure of NrdE suggests the putative dAMP binding site is solvent exposed. 302
4.4.3. dAMP increases the susceptibility of *B. subtilis* RNR to dATP inhibition. 306
4.4.4. dAMP may cause NrdE to form an atypical dimer with its N-terminal domain. 308
4.5. References 313

Chapter 5. Investigation of the Role of YmaB in the Regulation of the *Bacillus subtilis* Class Ib RNR 319
5.1. Introduction 320
5.2. Experimental 323
5.2.1. General methodology 323
5.2.2. Cloning, expression, and purification of tagless YmaB 324
5.2.3. Examination of YmaB for Nudix hydrolase activity by phosphate detection 325
5.2.4. Qualitative examination of YmaB for Nudix hydrolase activity by PEI-cellulose TLC 327
5.2.5. Determination of YmaB quaternary structure by SV-AUC 327
5.2.6. Determination of the ability of YmaB to affect the dAMP load of NrdE *in vivo* 329
5.3. Results 330
5.3.1. Expression and purification of tagless YmaB. 330
5.3.2. YmaB is a Nudix hydrolase with specificity towards dNTPs. 331
5.3.3. YmaB exists as a monomer in the absence of substrate. 336
5.3.4. Testing of the ability of YmaB to increase *in vivo* loading of NrdE with dAMP. 337
5.4. Discussion 341
5.4.1. dNTPs link YmaB activity with RNR regulation. 341
5.4.2. YmaB homologs form a structurally distinct subgroup of Nudix hydrolases. 342
5.4.3. Effort to test the proposed role for YmaB as a dAMP insertase in *E. coli*. 345

5.4.4. Can YmaB help insert dAMP into folded NrdE? 347
5.4.5. Redundancy of functions in *B. subtilis* will make establishing an *in vivo* role for YmaB challenging. 349
5.5. References 350

Appendix 1, Section 1: *In silico* Structural Studies of the *E. coli* Class Ia RNR 355
A1.1 Introduction. 356
A1.2. General considerations. 359
A1.3. E_{350} in β may play a structural role in the conformational gating of radical transfer. 362
A1.4. Mapping plausible locations of Y_{356} (β subunit) in the $\alpha_2\beta_2$ complex structural model. 366
A1.5. Predicting the function of E_{52} (β subunit) in RNR activity. 380
A1.6. Conclusions. 386
A1.7. References. 386
Appendix 1, Section 2: Class Ia RNR, Group G Sequence Alignments – NrdA proteins 389
Appendix 1, Section 3: Class Ia RNR, Group G Sequence Alignments – NrdB proteins 405
Appendix 2: Supplemental Information for Chapter 2 415
Appendix 3: Supplemental Information for Chapter 3 431
Appendix 4: Supplemental Information for Chapter 4 453
Appendix 5: Miscellaneous Studies with the *B. subtilis* Class Ib RNR 465
A5.1. Introduction 466
A5.2. Investigation of the role of the ferredoxin/flavodoxin reductase YumC in assembly of the $Mn(III)_2$-Y• cofactor in NrdF 466
A5.2.1. Introduction 466
A5.2.2. Experimental 468
A5.2.3. Results and Discussion 470
A5.3. Characterization of NrdEB, the α subunit of the SPβ bacteriophage Ib RNR encoded in the *B. subtilis* genome 476
A5.3.1. Introduction 476
A5.3.2. Experimental 479

A5.3.3. Results and Discussion 482
A5.4. References 486

List of Figures

Figure 1.1.	Schematic of the reaction catalyzed by RNRs.	34
Figure 1.2.	Comparison of the active site structures of the three classes of RNRs.	42
Figure 1.3.	Classification of RNRs based on the metallocofactor utilized to generate a transient cysteine thiyl radical.	42
Figure 1.4.	Schematic for RNRs that use two cysteine residues as the source of electrons for nucleotide reduction.	45
Figure 1.5.	Endogenous reducing systems for RNRs.	47
Figure 1.6.	The proposed mechanism for nucleotide reduction using two cysteine residues as the source of reducing equivalents.	48
Figure 1.7.	The proposed mechanism for nucleotide reduction using formate as the source of reducing equivalents.	49
Figure 1.8.	Crystal structures of the α subunit of the *E. coli* Ia and *S.* Typhimurium Ib RNRs.	51
Figure 1.9.	Crystal structures of the β_2 subunit of the *E. coli* Ia, *S.* Typhimurium Ib (Fe-loaded), and *C. ammoniagenes* Ib (Mn-loaded) RNRs.	52
Figure 1.10.	Structural and spectroscopic characteristics of $Fe(III)_2$-Y• and $Mn(III)_2$-Y• from class Ia and Ib RNRs.	53
Figure 1.11.	Docking model of the *E. coli* class Ia RNR.	57
Figure 1.12.	Comparison of the binding of C-terminal peptides or the tail of β to α in the *E. coli* Ia RNR.	58
Figure 1.13.	Crystal structure of the *S.* Typhimurium Ib holo-RNR.	59
Figure 1.14.	Comparison of the binding of the C-terminal tail of β to α in the *E. coli* Ia and *S.* Typhimurium Ib RNRs.	60
Figure 1.15.	Proposed radical transfer pathway in the *E. coli* Ia RNR.	62
Figure 1.16.	Schematic showing the allosteric regulation of dNTP production in *E. coli*.	66

Figure 1.17. Crystal structure of the *E. coli* α subunit with an ATP analog bound in the overall activity allosteric site (ATP-cone domain). 70

Figure 1.18. Quaternary structure interconversions that regulate the overall activity of the *E. coli* Ia RNR. 73

Figure 1.19. Types of dATP-inhibited quaternary structures observed in class Ia RNRs 74

Figure 1.20. Loops mediate the allosteric regulation of substrate specificity in the *E. coli* Ia α subunit. 79

Figure 1.21. Sequence alignment of Loop 2 from characterized class Ia, Ib, and II RNRs 80

Figure 1.22. Protein-nucleotide interactions in the S-site of the class Ia RNR α subunit from *E. coli*. 80

Figure 1.23. Structural basis for communication between the catalytic and specificity sites in the *E. coli* Ia α subunit. 83

Figure 1.24. Proposed *in vitro* mechanism and summary of kinetics for $Mn(III)_2$-Y• assembly in *B. subtilis*. 89

Figure 1.25. The class Ib RNR operons in the genome of *B. subtilis*. 95

Figure 1.26. Purine biosynthetic pathway in *B. subtilis*. 96

Figure 1.27. Pyrimidine biosynthetic pathway in *B. subtilis*. 97

Figure 1.28. The nucleobase and nucleoside transporters of *B. subtilis*. 98

Figure 1.29. Nucleotide metabolic pathways in *B. subtilis*. 99

Figure 1.30. Summary of the allosteric regulation of the *B. subtilis* Ib RNR and deoxycytidine analog deaminases. 103

Figure 1.31. Purine degradation pathway in *B. subtilis*. 105

Figure 1.32. General reaction scheme for Nudix hydrolases. 107

Figure 1.33. Common structural elements of Nudix hydrolases using *E. coli* MutT as a model. 110

Figure 2.1. Schematic of the reaction catalyzed by class I RNRs and the metallo-cofactors of the Ia, Ib, and Ic enzymes. 141

Figure 2.2. SDS-PAGE (10% (w/v)) showing co-purification of NrdE and NrdF isolated from the 1B-UP strain of *B. subtilis*. 142

Figure 2.3. Anion exchange chromatograms showing the separation of apo- from holo-Me(III)$_2$-Y• His$_6$-NrdF (Me = Fe or Mn). 155

Figure 2.4. SDS-PAGE of purified His$_6$-tagged RNR and redoxin proteins. 162

Figure 2.5. Spectroscopic characterization of MonoQ-separated apo-His$_6$-NrdF and holo-Me(III)$_2$-Y• His$_6$-NrdF. 164

Figure 2.6. Optimization of Mg^{2+}, ATP, and dATP concentrations for RNR activity assays. 168

Figure 2.7. Reconstitution of a "tight" RNR complex as demonstrated by anion exchange chromatography and SDS-PAGE densitometry. 171

Figure 2.8. Estimating the K_m of the interaction for the His$_6$-tagged RNR subunits. 172

Figure 2.9. Concentration dependent SEC characterization of His$_6$-tagged RNR subunits and complexes. 174

Figure 2.10. SV-AUC characterization of the His$_6$-tagged RNR subunits and complexes at physiological concentrations. 178

Figure 3.1. Comparison of the crystal structures of *E. coli* NrdA and *S.* Typhimurium NrdE with nucleotides soaked into the catalytic and/or specificity sites. 195

Figure 3.2. Schematic illustrating His$_6$-Smt3-tag removal by SUMO protease. 209

Figure 3.3. SDS-PAGE showing preparation of tagless NrdE and NrdF. 211

Figure 3.4. SDS-PAGE showing preparation of tagless NrdI. 212

Figure 3.5. Correlation of NADPH consumption and [5-^{3}H]-dCDP production in spectrophotometric assays of RNR. 213

Figure 3.6. Comparison of fits to the Michaelis-Menten and Hill equations to steady-state data obtained with RNR using variable ADP and GDP concentrations. 219

Figure 3.7. Allosteric regulation of substrate specificity of RNR by single effector nucleotides. 221

Figure 3.8. dATP inhibition of UDP and CDP reduction catalyzed by RNR. 222

Figure 3.9. Enhanced GDP reduction by RNR in the presence of ATP and TTP. 224

Figure 3.10. Product inhibition of *B. subtilis* RNR by dADP and the inability of ATP to reverse this effect. 227

Figure 3.11. Concentration dependent SV-AUC characterization of NrdE and NrdF in the absence of nucleotides. 228

Figure 3.12. SV-AUC characterization of 2.1 μM NrdE monomer with 0.5 – 50 μM dATP. 231

Figure 3.13. SV-AUC characterization of 2.3 μM NrdE monomer with TTP (0.12 – 12 μM) or GDP (2.5 – 250 μM). 234

Figure 3.14. SV-AUC characterization of 2.3 μM NrdE monomer with 12 μM TTP and 250 μM GDP. 236

Figure 3.15. SV-AUC characterization of 1:1 mixtures of 0.5 – 3.5 μM His_6-NrdE: $Mn(III)_2$-Y• His_6-NrdF with 40 μM dATP. 238

Figure 3.16. Schematic of pyrimidine nucleotide biosynthetic and salvage pathways in *B. subtilis*. 243

Figure 4.1. Schematic of the reaction catalyzed by formylglycinamide ribo-nucleotide aminotransferase and the role of ADP in large and small PurLs. 260

Figure 4.2. UV-visible spectrum of the compound(s) isolated from His_6-NrdE and spectral characteristics of common adenine-containing nucleotides. 272

Figure 4.3. PEI-cellulose TLC analysis of the compounds isolated from His_6-NrdE before and after separation by anion exchange chromatography. 273

Figure 4.4. ^{1}H-NMR comparison of the deoxynucleoside monophosphate isolated from His_6-NrdE and dAMP in the regions with characteristic peaks for adenine and 2′-deoxyribose. 275

Figure 4.5. ^{1}H-NMR comparison of the 3′, 4′, and 5′ proton peaks of the deoxy-nucleoside monophosphate isolated from His_6-NrdE with dAMP, dADP, and dATP. 277

Figure 4.6. Anion exchange chromatogram demonstrating resolution of as-isolated NrdE into apo-and holo-protein and rechromatographing of apo- and holo-NrdE. 283

Figure 4.7. PEI-cellulose TLC analysis of the nucleotide isolated from a homogeneous preparation of holo-NrdE. 284

Figure 4.8. Characterization of apo- and holo-NrdE by SEC and SV-AUC. 286

Figure 4.9. Activity assays comparing the ability of dATP to inhibit apo- and holo-NrdE. 289

Figure 4.10. Saturation curves for [2,8-^{3}H]-dATP binding to apo- and holo-NrdE. 290

Figure 4.11. SDS-PAGE showing the results of pull-down assays that qualitatively probe the interaction of apo- and holo-NrdE with $Fe(III)_2$-Y• His_6-NrdF. 293

Figure 4.12. SV-AUC characterization of 1:1 mixtures of $Mn(III)_2$-Y• NrdF and either as-isolated wildtype NrdE or holo-NrdE-$C_{382}S$ in the absence of nucleotides. 296

Figure 4.13. SV-AUC characterization of a 1:1 mixture of $Mn(III)_2$-Y• NrdF and holo-NrdE-$C_{382}S$ in the presence of 12 μM TTP and 250 μM GDP. 297

Figure 4.14. Schematic illustrating the linkage of dATP to the resin backbone of dATP-Sepharose affinity resin. 300

Figure 4.15. Comparison of the preliminary crystal structure of *B. subtilis* NrdE (resolution: 3.1 Å) with the structure of *S.* Typhimurium NrdE (3.0 Å). 304

Figure 4.16. Examination of the dAMP binding site in the preliminary crystal structure of *B. subtilis* NrdE. 305

Figure 4.17. A possible alternative dimeric structure formed by dAMP-loaded *B. subtilis* NrdE as illustrated by crystallographic symmetry mates. 307

Figure 4.18. Close-up of secondary structural elements responsible for forming the dAMP-mediated alternative N-terminal dimer of *B. subtilis* NrdE. 309

Figure 4.19. Potential inter-protomer hydrogen bonding interactions with bound dAMP in the N-terminal *B. subtilis* NrdE dimer. 310

Figure 4.20. Potential inter-protomer π-cation interaction between R_{117} and F_{47} mediated by dAMP binding in the N-terminal *B. subtilis* NrdE dimer. 310

Figure 4.21. Possible structural explanation for dAMP and/or dATP-mediated inhibition of *B. subtilis* RNR. 312

Figure 5.1. Schematic of the class Ib RNR operon from *B. subtilis*. 321

Figure 5.2. General reaction scheme for substrate hydrolysis by Nudix hydrolases. 322

Figure 5.3. Logic behind the use of phosphate detection to measure the activity of Nudix hydrolases. 326

Figure 5.4. SDS-PAGE illustrating the preparation of tagless YmaB. 331

Figure 5.5. Activity assay results for hydrolysis of dATP and TTP by YmaB as monitored by phosphate release. 333

Figure 5.6. Qualitative demonstration by PEI-cellulose TLC of the ability of YmaB to hydrolyze all four canonical dNTPs. 334

Figure 5.7. Concentration-dependent SV-AUC characterization of YmaB (10 – 58 μM). 336

Figure 5.8. SDS-PAGE confirming co-expression of YmaB and His_6-Smt3-NrdE in *E. coli* cells. 339

Figure 5.9. Crystal structure comparison of *Thermotoga maritima* YmaB and *E. coli* MutT and sequence alignment of the Nudix boxes of several YmaB and MutT homologs. 344

Figure 5.10. A model for YmaB-facilitated dAMP insertion into folded NrdE. 348

List of Tables

Table 1.1.	Non-exhaustive list of bacteria pathogenic to humans and, when present, their main class I RNR.	40
Table 1.2.	Important residues in the *B. subtilis* class Ib RNR.	86
Table 1.3.	Sequence motifs distinguishing subfamilies of the Nudix hydrolases.	109
Table 2.1.	Sequences of primers used to clone *B. subtilis* redoxins.	146
Table 2.2.	SV-AUC sample concentration ranges and monitoring wavelengths for concentration dependent experiments with His_6-tagged subunits alone and in 1:1 mixtures.	159
Table 2.3.	Molecular weights and calculated partial specific volumes of His_6-tagged RNR components.	159
Table 2.4.	Optimized CDP reduction activities of $Me(III)_2$-Y• His_6-NrdF (Me = Fe or Mn).	166
Table 2.5.	RNR activity as a function of α:β ratio.	167
Table 2.6.	Biophysical characteristics of His_6-tagged subunits alone and in 1:1 mixtures in the absence of nucleotides.	173
Table 2.7.	HYDROPRO predicted hydrodynamic properties of His_6-tagged subunits and $\alpha_2\beta_2$ complex.	175
Table 3.1.	Primers used to clone *nrdE*, *nrdF*, and *nrdI* into pE-SUMO.	201
Table 3.2.	Molecular weights and calculated partial specific volumes of tagless RNR components.	208
Table 3.3.	Measured steady-state kinetic parameters with 1 mM substrate and variable effector concentrations.	216
Table 3.4.	Measured steady-state kinetic parameters with constant effector and variable substrate concentrations.	217
Table 4.1.	Total dAMP recoveries after dialysis of as-isolated His_6-NrdE into water.	263

Table 4.2.	R_f values for TLC analyses of compounds recovered from as-isolated His_6-NrdE and adenosine/deoxyadenosine analog standards shown in **Figure 4.3**.	274
Table 4.3.	Chemical shifts for dAMP and the nucleoside monophosphate recovered from as-isolated His_6-NrdE.	275
Table 4.4.	J^2 and J^3 proton coupling constants for dAMP and the nucleoside monophosphate recovered from as-isolated His_6-NrdE.	277
Table 4.5.	Results of attempting to increase or decrease the nucleotide loaded of as-isolated NrdE *in vitro*.	281
Table 4.6.	Steady-state kinetic parameters for RNR using apo- and holo-NrdE.	288
Table 4.7.	Best fit binding parameters for [2,8-^{3}H]-dATP binding to apo- and holo-NrdE.	291
Table 5.1.	R_f values for TLC analyses of hydrolysis products from incubation of YmaB with the canonical dNTPs.	335
Table 5.2.	Quantitation of dAMP bound to NrdE with or without co-expression of YmaB.	340
Table 5.3.	Additional predicted and confirmed Nudix hydrolases in *B. subtilis*.	350

Abbreviations

α_2	Class I ribonucleotide reductase large subunit, contains the catalytic and effector binding sites
β_2	Class I ribonucleotide reductase small subunit, contains the metallocofactor essential for NDP reduction
ρ	Solvent density
η	Solvent viscosity
$\bar{v}$	Partial specific volume
5′-dA	5′-deoxyadenosyl radical
8-oxo-dGTP	8-oxo-2′-deoxyguanosine 5′-triphosphate
8-oxo-dGMP	8-oxo-2′-deoxyguanosine 5′-monophosphate
ADP	Adenosine 5′-diphosphate
AMP	Adenosine 5′-monophosphate
Amp	Ampicillin
Apo-NrdE	NrdE without a tightly bound equivalent of nucleotide
As-isolated NrdE	Mixture of apo- and holo-NrdE
A-site	Allosteric overall activity site
ATP	Adenosine 5′-triphosphate
AUC	Analytical ultracentrifugation
BME	β-mercaptoethanol
Cm	Chloramphenicol
CDP	Cytidine 5′-diphosphate
CI	Confidence interval
CTP	Cytidine 5′-triphosphate
CV	Column volume
dADP	2′-deoxyadenosine 5′-diphosphate
dAMP	2′-deoxyadenosine 5′-monophosphate
dATP	2′-deoxyadenosine 5′-triphosphate
dC	2′-deoxycytidine

dCMP	2′-deoxycytidine 5′-monophosphate
dCTP	2′-deoxycytidine 5′-triphosphate
dGMP	2′-deoxyguanosine 5′-monophosphate
dGTP	2′-deoxyguanosine 5′-triphosphate
dNDP	2′-deoxynucleoside 5′-diphosphate
dNMP	2′-deoxynucleoside 5′-monophosphate
dNTP	2′-deoxynucleoside 5′-triphosphate
DTNB	5,5′-dithio-(2-nitrobenzoic acid)
DTT	DL-dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
EPR	Electron paramagnetic resonance
FAD	Flavin adenine dinucleotide
Ferrozine	3-(2-pyridyl)-5,6-diphenyl-1,2,4-triazine-*p,p′*-disulfonic acid
$Fe(III)_2$-Y•	Diferric-tyrosyl radical
FMN	Flavin mononucleotide
FPLC	Fast protein liquid chromatography
G•	Glycyl radical
GDP	Guanosine 5′-diphosphate
GEMMA	Gel electrophoretic mobility mass analysis
GTP	Guanosine 5′-triphosphate
HEPES	(2-hydroxyethyl)-1-piperazineethanesulfonic acid
Holo-NrdE	NrdE with one equivalent of tightly bound nucleotide
HPLC	High pressure liquid chromatography
IMAC	Immobilized metal affinity chromatography
IPTG	Isopropyl β-D-1-thiogalactopyranoside
Kan	Kanamycin
LB	Luria-Bertani medium
$Mn(III)_2$-Y•	Dimanganic-tyrosyl radical
MWCO	Molecular weight cut off
NADH	Reduced nicotinamide adenine dinucleotide
NADPH	Reduced nicotinamide adenine dinucleotide phosphate

NDP	Nucleoside 5′-diphosphate
Ni-NTA	Nickel nitrilotriacetic acid
NMR	Nuclear magnetic resonance
$NrdI_{ox}$	NrdI with oxidized FMN bound
$NrdI_{hq}$	NrdI with FMN hydroquinone bound
$NrdI_{sq}$	NrdI with FMN semiquinone bound
NTP	Nucleoside 5′-triphosphate
Nudix	Nucleoside diphosphate linked to moiety X
PEI	Polyethyleneimine
PMSF	Phenylmethylsulfonyl fluoride
R_S	Stokes radius
RT	Radical transfer
$s_{20,w}$	Sedimentation coefficient at standard state (20 °C, water)
$s_{<Wt>}[20,w]$	Weight-averaged sedimentation coefficient at standard state
SAXS	Small angle X-ray scattering
SEC	Size exclusion chromatography
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
S-site	Allosteric specificity site
Strep	Streptomycin sulfate
SV-AUC	Sedimentation velocity analytical ultracentrifugation
TCEP	Tris(carboxyethyl)phosphine
TEABC	Triethylammonium bicarbonate
TLC	Thin-layer chromatography
TMP	2′-deoxythymidine 5′-monophosphate
Tris	Tris(hydroxymethyl)aminomethane
TTP	2′-deoxythymidine 5′-triphosphate
UDP	Uridine 5′-triphosphate
Y•	Tyrosyl radical

Chapter 1

Introduction to Ribonucleotide Reductases

1.1. SYNOPSIS

Ribonucleotide reductases (RNRs) catalyze the reduction of the four common ribonucleoside 5′-di- or triphosphates (NDPs or NTPs) to the corresponding 2′-deoxyribonucleoside 5′-di- or triphosphates (dNDPs or dNTPs) and serve as the only *de novo* source of dNTPs for DNA replication and repair in all organisms (**Figure 1.1.**).[1] Furthermore, these enzymes contribute to the high fidelity of these processes by being subjected to multiple levels of regulation (transcriptional to post-translational) so that proper dNTP pool sizes and ratios are maintained. As a part of their post-translational level of control, RNRs have one of the most sophisticated systems of allosteric regulation observed in nature which allows one enzyme to reduce four different substrates in their appropriate proportions. Allosteric effects, on the molecular level, are the result of changes in the secondary, tertiary, and/or quaternary structure of a protein, and RNRs, because of their rich allostery, adopt a diversity of tertiary and quaternary structures, many of which have only been recently understood.[2-5] Despite being investigated for over 60 years, studies of the allosteric regulation of RNR continues to yield fascinating and surprising results, which supports the role of these enzymes as paradigms for understanding the links between allostery and enzyme quaternary structure. The work presented in this thesis continues this tradition by describing the first detailed characterization of the structure and regulation of a class Ib RNR, which has led to the discovery of the first case of overall activity allosteric regulation for this group of enzymes.

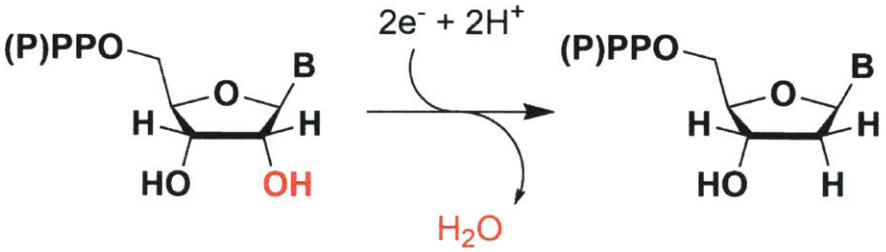


Figure 1.1. The reaction catalyzed by RNRs. **B** = Nucleobase: adenine, cytosine, guanine, and uracil.

1.1.1. Similarities and differences of class Ia and Ib RNRs. Class I RNRs are composed of two subunits termed α (typically 80 – 100 kDa) and β (80 – 90 kDa), that are proposed to form an $\alpha_2\beta_2$ complex that is active in NDP reduction. The class Ia RNRs (α = NrdA, β = NrdB), exemplified by the *Escherichia coli*, yeast, mouse, and human enzymes, are found in most eukaryotic organisms and in some strictly aerobic and facultative anaerobic bacteria.[6] These enzymes are the best characterized in terms of their chemical mechanisms, regulation, and structure. As such, the Ia RNRs (in particular, the *E. coli* enzyme) are often used as a benchmark against which all other class I enzymes are compared, and, therefore, will be frequently referred to throughout this thesis. The class Ib RNRs (α = NrdE, β = NrdF), which are the focus of this thesis, are found in most of the remaining strictly aerobic and facultative anaerobic bacteria with sequenced genomes.[6] The current prototype of the Ib RNRs is from *Salmonella enterica* Serovar Typhimurium (referred to as *S.* Typhimurium throughout the rest of this dissertation).

As described in **Section 1.2**, RNRs are divided into classes and subclasses on the basis of the type of metallocofactor employed by the enzyme.[7] The class Ia RNRs utilize a diferric-tyrosyl radical cofactor ($Fe(III)_2$-Y•) housed in β, which can spontaneously self-assemble from Fe^{2+} and O_2 in the presence of a one electron reductant.[8] The class Ib RNRs were long-thought to also employ $Fe(III)_2$-Y• as a result of the aforementioned cofactor self-assembly, which yielded NrdF preparations that were catalytically active.[9-11] However, conflicting reports in the literature suggested that, at least in the case of the *Corynebacterium ammoniagenes* enzyme (formerly *Brevibacterium ammoniagenes*), it was possible that a dimanganic-tyrosyl radical ($Mn(III)_2$-Y•) could be formed.[12-14] A major breakthrough explaining the discrepancy of these studies came in 2010 when it was demonstrated that $Mn(III)_2$-Y• could be reconstituted in *E. coli* NrdF *in vitro*.[15] The critical component that was absent in previous studies was found to be NrdI, a flavodoxin-like

protein conserved in the class Ib RNR operons in most sequenced bacterial genomes, which, in its two electron-reduced (hydroquinone) state ($NrdI_{hq}$), was responsible for generating an oxidant capable of oxidizing Mn^{2+}.[6, 15] Subsequent structural and kinetic studies showed that *E. coli* $NrdI_{hq}$ and $Mn(II)_2$-loaded NrdF formed a complex in which O_2 is reduced and then channeled to the dimanganous site of NrdF.[16] Current evidence suggests that $O_2^{\bullet-}$ is the oxidant generated by $NrdI_{hq}$, which reacts with the two Mn^{2+} ions to generate a putative $Mn^{4+}(\mu\text{-}O)(\mu\text{-}OH)Mn^{3+}$ intermediate capable of oxidizing tyrosine to a tyrosyl radical (Y•).[17] The physiological relevance of the cofactor was subsequently demonstrated by the isolation of endogenous $Mn(III)_2$-Y•-loaded NrdF from *E. coli*, *C. ammoniagenes*, and *Bacillus subtilis*.[18-21]

As a result of the long-held belief that class Ib RNRs utilized a $Fe(III)_2$-Y• cofactor and the structural homology of NrdE[22] and NrdF[23] to the class Ia subunits,[24-26] it has been assumed that the chemistry, allosteric regulation, and structure of the Ib RNRs is the same as the class Ia enzymes in all respects except one: overall activity regulation. The α subunits of class Ia RNRs have an N-terminal domain called the ATP-cone[27] to which ATP and dATP bind and allosterically regulate the overall activity of the enzyme; ATP binding stimulates NDP reduction whereas dATP binding inhibits it.[1] In contrast, NrdE does not possess an ATP-cone, which has manifested in the absence of overall activity regulation in the class Ib RNRs that have been studied up until now.[9, 11, 28-31] The belief that the Ib RNRs would behave the same as the Ia enzymes in all other regards was reinforced by only one study of the allosteric regulation of a Ib RNR, which was carried out with the Fe-loaded *S.* Typhimurium enzyme and is discussed further in Chapter 3.[9]

1.1.2. Why should the allosteric regulation of class Ib RNRs be reinvestigated? While the *S.* Typhimurium Ib RNR has provided some valuable insights into the structure and allosteric regulation of this subclass, in general it is not a good system to use as a prototype for the class Ib

enzymes. The reason this is true is that *S.* Typhimurium and related *Enterobacteriaceae*, such as *Escherichia*, *Klebsiella*, and *Yersina* species, are a unique family of bacteria in that they possess the genes for both class Ia and Ib RNRs.[6] Under standard laboratory growth conditions (37 °C in rich media such as Luria-Bertani (LB) broth), the Fe-dependent Ia enzyme in these bacteria is the primary RNR used in *de novo* deoxynucleotide synthesis for DNA replication and repair.[32, 33] In contrast, the genes for the Ib RNR are poorly expressed, and the enzyme, if it is even produced under these conditions, cannot support the growth of conditional mutants of the Ia enzyme.[32, 34] An increase in transcription of the Ib genes is observed during growth of *E. coli* and *S.* Typhimurium on minimal medium as well as during oxidative stress and severe iron limitation,[32, 33, 35] suggesting that the Mn-dependent Ib RNR may be important for cell viability under certain conditions. A scenario in which both iron limitation and oxidative stress are relevant is during tissue invasion and infection, and, therefore, having a Ib RNR might be seen as conferring an advantage to *S.* Typhimurium under such conditions. However, *in vitro* infection assays using tissue cultures of macrophages and epithelial cells revealed that the Ib RNR may help *S.* Typhimurium outlast the initial oxidative burst of the innate immune response, but was not essential for the pathogenesis of the bacterium.[33]

In contrast with *S.* Typhimurium, many other families of bacteria utilize a class Ib RNR as their sole aerobic source of dNTPs for DNA replication and repair.[6] Aside from the inability to be inhibited by dATP as noted above,[28-31] very little is known about the allosteric regulation and structure of these enzymes, and such studies with the $Mn(III)_2$-Y• cofactor are virtually non-existent.[28] As RNR regulation plays a major role in the maintenance of the proper amounts and ratios of the dNTP pools, it can be expected that these Ib RNRs might exhibit differences in their allosteric regulation, relative to the Ib enzymes of *S.* Typhimurium and other *Enterobacteriaceae*,

that give the cell greater control of the activity of these enzymes. Noteworthy, too, is the fact that Ib RNRs are utilized as the sole aerobic RNR by many human pathogens (e.g. *Enterococcus* species, *Mycobacterium tuberculosis*, *Staphylococcus aureus*, and *Streptococcus* species, **Table 1.1**) that the Center for Disease Control and Prevention have classified as concerning and serious threats to public health due to the increasing incidences of antimicrobial resistance.[36] The discovery of the $Mn(III)_2$-Y• cofactor and its associated biosynthetic pathway[15, 17] makes the class Ib RNRs in these pathogens an attractive new target for antimicrobial therapeutics. Gaining an understanding of any novelties in the regulation and structure of the Ib RNRs will further help in the development of potential antimicrobial compounds by broadening the scope of possible ways to disrupt and/or inactivate these enzymes.

The work presented in this dissertation lays foundations for understanding the allosteric regulation and structure of class Ib RNRs that serve as the aerobic source of dNTPs for a bacterium. The focus of these studies is the Ib RNR from *B. subtilis*, which was chosen over other bacteria for several reasons. First, the bacterium is the model organism for Gram-positive bacteria in general as well as for more specialized physiological processes such as sporulation,[37] competence development,[38] biofilm formation,[39] and bacterial cell differentiation.[40] As such, a wealth of information is available about the *B. subtilis* genome, transcriptome, proteome (e.g. *in vivo* protein concentrations as reported in Appendix 2), and metabolome (e.g. *in vivo* nucleotide concentrations as reported in Appendix 3), all of which can help make sense of the results presented in this thesis. Second, *B. subtilis* is genetically tractable, thus the *in vivo* significance of some of the results reported in this dissertation may be testable in the future using knockout or conditional mutants. Third, previous studies by Zhang *et al.* demonstrated that the subunits of this bacterium's RNR co-purified in an ~1:1 ratio, indicating that the subunit interaction was much "tighter" than in other

class I RNRs.[21] This observation suggested that the *B. subtilis* RNR might, therefore, yield to crystallographic characterization of the long-sought active $\alpha_2\beta_2$ complex. Last, but not least, *B. subtilis* NrdF is the first well characterized class Ib β subunit with both the $Fe(III)_2$-Y• and $Mn(III)_2$-Y• cofactors.[17, 21, 41, 42] As will be shown in the following chapters, studies of the allosteric regulation and structure of the *B. subtilis* Ib RNR have been fruitful as several new discoveries were made that will lead to a shift in thinking about the regulation of all class Ib RNRs.

1.2. A GENERAL INTRODUCTION TO RIBONUCLEOTIDE REDUCTASES

1.2.1. The classes of RNRs. In contrast with most other enzymes of primary metabolism, which use heterolytic bond-making and breaking mechanisms to complete the desired chemical transformations, RNRs utilize highly reactive, unpaired valence electrons, or radicals, as catalysts for the reduction of ribonucleotides via homolytic chemistry.[43, 44] In all RNRs, it has been demonstrated[45] or inferred[46] based on the structural homology of the catalytic sites of the large subunits (**Figure 1.2**) that the catalytic species that initiates the reaction is a cysteine thiyl radical transiently generated by the direct or indirect action of a metallocofactor. Three classes of RNRs have been defined based on the nature of the metallocofactor used to generate this transient radical during turnover (**Figure 1.3**).[44] The categorization also reflects the requirement and/or tolerance of O_2 for a particular class.

Class III RNRs utilize a captodatively-stabilized glycyl radical (G•) on the backbone of the large subunit (NrdD) to directly abstract the hydrogen atom from the conserved Cys residue each turnover.[44] The introduction of G• requires a second protein, NrdG, which uses a reduced $[4Fe4S]^+$ cluster to cleave S-adenosylmethionine into methionine and a putative 5′-deoxyadenosyl radical (5′-dA•) that abstracts a hydrogen atom from the conserved Gly residue on NrdD.[47] G• is stable, but extremely sensitive to O_2, thus class III RNRs are found only in strict or facultative anaerobic bacteria and archaea.[6] All characterized class III RNRs utilize NTPs as substrates.[48-52]

Table 1.1. Non-exhaustive list of bacteria pathogenic to humans and, when present, their main class I RNRs.

Genus	Species	Primary aerobic RNR	Other RNRs
Acinetobacter	*Acinetobacter baumannii*	Ia	None
Actinomyces	*Actinomyces israelii*[a]	Ib	None
Bacillus	*Bacillus anthracis*	Ib	III
	Bacillus cereus	Ib	III
Bacteroides	*Bacteroides fragilis*[a]	Ia	II, III
Bartonella	*Bartonella henselae*	Ib	None
	Bartonella quintana	Ib	None
Bordetella	*Bordetella pertussis*	Ia	None
Borrelia[b]	*Borrelia burgdorferi*	None	None
	Borrelia recurrentis	None	None
Brucella	*Brucella abortus*	Ib	II
	Brucella canis	Ib	II
	Brucella melitensis	Ib	II
	Brucella suis	Ib	II
Campylobacter	*Campylobacter jejuni*	Ia	None
Chlamydia	*Chlamydia pneumonaie*	Ic	None
	Chlamydia trachomatis	Ic	None
Chlamydophila	*Chlamydophila psittaci*	Ic	None
Clostridium	*Clostridium botulinum*	Ia	III
	Clostridium difficile	Ib (?)[c]	II, III
	Clostridium perfringens	Ia	III
	Clostridium tenani	None	II, III
Corynebacterium	*Corynebacterium diphtheriae*	Ib	III
Enterococcus	*Enterococcus faecalis*	Ib	III
	Enterococcus faecium	Ib	III
Escherichia	*Escherichia coli*	Ia	Ib, III
Ehrlichia	*Ehrlichia chaffeensis*	Ia	None
Francisella	*Francisella tularensis*	Ia	None
Haemophilus	*Haemophilus influenza*	Ia	III
Helicobacter	*Helicobacter pylori*	Ia	None
Klebsiella	*Klebsiella pneumoniae*	Ia	Ib, III

Legionella	*Legionella pneumophila*	Ia	None
Leptospira	*Leptospira interrogans*	None	II
	Leptospira santarosai	None	II
Listeria	*Listeria monocytogenes*	Ia	III
Mycobacterium	*Mycobacterium leprae*	Ib	None
	Mycobacterium tuberculosis	Ib	II (?)[d]
	Mycobacterium ulcerans	Ib	None
Mycoplasma	*Mycoplasma pneumoniae*	Ib	None
Neisseria	*Neisseria gonorrhoeae*	Ia	None
	Neisseria meningitidis	Ia	None
Nocardia	*Nocardia farcinia*	Ib	II
Pseudomonas	*Pseudomonas aeruginosa*	Ia	II, III
Rickettsia	*Rickettsia rickettsii*	Ia	None
	Rickettsia typhi	Ia	None
Salmonella	*Salmonella enterica*	Ia	Ib, III
Shigella	*Shigella sonnei*	Ia	Ib, III
Staphylococcus	*Staphylococcus aureus*	Ib	III
	Staphylococcus epidermidis	Ib	III
	Staphylococcus saprophyticus	Ib	III
Streptococcus	*Streptococcus agalactiae*	Ib	III
	Streptococcus pneumoniae	Ib	III
	Streptococcus pyogenes	Ib	III
	Streptococcus sanguinis	Ib	III
Treponema	*Treponema pallidum*	Ia	None
Ureaplasma[b]	*Ureaplasma urealyticum*	None	None
Vibrio	*Vibrio cholerae*	Ia	III
Yersinia	*Yersinia pestis*	Ia	Ib, III
	Yersinia enterocolitica	Ia	Ib, III
	Yersinia pseudotuberculosis	Ia	Ib, III

[a] These bacteria are typically commensal, but have been reported to be opportunistic pathogens.

[b] *Borrelia* and *Ureaplasma* species do not encode any RNRs and are thought to acquire (deoxy)nucleosides solely from the host.[53, 54]

[c] The class I genes in *Clostridium difficile* strains are annotated as *nrdE* and *nrdF*, but only in a few published genomes is there a *nrdI* gene. The class I RNR maybe a remnant since the organism uses a class II RNR.

[d] A putative *nrdZ* gene is encoded in the genome, but data are inconclusive on whether or not it is active.[55]

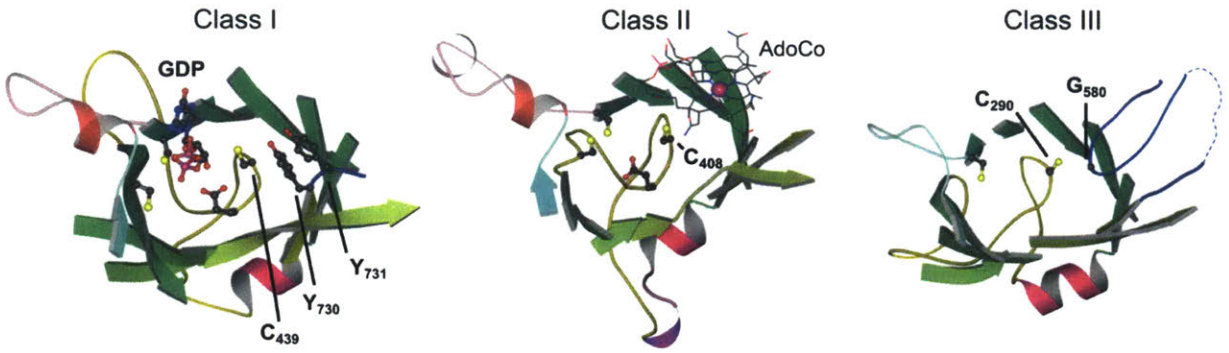


Figure 1.2. Comparison of the catalytic site structures from the three classes of RNRs. The structures are from the *E. coli* class Ia RNR (PDB 4R1R[24]), *Lactobacillus leichmannii* class II RNR (PDB 1L1L[56]), and the T4 phage class III (PDB 1H7B[57]). The conserved catalytic cysteine residue and the spatially conserved oxidants of this residue in the class I and III RNRs are indicated in ball and stick representation. Adenosylcobalamin (AdoCo) is shown as a stick model.

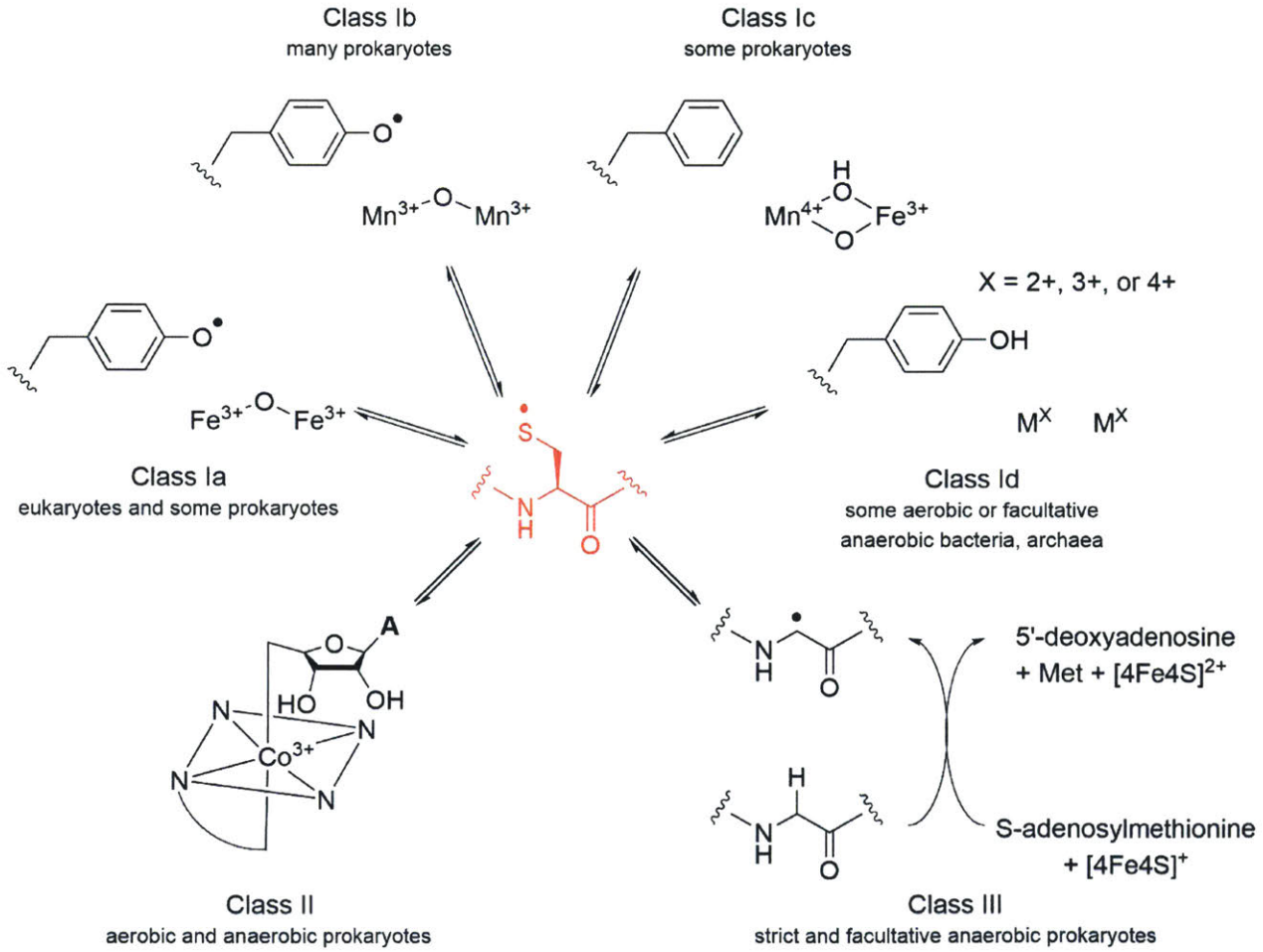


Figure 1.3. The different classes and subclasses of RNRs based on the metallocofactor used in the generation of the transient cysteine thiyl radical. The Id system is poorly characterized and preliminary results are unpublished (Boal, Kerbs, Bollinger labs, Pennsylvania State University). The physiological relevance of the Ic and Id clusters have not been established.

Class II RNRs (NrdJs) utilize adenosylcobalamin as the metallocofactor required for Cys oxidation. Homolytic cleavage of the Co – C bond generates a putative 5′-dA• that directly abstracts the hydrogen atom from the conserved Cys residue.[44] This cofactor is O_2 tolerant, thus class II RNRs are found in both aerobic and strict or facultative anaerobic bacteria and archaea.[6] The class II RNRs characterized to date have been found to reduce either NTPs (*Lactobacillus leichmannii*[58] and *Pseudomonas aeruginosa*[59]) or NDPs (*Thermotoga maritima* and *Thermoplasma acidophilum*).[60]

Class I RNRs, as described earlier in this chapter, utilize dinuclear metal clusters composed of Mn or Fe housed in β (NrdB = Ia, NrdF = Ib) to generate stable Y•s that oxidize the conserved Cys in α (NrdA = Ia, NrdE = Ib) each turnover.[7] Molecular oxygen, either as O_2, $O_2^{\bullet-}$, or H_2O_2,[7, 8, 17, 61] is required for cofactor assembly in these enzymes, thus class I RNRs are found in strict aerobic and facultative anaerobic organisms, including some bacteria and most eukaryotes.[6] All class I RNRs characterized to date reduce NDPs to the corresponding deoxynucleotides. The class I enzymes have been further categorized into four subclasses on the basis of the metal composition and the presence or absence of a Y• in the resting state (**Figure 1.3**). As described earlier, the class Ia and Ib RNRs can utilize $Fe(III)_2$-Y•, whereas $Mn(III)_2$-Y• can also be generated in the class Ib enzymes with the aid of NrdI. Mounting evidence suggests that $Mn(III)_2$-Y• is the physiologically relevant cofactor in the Ib RNRs.[18-21, 62] The focus of the work presented in this thesis is mainly on the class Ib RNR from *B. subtilis*, but the *E. coli* Ia RNR, being the paradigm of all class I systems, will also be discussed frequently.

The remaining two subclasses of the class I RNRs are not as well understood as the Ia and Ib systems and, thus, will be mentioned only briefly (**Figure 1.3**). The class Ic RNR from *Chlamydia trachomatis* is proposed to use a Mn(IV)Fe(III) cofactor as the Tyr found adjacent to

the metal cluster in other class I enzymes is replaced with phenylalanine.[7] The fourth group of class I RNRs, tentatively designated class Id, have only recently been discovered, with the RNR from *Flavobacterium johnsoniae* being the only member of this subclass that has been characterized at this point in time. A Tyr residue, sequentially and spatially homologous to those found adjacent to the metal binding sites in class Ia and Ib enzymes, is present, but preliminary studies indicate that the dinuclear metal cluster bypasses forming Y• and instead acts as the direct oxidant for thiyl radical formation (Boal, Bollinger, and Krebs laboratories, Pennsylvania State University). In both cases, the physiological relevance of these cofactors has not been established as the endogenous RNRs have not been purified from their hosts.

The genomes of many microoganisms encode more than one class of RNR, allowing for growth and proliferation in environments with differing oxygen tensions and resource availability, and in exceptional cases, such as in the opportunistic pathogen *P. aeruginosa*, genes for all three classes can be found.[6] All three classes of RNRs in *P. aeruginosa* are apparently active *in vivo*, and in a *Drosophila melanogaster* model system for pathogenesis, the transcription and translation of the class II and class III enzymes were found to be upregulated during tissue invasion and colonization.[59, 63] Furthermore, deletion of either of these two RNRs resulted in attenuated virulence of the *P. aeruginosa* mutants.[63] Noteworthy is the fact that many human pathogens encode multiple classess of RNRs (**Table 1.1**), suggesting that their presence can give these bacteria an advantage for infiltration and colony establishment during infection. Lastly, the *Enterobacteriaceae* are an aberrant group in that they encode two class I RNRs (Ia and Ib) as well as a class III (**Table 1.1**),[6] and as discussed previously, the Ib RNRs are only expressed under defined sets of conditions.[32, 35] In the case of *S.* Typhimurium, the class Ib RNR appears to not confer an apparent advantage to the organism during *in vitro* infection assays of macrophage and

epithelial cell tissue cultures,[33] suggesting that the enzyme may not be essential for pathogenesis. Further studies are required, including with *in vivo* models of pathogenesis, in order to definitively establish the essentiality and role of the Ib RNRs in the *Enterobacteriaceae*.

1.2.2. The sources of reducing equivalents. Given the structural homology of the large subunits noted earlier (**Figure 1.2**),[46] all RNRs utilize essentially the same chemical mechanism for ribonucleotide reduction, although small differences do occur depending on the source of reducing equivalents. In the vast majority of RNRs,[6, 51] the reducing equivalents for the reaction are supplied by two additional conserved Cys residues (C_{225} and C_{462} in *E. coli* α) located on the bottom face of the ribofuranose ring of the substrate in the catalytic site of the α subunit (**Figure 1.4**). After each turnover, the resulting disulfide bond formed in the catalytic site during the course of the reaction is reduced by two Cys residues in the C-terminal tail of α (C_{754}/C_{759} in *E. coli*). The tail is proposed to enter the catalytic site after product release to complete the reaction via a thiol/disulfide exchange reaction.[64] The C-terminal tail disulfide bond is subsequently reduced by an endogenous system that draws its electrons from the cellular NADPH pool.

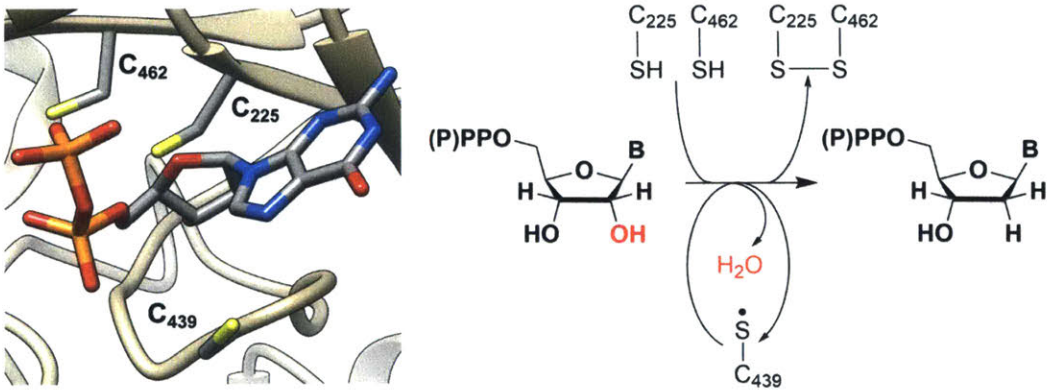


Figure 1.4. The most common source of reducing equivalents for ribonucleotide reduction catalyzed by RNRs is a pair of cysteine residues located in the catalytic site of α that sit on the bottom face of the ribofuranose ring of the substrate. Shown in the figure are the catalytic (C_{439}) and reducing cysteines (C_{225} and C_{463}) of the *E. coli* Ia RNR. **B** = nucleobase.

A part of characterizing RNRs is identifying the endogenous reducing system that supports dNTP production *in vivo*. In bacteria, the most commonly reported system involves thioredoxin (TR) and thioredoxin reductase (TRR) (**Figure 1.5A**), which was the first to be characterized[65] and is the most well studied due to its involvement in many other physiological processes.[66] In many prokaryotes, including nearly all those encoding a class Ib RNR, the requirement for TR can be substituted by another dithiol redoxin, NrdH, which receives its electrons from TRR and is often encoded in the same operon as the RNR genes (**Figure 1.5A**).[67, 68] Finally, in those organisms that can produce glutathione (GS), RNR can draw electrons from the pool of this small molecular weight thiol compound via the action of glutaredoxin (Grx) and glutathione reductase (GSR) (**Figure 1.5B**).[66, 69] Many prokaryotes encode multiple reducing systems that can service RNRs during turnover, making it difficult to determine which is likely to be the most physiologically relevant from genomic and transcriptomic analysis. Proteins like thioredoxin and thioredoxin reductase, however, are becoming effective targets of antimicrobial therapeutics,[70, 71] therefore increasing efforts are being made to identify the endogenous reducing systems of RNRs so that they can be directly targeted for inactivation, which, consequently, also disrupts dNTP metabolism in pathogens.[30, 42, 72] Studies to identify the endogenous reducing system for *B. subtilis* are described in Chapter 2.

In the remaining RNRs, all of which are class III enzymes, the reducing equivalents for ribonucleotide reduction are not supplied by oxidation of two spatially conserved Cys thiols to a disulfide bond since one Cys (equivalent to C_{462} in *E. coli* α) is missing.[6] For these RNRs, it has been found that oxidation of formate to CO_2 supplies the reducing equivalents for nucleotide reduction (**Figure 1.5C**), thus bypassing the requirement for disulfide re-reduction after each turnover.[73] Interestingly, this sub-type of class III RNR is found in all of the pathogens listed in

Table 1.1, suggesting that formate-utilizing RNRs might be advantageous for these bacteria during pathogenesis. Tissues at the sites of wounds or infection are reported to have low oxygen tension,[74] therefore the use of a formate-requiring RNR likely allows invading pathogens to persist at these anoxic sites during the initial stages of tissue invasion and colonization.

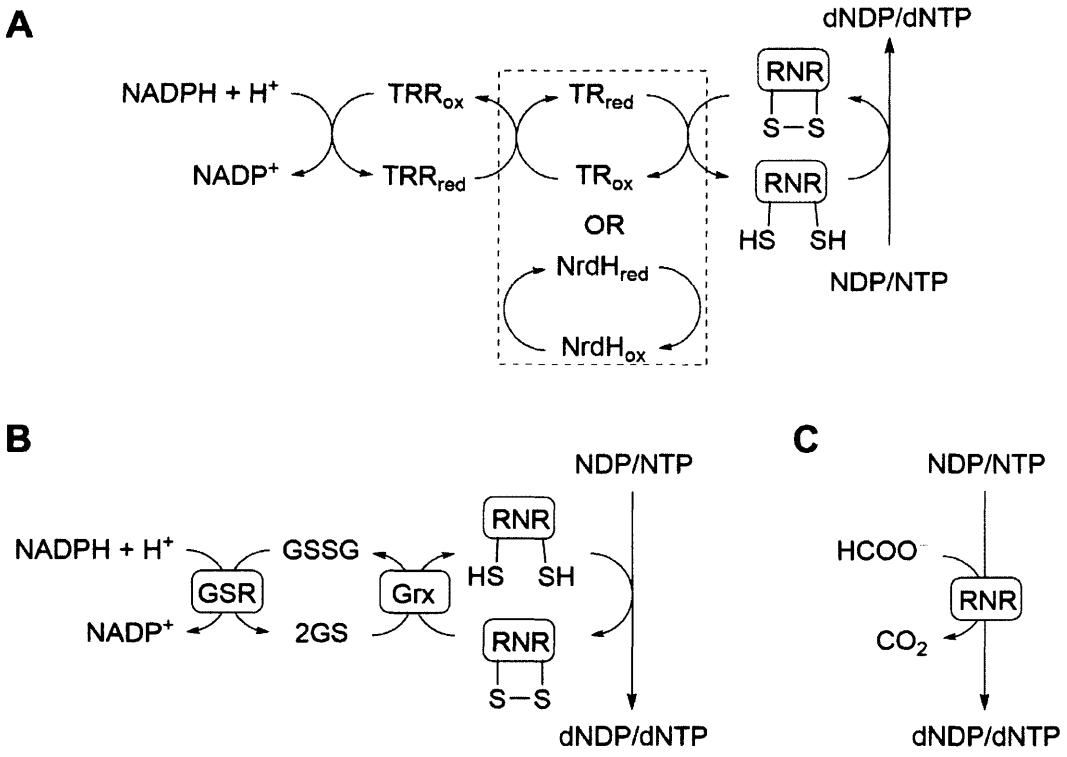


Figure 1.5. Endogenous reducing systems for RNRs. (**A**) The thioredoxin (TR)/thioredoxin reductase (TRR) system. In many prokaryotes, particularly those encoding a class Ib RNR, the requirement for TR can substituted with NrdH. (**B**) The glutaredoxin (Grx)/glutathione reductase (GSR) system. Grx supplies RNR with electrons at the expense of oxidizing two equivalents of glutathione (GS) to generate glutathione disulfide (GSSG), which is subsequently re-reduced by the action of GSR. (**C**) Formate oxidation supplies the reducing equivalents for some class III RNRs.

1.2.3. The mechanism of ribonucleotide reduction. As described above, the chemical reaction that occurs subsequent to Cys thiyl radical generation is similar for all classes. The mechanism of ribonucleotide reduction catalyzed by the *E. coli* class Ia enzyme has been the best characterized and, as mentioned above, serves as paradigm for other RNRs utilizing Cys residues

as the source of reducing equivalents (**Figure 1.6**). Thus, the residue numbering of this enzyme will be used as an example in the following brief synopsis of key points of the proposed mechanism. For a more detailed discussion on this topic, see the review by Licht and Stubbe.[43]

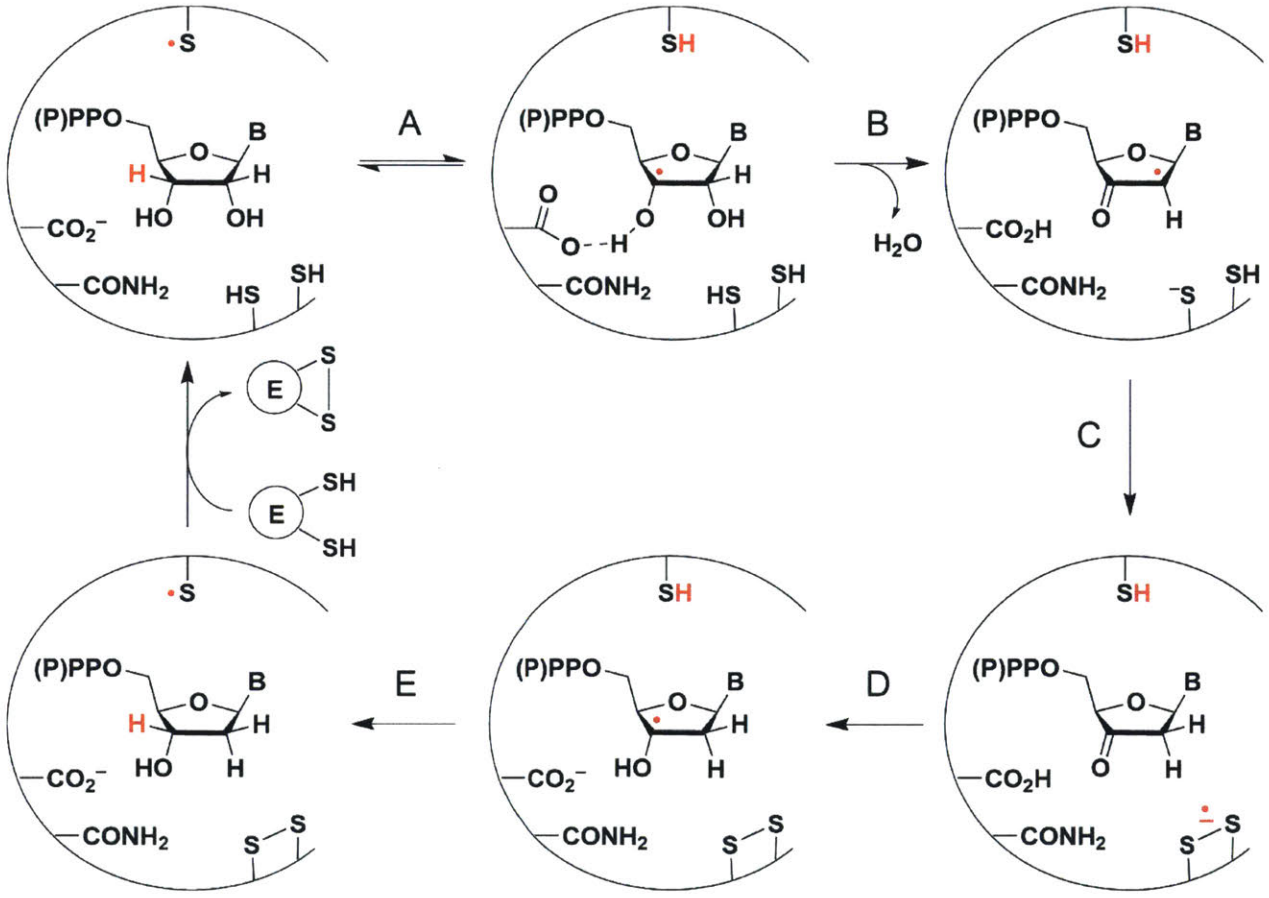


Figure 1.6. The proposed mechanism for ribonucleotide reduction by RNRs using cysteine residues as the source of reducing equivalents. **B** = nucleobase.

The Cys thiyl radical (C_{439}), upon generation, initiates chemistry on the substrate by abstracting the 3′-hydrogen atom from the sugar moiety (**Figure 1.6**, step A). This is followed by rapid and irreversible loss of the 2′-OH group as a water molecule, driven in part by deprotonation of the 3′-OH group by a conserved Glu (E_{441}), as the substrate-based radical rearranges to a 3′-keto-2′-deoxynucleotidyl radical (**Figure 1.6**, step B). Subsequent hydrogen atom abstraction

from C_{225} yields a 3′-keto-2′-deoxyribonucleotide intermediate and a protein-based disulfide radical anion between C_{225} and C_{462} (**Figure 1.6**, step C). Proton-coupled electron transfer (PCET) using protonated E_{441} is proposed to reduce the intermediate ketonucleotide to give the 2′-deoxy-3′-nucleotidyl radical (**Figure 1.6**, step D) where it abstracts the hydrogen atom from C_{439} (step E), thus regenerating the catalytic thiyl radical to yield the 2′-deoxyribonucleotide product. After each turnover, the radical is returned to its resting state (class I = $Me(III)_2$-Y• in β, class II = adenosylcobalamin, class III = G•) and reduction of the disulfide bond in the catalytic site resets the enzyme for another round of catalysis.

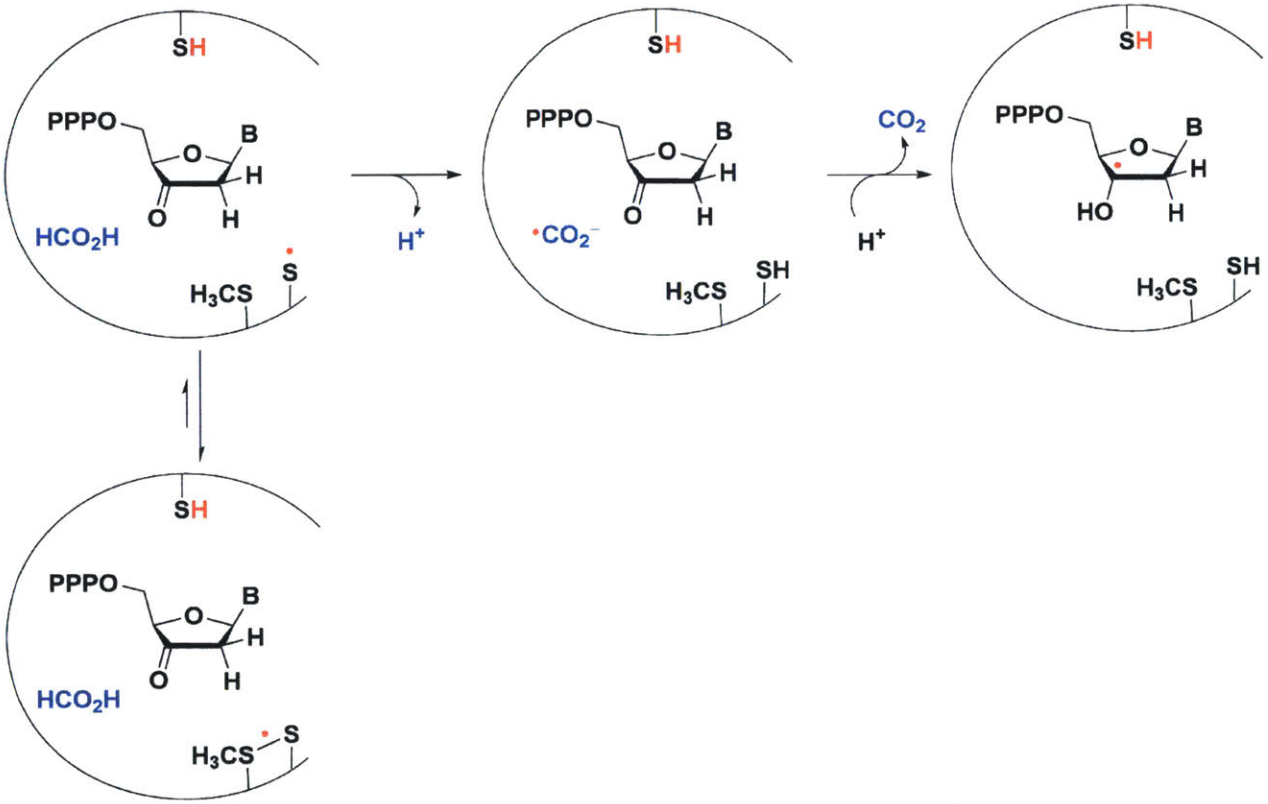


Figure 1.7. The proposed mechanism for class III RNRs that utilize formate as the source of reducing equivalents. The steps up to generation of the 3′-keto-2′-deoxynucleotidyl radical are thought to be the same as in the mechanism shown in **Figure 1.6**. **B** = nucleobase.

For the class III RNRs that use formate, the mechanism of nucleotide reduction is proposed to be the essentially the same as in **Figure 1.6** up through step C. The proton acceptor from the

3′-OH group is different in these enzymes because a residue spatially equivalent to E_{441} in the *E. coli* Ia RNR is not present. Instead it is proposed that formate is transiently protonated to formic acid by the 3′-OH to drive the formation of the 3′-keto-2′-deoxynucleotidyl radical, and then regenerated with loss of the proton to solvent. Hydrogen atom abstraction from the bottom-face Cys (**Figure 1.7**) generates the 3′-keto-2′-deoxynucleotide and a thiyl radical which, recently, was demonstrated to be able to form a thiosulfuranyl radical with an adjacent Met residue (**Figure 1.7**).[75] This species is very reminiscent of the disulfide radial anion observed with other RNRs (**Figure 1.6**) and, thus, provides further support for the involvement of sulfur-based three electron bonds in the mechanism of ribonucleotide reduction. The remainder of the reaction is not as well understood, but involves oxidation of formate to CO_2 and regeneration of the 2′-deoxy-3′-nucleotidyl radical (**Figure 1.7**), the latter of which abstracts the hydrogen atom from the top face Cys to yield the 2′-deoxynucleotide product.

1.3. GENERAL INTRODUCTION TO CLASS I RNR STRUCTURE

1.3.1. Characteristics of class I RNR subunits. Class I RNRs are assembled from two different proteins termed α (NrdA (Ia) or NrdE (Ib)) and β (NrdB (Ia) or NrdF (Ib)). α is a large protein of typically 85 – 100 kDa that houses the catalytic site for reduction of NDPs (ADP, CDP, GDP, UDP) and up to two classes of sites to which dNTPs (dATP, dGTP, TTP) and ATP can bind to allosterically effect the activity of the enzyme. The α subunits of Ia and Ib RNRs share very limited sequence homology (*E. coli* NrdA vs. *S.* Typhimurium NrdE: 24% identity, 41% with positive substitutions; *E. coli* NrdA vs. *B. subtilis* NrdE: 23% identity, 44% with positive substitutions), with only the most critical residues for activity (e.g. the catalytic and reducing Cys residues in the catalytic site) and regulation conserved between the two proteins. Despite this limited sequence similarity, the α subunits of Ia and Ib RNRs are structurally homologous (**Figure**

1.8). *E. coli* Ia and *S.* Typhimurium Ib α subunits are reported to exist as tightly associated dimers (α_2).[9, 76] Further characterization of the *E. coli* enzyme revealed the subunit dissociated into monomers upon removal of Mg^{2+} or if an unknown number of Cys residues on the protein were oxidized.[76] In other organisms, such as *P. aeruginosa* (Ia) and *Bacillus anthracis*/*Bacillus cereus* (Ib), α as-isolated is a monomer, but dimerizes in the presence of nucleotides and β_2.[28, 77]

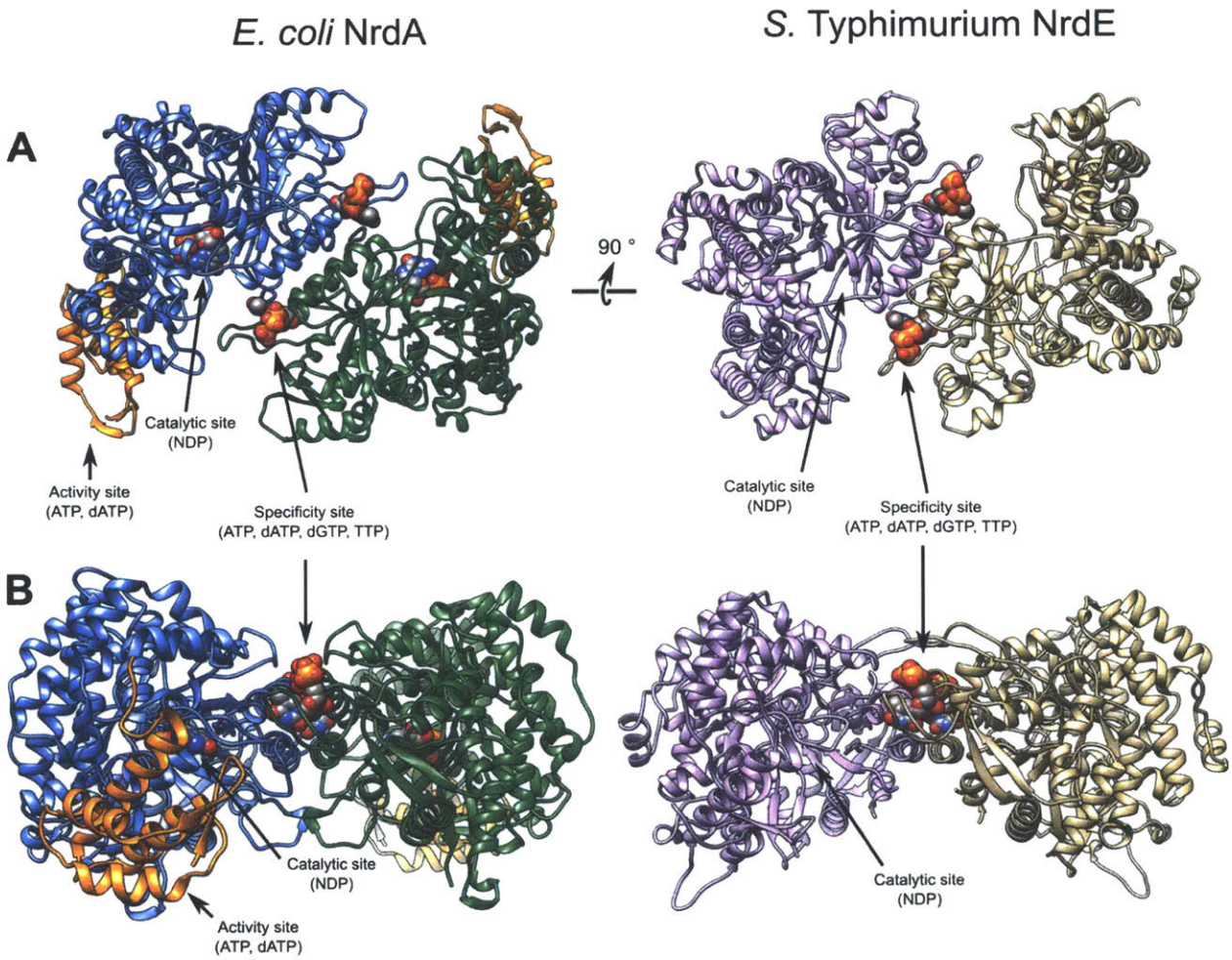


Figure 1.8. Crystal structures of the *E. coli* class Ia (left panels, PDB 4R1R[24], 3.2 Å) and *S.* Typhimurium Ib (right panels, PDB 1PEQ[22], 2.8 Å) α_2 subunits in two different views: (**A**) top and (**B**) front. To get from the top view to front view, the structure was rotated 90 ° about the axis indicated in (**A**). Each protomer of the dimer is colored differently. In the *E. coli* structure, the ATP-cone domain is colored orange. The effector TTP and substrate GDP are shown as space-filling models and colored according to element: gray = C, red = O, blue = N, orange = P.

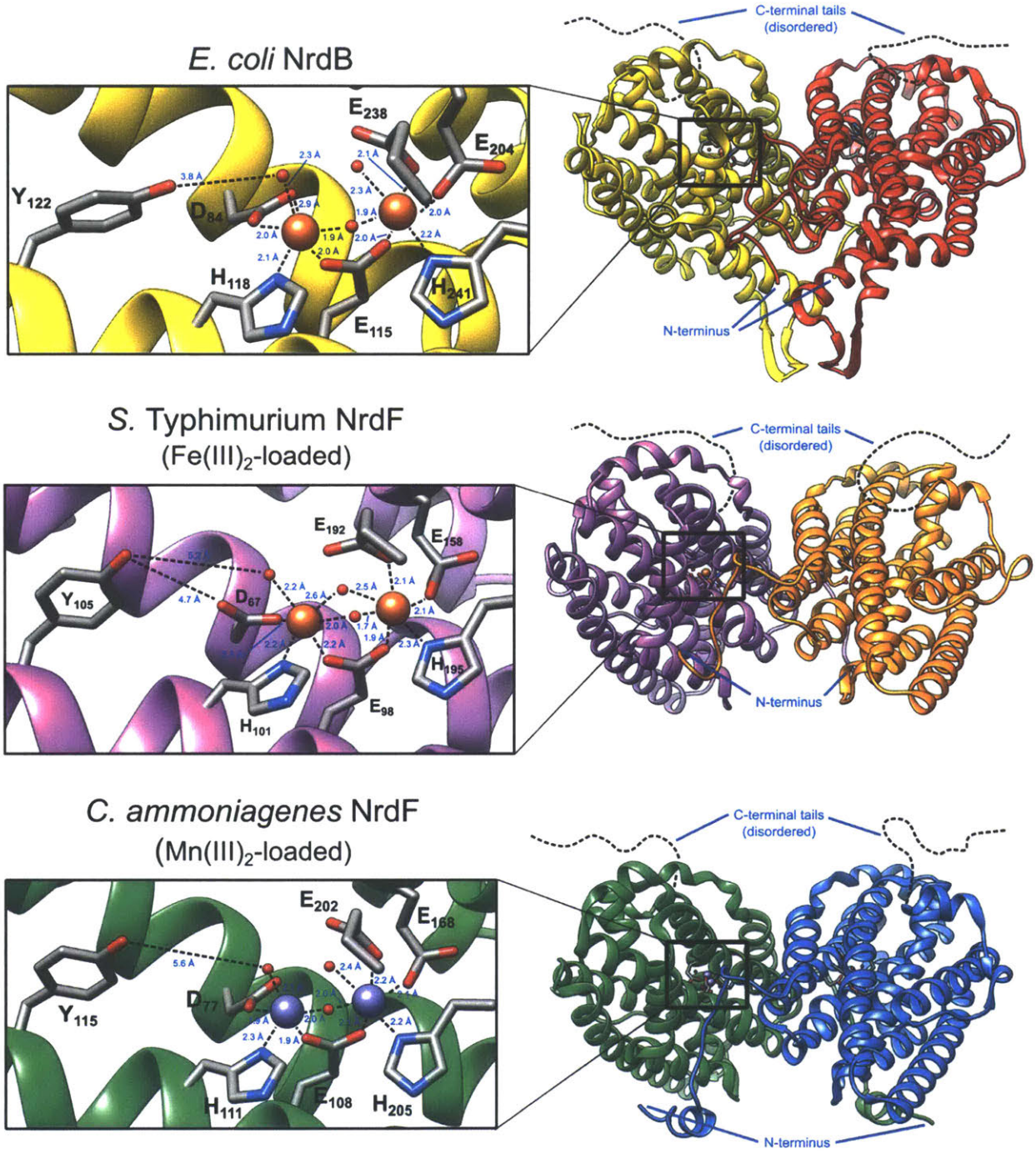


Figure 1.9. Crystal structures of *E. coli* NrdB (top, PDB 1MXR,[78] 1.4 Å), $Fe(III)_2$-loaded *S.* Typhimurium NrdF (middle, PDB 2R2F,[23] 2.3 Å), and $Mn(III)_2$-loaded *C. ammoniagenes* NrdF (bottom, PDB 3MJO,[19] 1.8 Å). Each protomer of the dimer is colored differently. Atoms are colored according to element: gray = C, red = O, blue = N, orange = Fe, purple = Mn. Distances between metal ions: *E. coli* NrdB = 3.3 Å, *S.* Typhimurium NrdF = 3.4 Å, and *C. ammoniagenes* NrdF = 3.3 Å.

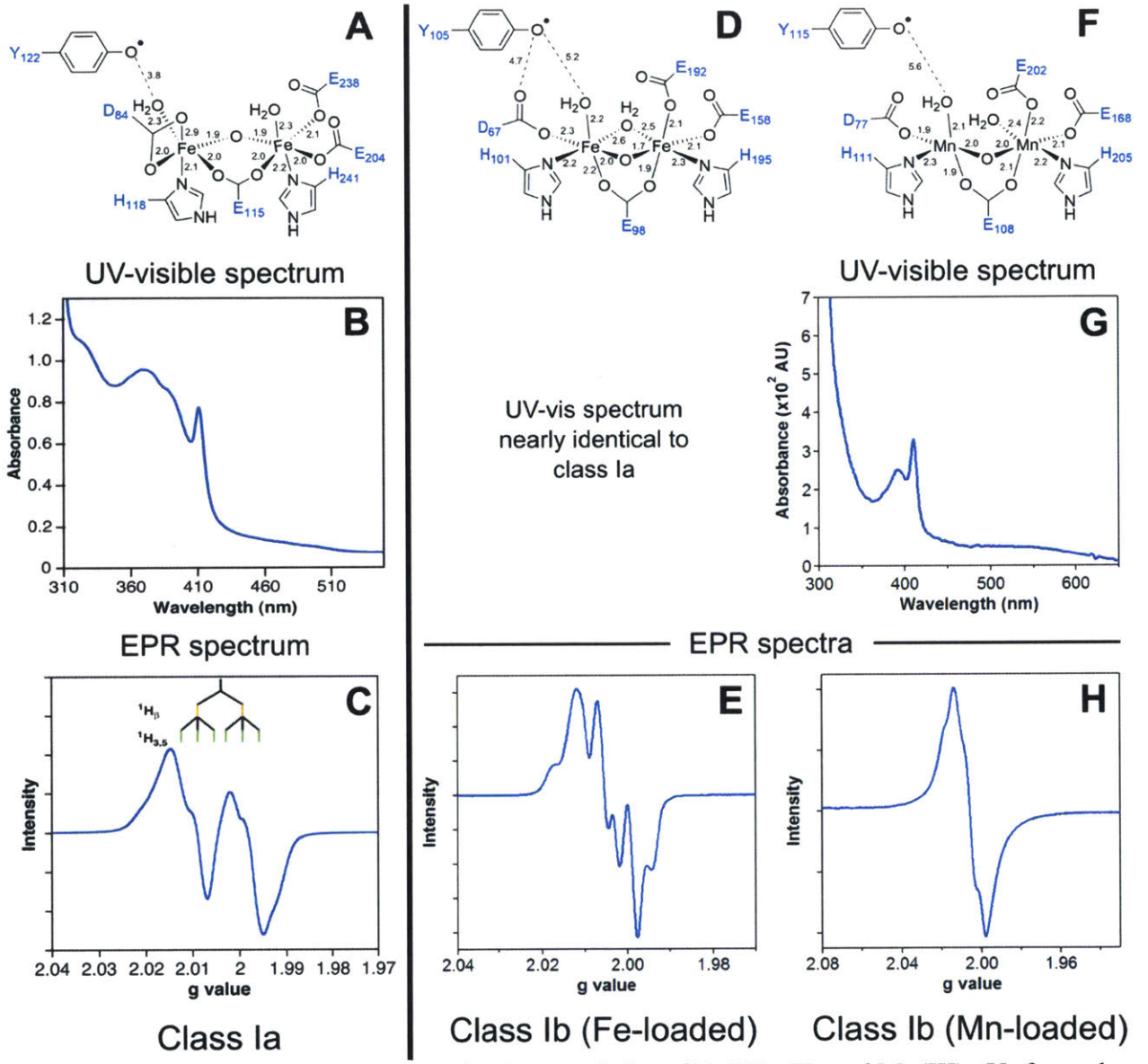


Figure 1.10. Structural and spectroscopic characteristics of $Fe(III)_2$-Y• and $Mn(III)_2$-Y• from class Ia and Ib RNRs. The numbers in the structures indicate distances in Å. (**A**) Structure of $Fe(III)_2$-Y• in the *E. coli* Ia β_2 subunit based on the X-ray crystal structure of the oxidized "met" (Y• reduced) form of the subunit (PDB 1MXR[78]). (**B**) The UV-visible spectrum of the cofactor in *E. coli* Ia β_2 has contributions from $Fe(III)_2$ (325 nm and 365 nm) and Y• (410 nm). (**C**) The X-band electron paramagnetic resonance (EPR) spectrum of Y_{122}• in *E. coli* Ia β_2 at 77K. (**D**) Structure of $Fe(III)_2$-Y• in the *S.* Typhimurium Ib β_2 subunit based on the "met" X-ray crystal structure (PDB 2R2F[23]). Note that the UV-visible spectrum of $Fe(III)_2$-Y• in class Ib RNRs looks identical to that observed in Ia enzymes. (**E**) X-band EPR spectrum of $Fe(III)_2$-Y• in *B. subtilis* Ib β_2 at 77K. (**F**) Structure of $Mn(III)_2$-Y• in the *C. ammoniagenes* Ib β_2 subunit based on PDB 3MJO.[19] (**G**) The UV-visible spectrum of $Mn(III)_2$-Y• from *B. subtilis* Ib β_2 shows absorption features from Y• (390 nm and 410 nm) and $Mn(III)_2$ (broad feature from 450 – 650 nm[15]). (**H**) X-band EPR spectrum of $Mn(III)_2$-Y• in *B. subtilis* Ib β_2 at 77K.

The allosteric specificity sites are located at the α_2 interface[22, 24] (**Figure 1.8**) and, as the name implies, control the specificity of the enzyme for its four different substrates. As discussed earlier in the chapter, class Ia RNRs have an additional domain on the N-terminus of their α subunits called the ATP-cone,[27] which houses a second allosteric regulatory site, the overall activity site (**Figure 1.8**). Generally, ATP (stimulatory) and dATP (inhibitory) bind to this site to control the overall activity of the enzyme. This domain is absent from NrdE, which makes it smaller relative to NrdAs (80 – 83 kDa versus ≥85 kDa, respectively) and, in all Ib RNRs characterized up to this point in time, results in no overall activity regulation.[9, 11, 29-31] The results presented in Chapters 2 and 3 demonstrate that the *B. subtilis* Ib RNR is the first enzyme of the Ib RNRs to exhibit overall activity regulation in the form of inhibition by dATP.

The β subunit of Ia and Ib RNRs is an obligate dimer (β_2) of ~80 – 90 kDa (**Figure 1.9**) and, as previously mentioned, houses the $Me(III)_2$-Y• (Me = Fe or Mn) cofactor that is essential for catalysis. Like the α subunit, very little sequence homology exists between NrdBs and NrdFs despite both proteins adopting a similar tertiary structure (*E. coli* NrdB vs. *B. subtilis* NrdF: 21% identity, 44% with positive substitutions; *E. coli* NrdB vs. *S.* Typhimurium NrdF: 26% identity, 42% with positive substitutions). All class I βs supply the same ligand set for coordinating the two metal ions, consisting of the side chains of two His, three Glu, and one Asp residue (**Figure 1.10A**, **D**, and **F**). The $Fe(III)_2$-Y• and $Mn(III)_2$-Y• cofactors in both Ia and Ib RNRs exhibit spectroscopic signatures that are distinctly characteristic of the subclass and metal cluster (**Figure 1.10**). By UV-visible spectrophotometry, Y• has characteristic absorption features at 410 nm (sharp peak), 390 nm, and ~600 nm (broad feature, not shown), of which the former two can be seen clearly in the spectrum of Mn-loaded β_2 (**Figure 1.10B** and **G**).[8, 79] The diferric clusters in Ia and Fe-loaded Ib β_2 have characteristic features at 325 nm and 365 nm (**Figure 1.10B**) whereas

broad features spanning 500 – 650 nm (**Figure 1.10G**) in spectra of Mn-loaded β_2 have been attributed to the dimanganic cluster based on similarity to the features observed in spectra of oxidized Mn catalases.[15, 79]

At 77 K, the electron paramagnetic resonance (EPR) signal arises from spin delocalization on the phenolic oxygen; carbons 1, 3, and 5 of the aromatic ring; and the β carbon of the Tyr residue.[80, 81] As a result of this delocalization, hyperfine splitting of the signal results from coupling of the radical to the β-methylene protons, which contribute to the majority of the splitting, and the protons on C3 and C5.[80] The spectra shown in **Figure 1.10C**, **E**, and **H** reveal striking differences between the Ia, Fe-loaded Ib, and Mn-loaded Ib β subunits. These differences primarily arise due to two sources. (1) Depending on the dihedral angles between each β-methylene proton and the p_z-orbital on C1 of the aromatic ring, coupling of the radical to one proton may be stronger than the other, such as is the case in the *E coli* Ia β subunit.[80] (2) The relaxation properties of the radicals will be different depending on the distance and any hydrogen bonding interactions between Y• and the metal cluster.

1.3.2. Characteristics of class I RNR $\alpha_2\beta_2$ complexes. During turnover, the two subunits of class I RNRs associate into the catalytically active form of the enzyme for which mounting evidence strongly indicates is $\alpha_2\beta_2$ in composition. The equivalent stoichiometry of α and β was first suggested in studies by Brown and Reichard using sucrose gradient ultracentrifugation to monitor the sedimentation of mixtures of the *E. coli* Ia RNR subunits alone and in the presence of effector nucleotides (dATP, dGTP, TTP).[82] In all cases except for the experiment run with dATP (discussed later), a species with a standard state sedimentation coefficient ($s_{20,w}$) equal to ~9.7 S was detected, which is distinctly different from the sedimentation of the individual subunits (α_2 = 7.8 S, β_2 = 5.5 S).[83] Quantitation of the subunit composition of the 9.7 S complex using a

combination of UV-visible spectrophotometry and radiolabeling of β_2 with ^{59}Fe revealed that both were present in an approximately 1:1 molar ratio. Subsequent biophysical experiments performed by Thelander[76] revealed that both subunits of the *E. coli* Ia RNR were dimers, thus leading to the proposal that the active form of the enzyme was $\alpha_2\beta_2$.

Formation of the active RNR complex is mostly dependent on the last 20 amino acid residues of the C-terminal tail of β. This first became evident in glycerol gradient ultracentrifugation studies performed by Sjöberg and coworkers,[84] who demonstrated that mixtures of *E. coli* Ia RNR β_2 subunit truncated by 30 residues at the C-terminus ($V_{346} - L_{375}$) and wildtype α_2 sedimented as discrete species under the same conditions in which wildtype subunits associated into a complex as reported previously by Brown and Reichard.[82] Subsequently it was shown that synthetic peptides corresponding to the C-terminal tail of β could competitively inhibit the *E. coli* RNR, with the last 20 residues ($Y_{356} - L_{375}$) appearing to be the most critical component for binding to α.[85] This limited contact between α and β may explain the overall weak subunit affinity of the *E. coli* Ia RNR under steady-state conditions, with a K_d of ~0.4 μM in the absence of nucleotides[86] and ~0.2 μM in the presence of ATP and CDP.[85] In general, the addition of substrate and effector nucleotides tightens the interaction between the subunits of class Ia RNRs, as has been reported for the *E. coli*,[87] *P. aeruginosa*,[88] and mouse[89] enzymes. The interaction, however, is still insufficiently weak to allow for structural characterization of the complexes.

As a result of the weak interactions between α_2 and β_2, an atomic structure of the active form of a class I RNR does not currently exist. However, the subunits individually have been characterized extensively by X-ray crystallography in Ia and Ib systems, and in the case of the *E.coli* Ia RNR, the subunit structures allowed Uhlin and Eklund to propose a docking model for the active $\alpha_2\beta_2$ complex based on shape and charge complementarity (**Figure 1.11**).[26]

Hydrodynamically, the model is consistent with the $s_{20,w}$ value of 9.7 S reported by Brown and Reichard,[82] and has received strong support by further biophysical,[2, 90] pulsed electron-electron double resonance (PELDOR) spectroscopy,[91-93] and cryo-electron microscopy[94] studies.

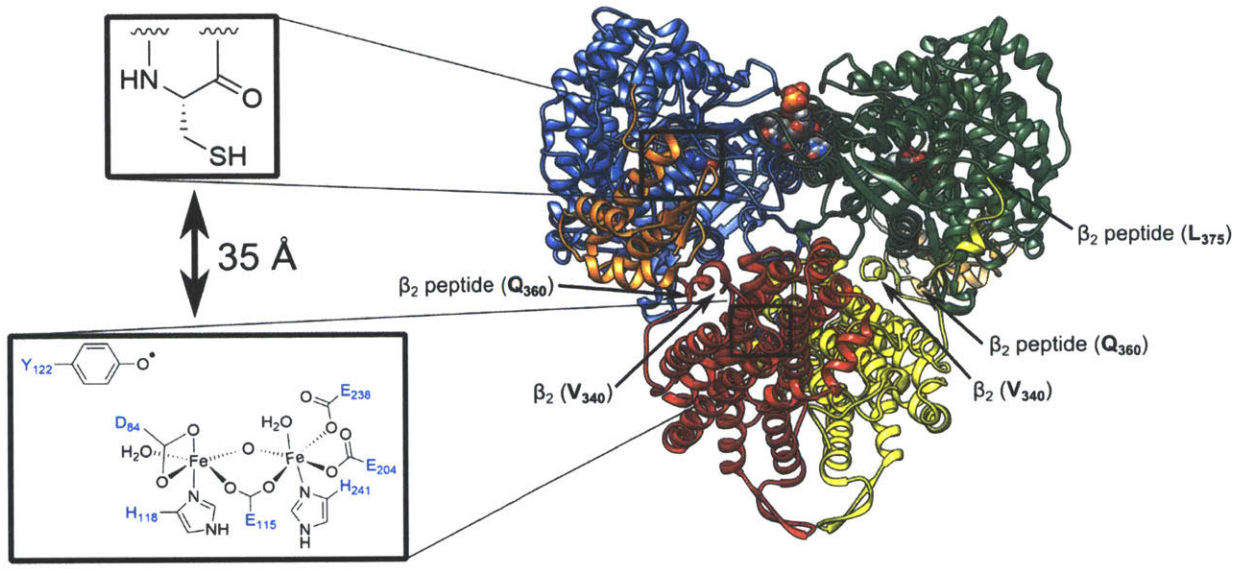


Figure 1.11. Docking model of the *E. coli* class Ia $\alpha_2\beta_2$ complex proposed by Uhlin and Eklund.[26] The α subunits (PDB 4R1R[24]) are colored green and blue and the β subunits (PDB 1MXR[78]) red and yellow. The ATP-cone domains on α are colored orange. TTP and GDP are shown as space filling models and colored according to element. α was crystallized with a peptide corresponding to residues 360 – 375 of β. The peptide shown in the foreground is colored yellow.

The docking model of the *E. coli* Ia RNR (**Figure 1.11**) has several noteworthy features. In all structures of β_2 solved to date, the final 30+ residues of the C-terminal tails are disordered due to high flexibility of this region. This flexibility is also a likely contributor to the relatively weak affinities observed with most RNRs. The C-terminal tails of α_2, like β, are also flexible and disordered, thus the location of the two Cys residues that reduce the catalytic site disulfide bond after each turnover[64] is unknown in the active complex. The binding site on *E. coli* α for the β C-terminus (**Figure 1.11** and **12**) was initially inferred from co-crystallization of the protein with

peptides corresponding to the last 20 residues, $Y_{356} - L_{375}$, although only $Q_{360} - L_{375}$ were visible in the structure.[24, 26] More recently, $Q_{360} - L_{375}$ were bound to the same place on α in structures in which both subunits co-crystallized as dATP-inhibited $\alpha_4\beta_4$ complexes (**Figure 1.12**),[2, 5, 95] thus confirming the site as the actual β binding site.

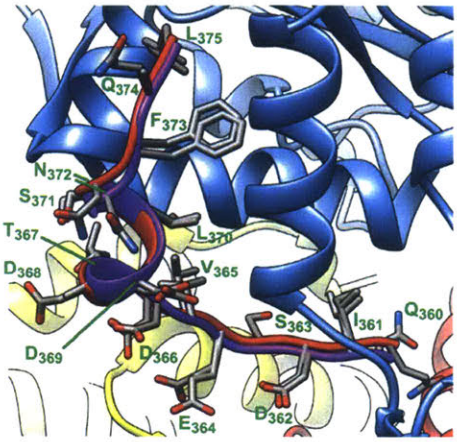


Figure 1.12. Comparison of the position of residues $Q_{360} - L_{375}$ of the β C-terminal *peptide* (red) used in the early crystal structure studies of the *E. coli* Ia α subunit (PDB 4R1R),[24, 26] and the β C-terminal *tail* (purple) observed in the recent structures of the $\alpha_4\beta_4$ dATP-inhibited complex (PDB 5CNV).[5]

The hypothesis that an $\alpha_2\beta_2$ complex is the active form of class I RNRs has been extended to include all subclasses as well. For the Ib RNRs, the strongest support for this hypothesis was provided using sucrose gradient ultracentrifugation experiments with the Fe-loaded *S.* Typhimurium Ib RNR.[9] These experiments showed a species with $s_{20,w}$ = 9.6 S formed independent of nucleotides when NrdE and NrdF were mixed together in approximately a 1:1 ratio (~2 nmol of each subunit). The $s_{20,w}$ value is similar to that reported for the *E. coli* Ia RNR,[82] and with the predicted $s_{20,w}$ (10.6 S) from a docking model of the $\alpha_2\beta_2$ complex prepared by superimposing the published structures of *S.* Typhimurium NrdE (PDB 1PEM[22]) and NrdF (PDB 2R2F[23]) onto the *E. coli* RNR docking model (**Figure 1.11**). Further evidence supporting $\alpha_2\beta_2$ as the active form of Ib RNRs was provided by gel electrophoretic mobility mass analysis (GEMMA)

experiments with the *B. anthracis*/*B. cereus* enzymes,[28] which detected complexes with molecular weights (202 – 212 kDa) that were consistent with those predicted for the $\alpha_2\beta_2$ form of the enzyme (234 kDa).

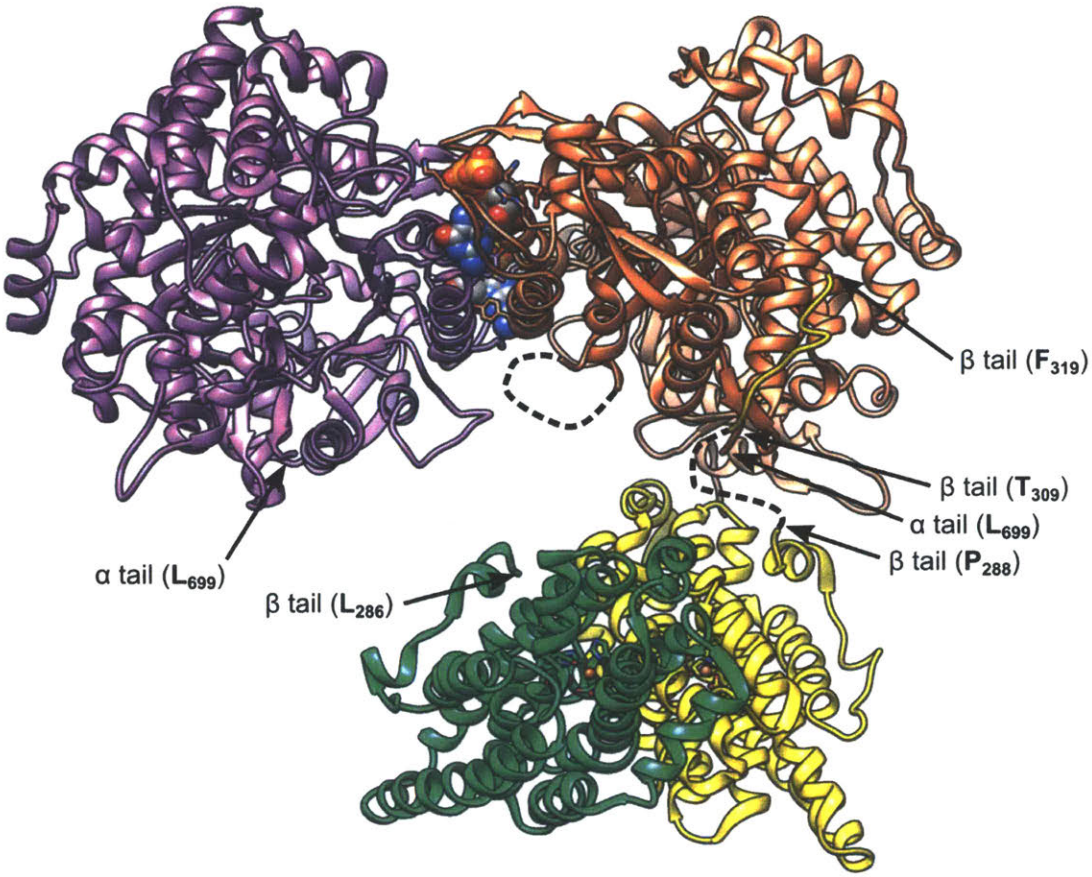


Figure 1.13. Crystal structure of the Fe-loaded *S.* Typhimurium Ib RNR holo complex (PDB 2BQ1, 4 Å).[96] The NrdE protomers are colored purple and light orange, and the NrdF protomers green and yellow. The specificity sites are occupied by dGTP. The dashed lines indicate disordered residues in the crystal structure.

The ultracentrifugation results with the *S.* Typhimurium Ib RNR complex were distinct from those for the *E. coli* Ia RNR[82] in that it was reported that the subunits of the former system did not readily dissociate during centrifugation,[9] even in the absence of nucleotides. This interesting observation suggested the interaction of the *S.* Typhimurium RNR subunits was tighter

and, therefore, might be amenable to crystallographic characterization. Efforts to crystalize the *S.* Typhimurium holo-RNR complex resulted in the first and only published structure of a class I RNR $\alpha_2\beta_2$ complex, albeit in an inactive state (**Figure 1.13**).[96] The surfaces of α_2 and β_2 predicted to interact with each other in the docking model (**Figure 1.11**) are facing each other in this elongated structure, and the authors of the report suggested that it represented an intermediate state in the formation of the active $\alpha_2\beta_2$ complex from the individual subunits. The last 11 residues of one NrdF protomer were found to be bound to NrdE in a hydrophobic cleft (**Figure 1.13**). Superimposition of this NrdE protomer with the structure of *E. coli* NrdA bound to the respective β C-terminal peptide reveals that the binding sites are similarly located on both proteins (**Figure 1.14**), although the interactions are different due to the sequence differences between *E. coli* NrdB and *S.* Typhimurium NrdF. The structure of the *S.* Typhimurium complex highlights the challenge of obtaining a complete atomic model of the active $\alpha_2\beta_2$ complex, in part due to the flexibility of the C-terminal tails of β and the weak interactions of the subunits in the resting state.

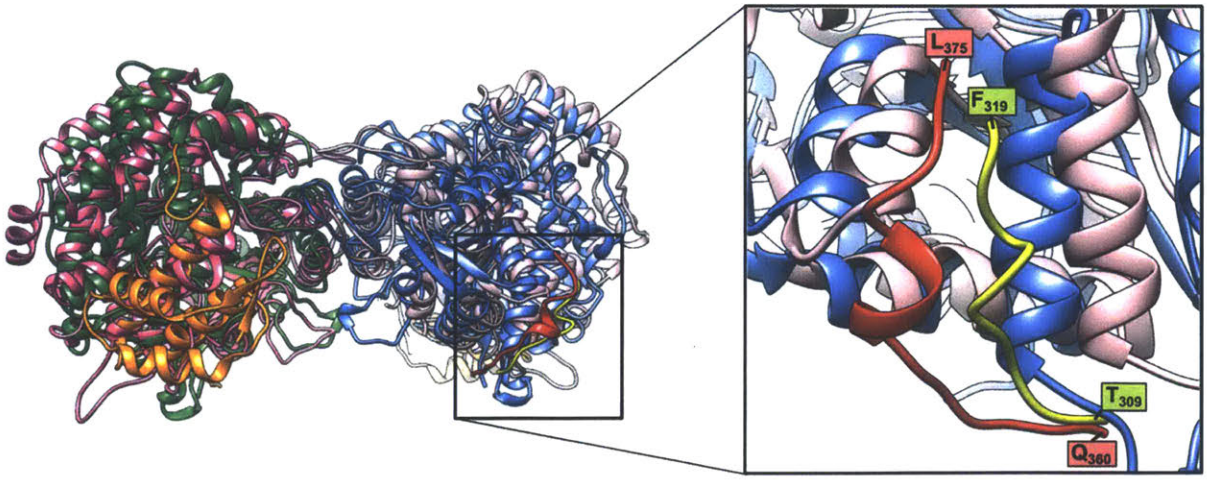


Figure 1.14. Comparison of the binding sites for the β subunit C-terminus on *E. coli* NrdA (PDB 4R1R[24], red) and *S.* Typhimurium NrdE (PDB 2BQ1[96], yellow). *E. coli* α_2 is colored as described in the legend of **Figure 1.9**, whereas the two protomers of *S.* Typhimurium NrdE are colored different shades of pink.

1.3.3. The radical transfer pathway of the E. coli class Ia RNR. One of the most noteworthy features observed in the *E. coli* Ia docking model is that the Cys residue (C_{439}) that is transiently oxidized to a thiyl radical in the catalytic site of α and Y_{122}, the site of the stable Y•, in β are over 35 Å apart (**Figure 1.11**). The measured turnover numbers of this enzyme ($2 - 10\ s^{-1}$)[97] and the expected rate constant for electron transfer over this distance via a tunneling mechanism ($10^{-7} - 10^{-9}\ s^{-1}$)[98] indicate that Y_{122}• cannot directly oxidize C_{439} in a single step. Therefore, additional redox-active intermediates are required in order to accelerate electron transfer rates to levels equal to or exceeding the turnover numbers of the enzyme. Based on the docking model and sequence alignments of class Ia RNRs, Uhlin and Eklund proposed that radical transfer (RT) from Y_{122}• in β to C_{439} in α involved a pathway of conserved aromatic amino acids linking the two sites in the $\alpha_2\beta_2$ complex;[26] these residues are W_{48} and Y_{356} in β and Y_{730} and Y_{731} in α. Under physiological conditions, Tyr and Trp oxidation at neutral pH requires the loss of both a proton and an electron, and in the case of RNR, these two processes occur concomitantly to avoid the buildup of high energy intermediates, thus implicating multiple PCET steps as the mechanism of RT between the subunits.[98] The current manifestation of the Stubbe-Nocera PCET model for RT in the *E. coli* class Ia RNR is shown in **Figure 1.15**.

The importance of the three pathway Tyr residues in RNR function was first suggested using site-directed mutagenesis.[99, 100] Inactive mutants, however, cannot definitively establish that the residues are transiently oxidized to Y• during turnover nor distinguish the mechanistic possibilities for radical transfer between the residues. Efforts by many groups to directly investigate the pathway revealed that the spectroscopic signatures of Y_{122}• (**Figure 1.10B** and **C**) did not change under single and multiple turnover conditions.[97, 100] These results indicated that nucleotide reduction and RT were kinetically masked by a protein conformational change that

gated radical injection onto the pathway. However, with improvements in expressed protein ligation technology and the advent of the stop codon suppression method, it became possible to incorporate unnatural Tyr analogs with different pK_as and reduction potentials site-specifically into pathway positions in RNR.[98] Utilization of these technologies in conjunction with rapid kinetic and spectroscopic techniques have allowed the protein conformational gate to be bypassed, radical intermediates to be trapped on pathway, and have demonstrated the reversibility of radical transfer. These studies and the use of photo-RNRs have substantially increased the understanding of mechanism of RT in the *E. coli* Ia RNR.[98, 101-105] The involvement of W_{48} in the pathway, however, has been difficult to assess due to its requirement for assembly of $Fe(III)_2$-Y•,[106, 107] and extensive work by former members of the Stubbe group indicates the residue likely does not participate directly in radical transfer as a transiently oxidized species (Yokoyama, Stubbe, unpublished).[108]

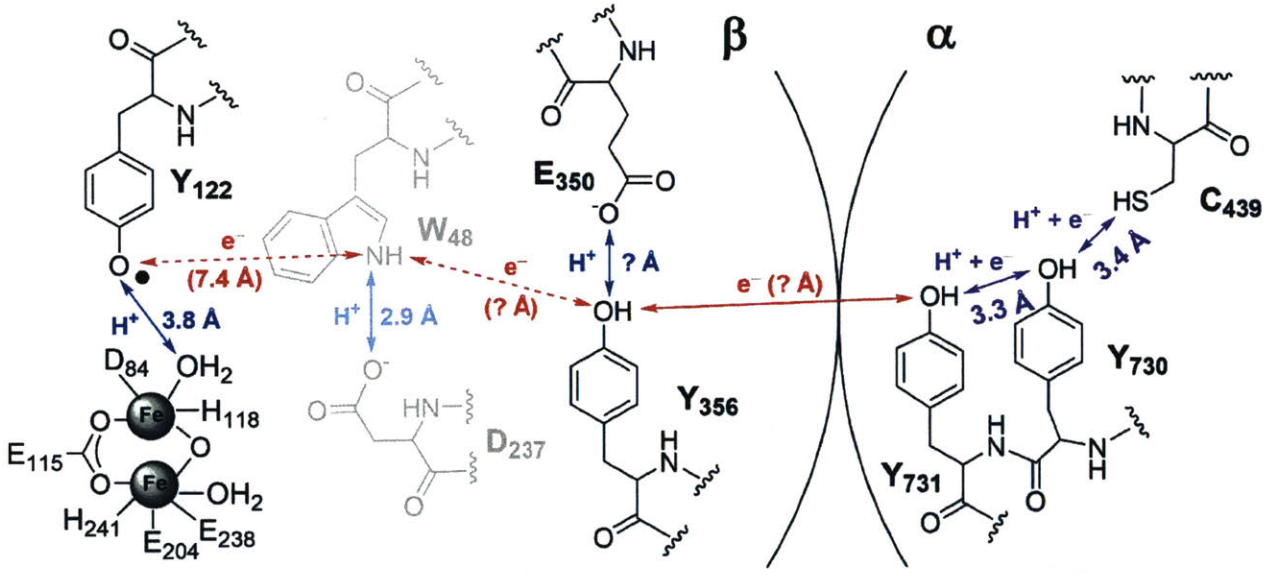


Figure 1.15. Proposed PCET mechanism for radical transfer in the *E. coli* class Ia RNR. Electron movement is shown as red arrows, proton movement as blue arrows, and co-linear movement of both proton and electron as purple arrows. Y_{356} and its putative proton acceptor E_{350} are located on the flexible C-terminal tail of β and are disordered in all published structures. W_{48} and its putative proton acceptor D_{237} are shown in gray as there currently is no evidence supporting their involvement in the radical transfer mechanism.

An extremely important observation to come out of the studies of PCET in the *E. coli* Ia RNR is that the subunit interaction strengthens dramatically when the radical is on the pathway.[94] This was demonstrated by inserting 3-aminotyrosine (NH_2Y) at position 730 in α. NH_2Y is easier to oxidize than Y and, therefore, serves as a radical "trap" on the RT pathway. A K_d of 7 nM was measured for the interaction of wildtype β_2 with the α_2-NH_2Y_{730} mutant using a competitive inhibition assay, which is ~25-fold tighter than was measured for the wildtype subunits using the same technique (0.18 μM).[85] Furthermore, stopped-flow fluorescence spectroscopy using α_2-NH_2Y_{730} and β_2-$V_{365}C$ labeled with an environmentally-sensitive fluorophore[86] revealed the dissociation rate constant was on the order of 10^{-3} s^{-1}, which is 10^4 slower than observed for the wildtype subunits. This amazingly tight complex was analyzed by cryo-electron microscopy and revealed a compact structure similar to the active $\alpha_2\beta_2$ complex predicted by the docking model (**Figure 1.11**). The fact that a difference in positon of one hydrogen atom can cause such a tight complex to form demonstrates the level of control that nature has evolved in order to harness the power of free radicals for difficult chemical reactions.

Sequence alignments of class Ib, Ic, and Id RNRs reveals conservation of the pathway residues in their respective α and β subunits, thus indicating that a PCET mechanism for RT is likely operative in these systems as well. However, the only other class I RNR in which the pathway has been studied in any detail thus far is the class Ic RNR from *C. trachomatis*.[109-111] Future studies of the radical propagation pathway in the class Ib RNRs would likely be the most informative since both $Fe(III)_2$-Y• and $Mn(III)_2$-Y• can be assembled in β_2, but the latter can exhibit up to 10-fold greater activity as reported for the *B. subtilis* enzyme in Chapter 2 of this thesis. This observation indicates that the initial triggering mechanism for injection of the radical

onto the pathway will be different for the two clusters and, thus, obtaining insight into these differences should provide valuable insights for radical transfer pathways in all class I RNRs.

1.4. THE ALLOSTERIC REGULATION OF RNR ACTIVITY

As mentioned at the beginning of this chapter, RNRs, being the only *de novo* source of dNTPs for DNA replication and repair, are subjected to many levels of regulation to ensure that the proper amounts and ratios of these building blocks are maintained. A loss of balance in the dNTP pools or too high concentrations in general results in increased incidences of mutation, recombination, and genomic abnormalities/break down, for which the ultimate price is cell death or, in higher eukaryotes, immortalized cells in the form of cancer.[112, 113] The dysregulation of dNTP pools also provides a tool for combating cancer and other human diseases, and RNRs, because of their central role in maintaining dNTP balances, have been the successful targets of antineoplastic[114] and antiviral therapeutics,[115] and, as described earlier, are new attractive targets for antimicrobial compound development.

Known control mechanisms of RNR activity include transcriptional[116] and allosteric regulation,[1] post-translational modification (e.g. cofactor assembly),[7] sub-cellular location (in yeast),[117] cell-cycle[118] or growth phase-dependent proteolysis,[119] and binding of small inhibitory proteins.[118, 120] The focus of the following discussion will be on the allosteric regulation of prokaryotic class I RNRs although many of the points are pertinent to eukaryotic systems as well. However, because of the increased complexity of the eukaryotic cell, the regulation of RNRs in these systems is more complicated, especially in the case of their quaternary structure,[4, 121] and will, therefore, be minimally discussed.

1.4.1. General scheme for allosteric substrate specificity and overall activity regulation of RNRs. Compared with most of the enzymes of primary metabolism, RNRs are very unique

systems because one enzyme is able to reduce four different substrates. This ability is the result of an intricate system of allosteric regulation in which the binding of effector nucleotides to the enzyme causes changes in its tertiary/quaternary structure that affect enzymatic activity. Prokaryotic RNRs thus far characterized can have up to two different types of allosteric sites. All enzymes have specificity sites (S-sites) to which ATP and dNTPs reversibly bind and control which substrates are preferentially reduced. The second class of allosteric sites that control the overall activity of RNR (A-sites) have been mostly observed in class I and class III systems that have an ATP-cone domain[27] fused to the N-terminus of α.

Classic experiments by Peter Reichard and various coworkers, mainly using the *E. coli* Ia RNR, laid the foundations for the currently accepted model of the regulation of RNR substrate specificity (**Figure 1.16**).[82, 83, 122-129] The results of kinetic studies revealed that NDP reduction catalyzed by the *E. coli* enzyme could be stimulated by the addition of ATP or dNTPs, with each substrate having a prime effector(s) that maximized its rate of turnover. For both CDP and UDP, the prime effectors were ATP (high concentrations) and dATP (low concentrations, ~1 μM), whereas for ADP and GDP, the prime effectors were dGTP and TTP, respectively. In contrast, concentrations of dATP greater than 10 μM inhibited the reduction of all four substrates. Nucleotide binding and competition experiments revealed that only α_2 binds nucleotides and possesses two sites for substrates and four sites for allosteric effectors. Two of the allosteric sites could be occupied by dATP, dGTP, and TTP with high affinity (K_d = 0.3 – 3 μM) and ATP with low affinity ($K_d \approx$ 100 μM), and were thus termed the S-sites. The remaining two allosteric sites bound either dATP at lower affinity (K_d = 1.2 – 5 μM) or ATP at a similar affinity to the S-sites. Based on the ability of increasing ATP concentrations to competitively alleviate dATP inhibition, these allosteric sites were termed the A-sites.

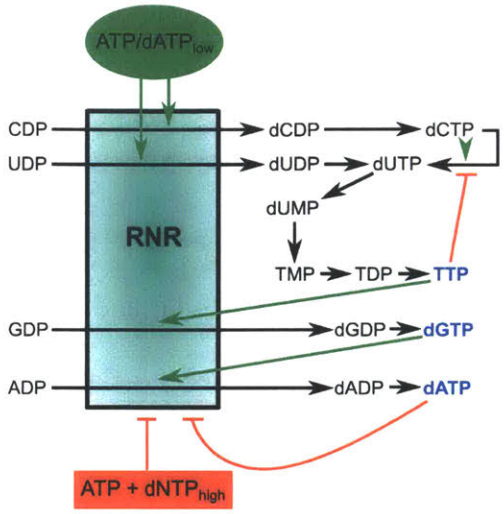


Figure 1.16. Overview of the allosteric regulation of dNTP production in *E. coli*. Four effector nucleotides (ATP, TTP, dGTP, dATP) control the specificity and overall activity of RNR. Green arrows indicate activator attributes and red lines inhibitory attributes of the nucleotides. The relative ratios and amounts of TTP:dCTP are controlled in *E. coli* by dCTP deaminase. This enzyme is allosterically activated by dCTP and inhibited by TTP.

1.4.2. All RNRs have the same allostery for substrate specificity regulation, but diversity in overall activity regulation. Surprisingly, the general attributes of the allosteric regulation of substrate specificity uncovered for the *E. coli* Ia RNR (**Figure 1.16**) have been found to be universally conserved among all RNR classes studied to date.[4, 9, 49, 50, 60, 77, 88, 121, 130-134] This seems like a rather strong conclusion given the limited number of systems that have been characterized, and it is noteworthy that divergences from the scheme presented in **Figure 1.16** have been reported in the past for the *E. coli* Ia[124, 125] and other RNRs from different species.[133, 135] In these older studies, substrates were tested one at a time due to technical limitations of the methods available. Technological advances have allowed for the recent development of assay procedures in which the turnover of all four substrates is monitored simultaneously.[90, 136-138] Product deoxyribonucleotides are first separated from ribonucleotides using boronate affinity chromatography, and then individual dNDPs are isolated and quantitated via column calibration or radioactivity using high pressure liquid chromatography (HPLC). The results of these recent studies have revealed a role

for competitive binding of substrates to the catalytic site of RNR in the allosteric regulation of the enzyme. Under the conditions of four substrate assays, the substrate specificity allosteric regulation of RNRs conforms to the general scheme originally proposed by Reichard (**Figure 1.16**), and reassessment of systems diverging from this scheme, such as the *E. coli* Ia enzyme,[90] with such assay procedures revealed that the unusual allostery was likely an artifact of studying substrates and effectors one at a time.

More often than not, the allostery of an RNR diverges from the general scheme shown in **Figure 1.16** as a result of differences in the overall activity regulation of the enzyme by ATP and dATP. Previously, the presence of this form of allosteric regulation only coincided with the presence of at least one ATP-cone domain on the N-terminal end of the α subunit, which can be discerned by the signature sequence, VxKRDG.[1] The class Ib RNRs are noteworthy for not possessing an ATP-cone domain at all,[6] which manifests in the absence of dATP inhibition at high concentrations (up to 1 mM) in the previously studied members of this subclass.[9, 11, 29-31] Similar results were obtained with class II RNRs from *Lactobacillus leichmannii* and *T. maritima*, which also lack the ATP-cone signature sequence,[60] and, thus, provide further support for ATP-cone domains as a major player in regulating the overall activity of RNRs in general.

For those RNRs that do have ATP-cones, a remarkable diversity in the overall activity allosteric regulation has been observed. The presence of the signature motif alone does not necessarily indicate that the ATP-cone domain is functional. This scenario is often seen with viral enzymes, which are notable for encoding potential ATP-cones, some of which that can even bind ATP and dATP, but do not exhibit overall activity regulation.[130, 137, 139-142] Inactivation of the ATP-cone domain can be regarded as evolutionarily advantageous for viruses since unregulated dNTP production by RNR increases the rates of spontaneous mutation and, therefore, helps drive the

evolution of the viral genome.[143] However, there are also non-viral RNRs with ATP-cone domains that exhibit no overall activity allosteric regulation, such as the class II RNR from *Thermoplasma acidophilum*.[60] Other systems, like the *P. aeruginosa* Ia RNR,[88] have two or three ATP-cone domains fused to the N-terminus of α (695 sequences with two domains and 50 sequences with three domains).[6, 77] Only the first domain in the *P. aeruginosa* α subunit was reported to be capable of binding dATP and inhibiting the enzyme.[77, 88, 144] Finally, in certain systems the inhibition of RNR activity can be induced not only by dATP, but by ATP plus a different dNTP effector, such as TTP in the case of the *E. coli* Ia enzyme.[82]

1.4.3. Pyrimidine dNTP pools are additionally regulated by deaminases of deoxycytidine analogs. One notable feature of the allosteric regulation of RNRs is that dCTP is not an effector of enzymatic activity, thus making RNRs unresponsive directly to elevated levels of this nucleotide. To compensate for this potentially detrimental situation, cells have additional regulatory mechanisms in place that control the source of dUMP, the precursor of deoxythymidine nucleotides. In *E. coli* (**Figure 1.16**), dUMP originates from either UDP reduction or from the deamination of dCTP by dCTP deaminase. In other organisms, such as *B. subtilis*, dC and dCMP are deaminated to their deoxyuridine derivatives instead of dCTP.[145] The deoxycytidine deaminases, like RNR, are also allosterically regulated, with dCTP stimulating the deamination of the designated deoxycytidine analog(s) and TTP inhibiting the enzyme. Therefore, the cell can effectively control the pyrimidine dNTP pools by adjusting the flux of dUMP derived from the deoxycytidine pool via regulation of the cytidine deaminases. RNRs may have a role in this regulation since UDP is reported to be the least efficient substrate for several systems,[124, 132, 133] thus in these cases the relative contribution of this route to dUMP generation should be minimal compared to deamination of dC analogs. Ultimately, however, the primary source of dUMP for

TTP biosynthesis varies between organisms, with ~75 – 80% of the dUMP in *E. coli* reported to be derived from dCTP pools[146] versus 100% in *Trypanosoma brucei* being derived from UDP reduction.[133] In the latter case, the use of RNR for generating all of the dUMP for TTP biosynthesis results from the fact that the *T. brucei* CTP synthase is a very inefficient enzyme, thus the CTP/dCTP pools in this organism are very low.[147] Additionally, *T. brucei* also does not encode a deaminase for deoxycytidine analogs. In *B. subtilis*, it is estimated that roughly half of the dUMP pool is derived from UDP reduction and the other half from dC analog deamination.[145]

1.4.4. Structural basis for allosteric regulation of overall activity in class I RNRs. Early experiments by Reichard and Brown[82] revealed that overall activity regulation of the *E. coli* class Ia RNR was caused by substantial changes in the quaternary structure of the enzyme. As described earlier in this chapter, the $\alpha_2\beta_2$ complex ($s_{20,w}$ = 9.7 S) was observed to form in sucrose gradient ultracentrifugation experiments when the subunits (2:1 mixtures of both α:β and β:α) were mixed both in the presence and absence of the effectors ATP, dGTP, and TTP. In contrast, the presence of dATP or ATP/TTP caused an additional larger complex with an $s_{20,w}$ = 15 S to form. The 15 S species could be fully converted into $\alpha_2\beta_2$ complexes by addition of ATP at concentrations that reversed dATP inhibition in activity assays. Quantitation of the subunit stoichiometry revealed that the 15 S species was still composed of a 1:1 ratio of α:β. Based on the observed formation of the 15 S species at inhibitory concentrations of dATP, the authors concluded that this complex was an inhibited form of the *E. coli* Ia RNR and that structurally it was a dimer of the 9.7 S species (i.e. the $\alpha_2\beta_2$ complex). Since the subsequent biophysical studies of the enzyme by Thelander revealed that each subunit was a dimer,[76] the stoichiometry of the inhibited complex was suggested to be $\alpha_4\beta_4$, a conclusion that was later supported by the results from GEMMA experiments.[90]

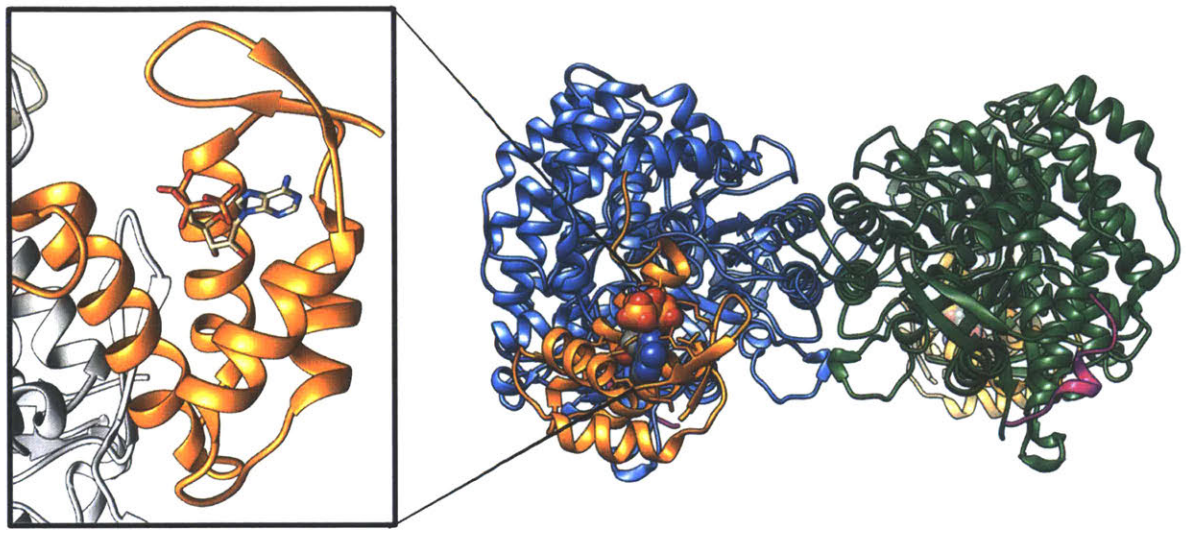


Figure 1.17. Crystal structure (PDB 3R1R[24]) of the *E. coli* Ia RNR α_2 subunit with adenosine 5′-(β,γ-imido)triphosphate bound in the A-site.

An important advancement in the understanding the structural basis of the dATP-mediated inhibition of the *E. coli* Ia RNR was reported in 1997 with the crystal structures of α_2 bound to substrate and effector nucleotides/analogs that were co-crystallized or soaked in.[24] These structures showed for the first time the locations of the catalytic, S-, and A-sites on α_2. The structure obtained from soaking crystals with CDP and the ATP analog adenosine 5′-(β,γ-imido)triphosphate (AMP-PNP) revealed only AMP-PNP, which was bound in the ATP-cone domain, a bundle of four α-helices at the N-terminus of the protein (**Figure 1.17**) Given that the domain was located close to the last few visible C-terminal residues of β in the docking model (**Figure 1.11**),[26] it was speculated that ATP or dATP binding to the ATP-cone would shift the C-terminal tail of β to adjust the location of the radical transfer pathway residues between the subunits to favor or disfavor turnover, respectively. However, the structure offered no hints as to the nature of the $\alpha_4\beta_4$ complex, and as AMP-PNP was not well ordered due to low occupancy and poor correlation of the model with the electron density, no information was gleaned on how the binding of ATP versus dATP is discerned.

The breakthrough in structurally understanding the overall activity regulation of the *E. coli* Ia RNR was reported in 2011 in a joint study by the Stubbe and Drennan groups[2] which revealed, using a combination of biophysical techniques, that the $\alpha_4\beta_4$ complex was a ring of alternating α_2 and β_2 subunits (**Figure 1.18**). A highly unusual structure for the complex was first indicated by disagreement between the experimentally measured $s_{20,w}$ (15.6 S) and the expected value (19 S) for a globular complex with a molecular weight of 510 – 533 kDa, as estimated by analytical ultracentrifugation (AUC), small angle X-ray scattering (SAXS), and, in an earlier study, by GEMMA.[2, 90] A 23 Å cryo-electron microscopy (cryo-EM) reconstruction of the dATP-inhibited complex revealed a ring-like shape into which two copies of the crystal structures of the individual subunits could be unambiguously modeled. This pseudo-atomic model was found to be nearly superimposable with a 5.7 Å crystal structure of the $\alpha_4\beta_4$ complex, and the predicted $s_{20,w}$ and other hydrodynamic properties of the ring agreed with the experimental data obtained with AUC and SAXS. Taken together, the data suggest the $\alpha_4\beta_4$ ring can form in solution at physiologically relevant protein and dATP concentrations and is, therefore, the inhibited complex originally observed by Reichard and Brown in 1969.[82]

The mechanism of activity inhibition is evident in the structure of the *E. coli* Ia RNR $\alpha_4\beta_4$ complex. In the docking model of the $\alpha_2\beta_2$ complex,[26] the RT pathway residues in α and β are fairly close to one another (**Figure 1.15**), with a distance between Y_{731} in α and W_{48} in β of ~25 Å. Placing Y_{356} (which, as mentioned previously is disordered in all β_2 structures acquired to date) between these residues completes a conduit from Y_{122} to C_{439} with reasonable distances (at most 10 – 15 Å) between pathway residues such that PCET can take place. In the $\alpha_4\beta_4$ ring, the distance between Y_{731} and W_{48} has increased to ~55 Å,[2] too large of a distance for PCET to occur at a reasonable rate, even with the placement of Y_{356} between them. Furthermore, the pathway is now

solvent exposed, hence a radical on pathway would easily be quenched by any reductant present in solution. It can, therefore, be concluded that the mechanism of overall activity regulation of the *E. coli* class Ia RNR is the formation or disruption of the RT pathway between the subunits by quaternary structural interconversions.

From these results, a three-state model for the overall activity regulation of the *E. coli* Ia RNR has been proposed (**Figure 1.18**).[148] The enzyme exists in an equilibrium consisting of dissociated subunits, $\alpha_2\beta_2$ complex, and $\alpha_4\beta_4$ complex. The binding of substrate and effector (ATP, dATP (low concentrations), TTP, or dGTP) shifts the equilibrium towards formation of the $\alpha_2\beta_2$ complex, which puts RNR in an active state. When dNTP concentrations accumulate, the binding of dATP or ATP/TTP to the activity site shifts the quaternary structural equilibrium toward the $\alpha_4\beta_4$ complex, thereby shutting RNR off to prevent over-production of dNTPs. Under physiological conditions, the *in vivo* ATP concentration range in *E. coli* is estimated to be 0.5 – 4 mM, [149] and given the reported K_d of the ATP-cone for this nucleotide (~100 – 200 μM),[150-152] the A-site should typically be occupied by ATP. This observation suggests that RNR activity may mainly be adjusted by ATP/TTP- or, perhaps, ATP/dATP-mediated inhibition in response to fluctuations in the *in vivo* dNTP pools.

Formation of inhibited complexes as a result of dATP binding to the A-site has been found to be a fundamental aspect of overall activity regulation of class Ia RNRs. The *E. coli* $\alpha_4\beta_4$ complex (**Figure 1.19A**), however, represents only one of three different inhibited structures that have recently been characterized. It is noteworthy that the ATP-cone domain is the key component that makes formation of the inhibitory complexes possible. However, the different structures that have been observed result from interactions on different regions of this domain (**Figure 1.19**).

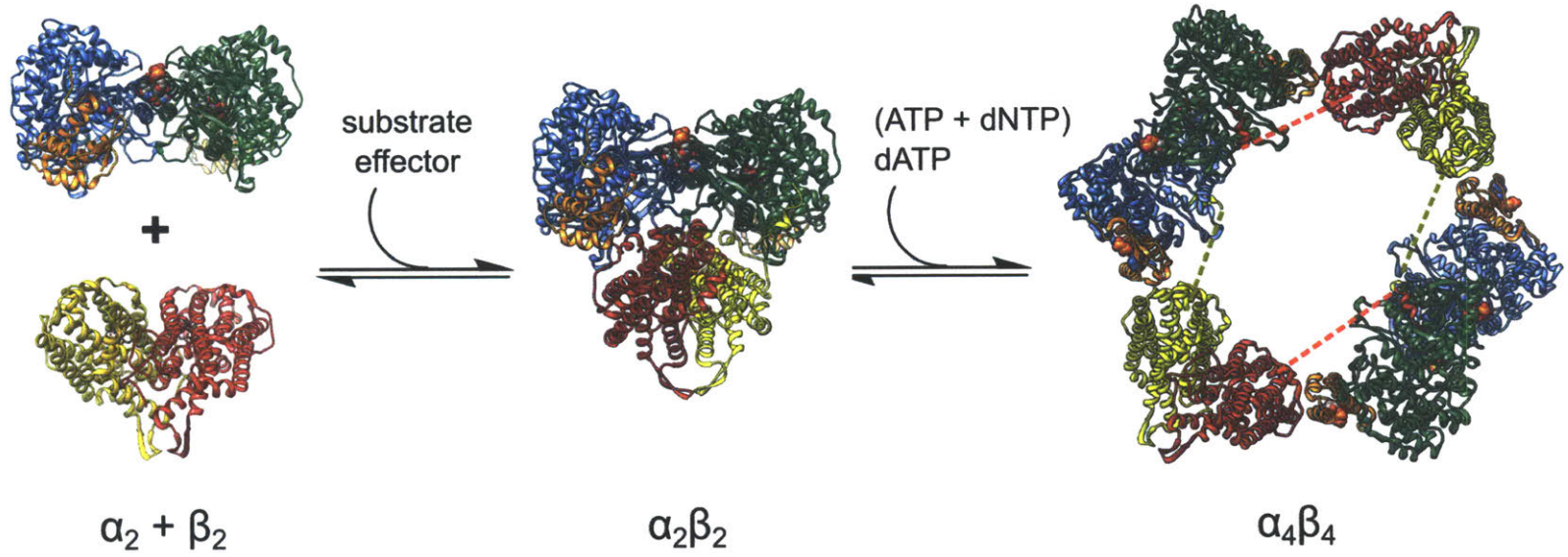


Figure 1.18. Quaternary structural interconversions regulating the overall activity of the *E. coli* class Ia RNR. Each protomer of α_2 and β_2 are colored as described in previous figure legends. Crystal structures used for the individual subunits and for the $\alpha_2\beta_2$ complex were PDB 4R1R[24] for α_2 and PDB 1MXR[78] for β_2. The structure of the $\alpha_4\beta_4$ complex was from PDB 3UUS.[2] In the α_2 and $\alpha_2\beta_2$ complex, the S-site is occupied by TTP and the catalytic site by GDP. In the $\alpha_4\beta_4$ complex, the S- and A-sites are occupied by dATP. All nucleotides are shown as a space filling model and colored according to element: gray = C, red = O, blue = N, orange = P. The dashed lines indicate the disordered regions of the C-terminal tails of β. Note that the C-terminal tails of α (not shown) are pointing into the center of the ring

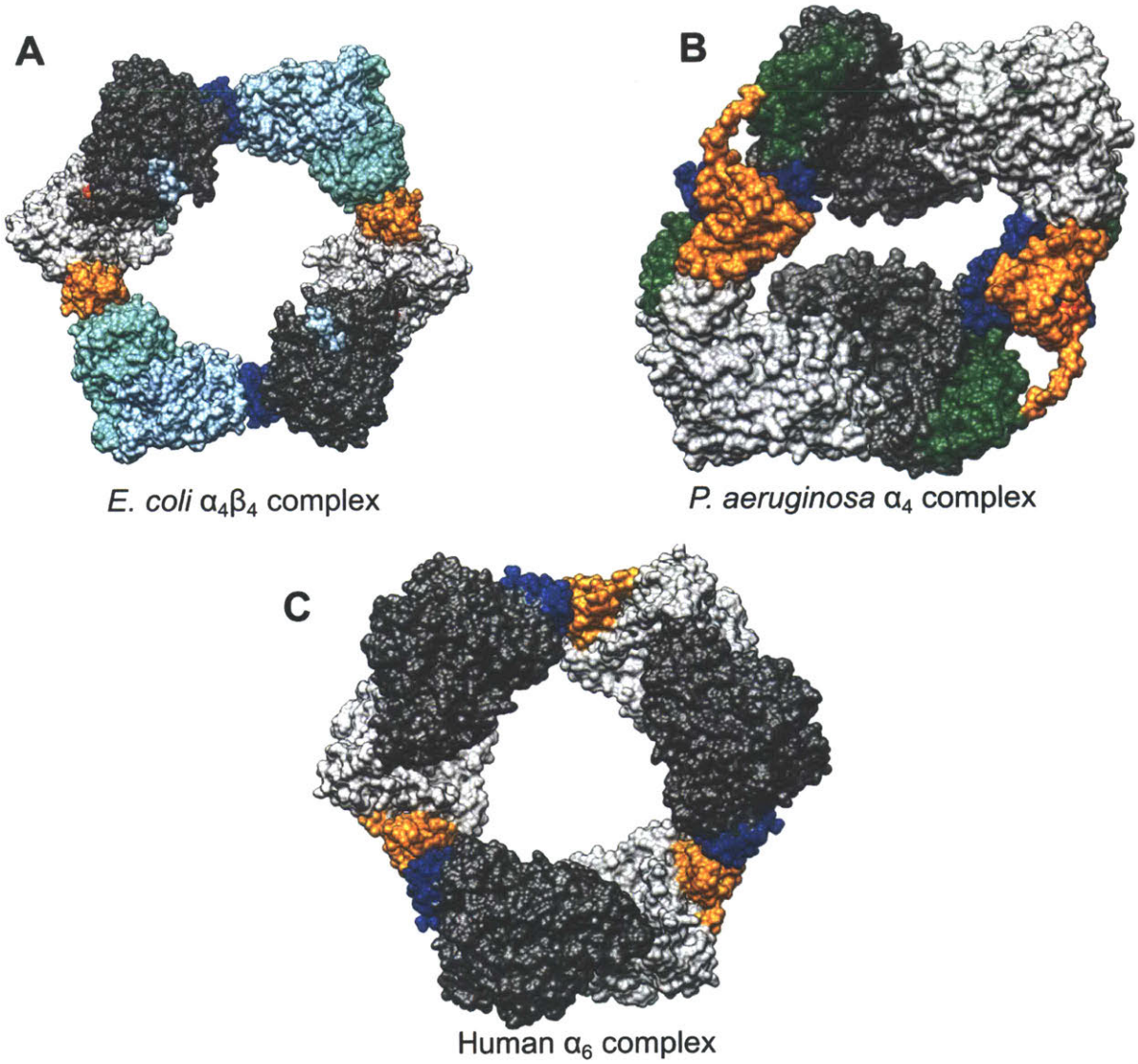


Figure 1.19. Types of dATP-inhibited quaternary structures observed with class Ia RNRs. The formation of these structures is dependent on the presence of an ATP-cone domain (orange and blue), which binds dATP and mediates formation of the complex through key interactions with ATP cones on adjacent α protomers (light and dark gray) or with the β subunit (light blue). (**A**) The $\alpha_4\beta_4$ complex of the *E. coli* Ia RNR (PDB 3UUS).[2] (**B**) The α_4 complex of the *P. aeruginosa* Ia RNR (PDB 5IM3).[144] The second, non-functional ATP-cone is colored green in this structure. (**C**) The α_6 complex of the human Ia RNR (PDB 5D1Y).[3] In (**A**) and (**B**), dATP occupying the S- and A-sites are shown as spheres and colored according to element: gray = C, blue = N, red = O, P = orange. Only portions of the nucleotides are visible in the panels above.

GEMMA experiments initially demonstrated that the α subunit of the *P. aeruginosa* Ia RNR formed a tetramer (α_4) at inhibitory concentrations of dATP, the structure of which was

recently published (**Figure 1.19B**).[77, 144] Recall that this α subunit has two tandem ATP-cones on its N-terminus. Deletion of the first of these domains resulted in the inability of the *P. aeruginosa* α subunit to form the inhibited α_4 complex, thus demonstrating that only the first ATP-cone was required for overall activity regulation.[77] This result is supported by the recent structure (**Figure 1.19B**) which shows that all of the contacts required for formation of α_4 are mediated by the N-terminal ATP-cones of adjacent subunits.[144] The third type of inhibitory complex has been observed with the yeast and human Ia RNRs (one ATP-cone at the N-terminus of α), where dATP at inhibitory concentrations causes the formation of an α_6 complex (**Figure 1.19C**).[3, 4] In these cases, activity inhibition again results from disrupting the radical transfer pathway due to the fact that the space inside the α_4 and α_6 rings is too small to allow β_2 to bind competently.[3, 144]

Evolutionary mobility of the ATP-cone likely explains the variety of different inhibitory structures that RNRs containing these domains can form.[27] A comparison of the architecture of the domains from the *E. coli*, *P. aeruginosa*, yeast, and human enzymes reveals that all have a homologous tertiary structure. Furthermore, most of the conserved residues that can be identified in sequence alignments of multiple ATP-cones from different class Ia and III RNRs are responsible for stabilizing the domain's architecture.[27] This high structural and sequence conservation relative to the distinct catalytic domains of the different classes of RNRs suggests that the ATP-cone was acquired by each class independently, perhaps through horizontal transfer of genetic information. This would allow the ATP-cones of different RNRs to evolve individually, but in parallel, thus leading to the appearance of a variety of structures that can form from the same general domain architecture. Further support for evolutionary mobility of ATP-cones is the observation of these domains fused with 2-phosphoglycerate kinases, which catalyzes a different reaction compared to RNR, and with Zn-mediated DNA-binding motifs in the transcription factor NrdR.[27, 153]

It will be interesting to see if other types of ATP-cone-mediated inhibited RNR quaternary structures are observed, especially in the case of those systems with multiple ATP-cones fused to the α subunit. Modern bioinformatic techniques that can distinguish functional from non-functional ATP-cones have revealed that most of these cases are similar to the *P. aeruginosa* RNR, where the terminal domain appears to be functional while the interior domain(s) have degenerated evolutionarily and are likely not functional.[77] However, there are a few systems, such as the *C. trachomatis* class Ic RNR α-subunit (which has three ATP-cones), where the bioinformatics analysis suggests that more than one of these domains will be functional.[77] The inhibitory effects of dATP were demonstrated in an early study of the *C. trachomatis* enzyme,[154] but the quaternary structure(s) adopted by the system are currently unexplored. It also remains to be determined in RNRs possessing multiple, potentially active ATP-cones if the allosteric regulation, both in terms of specificity and overall activity, is distinctly different relative to currently characterized systems.

1.4.5. Structural basis for allosteric regulation of substrate specificity in class I RNRs. The first crystal structures of *E. coli* α_2 with substrate and effector nucleotides bound revealed that the catalytic and S-sites were approximately 15 Å apart, therefore ruling out a direct interaction between the nucleotides.[24] Instead the protein mediates communication between the two sites through three loops close to the α_2 interface.[24] These loops, termed loop 1, loop 2, and loop 3, were originally disordered in the first α_2 structure (without nucleotides),[26] but had become partially ordered upon occupancy of the binding sites with the soaked-in nucleotides (**Figure 1.20A**) and appeared to be good candidates for linking the catalytic and S-sites. Support for this hypothesis came from sequence alignments, which revealed loop 1 (not shown) and loop 2 (**Figure 1.21**) had several residues that were conserved among other class I RNRs. Furthermore, loops 1 and 3 from one α protomer and loop 2 from the other were connected through hydrogen bonding and van der

Waal interactions, indicating that each loop was able to influence the position of the others.[24] However, none of the loops in these older structures were positioned in a manner that would facilitate communication between the catalytic and S-sites, and the substrate and effector nucleotides were poorly defined due to partial occupancy and poor electron density. Therefore, the structural basis for the substrate specificity regulation of the *E. coli* Ia RNR was unclear.

Details of how the effector nucleotide in the S-site of the *E. coli* Ia RNR is read and translated into protein conformational changes in the catalytic site that lead to preferential binding of one substrate have recently been reported by the Drennan group.[5] These structures were obtained by soaking crystals of the $\alpha_4\beta_4$ inhibited complex discussed previously (**Figure 1.18**) with the different canonical substrate/effector nucleotide pairs. The structures of the catalytic and S-sites in the α_2 subunits of the $\alpha_4\beta_4$ complexes appear to be minimally perturbed relative to previously released structures of the protein, despite the enzyme being in an inhibited state.[2, 5, 95] Therefore, it was concluded that in these $\alpha_4\beta_4$ complexes, the positions of the loops mediating communication between the catalytic and S-sites are reflective of that in the active $\alpha_2\beta_2$ complex. In the report by Zimanyi *et al.*,[5] the earlier structures of *E. coli* α_2 by Eklund *et al.* with substrate and effectors[24] were referred to as "substrate-free" and their structures as "substrate-bound." This terminology will be used in the following discussion.

Compared with the substrate-free structures, two large-scale changes were observed in the architecture of *E. coli* α_2 in the substrate-bound state. The first change was a large contraction of the tertiary structure which causes the catalytic site to clamp down onto the substrate nucleotide and, therefore, bind it tightly. This structural change makes sense given that it is not desirable to inadvertently release the highly reactive intermediates that form during the catalytic cycle of nucleotide reduction (**Figure 1.6**). The second notable change that occurs is that loop 2 flips up

from its position in the substrate-free structure and becomes positioned between the catalytic and S-sites (**Figure 1.20B**). This motion breaks any contact made between this loop and loop 3 in the opposite protomer, as observed in the substrate-free state, and, therefore, may act as a switch that allows β_2 to bind once substrate and effector have associated with α_2.

Regardless of which effector is bound, the nucleotide is anchored in the S-site by similar non-covalent interactions, involving hydrogen bonding and hydrophobic interactions between the sidechains and peptide amides of the protein and the ribose and phosphate moieties of the nucleotide (**Figure 1.22**). Loop 1 plays a large role in the binding of effectors, and in all characterized class I α subunits from bacteria and eukaryotes, this loop contains a conserved Arg and Ile (R_{262} and I_{268} in *E. coli* α).[24] The conserved Arg residue likely plays an important role in the preferential binding of nucleoside/deoxynucleoside triphosphates to the S-site by forming two hydrogen bonding interactions with the γ-phosphate of the bound effector.[5, 24] Similar hydrogen bonds with nucleoside mono- or diphosphates could not be made without significant re-organization of loop 1. The conserved Ile residue stacks against the nucleobase of the effector and likely helps stabilize its position so that the proper non-covalent interactions can be made between the base and loop 2.

As noted earlier, both ATP and dNTPs can bind to the S-site, albeit with starkly contrasting affinities. The crystal structures of α provide a possible explanation for this phenomenon in that the S-site is quite crowded and often has a large bulky residue that would sterically clash with a 2′-OH of a bound ATP molecule. In the *E. coli* Ia α structures, this residue is F_{281}, whereas in the *S.* Typhimurium NrdE (and likely other class Ib RNRs due to sequence conservation at the equivalent position), the bulky residue is Y_{244}.[22, 24] Therefore, ATP binding to the S-site is not prevented, but also is not as favorable as the binding of dNTPs.

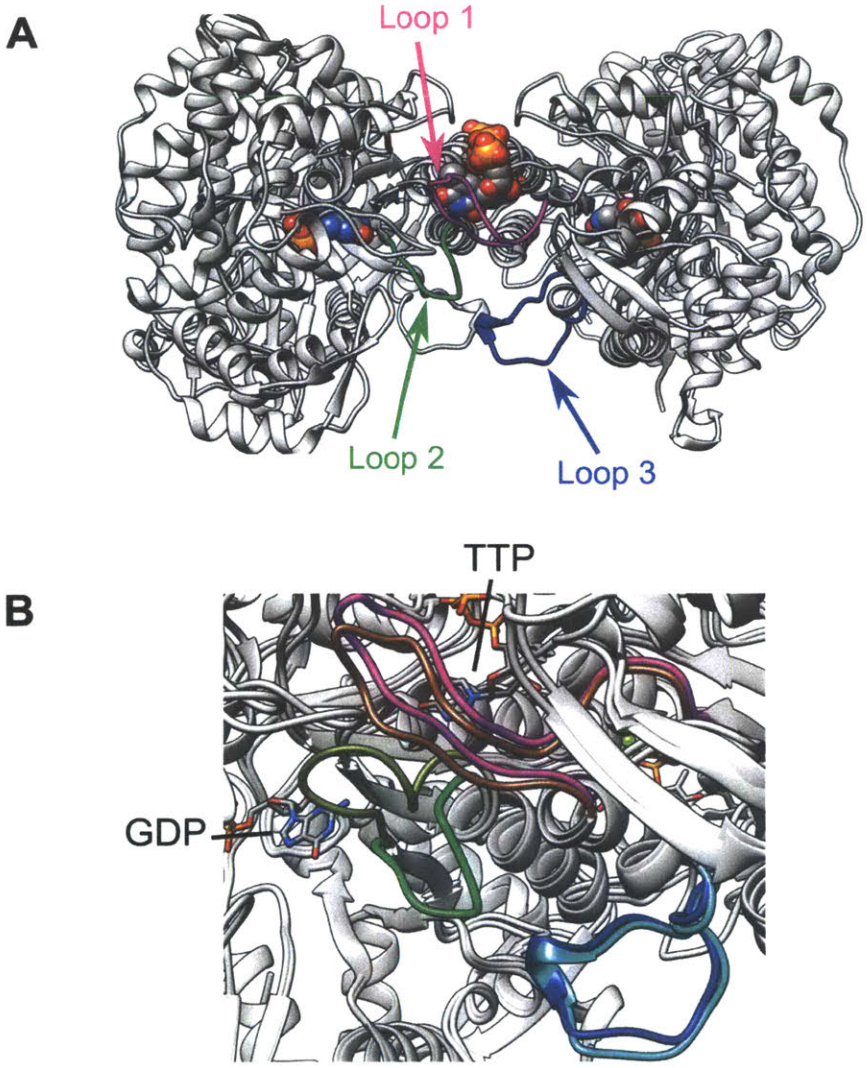


Figure 1.20. Secondary structural elements involved in the communication between the catalytic and S-sites on the *E. coli* α_2 subunit. (**A**) Overview of the locations of loop 1 (magenta), loop 2 (green), and loop 3 (blue) on α_2 (PDB 4R1R).[24] Loops 1 and 3 from one α protomer interact with loop 2 of the second protomer. The catalytic and S-sites are occupied by GDP and TTP, respectively, which are shown as space filling models colored according to element: gray = C, red = O, blue = N, orange = P. (**B**) Superimposition of the "substrate-free" and "substrate-bound" crystal structures of α_2 (PDB 4R1R and 5CNV,[5] respectively) illustrating the conformational changes that move the three loops into positions that mediate communication between the catalytic and S-sites. The loops are colored as in (**A**) for the "substrate-free" structure whereas loops 1, 2, and 3 in the "substrate-bound" structure are colored brown, dark green, and cyan, respectively. TTP and GDP are shown as stick models and colored as in (**A**).

```
Escherichia coli*          (289)VKSCSQGG-VRGGAAT(303)
Trypanosoma brucei         (293)ARYVDQGGGKRKGAFA(308)
Bos taurus                 (283)ARYVDQGGNKRPGAFA(298)
Mus musculus               (283)ARYVDQGGNKRPGAFA(298)
Homo sapiens*              (283)ARYVDQGGNKRPGAFA(298)
Saccharomyces cerevisiae*  (283)ARYVDQGGNKRPGAFA(298)
Salmonella enterica*       (242)FSYANQLG-ARQGAGA(256)
Thermotoga maritima*       (198)ISVVKQGY-RRRGALM(212)
Thermoplasma acidophilum   (314)TDVIKQGG-KRRGANM(328)
                                .*    * **
```

Ia
Ib
II

Figure 1.21. Sequence alignment of loop 2 from characterized class Ia, Ib, and II RNRs. An asterisk next to the organism name indicates a crystal structure of the α subunit is available. Residues highlighted in yellow are conserved in all characterized RNRs. Note that in some class Ib RNRs, in comparison to *S. enterica*, the position equivalent to A_{254} is substituted with serine. This is the case with the RNR from *B. subtilis*. This figure was reproduced with permission under the Creative Commons Attribution License from Zimanyi, C. *et al.* (**2016**) *eLife* **5**: e07141. DOI: 10/7554:eLife07141.

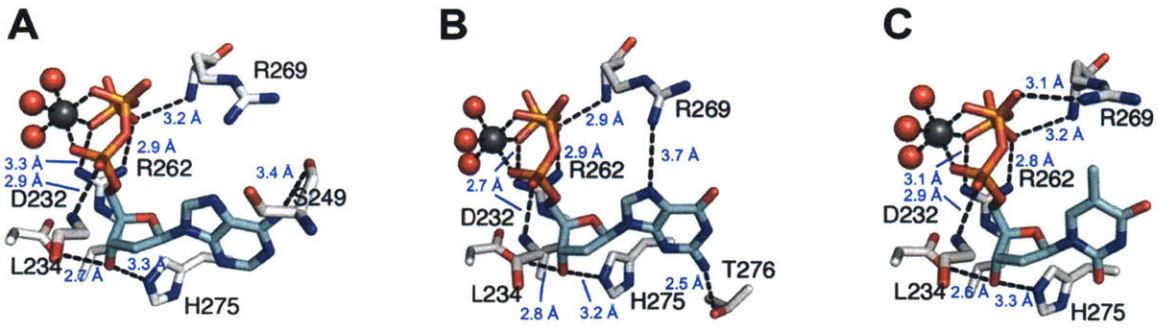


Figure 1.22. Hydrogen bonding interactions between the effectors (**A**) dATP, (**B**) dGTP, and (**C**) TTP and the *E. coli* Ia α subunit in the S-site. Nucleotide C atoms are colored cyan and protein C atoms light gray. Mg^{2+} ions are colored dark gray in all three panels. The remaining atoms are colored according to element: blue = N, red = O, and orange = P. Note that the sidechain of L_{234} is hidden for clarity, but it packs against the nucleobase for all effector nucleotides. This figure was reproduced with permission under the Creative Commons Attribution License from Zimanyi, C. *et al.* (**2016**) *eLife* **5**: e07141. DOI: 10/7554:eLife07141.

In contrast, loop 2, likely due to its high Gly content (**Figure 1.21**), adopts a different conformation depending on whether dATP, dGTP, or TTP is bound in the S-site. Surprisingly, the identity of the substrate and effector bases is read mostly by hydrogen bonding interactions with backbone amides and carbonyl groups rather than the sidechains of the residues in loop 2 (**Figure 1.23**). The observation likely explains the limited sequence conservation of loop 2 from different organisms (**Figure 1.21**). The conserved residues, however, play important roles in the

specificity and affinity of the catalytic site for the NDP substrates. The conserved Gln residue (**Figure 1.21**, Q_{294} in *E. coli* α) distinguishes pyrimidine from purine substrates bound to the catalytic site. In the case of CDP/dATP and UDP/dATP, the sidechain of Q_{294} points into the catalytic site and forms a hydrogen bond with the C2 carbonyl group of either pyrimidine substrate (**Figure 1.23A** and **B**), whereas it is displaced out of the catalytic site to make room for the binding of either ADP or GDP (**Figure 1.23C** and **D**). The conserved Gly (G_{300} in *E. coli* α) makes no direct interactions with either the substrate or effector nucleotide and instead appears to impart flexibility in a key region of loop 2 that allows RNR to distinguish between ADP and GDP when dGTP or TTP, respectively, are bound in the S-site (discussed below), The conserved Arg residue in loop 2 (R_{298} in *E. coli*) reaches across the face of the substrate nucleobase, regardless of its identity, to make a hydrogen bonding/salt bridge interaction with β-phosphate of the substrate. The sidechain methylene groups of R_{298} provide additional hydrophobic interactions with the substrate nucleobase and, thus, further stabilize the binding of the NDP to α.

The substrate/effector-driven movements of Q_{294} and G_{300} in *E. coli* α are largely due to changes in the position and hydrogen bonding interactions of the backbone carbonyl and amide group of the non-conserved residue, S_{293}.[5] When dATP is bound in the S-site, the conformation of loop 2 allows hydrogen bonding interactions between N1 and exocyclic N6 of dATP and the backbone amide and carbonyl groups of S_{293}, respectively (**Figure 1.23A** and **B**). This hydrogen bond network stabilizes Q_{294} with its sidechain pointed into the catalytic site. Given that the conformation of loop 2 results from two hydrogen bonds between the adenine base of dATP and the protein, it is likely that this loop will adopt a similar conformation when ATP is bound to the S-site. In the case of dGTP occupying the S-site, the conformation of loop 2 is altered so that N1 and exocyclic N2 on dGTP donate hydrogen bonds to two backbone carbonyls on C_{292} and G_{294}

(**Figure 1.23C**). This new network forces S_{293} to move closer towards the catalytic site, which allows its sidechain OH group and backbone carbonyl to form hydrogen bonds with the backbone carbonyl of G_{300}. This, in turn, forces G_{299} to move into the catalytic site where its backbone carbonyl accepts a hydrogen bond from exocyclic N6 of ADP (**Figure 1.23C**). With TTP bound in the S-site, the smaller nucleobase causes loop 2 to shift towards the effector nucleotide, which allows N2 of thymine to hydrogen bond with the backbone carbonyl of C_{292} (**Figure 1.23.D**). This interaction essentially prevents any of the mainchain groups of C_{292}, S_{293}, and Q_{294} from hydrogen bonding to G_{299} and G_{300}, allowing the latter two residues to move towards the S-site as well, which opens up the catalytic site to incorporate the guanine base of GDP (**Figure 1.23D**). The binding of GDP is stabilized by hydrogen bonds donated from N1 and exocyclic N2 of the nucleobase to the backbone carbonyl of G_{299} (**Figure 1.23D**).

The structures reported by Zimanyi *et al.*[5] provide the first complete picture to date for the structural basis of specificity allosteric regulation for the *E. coli* class Ia and, perhaps, other class I RNRs in general. Prior to this study, successful attempts to delineate structurally the mechanism of substrate specificity allosteric regulation were reported for the *T. maritima* class II (acquired in the absence of adenosylcobalamin)[155] and yeast Ia RNRs.[156] Structures with all four substrate/ effector pairs bound to the catalytic and S-sites were obtained for both enzymes. In the case of the *T. maritima* II RNR, however, loop 2 in all but the TTP/GDP structure (PDB 1XJE) was partially disordered and, therefore, many of the interactions between the protein and the nucleobases were missing.[155] Despite this issue, the *T. maritima* structures revealed a similar placement, in comparison to the *E. coli* Ia α structures, of the conserved Gln sidechain (Q_{203}) into or out of the catalytic site that was dependent on the identity of the effector bound in the S-site.[155] Furthermore, the hydrogen bonding network established between loop 2 and the nucleobases of TTP and GDP,

as well as the conserved Arg (R_{207}) reaching across the face of the substrate nucleobase to hydrogen bond with its β-phosphate, was the same in both the *E. coli* and *T. maritima* structures, despite the differences in sequence of loop 2 in these enzymes (**Figure 1.21**).[5, 155]

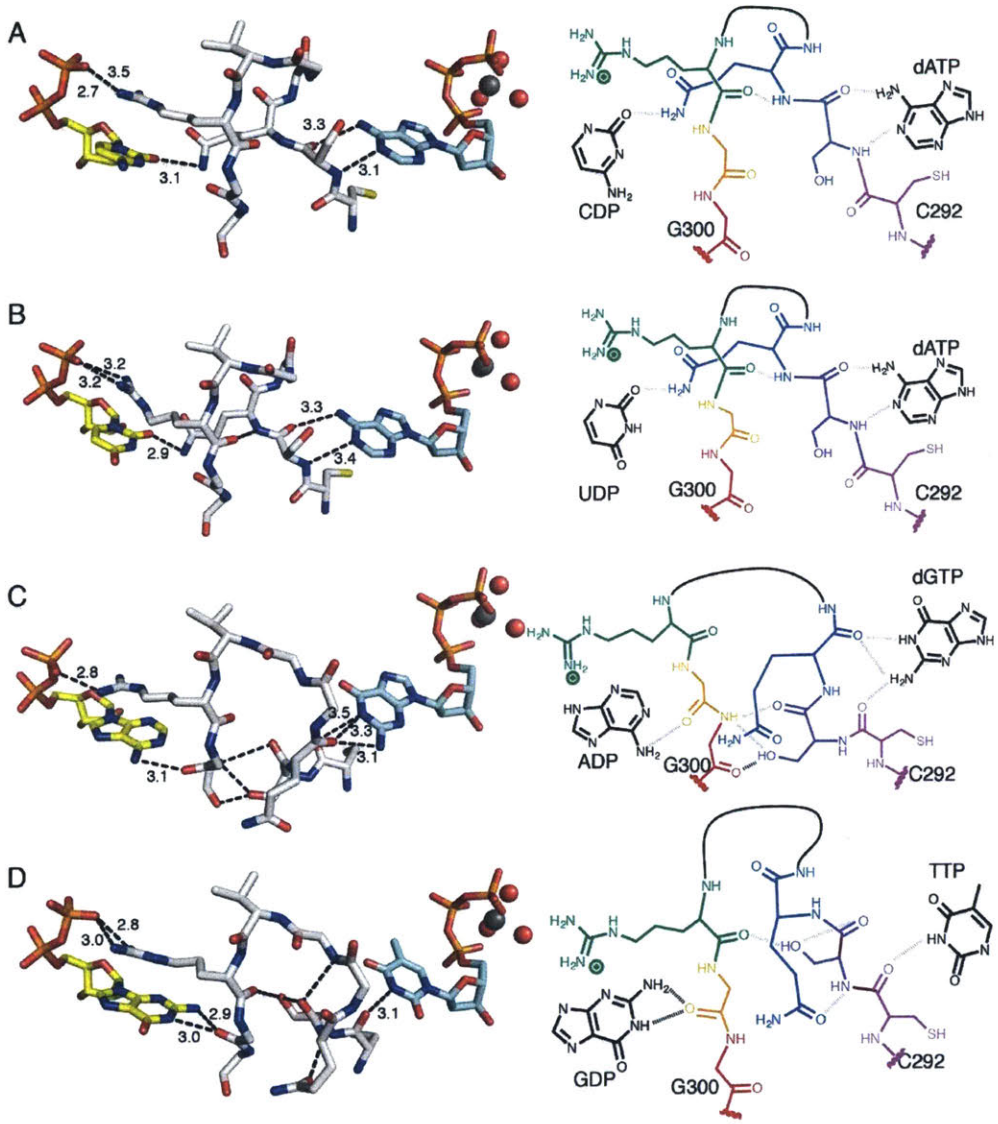


Figure 1.23. Structural basis for communication between the catalytic and S-sites via loop 2 in the *E. coli* class Ia RNR for (**A**) CDP/dATP, (**B**) UDP/dATP, (**C**) ADP/dGTP, and (**D**) GDP/TTP. Three dimensional representations of the structure are shown to the left and two dimensional representations to the right. Carbon atoms in the 3D representations are colored as follows: substrate nucleotide = yellow, protein = white, and effector nucleotide = cyan. All other atoms are colored according to element: red = O, blue = N, orange = P, and gray = Mg. Numbers in the panels on the left are distances in Å. This figure was reproduced with permission under the Creative Commons Attribution License from Zimanyi, C. *et al.* (**2016**) *eLife* **5**: e07141. DOI: 10/7554:eLife07141.

In contrast, loop 2 was ordered in all of the yeast Ia α subunit structures and adopts different conformations that appeared to vary with the identity of the substrate nucleotide rather than the effector.[156] Also noteworthy is the fact that the conserved loop 2 residues Q_{288} and R_{293} of the yeast enzyme appear to not be playing the same roles as they do in the *E. coli* and *T. maritima* RNRs. The sidechain of Q_{288} is pointing into the catalytic site in all cases, and in the dGTP/ADP structure (PDB 2CVX), the amide group forms a hydrogen bond with N1 of the substrate.[156] The sidechain of R_{293} adopts different conformations in most structures that places the guanidinium group at distances too great to allow hydrogen bonding between it and the β-phosphate of the substrate nucleotide. Only in the dGTP/ADP structure does R_{293} reach across the face of the adenine base, but it is still too far to allow a direct hydrogen bond between the sidechain and the substrate β-phosphate.[156] The final noteworthy feature of the yeast Ia α subunit structures relative to those of the *E. coli* and *T. maritima* enzymes is that no or very few direct non-covalent interactions between loop 2, either by residue sidechains or mainchain amides and carbonyls, and the substrate nucleobase are discernable. In some cases, such as the TTP/GDP structure (PDB 2CVW), interactions between the loop and the nucleobase are instead mediated by intervening water molecules.[156]

The differences in the structures of the *E. coli* and *T. maritima* RNRs versus those of the yeast α subunit suggest that there may be different modes by which loop 2 can mediate communication between the catalytic and S-sites for substrate specificity allosteric regulation. Possible support for this hypothesis could be the fact that loop 2 in yeast and other eukaryotic α subunits is longer by one residue compared to the loops of prokaryotic RNRs (**Figure 1.21**). Furthermore, the eukaryotic loops contain a proline residue which possibly may reduce their flexibility and, therefore, force the formation of different interactions between loop 2 and the

substrate and effector nucleotides. However, factors that arose during crystallization or data collection could have prevented loop 2 in the yeast α structures from adopting conformations that are similar to those seen in the *E. coli* and *T. maritima* enzymes. For example, crystal packing might have prevented the yeast α subunit from clamping down on the substrate nucleotide, as was seen with the *E. coli* enzyme,[5] and, therefore, any additional interactions between loop 2 and other regions of the protein that would hold the loop in place were absent. More high-resolution structures of RNRs from other organisms with substrate and effectors bound are required in order to assess whether or not there is a universal structural basis for substrate specificity allosteric regulation in RNRs.

1.5. THE *BACILLUS SUBTILIS* CLASS IB RNR

1.5.1. Important residues in the class Ib RNR from B. subtilis. Arguably the next best characterized Ib RNR after the *S.* Typhimurium system is from *B. subtilis* and, to a lesser extent, *B. anthracis* and *B. cereus*. The latter two bacteria are pathogenic to humans (**Table 1.1**), with *B. anthracis* causing the most concern due to the possibility of its use as a biological weapon. All three species have the ability to form very robust endospores when faced with nutrient limitation.[157] RNR activity is expected to cease since sporulating cells no longer require dNTPs for DNA replication and repair processes, although no studies of this hypothesis have been reported to date. The regulation of the Ib RNRs from these Bacilli should be similar in response to the commitment to sporulation as well as other shared physiological processes and, therefore, the studies of a RNR from one species can provide valuable insights into the function and regulation of the others. The focus of the following discussion will be on the previous studies of the *B. subtilis* RNR as well as what possible influences that this bacterium's lifestyles and nucleotide metabolism might have on the regulation of the enzyme.

Table 1.2. Important residues in the *B. subtilis* class Ib RNR.

Function	Residue number in *B. subtilis*	Equivalent position in *E. coli*
Resting state Y•	Y_{105} (NrdF)	Y_{122} (NrdB)
Metal cluster ligands	D_{66} (NrdF)	D_{84} (NrdB)
	E_{97} (NrdF)	E_{115} (NrdB)
	H_{101} (NrdF)	H_{118} (NrdB)
	E_{164} (NrdF)	E_{204} (NrdB)
	E_{198} (NrdF)	E_{238} (NrdB)
	H_{201} (NrdF)	H_{241} (NrdB)
RT pathway	W_{30} (NrdF)	W_{48} (NrdB)
	E_{34} (NrdF)	E_{52} (NrdB)
	R_{196} (NrdF)	R_{236} (NrdB)
	D_{197} (NrdF)	D_{237} (NrdB)
	Y_{307} (NrdF)	Y_{356} (NrdB)
	D_{298} (NrdF) (?)	E_{350} (NrdB)
	Y_{683} (NrdE)	Y_{730} (NrdA)
	Y_{684} (NrdE)	Y_{731} (NrdA)
Catalytic Cys	C_{382} (NrdE)	C_{439} (NrdA)
Reducing Cys	C_{170} (NrdE)	C_{225} (NrdA)
	C_{409} (NrdE)	C_{462} (NrdA)
C-terminal Cys	C_{695} (NrdE)	C_{754} (NrdA)
	C_{698} (NrdE)	C_{759} (NrdA)
Mechanistically important	N_{380} (NrdE)	N_{437} (NrdA)
	E_{384} (NrdE)	E_{441} (NrdA)
Loop 1 conserved residues	R_{207} (NrdE)	R_{262} (NrdA)
	I_{213} (NrdE)	I_{268} (NrdA)
Loop 2 conserved residues	Q_{239} (NrdE)	Q_{288} (NrdA)
	R_{243} (NrdE)	R_{293} (NrdA)
	G_{245} (NrdE)	G_{300} (NrdA)

As mentioned earlier in this chapter, sequence alignments of the subunits of all class I RNRs reveals very limited conservation. Those residues conserved in all class I systems, however, are absolutely required for RNR activity and substrate specificity allosteric regulation, including the di-metal cluster ligands and radical-harboring Tyr residue in β, the RT pathway residues in both subunits, and the mechanistically important residues for NDP reduction and specificity regulation in α.[6] The residues that are critical for the function and regulation of the *B. subtilis* Ib RNR, based on sequence alignment, are shown in **Table 1.2** along with their equivalents in the paradigm *E. coli* Ia enzyme. The functions of some residues have been verified experimentally, such as the metal ligands by the X-ray structure of NrdF (PDB 4DRO),[41] Y_{105} as the site of the Y• in the resting state using unnatural amino acids (Ravichandran and Stubbe, unpublished), and C_{382} as the catalytic Cys by site-directed mutagenesis (see Chapter 4). Sequence alignment of loop 2 from *B. subtilis* NrdE with other α subunits reveals that it more closely resembles the loops of the *E. coli* class Ia and *T. maritima* class II enzymes (**Figure 1.21**)[5, 24] and, therefore, will likely adopt similar conformations to those shown in **Figure 1.23** for substrate specificity regulation.

1.5.2. Assembly of the Mn(III)$_2$-Y• cofactor in B. subtilis NrdF. The *B. subtilis* RNR has been characterized the best in terms of the assembly and properties of its Mn(III)$_2$-Y• cofactor. An initial report examining ribonucleotide reduction in *B. subtilis* suggested that its RNR might be a Mn-dependent enzyme,[158] which was later confirmed[21] following the discovery that NrdI was required for Mn(III)$_2$-Y• assembly in *E. coli* NrdF *in vitro*[15] and the isolation of endogenous *E. coli* and *C. ammoniagenes* NrdF loaded with the Mn(III)$_2$-Y• cofactor.[18, 19] In the 2011 study by Zhang and Stubbe,[21] a derivative of *B. subtilis* JH624 (termed 1B-UP) was engineered to overexpress the Ib RNR by cloning the operon into the *amyE* locus and placing it under the control of an isopropyl-β-D-thiogalactopyranoside (IPTG)-inducible promoter. Quantitative Western

blotting analysis, using antibodies raised against the recombinant His_6-tagged proteins, revealed that, relative to wildtype, NrdE and NrdF expression was increased 35-fold and NrdI expression 80 – 160-fold in the 1B-UP strain grown in the presence of 1 mM IPTG. Isolation and subsequent characterization of endogenous *B. subtilis* NrdF from the 1B-UP strain by EPR and atomic absorption spectroscopy revealed that the enzyme did indeed incorporate $Mn(III)_2$-Y•, although the protein, most likely due to non-physiological overexpression, was substoichiometrically loaded with the cofactor (~0.45 Y• and ~2 Mn/β_2).

Several observations made in the study by Zhang and Stubbe are worth mentioning.[21] In addition to examining the 1B-UP strain, quantitative Western blotting analysis was also performed with the wildtype bacterium. The results of these analyses revealed that NrdE and NrdF were expressed in an approximately 1:1 ratio in both strains, whereas NrdI levels were sub-stoichiometric relative to the larger subunits (~8-fold less in 1B-UP grown with 1 mM IPTG). This observation indicates that NrdI acts catalytically during cluster biosynthesis. The Western blotting results with wildtype *B. subtilis* also indicated that NrdE and NrdF were present at final concentrations of ~1 μM, which is consistent with more recently reported 2D gel electrophoresis and targeted mass spectrometry studies by Dörte Becher's group that found NrdE and NrdF concentrations fall within the range of 0.1 – 1 μM (see Appendix 2).[159-161] However, one of the most interesting observations made by Zhang and Stubbe, which served as the initial motivation for the studies presented in this dissertation, was the co-purification of endogenous *B. subtilis* NrdE and NrdF through a three-step purification procedure in roughly a 1:1 ratio.[21] Moreover, this ratio was maintained despite the fact that no exogenous Mg^{2+} or nucleotides were added to the buffers during the purification. The apparent strength of the interaction suggested that the *B. subtilis* RNR might be stable enough to allow crystallographic characterization of the active enzyme.

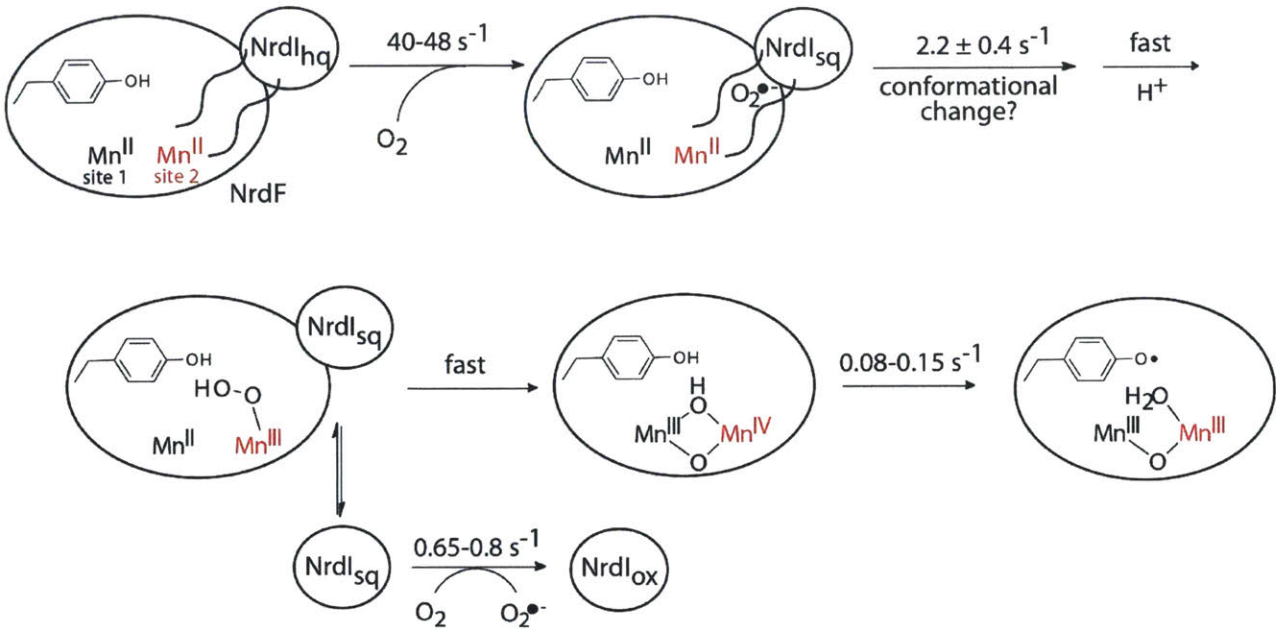


Figure 1.24. Proposed mechanism and a summary of the kinetics for $Mn(III)_2$-Y• assembly in *B. subtilis* NrdF *in vitro*. This figure is reproduced with permission from Cotruvo, J.A., Jr.; Stich, T.A.; Britt, R.D.; Stubbe, J. (**2013**) Mechanism of Assembly of the Dimanganese-Tyrosyl Radical Cofactor of Class Ib Ribonucleotide Reductases: Enzymatic Generation of Superoxide is Required for Tyrosine Oxidation via a Mn(III)Mn(IV) Intermediate. *J. Am. Chem. Soc.* **135**: 4027-4039. Copyright 2013 American Chemical Society.

In addition to the isolation of endogenously-sourced NrdF, the 2011 study of the *B. subtilis* RNR also reported the ability to reconstitute recombinant NrdF with $Mn(III)_2$-Y• using $NrdI_{hq}$.[21] A yield of ~0.6 Y•/β_2[21] was the highest *in vitro* radical loading achieved for a Mn-reconstituted NrdF at the time (0.3 – 0.4 Y•/β_2 had been reported previously for the *E. coli*[15] and *B. anthracis*/*B. cereus*[28] NrdFs). This result prompted former Stubbe group member Joseph Cotruvo to use *B. subtilis* NrdF to study the mechanism of $Mn(III)_2$-Y• assembly using stopped-flow UV-visible and rapid freeze-quench EPR spectroscopies to respectively monitor changes in the oxidation state of the NrdI-bound FMN cofactor and the formation and decay of paramagnetic intermediates.[17] To maximize the signal from different intermediates, the $NrdI_{hq}$-$Mn(II)_2$-NrdF interaction strength was estimated (K_d = 0.6 μM), and the subsequent rapid kinetic experiments were run under conditions in which > 95% of $NrdI_{hq}$ was in complex with NrdF. The proposed mechanism for

$Mn(III)_2$-Y• assembly that resulted from the study by Cotruvo *et al.*[17] is shown in **Figure 1.24**, and the data specifically support the following features. (1) $NrdI_{hq}$ rapidly reduces O_2 by one electron to generate $O_2^{\bullet -}$ and $NrdI_{sq}$. (2) Oxidation of the $Mn(II)_2$ cluster by $O_2^{\bullet -}$ to a Mn(III)Mn(IV) intermediate occurs at approximately an order of magnitude slower rate relative to $O_2^{\bullet -}$ generation. This process is likely slower due to channeling of the oxidant to the dimanganous site and re-arrangements of residue sidechains upon binding of $O_2^{\bullet -}$ to the metal cluster. (3) Oxidation of Y_{105} to Y_{105}• by the Mn(III)Mn(IV) intermediate occurs at approximately a two orders of magnitude slower rate relative to $O_2^{\bullet -}$ generation.

An outstanding issue regarding the assembly of $Mn(III)_2$-Y• in NrdF for all class Ib RNRs has been the identity of the endogenous reductant of NrdI. The utilization of this enigmatic protein to recycle NrdI during *in vitro* reconstitution of the $Mn(III)_2$-Y• cofactor could potentially improve the yields of fully-activated NrdF and, thus, facilitate further studies of the mechanism of cluster assembly as well as of the activity and regulation of the class Ib RNRs in general. In the case of *B. subtilis*, a recent microbiological study has tentatively established that the generic flavodoxin reductase, termed YumC, is the physiological reductant of NrdI.[162] This study, along with a preliminary experiment in which YumC was tested for its ability to support $Mn(III)_2$-Y• assembly in *B. subtilis* NrdF, are described in Appendix 5 of this dissertation. Furthermore, during the preparation of this dissertation, it was reported that the YumC homolog of *B. cereus* was able to support synthesis of $Mn(III)_2$-Y• in this bacterium's NrdF,[163] thus providing direct evidence that generic flavodoxin reductases, at least in the Bacilli, are the likely endogenous reductants of NrdI.

1.5.3. Influence of the lifestyles of B. subtilis on RNR regulation. B. subtilis can inhabit a diverse range of environments. Originally described as strict soil-dwelling bacterium, different strains of *B. subtilis* have been more recently isolated from aquatic environments, in biofilms

adhered to the roots of plants, and as vegetative cells in the gastrointestinal (GI) tracts of animals and humans.[164-166] Emerging evidence suggests that the latter two associations are likely symbiotic as *B. subtilis* has been shown to outcompete pathogens for colonization of GI tracts and root surfaces, and, in the case of plants, the bacterium is suggested to play a role in making certain nutrients more readily available for the plant host.[164, 167] Comparison of sequenced genomes of different strains of *B. subtilis* reveal high levels of genomic synteny and proteome conservation for core structural and metabolic processes, but strain-specific regions that encode different genes for developmental processes and biosynthesis of a diversity of secondary metabolites.[166, 168, 169] *B. subtilis* also has the remarkable ability to differentiate into subpopulations of phenotypically distinct, but genotypically identical cell types, including spore and biofilm formers, cells competent in the uptake of extracellular DNA, and motile cells.[40] The coexistence of these subpopulations allows *B. subtilis* to adapt to particular environmental niches within a community of other microorganisms.

From the perspective of RNR activity and regulation in *B. subtilis*, a couple of aspects of its lifestyle deserve comment. One of the most interesting observations is the ability of the bacterium to grow under anaerobic conditions using nitrate as the terminal electron acceptor for respiration, or by fermentation.[170, 171] Indeed, the capability of growing anaerobically is absolutely required in order for the bacterium to be able to survive and thrive in the low oxygen environment of animal GI tracts.[164-166] The surprising, yet puzzling aspect of the ability of *B. subtilis* to grow anaerobically is that its class Ib RNR, or at the very least NrdE, is absolutely required under these conditions.[172] Furthermore, in all reported genomes of this organism, the genes for either a class II or class III RNR are absent.[6]

The requirement for NrdE in the anaerobic growth of *B. subtilis* was established in a 2006 study by Elisabeth Härtig and coworkers.[172] They used a previously identified temperature-sensitive *B. subtilis* mutant, *ts*-A13, in which the causative mutation had been mapped to the *nrdE* gene; the mutation allows *ts*-A13 to grow at 30 °C, but not at 45 °C.[173, 174] When early exponential phase cultures of *ts*-A13 grown anaerobically in minimal medium at 30 °C were shifted to the non-permissive temperature, growth ceased and live cell counts steadily decreased over a 3 h timeframe. The growth of these cultures could be rescued and restored to wildtype levels by adding deoxynucleosides to the growth medium at any point after the temperature shift, thus indicating that NrdE was required for anaerobic growth of the bacterium. Härtig *et al.* moreover identified two partially conserved operator sequences for ResD, the master transcription factor regulating the expression of genes required for anaerobic growth,[175, 176] immediately upstream of the RNR operon.[172] Introduction of consensus residue mutations into the *ResD* operator closest to the site of transcriptional initiation abolished ResD-mediated regulation of the Ib RNR operon in response to environmental O_2 concentrations.

The results of the study by Härtig *et al.*[172] have left quite an interesting paradox in their wake: how does the aerobic class Ib RNR function in the absence in the absence of O_2? Once assembled, the $Mn(III)_2$-Y• cofactor could, in theory, catalyze an infinite number of rounds of NDP reduction without being inactivated. Realistically, however, the radical will be quenched, either inadvertently or by an as-yet uncharacterized maintenance pathway that controls RNR activity via reduction of the Y•. RNR will need to be reactivated once the demand for dNTPs increases and, thus, the requirement for O_2 will again present itself. Furthermore, cell division during growth will eventually dilute out all the activated RNR that was initially present in the first

generation of cells. Therefore, later generations of *B. subtilis* will need to be able synthesize more $Mn(III)_2$-Y•-loaded NrdF in order for culture growth to continue once this point is reached.

Currently, there are no obvious answers for how *B. subtilis* will obtain O_2 under anaerobic conditions in order to synthesize $Mn(III)_2$-Y•-loaded NrdF and, consequently, an active RNR. There are some possible alternatives that could allow *B. subtilis* to survive and reproduce under anaerobic conditions for relatively long periods of time without requiring an active RNR. Perhaps the most plausible route is that *B. subtilis* acquires its dNTPs using its extensive nucleotide salvage pathways (discussed later) to take up deoxynucleosides from the extracellular environment. The GI tracts of animals are likely rich in these metabolites and, therefore, the salvage pathways likely play a critical role in the ability of *B. subtilis* to inhabit such environments. Ultimately, if *B. subtilis* cannot acquire dNTPs from *de novo* or salvage pathways, it can undergo the process of sporulation and lie dormant until more favorable conditions are encountered.

As mentioned earlier, *B. subtilis* undergoes sporulation when its environment is depleted of nutrients,[157] and it is expected that RNR activity will cease once the bacterium has committed to this process. In support of the latter idea, a proteomics study identified NrdE as one of the 200 central metabolic proteins that undergoes significant proteolysis in *B. subtilis* cultures entering stationary phase with glucose exhaustion, conditions under which sporulation can be triggered.[119] Another facet of sporulation that could affect RNR regulation is the dependence of the former, in part, on the intracellular GTP/GDP concentrations.[177] Under starvation conditions, many bacteria will produce the alarmones guanosine-3′,5′-bis(diphosphate) (ppGpp) and guanosine-3′-di-phosphate-5′-triphosphate (pppGpp) from GDP or GTP, respectively, and ATP.[178] These alarmones initiate signaling cascades that shut off anabolic processes, such as ribosomal and purine/pyrimidine biosynthesis, and upregulate adaptive responses to nutrient limitation.[179] As a

consequence of producing the alarmones, intracellular GTP/GDP concentrations undergo a dramatic decrease, and in the case of *B. subtilis*, a sharp decrease in the guanosine pools are thought to be a major trigger for initiating the sporulation process.[177] From the perspective of RNR, the production of ppGpp depletes the cell of one of the substrates of the enzyme (GDP), which could lead to perturbed dNTP pools if RNR is not properly regulated. However, it is likely that additional layers of regulation are in place in *B. subtilis* such that RNR activity can be attenuated or abolished in response to the alarmones and/or other factors that cause cell differentiation processes such as sporulation. This area of RNR regulation is unexplored since very few enzymes have been characterized from organisms that can form spores, but warrants further investigation.

1.5.4. Some strains of B. subtilis encode a second, viral-associated class Ib RNR. One of the greatest sources of variation between *B. subtilis* genomes results from the presence versus absence of latent prophages or prophage-like regions, which are distinguished from the host genome by their local A:T rich content.[180] Such regions encode genes for which the function may be beneficial for the bacterium, such as antibiotic biosynthesis and/or detoxification. Often, however, these genes encode functions that are redundant in the host. Pertinent to the work presented in this dissertation is the unusual observation that the genomes of many *B. subtilis* strains, including the standard laboratory strain 168, encode two class Ib RNR operons.[6] These strains are host to a lysogenic prophage called SPβ that encodes one of these RNRs. To distinguish the two operons, the genes belonging to SPβ will be denoted with the letter "B" after the name of the gene and protein (i.e. *nrdEB* and NrdEB).

The organization and members of each operon are distinct as shown in **Figure 1.25**. Both encode α and β subunits as well as the flavodoxin NrdI (NrdIB) that is required for the assembly of $Mn(III)_2$-Y• in β. Notably, the *nrdEB* and *nrdFB* genes are interrupted by introns, and *nrdEB*

is further split by a segment of DNA that encodes an intein. Each operon has a different fourth member associated with it. NrdHB (also known as YosR), a glutaredoxin-like thioredoxin that associated with many class Ib RNRs in other organisms,[6] is encoded in the SPβ operon. The host's operon instead encodes 621 nt gene called *ymaB*, the protein product of which has limited homology with Nudix hydrolases (discussed in **Section 1.6**). Substantial evidence from previous genetic,[174, 181, 182] transcriptomic,[183, 184] and bioengineering[180] studies revealed *B. subtilis* is viable when the SPβ operon is inactivated or removed from the genome all together. Furthermore, the temperature-sensitive *B. subtilis* mutant described in the previous section had an intact SPβ RNR operon.[172-174] The fact that the SPβ Ib RNR was unable to support growth of the mutant at the non-permissive temperature further supports the non-essentiality of this enzyme for *B. subtilis* survival.

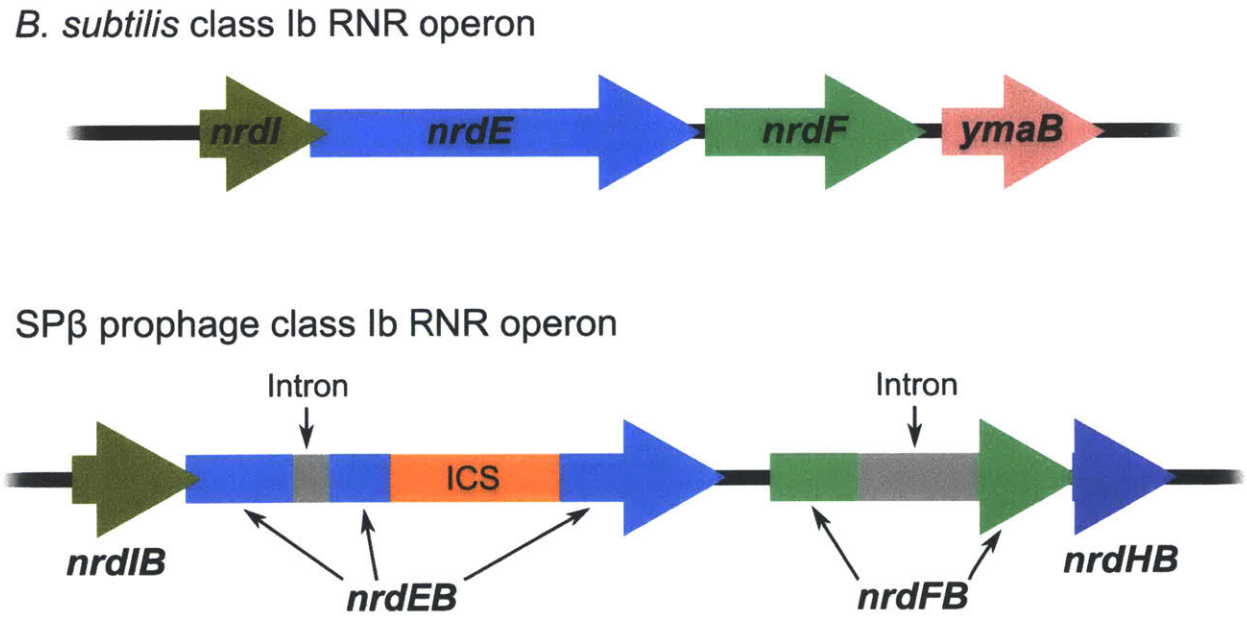


Figure 1.25. The class Ib RNR operons encoded in the genome of *B. subtilis*. ICS = intein coding sequence.

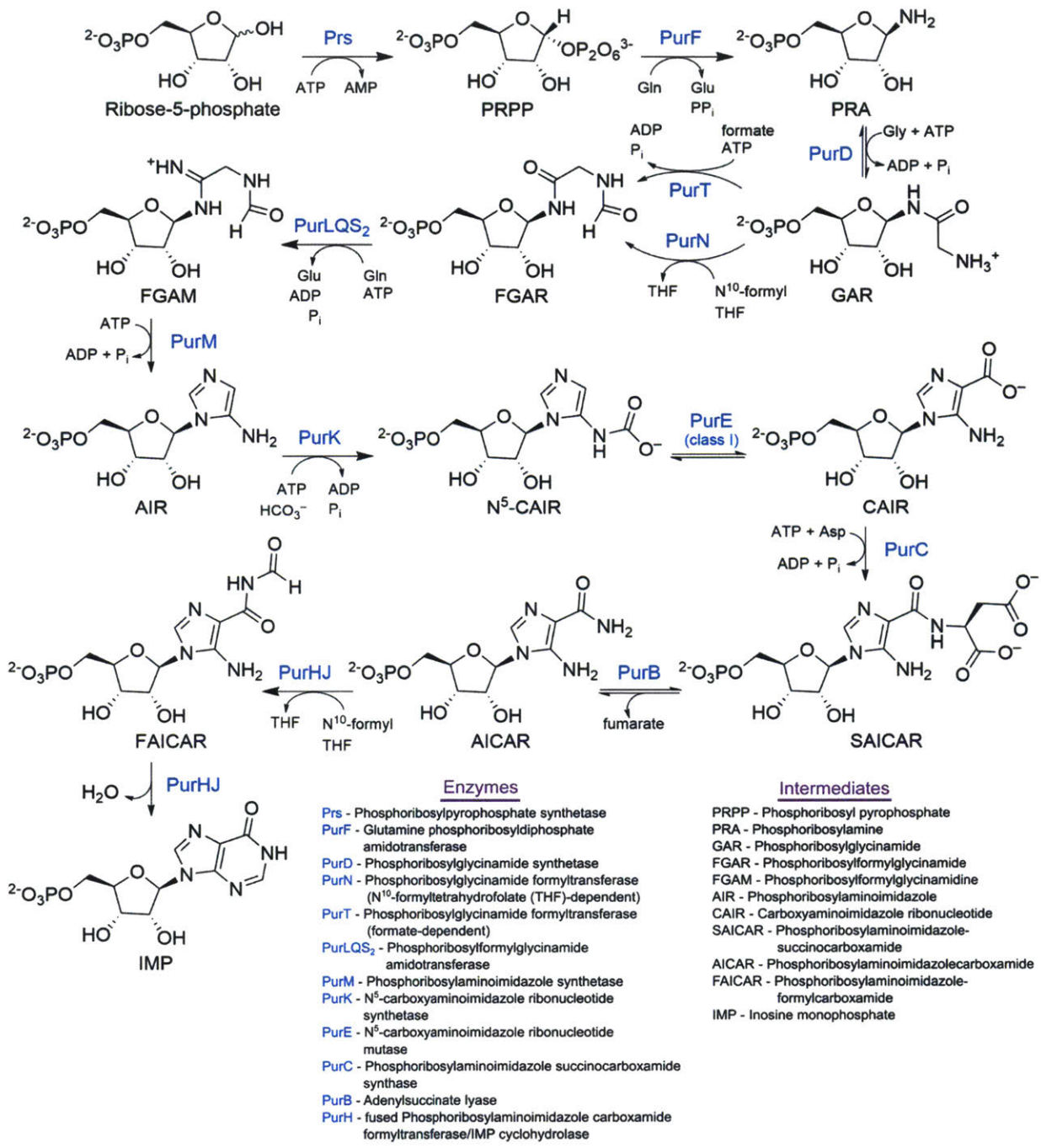


Figure 1.26. Purine biosynthetic pathway in *B. subtilis*.

1.5.5. Nucleotide metabolism in B. subtilis. Nucleotides utilized in DNA and RNA biosynthesis in *B. subtilis* can be derived from either *de novo* pathways or, as mentioned above, from salvage of nucleobases and (deoxy)nucleosides resulting from DNA/RNA turnover or import

from the extracellular environment. The eleven-step *de novo* purine and six-step *de novo* pyrimidine biosynthetic pathways present in *B. subtilis* are depicted in **Figure 1.26** and **1.27**, respectively. Like some other bacteria, such as *E. coli*,[185] *B. subtilis* encodes two different enzymes, N^{10}-formyltetrahydrofolate-dependent PurN and formate-dependent PurT, that can catalyze formylation of phosphoribosylglycinamide in the third step in purine biosynthesis (**Figure 1.26**).[186] *purN* is part of a 12 gene operon encoding the other purine biosynthetic enzymes.[187] The operon is regulated by the transcriptional repressor, PurR, and a guanine-responsive ribo-switch.[187-190] Purine nucleotides, adenine, adenosine, and/or guanine repress the expression of the *pur* operon, whereas phosphoribosyl pyrophosphate (PRPP) enhances its transcription. In contrast, *purT* is likely expressed as a monocistronic mRNA, and its regulation in response to elevated levels of purines or precursors (e.g. PRPP) has not been examined.

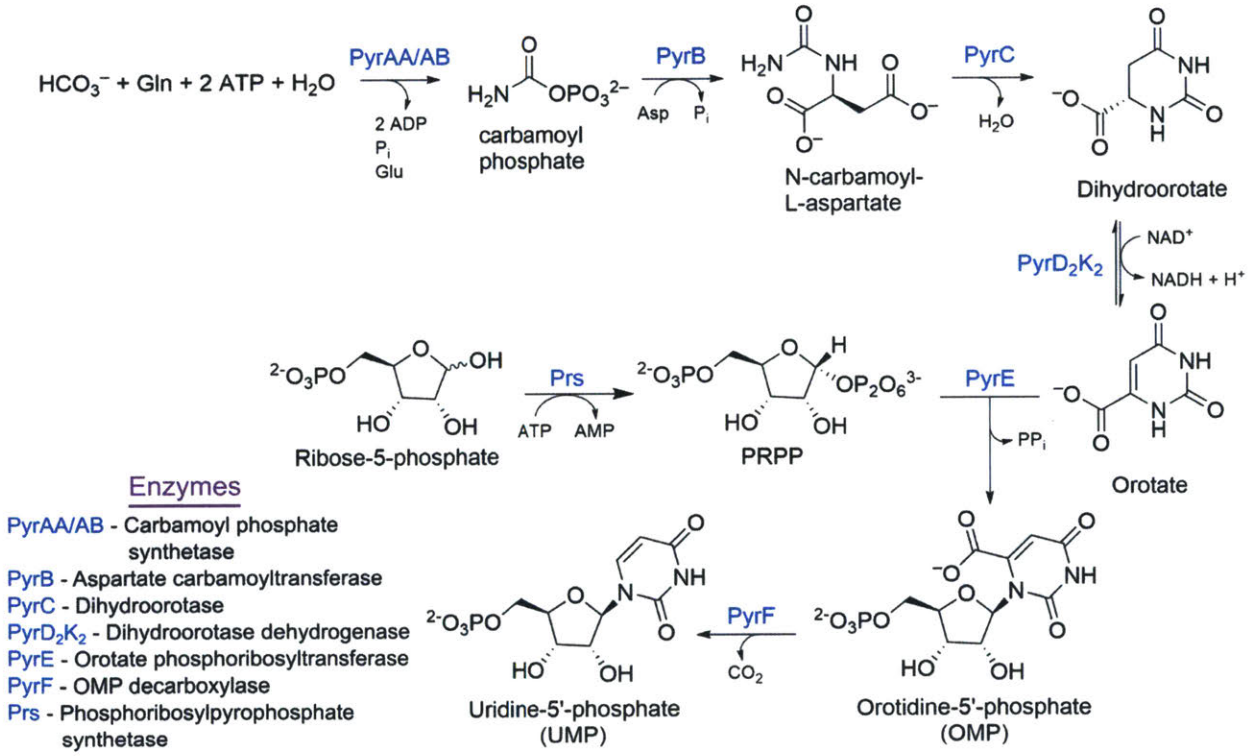


Figure 1.27. Pyrimidine biosynthetic pathway in *B. subtilis*.

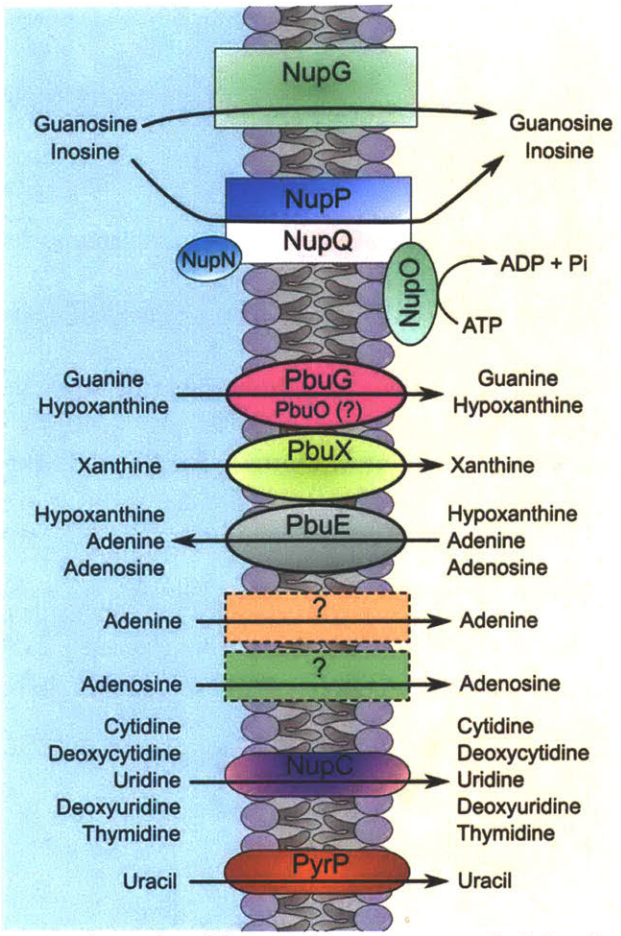
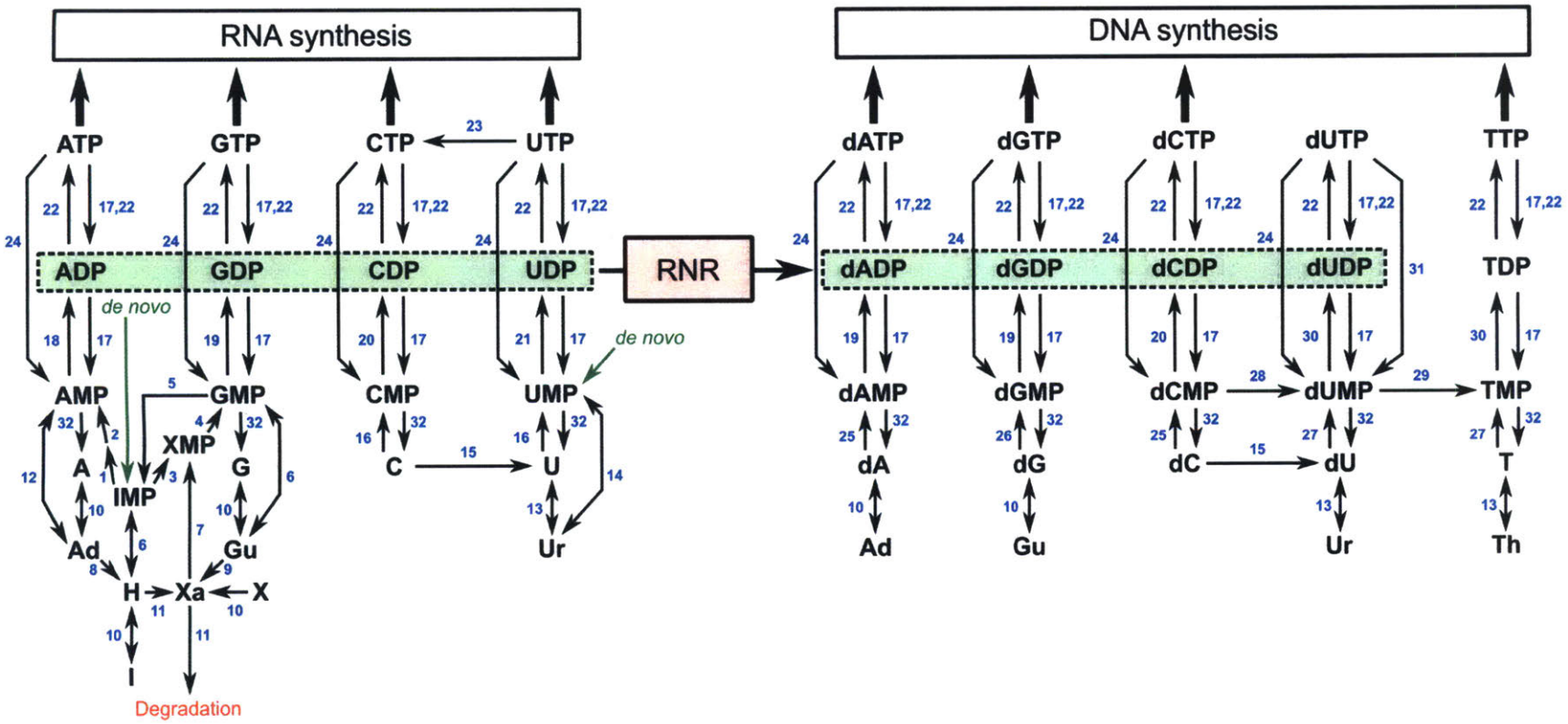


Figure 1.28. Nucleobase and nucleoside transporters encoded in the *B. subtilis* genome that participate in salvage pathways for nucleotide biosynthesis. The left side of the membrane represents the extracellular environment and the right side the cytosol of the cell. PbuO is predicted to be a second guanine/hypoxanthine transporter based on its similarity to PbuG (47% identical, 64% with positive substitution) as well as being regulated by the *pur* operon transcriptional repressor, PurR.[191] The transport of adenine and adenosine use as-yet unidentified proteins that are purportedly distinct from those listed for guanine- and guanosine-specific uptake.

Figure 1.29 (opposite page). Nucleotide metabolic pathways in *B. subtilis*. Except for RNR (red box in the middle of the figure), numbers are used to indicate the enzyme(s) that catalyze the given reactions. Legends for the enzymes and intermediates are listed at the bottom of the figure.



Nucleotides

NMP/dNMP = nucleoside/deoxynucleoside 5'-monophosphate
NDP/dNDP = nucleoside/deoxynucleoside 5'-diphosphate
NTP/dNTP = nucleoside/deoxynucleoside 5'-triphosphate

Nucleobases

Ad = adenine
Gu = guanine
H = hypoxanthine
Ur = uracil
Th = thymine
Xa = xanthine

Nucleosides

A/dA = adenosine/deoxyadenosine
C/dC = cytidine/deoxycytidine
G/dG = guanosine/deoxyguanosine
I = inosine
T = deoxythimidine
U/dU = uridine/deoxyuridine
X = xanthosine

Enzymes

1 - Adenylosuccinate synthetase (PurA)
2 - Adenylsuccinate lyase (PurB)
3 - IMP dehydogenase (GuaB)
4 - GMP synthetase (GuaA)
5 - GMP reductase (GuaC)
6 - Guanine/hypoxanthine phosphoribosyltransferase (HprT)
7 - Xanthine phosphoribosyltransferase (Xpt)
8 - Adenine deaminase (AdeC)
9 - Guanine deaminase (GuaD)
10 - Purine nucleoside phosphorylase (PupG, DeoD)
11 - Xanthine dehydrogenase (PucABCDE)
12 - Adenine phosphoribosyltransferase (Apt)
13 - Pyrimidine nucleoside phosphorylase (Pdp)
14 - Uracil phosphoribosyltransferase (Upp, PyrR)
15 - Cytidine/deoxycytidine deaminase (Cdd)
16 - Uridine kinase (Udk)
17 - Specific/non-specific phosphatases (YsnA (?))
18 - Adenylate kinase (Adk)
19 - Guanylate kinase (Gmk)
20 - Cytidylate kinase (Cmk)
21 - Uridylate kinase (PyrH)
22 - Nucleoside diphosphate kinase (Ndk)
23 - CTP synthase (PyrG)
24 - (Deoxy)nucleoside triphosphatase, Mn^{2+} dependent (YpgQ)
25 - Deoxyadenosine/deoxycytidine kinase (Dck)
26 - Deoxyguanosine kinase (Dgk)
27 - Deoxythymidine kinase (Tdk)
28 - dCMP deaminase (ComEB)
29 - Thymidylate synthase (ThyA, ThyB)
30 - Thymidylate kinase (Tmk)
31 - Deoxyuridine triphosphatase (YcnF, YosS)
32 - 5'-Nucleotidase (YcsE, YktC)

The entire set of pyrimidine biosynthetic genes are also transcribed on a single poly-cistronic mRNA along with *pyrR*, which encodes a transcriptional anti-terminator/uracil phosphoribosyltransferase, and *pyrP*, which encodes a putative uracil permease (**Figure 1.28**).[192] PyrR regulates the transcription of the *pyr* operon by binding to an anti-terminator stem-loop in the 5′-untranslated region of the mRNA, which is thought to allow other terminator stem-loops to form and, thus, leads to *pyr* mRNA degradation.[193] Silencing of the expression of the *pyr* operon by PyrR is enhanced by UMP and attenuated by PRPP.[193]

B. subtilis possesses an extensive system of transporters that allow the bacterium to import extracellular nucleobases and (deoxy)nucleosides for salvage or use as a source of nitrogen (**Figure 1.28**). The identity of most of the transporters was established in genetic and mutational studies, and currently there is very little biochemical or structural information available for any of the systems shown in **Figure 1.28**. For pyrimidine salvage, indirect evidence from several studies suggest that a single transporter, NupC, is responsible for the transport of all pyrimidine (deoxy)nucleosides,[194-196] whereas uracil is the only pyrimidine nucleobase for which uptake data is currently available.[145, 197] Transport of uracil into *B. subtilis* is currently hypothesized to be mediated by the annotated permease, PyrP, due to its sequence homology (66% identity, 81% with positive substitutions) to the characterized homologous protein from the thermophile *Bacillus caldolyticus*.[198]

For purine salvage, published data supports the presence of distinct systems for the transport of adenine, guanine, and xanthine as well as the nucleoside derivatives of the former two bases (**Figure 1.28**).[194, 197, 199-202] Furthermore, the results of experiments examining the uptake and competition between radiolabeled bases and nucleosides for *B. subtilis* transporters indicate that the guanine- and guanosine-specific systems can also efficiently take up the purine base

hypoxanthine and the nucleoside inosine, respectively.[197, 201] To date, the proteins responsible for mediating adenine and adenosine uptake have not been identified. A recent study by Belitsky and Sonenshein on the ability of *nupG* and *nupN* mutants to import [^{3}H]-guanosine led the authors to suggest that adenosine may be able to be taken up into *B. subtilis*, in part, by these two systems.[199] They observed that the uptake of [^{3}H]-guanosine (0.14 μM) by NupG (in a *nupN* knockout mutant) and NupNOPQ (in a *nupG* knockout mutant) was significantly reduced in the presence of 50 μM adenosine and inosine, and to a lesser extent by 50 μM xanthosine. Transport of [^{3}H]-guanosine was also slightly reduced by NupNOPQ in the presence of 50 μM cytidine and uridine. These results led Belitsky and Sonenshein to suggest that the ATP-binding cassette transporter NupNOPQ was a general nucleoside importer whereas NupG was specific for purine nucleosides.[199] The concentrations of the other nucleosides used to inhibit [^{3}H]-guanosine uptake in this study are likely not physiologically relevant for *B. subtilis* in its natural environments; thus, utilization of these transporters by the bacterium as general nucleoside or purine nucleoside uptake systems seems improbable.

Inosine and uridine 5′-monophosphates derived from *de novo* biosynthesis and any nucleobases or (deoxy)nucleosides taken up from the extracellular environment enter the general nucleotide and deoxynucleotide pools through the pathways shown in **Figure 1.29**. Since *B. subtilis* does not possess any ribonucleoside kinase activities,[203] the bases of imported ribonucleosides are cleaved from the sugar moiety by pyrimidine and purine nucleoside phosphorylases (PupG, DeoD, and Pdp),[204, 205] regardless of whether these molecules are needed for nucleotide synthesis or for nitrogen. For nucleotide synthesis, adenine, guanine/ hypoxanthine, xanthine, and uracil are introduced into the nucleotide pools by the actions of phosphoribosyltransferases (Apt, HprT, Xpt, and Upp or PyrR),[206, 207] which catalyze the transfer of a base

to the 1′ position of PRPP to yield the corresponding nucleoside monophosphate (NMP) and pyrophosphate. In contrast, the presence of deoxyadenosine/deoxycytidine (Dck),[208, 209] deoxyguanosine (Dgk),[208] and deoxythymidine kinases (Tdk)[210] allows imported deoxynucleosides to be phosphorylated to the corresponding deoxynucleoside 5′-monophosphates (dNMPs) and be added directly into their respective pools rather than being immediately broken down.

Once NMPs and dNMPs are generated, their phosphorylation state can be increased or decreased by the kinases and phosphatases indicated in **Figure 1.29**. As ribonucleotides, both purines and pyrimidines can be interconverted to help balance the intracellular ribonucleotide pools. Interconversion of purines between adenine and guanine is made possible by the presence of adenine deaminase (AdeC) and GMP reductase (GuaC),[203] whereas CTP synthase (PyrG) and cytidine/deoxycytidine deaminase (Cdd),[211] with its ability to deaminate cytidine, allows pyrimidines to be interchanged between cytosine and uracil. In contrast, no enzymes capable of interconverting purine deoxynucleotides are present, and Cdd and dCMP deaminase (ComEB) only produce deoxyuridine analogs to feed into TTP biosynthesis. These differences highlight the importance of the allosteric regulation of the *B. subtilis* Ib RNR in maintaining the amounts and ratios of intracellular dNTPs.

As described earlier in this chapter, dNTP pool maintenance in cells also involves the deaminases of deoxycytidine, which are allosterically stimulated by dCTP and inhibited by TTP. Shown in **Figure 1.30** is a summary of the current understanding of the allosteric regulation of the *B. subtilis* Ib RNR, as concluded from the studies presented in this dissertation, and the cytidine/deoxycytidine and dCMP deaminases that helps this bacterium maintain the proper amounts and ratios of intracellular dNTPs. Comparison of **Figure 1.30** with **Figure 1.16** shows that the *E. coli* Ia and *B. subtilis* Ib RNRs exhibit similar schemes in terms of substrate specificity

and overall activity regulation. The major difference between the two systems is that the *E. coli* enzyme can also be inhibited by the combination of ATP and TTP,[82] whereas *B. subtilis* RNR exhibited no such inhibition as shown in Chapter 3. In contrast with the *E. coli* dCTP deaminase, less is currently known about the deoxycytidine analog deaminases of *B. subtilis*. dCMP deaminase has been shown to be allosterically stimulated by dCTP,[145] but inhibition by TTP has not yet been reported. No studies of the allosteric regulation of dC deaminase have been reported.

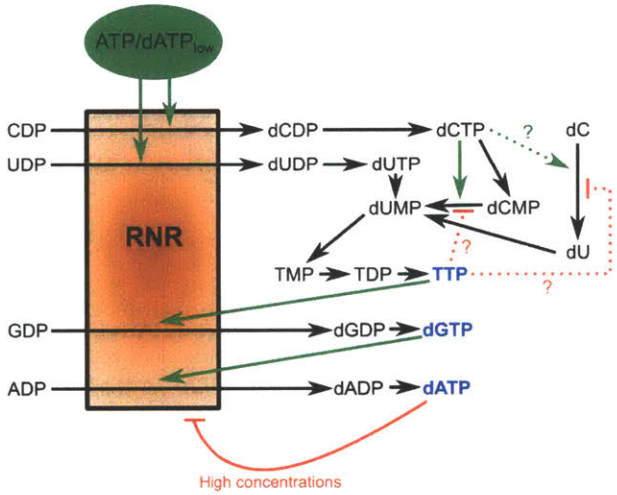


Figure 1.30. Summary of the allosteric regulation of the Ib RNR and cytidine analog deaminases that help control intracellular dNTP pools in *B. subtilis*.

Recently, studies of nucleotide metabolism in human cell lines indicate that cycling between (deoxy)nucleotides and (deoxy)nucleosides, via (deoxy)nucleoside kinases and 5′-nucleotidases, respectively, can also serve as a mechanism for regulating (d)NTP pools.[113] Although this mechanism has not yet been demonstrated in prokaryotes, *B. subtilis* encodes several characterized and putative enzymes, including the phosphoribosyl transferases and deoxy-nucleoside kinases described above, that suggest that (d)NTP pool cycling could also be operative in this bacterium. To remove nucleotides from the pool, a number of pyrophosphatases, such as the recently characterized Mn^{2+}-dependent (d)NTP pyrophosphatase YpgQ,[212] are present that could hydrolyze (d)NTPs to their corresponding monophosphates. As shown in Chapter 5 of this

dissertation, the fourth protein encoded in the *B. subtilis* RNR operon, YmaB, was found to possess dNTP pyrophosphohydrolase activity and, thus, could be involved in dNTP pool regulation as well. Subsequent conversion of (d)NMPs to the corresponding nucleosides can be accomplished by one of the several recently characterized 5′-nucleotidases.[213] Finally, nucleotide surpluses, at least in the case of purines, can further be dealt with by efflux of the bases adenine and hypoxanthine and the nucleoside adenosine into the extracellular environment by the transporter PbuE (**Figure 1.28**).[214, 215]

B. subtilis can also break down nucleotides and use the sugar for carbon and/or energy and the nucleobases for nitrogen when other, more easily utilized sources have been exhausted or are unavailable. Cleavage of the base from nucleosides by phosphorylases (PupG, DeoD, and Pdp, **Figure 1.29**) releases ribose- or deoxyribose-1-phosphate, both of which can be converted to the corresponding 5-phosphate isomer by the action of phosphopentomutase (Drm).[205] Ribose-5-phosphate enters central carbon metabolism via the non-oxidative branch of the pentose phosphate pathway, whereas deoxyribose-5-phosphate is cleaved by deoxyribose aldolase (Dra) into acetaldehyde and glyceraldehyde-3-phosphate, the latter of which can then enter glycolysis/gluconeogenesis.[205] Nitrogen, in the form of ammonia, can be extracted from nucleotides by deamination of cytidine/ deoxycytidine analogs (Cdd and ComEB), adenine (AdeC), and guanine (guanine deaminase, GuaD), and if more is needed, *B. subtilis* can degrade purine, but not pyrimidine, bases (**Figure 1.31**). The pathway of purine degradation utilized by *B. subtilis* is similar to that of other bacteria.[216] The final step catalyzed by uredoglycine-glyoxylate aminotransferase allows *B. subtilis* to obtain glycine without deaminating another amino acid as well as to detoxify the glyoxylate that is produced from the spontaneous breakdown of uredoglycine.[217] All of the purine degradation enzymes listed in **Figure 1.31** are encoded in a cluster along with

permeases for uric acid (PucJ and PucK) and allantion (PucI), and a transcriptional regulator (PucR).[218] The expression of these enzymes is enhanced when *B. subtilis* is grown in the presence of uric acid and/or allantoin.[218]

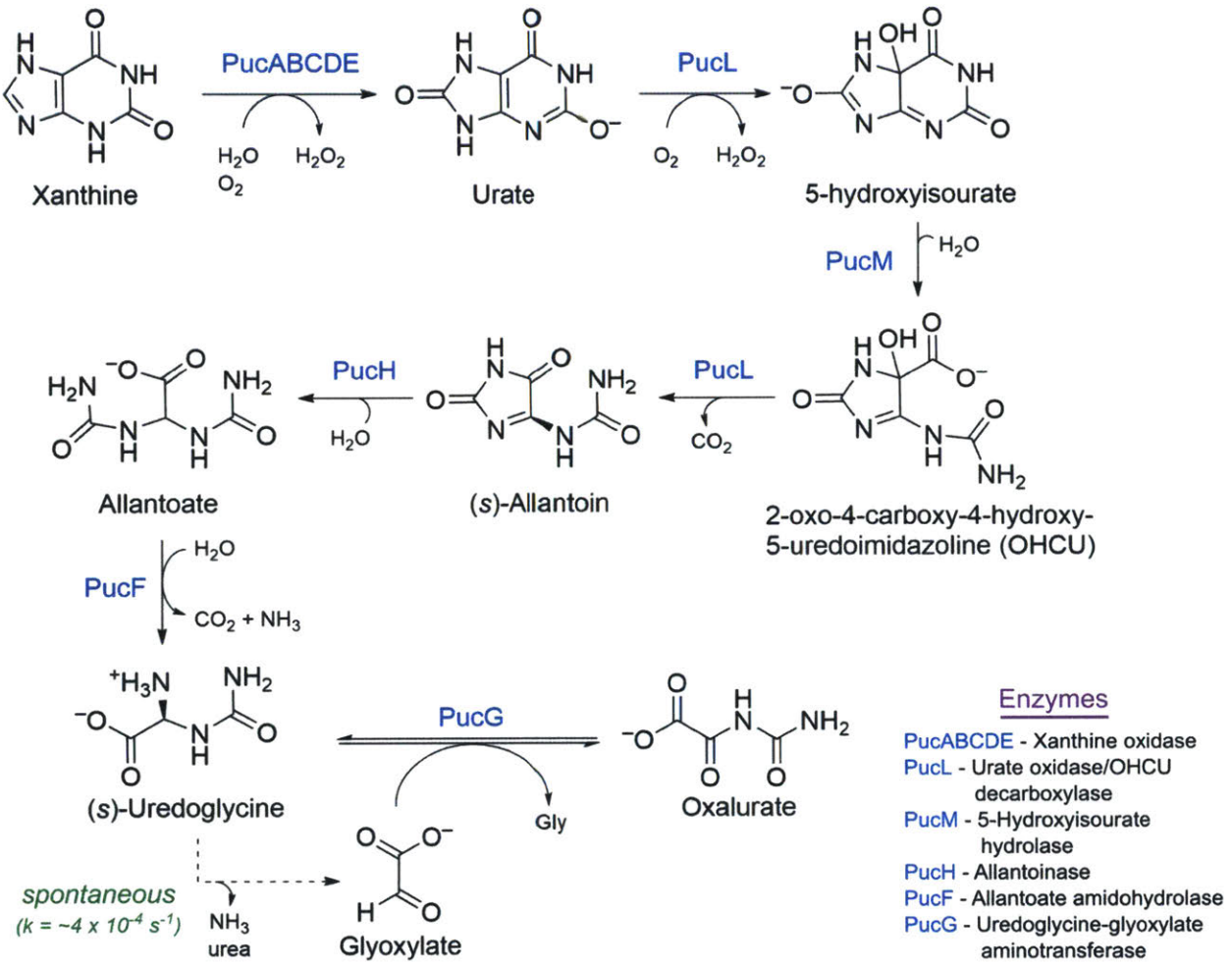


Figure 1.31. Purine degradation pathway in *B. subtilis*.

1.5.6. Influence of B. subtilis nucleotide metabolism on RNR activity. Of all the pathways just described, the ability of *B. subtilis* to import and utilize exogenous deoxynucleosides for DNA replication and repair processes will likely have the most impact on the regulation of RNR activity. If the bacterium finds itself in an environment particularly rich in deoxynucleosides, such as, perhaps, the GI tracts of animals, then it may not be necessary to reduce ribonucleotides in order

to meet the demands of DNA synthesis. The novel overall activity allosteric regulation described in this dissertation could be one mechanism by which the *B. subtilis* can attenuate RNR activity under such conditions. However, inhibition of the transcription and translation of the RNR genes would likely be the best method of reducing RNR activity in terms of energy efficiency. To this end, a novel transcription factor called NrdR has been identified in many prokaryotic genomes and is proposed to regulate the transcription of RNR genes in response to intracellular concentrations of dNTPs.[219]

NrdR is composed of an ATP-cone fused to the C-terminal end of a DNA-binding Zn ribbon domain.[153] The protein from *Streptomyces coelicolor* was reported to have one equivalent of tightly bound ATP or dATP bound,[220] whereas with the *E. coli* NrdR, adenosine and deoxyadenosine mono-, di-, and triphosphates were reported to be associated with the protein.[221] Mutation of conserved residues in the ATP-cone that abolished nucleotide binding also prevented the *S. coelicolor* NrdR from binding to its tandem 16-bp operator sequence, called the NrdR-box, indicating that this transcription factor regulates gene expression in response to at least intracellular adenosine/deoxyadenosine analog concentrations.[220] The current model for the regulatory mechanism is that the binding of ATP or dATP induces a protein conformational change that allows the Zn ribbon domain to bind to the NrdR-boxes in the upstream regulatory regions of RNR operons and, thus, inhibit transcription.[219, 220]

The YtcG protein of *B. subtilis* shows high homology to the *E. coli* (49% identity, 66% with positive substitutions) and *S. coelicolor* NrdRs (46% identity, 63% with positive substitutions), and a putative NrdR-box, based homology to the consensus sequence,[219] is located upstream of the *nrdI* gene of the Ib RNR operon. Therefore, the transcription of the *B. subtilis* RNR genes is likely regulated by YtcG (henceforth called NrdR) in response to *in vivo* nucleotide

concentrations, although experimental verification has not yet been reported. Possibly in support of this hypothesis is the observation of a second putative NrdR-box in the regulatory region of the gene *yvdC*, which encodes a putative pyrophosphohydrolase.[219] Although the regulation of *yvdC* transcription by NrdR has also not been demonstrated, an intriguing possibility is that NrdR binding induces the expression of YvdC, which could help in the maintenance of the dNTP pools by hydrolyzing any surpluses of nucleotide.

1.6. NUDIX HYDROLASES

As mentioned in the previous section, *B. subtilis* has several mechanisms in place, such as phosphatases and an efflux system, to rectify any imbalances in the overall amounts and ratios of intracellular dNTP pools should such a situation arise naturally. In addition to imbalanced dNTP pools, however, genome integrity can also be reduced by the presence of chemically modified or oxidized nucleotide derivatives which, when incorporated into DNA, cause mutations and/or double strand breaks that can be fatal to the cell. One superfamily of enzymes that perhaps has a dual function in removing both mutagenic nucleotides and any dNTP surpluses are the Nudix hydrolases.[222] An introduction to this superfamily is provided here since the work described in Chapter 5 of this thesis involves a Nudix hydrolase known as YmaB that is encoded in the same operon as the *B. subtilis* class Ib RNR genes.

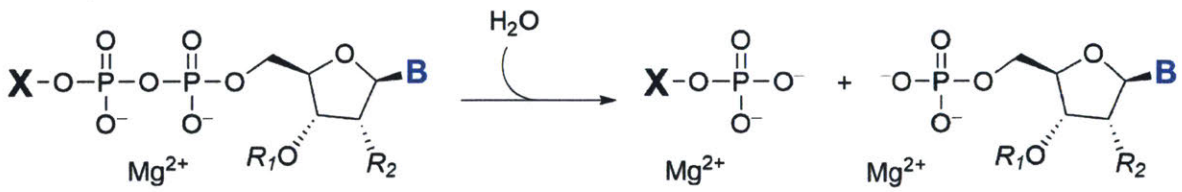


Figure 1.32. General reaction scheme for hydrolysis of (d)NDP-X substrates by Nudix hydrolases. **B** = base (A, C, G, T, U and derivatives thereof). R_1 = H, PO_3. R_2 = H, OH, OPO_3. **X** can be any of a number of chemical functional groups, including phosphate, nucleoside polyphosphates, sugars, and cofactors (for example, nicotinamide riboside or pantethine).

Nudix (nucleoside 5′-diphosphates linked to moiety X) hydrolases are generally small enzymes (16 – 21 kDa) with *in vitro* alkaline pH optima that catalyze the hydrolysis of (d)NDP-X substrates into the corresponding (d)NMP and P-X (**Figure 1.32**).[222, 223] These enzymes are broadly distributed in all domains of life, including viruses. Known substrates of these enzymes include the canonical (d)NTPs and their oxidized or chemically altered derivatives,[224-226] cofactors (e.g. NAD(P)H, FAD, coenzyme A),[227-229] nucleotide-sugar conjugates (e.g. GDP-mannose),[230] lipid precursors (e.g. CDP-choline),[231] dinucleoside 5′,5‴-polyphosphates (e.g. AP_nA, where n = 3 – 6),[232-234] and the 5′-triphosphate caps of mRNAs.[235, 236] Examples also exist of Nudix hydrolases acting on non-nucleotide substrates which contain a pyrophosphate moiety, such as thiamine pyrophosphate,[237] diphosphoinositol polyphosphates,[238] and dihydroneopterin tri-phosphate.[239, 240]

Nudix hydrolases can typically be identified by a characteristic sequence motif called the Nudix box that has the consensus sequence of Gx_5Ex_5[UA]xREx_2EExGU, where x can be any amino acid and U is a bulky hydrophobic residue.[222] The three conserved glutamates in this segment are key for catalysis as they help position the substrate and coordinate two or three Mg^{2+} or Mn^{2+} ions that further stabilize substrate positioning, neutralize the negative charges on the polyphosphate group of the substrate, and activate water for deprotonation by a general base to the active nucleophile (either hydroxide or oxide).[223] Depending on the substrate and the location of the general base in the tertiary structure (it can be inside or outside of the Nudix box), the nucleophile can attack at different positions on a (d)NDP-X, such as the α or β phosphorus atom or, as in the case of GDP-mannose hydrolase, the C1 atom of the mannosyl moiety.[223] The specificity of a particular enzyme is dictated by residues outside of the Nudix box, and several

subfamilies of Nudix hydrolases have been categorized based on conserved residues and motifs identified in homologs catalyzing the same reaction (**Table 1.3**).[241-243]

Table 1.3. Sequence motifs distinguishing subfamilies of the Nudix hydrolases.

Substrate	Conserved sequence identifiers†	Reference
ADP-ribose	Proline 15 – 16 residues downstream of Nudix box	241
AP_nA (n = 3 – 6)	Tyrosine 16 – 18 residues downstream of Nudix box	241
NADH	SQPWPFPxS 10 – 11 residues downstream of Nudix box	241
Coenzyme A	LLTxR[SA]x_3Rx_3Gx_3FPGG just upstream of Nudix box	242
5-methyl-UTP, UTP	L[VL]VRK 10 residues upstream of Nudix box AND AANE 14 – 16 residues downstream of Nudix box	243

† The letter "x" represents any amino acid.

In addition to the Nudix box, the Nudix hydrolases also share a core tertiary structure, known as the Nudix fold, consisting of an α/β/α sandwich that is structurally homologous to isopentenyl diphosphate isomerases and certain DNA glycosylases.[222] Two views of this fold are shown in **Figure 1.33A** and **B** for the enzyme MutT (8-oxo-2′-deoxyguanosine 5′-triphosphate (8-oxo-dGTP) pyrophosphohydrolase) from *E. coli*, the proto-typical member of the Nudix hydrolases.[244] The critical RExxEE segment of the Nudix box is located at the C-terminal end of a loop-helix-loop motif (**Figure 1.33C**) sitting at the edge of a groove formed by the curl of the β-sheet which serves as the site for substrate binding. Many members of the Nudix hydrolases have extensions of the N- and C-termii as well as insertions into the Nudix fold that further aid in determining the substrate specificity of the enzyme and/or participate in the formation of higher order oligomers.[231, 242, 245]

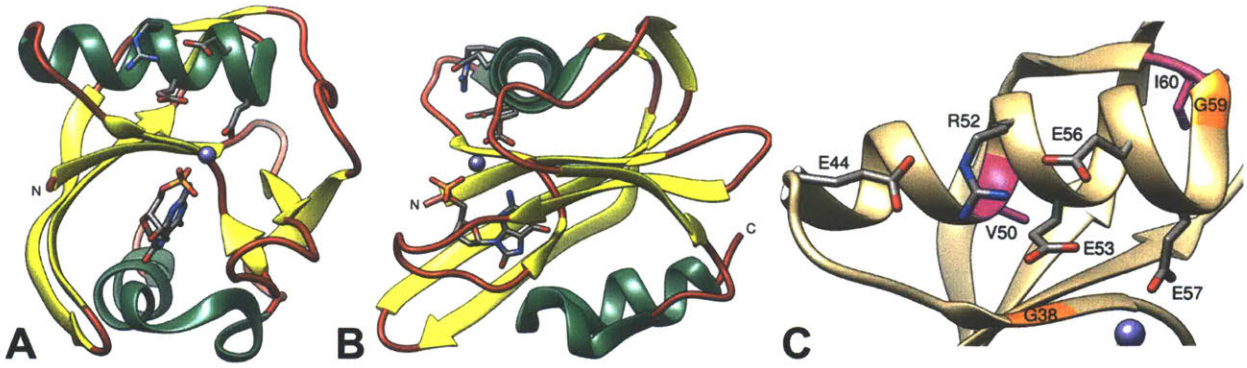


Figure 1.33. Common structural elements of Nudix hydrolases using *E. coli* MutT bound with product 8-oxo-2′-deoxyguanosine 5′-monophosphate (8-oxo-dGMP) as an example (PDB 3A6U).[244] (**A**) Tertiary structure of MutT illustrating the α/β/α core fold common to Nudix hydrolases. (**B**) Structure rotated 90 ° to the left of (**A**). Helices are colored green, sheets yellow, and coiled regions red. N and C indicate the N- and C-termini, respectively. The side chains of the conserved Glu and Arg residues of the RExxEE motif of the Nudix box and the atoms of 8-oxo-dGMP are depicted as stick models. (**C**) Close up of the loop-helix-loop motif of the MutT Nudix box with conserved glycine residues colored orange, bulky hydrophobic residues magenta, and side chains of conserved Arg and Glu residues depicted as sticks. Atom colors: C = gray, N = blue, O = red, P = orange, Mn = purple.

In general, because the substrates that these enzymes hydrolyze can be mutagenic or cytotoxic, or are transient signaling molecules and metabolic intermediates, the functional descriptor of "housecleaning enzymes," has been given to the Nudix hydrolases, and they are thought to be order keepers in the cell by removing toxins, ending signaling events, and monitoring/regulating the levels of metabolic intermediates, cofactors, and building blocks of processes such as DNA, RNA, and cell wall biosynthesis.[222, 246] However, definitively assigning the *in vivo* function of a particular Nudix hydrolase has been a relatively difficult task. Organisms tend to have several Nudix genes encoded in their genome,[222] and most, even those fitting into the subfamilies listed in **Table 1.3**, are fairly promiscuous in their substrate specificity. This often results in overlapping activities between two or more Nudix hydrolases and/or other (pyro)phosphatases that can mask any phenotypic consequences of single gene knock out mutants. In some cases, the physiological relevance of the best substrate for a Nudix enzyme *in vitro* can

be questionable, such as 5-methyl-UTP.[243] Ultimately, the assignment of an *in vivo* function to a Nudix hydrolase will require information from enzymology (i.e. k_{cat}, K_m, K_d), biology (i.e. *in vivo* metabolite concentrations and fluctuations, transcript and enzyme abundance, regulation), and ecology (habitat and lifestyles of the organism).

Arguably the best studied Nudix hydrolases for which a probable *in vivo* function has been established are MutT and RppH.[224, 235, 236] MutT is an antimutator that sanitizes the dNTP pool by degrading the mutagenic nucleotide 8-oxo-2′-deoxyguanosine 5′-triphosphate (8-OH-dGTP), and thereby prevents its incorporation into DNA during replication and repair.[222, 224, 247] In *E. coli*, *mutT* knockout mutants exhibit a ~1000-fold increase in the rate of spontaneous mutation due to AT:GC transversions, thus emphasizing the importance of this enzyme for cell viability.[224, 248] RppH catalyzes the first step of mRNA degradation in *E. coli* and *B. subtilis* by hydrolyzing the 5′-triphosphate of transcripts into the corresponding monophosphate and pyrophosphate.[235, 236] This decapping activity allows the downstream endo- and exonucleases to bind to the transcript and catalyze its decomposition. The lifetime of mRNAs in *E. coli* and *B. subtilis* RppH deletion mutants are dramatically increased, in some cases by over 10 min, as compared to wild type.[235, 236]

1.7. CHAPTER PREVIEW

At the outset of the studies reported in this thesis, it had been recently demonstrated that the subunits of the *B. subtilis* class Ib RNR, when isolated from their native host, co-purified through three columns in roughly a 1:1 ratio of α:β in the absence of Mg^{2+} and exogenous nucleotides.[21] This was an unprecedented and exciting result given that most RNRs that had been studied in detail at that point in time were class Ia systems that exhibited relatively weak subunit interactions.[85, 86, 89] Moreover, while the subunits of Ib enzymes had long been noted to form qualitatively "tight" interactions,[9] reports on their purification from the endogenous source often

describe the isolation of each protein separately.[11, 14, 20, 249] The apparent tight affinity of the *B. subtilis* Ib RNR subunits suggested that it might be the enzyme that would finally yield to characterization of the proposed $\alpha_2\beta_2$ active complex by X-ray crystallography, which is one of the overall goals of research in the Stubbe group. However, the activity of the endogenously isolated RNR (~80% homogeneous) was low (~160 nmol min^{-1} mg^{-1} β_2) as a consequence of substoichiometric amounts of Mn (2.1 Mn/β_2) and Y• (0.4 Y•/β_2) detected in NrdF, and efforts to reconstitute recombinant NrdF *in vitro* with a full complement of $Mn(III)_2$-Y• were also unsuccessful (1.9 Mn and 0.6 Y•/β_2) in improving CDP reduction activity (73 nmol min^{-1} mg^{-1} β_2).[21] In order to have the highest chance of success in crystallizing this RNR, it therefore was important to obtain highly active preparations of the individual subunits.

The initial goal of the work described here was to improve the catalytic activity of recombinant *B. subtilis* Ib RNR and to biophysically characterize the enzyme to better understand its quaternary structure prior to crystallization trials. These studies constitute the contents of **Chapter 2**. It is demonstrated that holo-$Me(III)_2$-Y• NrdF (Me = Fe or Mn, 4 Me and 1 Y•/β_2) can be separated from apo-protein using anion exchange chromatography on a MonoQ column. Concomitantly, candidate genes encoding plausible endogenous reductants for RNR were identified and cloned, and the proteins were overexpressed, purified, and tested for their ability to support CDP reduction activity; the results indicate thioredoxin (TrxA) and thioredoxin reductase (TrxB) are the physiological reductants. Utilization of these proteins and a 1:1 ratio of homogeneous holo-$Me(III)_2$-Y• NrdF and NrdE under optimized assay conditions resulted in a 7- to 9-fold improvement of the specific activity relative to the report by Zhang and Stubbe,[21] and in the case of the Mn-loaded enzyme, the activity (~1500 nmol min^{-1} mg^{-1} β_2) is one of the highest measured to date with a class Ib RNR (see Appendix 2, **Table A2.1**). Holo-NrdF and NrdE were

subsequently characterized individually and in 1:1 mixtures by anion exchange chromatography, size exclusion chromatography (SEC), and sedimentation velocity analytical ultracentrifugation (SV-AUC) in the absence of nucleotides. At physiologically relevant concentrations (~1 μM), NrdF behaved as a dimer and NrdE as a monomer. In contrast to expectations, 1:1 mixtures of NrdE:NrdF were composed of a complex mixture of structures rather than a single entity, indicating further studies were required to find conditions that would stabilize the $\alpha_2\beta_2$ complex for crystallization.

In addition to the quaternary structural complexity observed with the *B. subtilis* Ib RNR, one of the most surprising results of the studies described in **Chapter 2** was the discovery that dATP inhibited CDP reduction at physiologically relevant concentrations. Class Ib RNRs, as described earlier, do not have ATP-cone domains[27] and, thus, are thought not to have overall activity allosteric regulation by ATP and dATP. This unusual observation with *B. subtilis* RNR warranted further investigation. Furthermore, it seemed highly likely that a stable NrdE:NrdF complex could be obtained by including substrate and effector nucleotides. Therefore, as described in **Chapter 3**, an examination of the allosteric regulation of the Mn-loaded form of this RNR was made to gain a better understanding of the enzyme and to inform the design of subsequent biophysical experiments that probed the effects of nucleotides on the quaternary structure of NrdE and mixtures of NrdE and NrdF.

The work described in **Chapter 3** is the first study of the allosteric regulation of a class Ib RNR loaded with $Mn(III)_2$-Y•, which mounting evidence suggests is the physiologically relevant cofactor in this subclass,[18-21, 62] and, therefore, should be valuable for further studies of these enzymes from other organisms. The *B. subtilis* enzyme was found to exhibit the same general rules of allosteric regulation that have been observed with all RNRs:[1] ATP and dATP (at low

concentrations) stimulate pyrimidine NDP reduction, TTP stimulates GDP reduction, and dGTP stimulates ADP reduction. However, the enzyme also exhibited distinct attributes reminiscent of systems containing ATP-cone domains, suggesting that, in addition to the specificity site, the RNR possesses one other allosteric regulatory site. The inhibitory effects of dATP on RNR activity noted earlier were confirmed for both pyrimidine substrates, and SV-AUC analysis of 1:1 mixtures of NrdE:NrdF in the presence of dATP revealed the formation of large complexes with sedimentation coefficients as large as 35 S and M_WS larger than 1 MDa. Therefore, dATP appears to induce *B. subtilis* RNR to form alternate inhibitory quaternary structures. Further SV-AUC experiments probing the ability of GDP, TTP, and dATP to stabilize NrdE dimers revealed that only dATP could do this, whereas no effect was seen with TTP. GDP strangely appeared to cause the small amount of dimer that was present to convert to monomer. It therefore seemed that in order to stabilize a NrdE dimer and, consequently, generate an $\alpha_2\beta_2$ complex, all components (NrdE, NrdF, substrate, effector, and Mg^{2+}) would need to be present in solution.

Prior to attempting biophysical experiments with both RNR subunits in the presence of nucleotides, a curious feature in the sedimentation coefficient distributions of the NrdE samples collected in **Chapter 3** required further investigation. In these experiments, sedimentation was monitored by interference optics, a higher resolution detection method, which revealed NrdE sedimented as two different species near physiologically relevant concentrations. In contrast, similar experiments reported in **Chapter 2**, which were monitored by A_{280} detection, revealed only one species that was consistent with monomer. Before it was appreciated that the second species corresponded to NrdE dimer, it was interpreted as an alternate conformation of the monomer that was speculated to have been induced by the binding of dATP. Surprisingly, this speculation turned out to be partially correct as examination of supernatant recovered from a denatured sample of

NrdE revealed the presence of (a) tightly bound compound(s) with absorption at 260 nm. This discovery added a new layer of complexity to the already unique class Ib RNR and follow up studies were warranted.

The contents of **Chapter 4** detail the experiments performed to establish the identity of the compound(s) bound to NrdE and to determine what effects it (they) had on the activity and structure of the *B. subtilis* Ib RNR. The results of UV-visible spectrophotometry, ^{1}H-NMR spectroscopy, and polyethyleneimine-cellulose TLC analysis revealed a mixture of nucleotides that surprisingly was composed mostly of 2′-deoxyadenosine 5′-monophosphate (dAMP). NrdE as-isolated was substoichiometrically loaded with the nucleotides (~0.5 – 0.7 equivalents), and efforts to increase the loading by incubation of the protein with dAMP or dATP *in vitro* were unsuccessful. However, it was found that protein fully loaded with one equivalent of dAMP (holo-NrdE), as judged by TLC, could be separated from apo-protein using MonoQ anion exchange chromatography. Subsequent activity analysis revealed the presence of dAMP renders *B. subtilis* RNR more susceptible to dATP inhibition, and comparison of the dATP binding properties of apo- and holo-NrdE suggested the former has two binding sites per monomer with an apparent K_d = 6 μM and the latter a single site per monomer with an apparent K_d = 2 μM. These results thus confirm the presence of a second set of allosteric regulatory sites in *B. subtilis* NrdE, although it appears that dATP and dAMP can occupy this site, the latter with a much higher affinity.

Biophysical characterization of holo-NrdE revealed the protein itself had a higher propensity to oligomerize, but it did not stabilize a NrdE:NrdF complex in the absence of nucleotides. Therefore, a catalytically inactive mutant of NrdE ($C_{382}S$) was prepared and its holo-form isolated. The results of SV-AUC analysis of a 1:1 mixture of NrdE-$C_{382}S$:NrdF at ~2 μM in the presence of TTP and GDP showed the formation a complex mixture of structures with

sedimentation coefficients as large as ~37 S and M_Ws greater than 1 MDa. However, unlike the results with dATP, a majority of the protein sedimented at 10 S, consistent with an $\alpha_2\beta_2$ complex. This observation is encouraging and can serve as a foundation for continuing studies to find conditions that may fully stabilize a single quaternary structure of the *B. subtilis* Ib RNR.

Finally, in **Chapter 5** studies concerning the function of YmaB, encoded by the fourth gene of the *B. subtilis* Ib RNR operon (**Figure 1.25**), and its link to RNR activity are reported. The presence of *ymaB* in the RNR operon is unique to *B. subtilis* and its immediate phylogenetic relatives. In comparison, the Ib operons of other bacteria either do not have a fourth gene or, as is the case for the SPβ RNR (**Figure 1.25**), encode the glutaredoxin-like thioredoxin NrdH.[6, 67, 68] Given the unusual dATP inhibition and ability to tightly bind dAMP observed with the *B. subtilis* Ib RNR, it seems likely that YmaB plays some as-yet undefined role in the regulation of RNR. Homology searches revealed that YmaB is similar to Nudix hydrolases which, as described earlier, are small enzymes that hydrolyze substrates of the form (d)NDP-X, where X can be a number of different chemical groups (**Figure 1.32**). Screening potential substrates revealed YmaB is indeed a Nudix hydrolase with specificity toward the canonical dNTPs, thus providing the first potential link between the activity of YmaB and regulation of RNR. SV-AUC experiments were performed to acquire more fundamental information about the enzyme and showed that YmaB behaves as a monomer in the absence of substrate. Its ability to hydrolyze dATP to dAMP and pyrophosphate suggested that YmaB might be the factor responsible for installing dAMP in NrdE in *B. subtilis*. However, coexpression of YmaB and NrdE together in *E. coli* cells did not result in a significant increase in the dAMP loading of the latter. The result is inconclusive and additional experiments are warranted to investigate the connection between YmaB and RNR.

In relation to the goal of using the *B. subtilis* Ib RNR to obtain an X-ray crystal structure of the $\alpha_2\beta_2$ complex, it can be concluded from the work presented in this thesis that the enzyme is probably not the best candidate for crystallization trials unless additional studies are performed to find conditions that stabilize a single NrdE:NrdF complex. More importantly, the results demonstrate a hitherto uncharacterized mode of allosteric regulation by dATP of RNRs that lack traditional allosteric activity sites, implying that overall activity regulation by dATP is not solely mediated by the ATP-cone domain. Interestingly, the mechanism of dATP inhibition still involves quaternary structural interconversions, suggesting that this is likely a general principle in RNRs that have overall activity regulation, albeit the inhibitory structures formed will be distinctly different depending on the presence versus absence of an ATP-cone domain.[2, 95] Continuation of the studies on this novel mode of dATP inhibition, along with the potential involvement of NrdE-associated dAMP and YmaB in further modulating RNR activity, should provide valuable insights into the regulation of nucleotide metabolism in the model Gram-positive bacterium *B. subtilis* and, more generally, the allosteric regulation of class Ib RNRs.

1.8. REFERENCES

1. Hofer, A., Crona, M., Logan, D. T., and Sjöberg, B. M. (2012) DNA Building Blocks: Keeping Control of Manufacture, *Crit. Rev. Biochem. Mol. Biol. 47*, 50-63.
2. Ando, N., Brignole, E. J., Zimanyi, C. M., Funk, M. A., Yokoyama, K., Asturias, F. J., Stubbe, J., and Drennan, C. L. (2011) Structural Interconversions Modulate Activity of *Escherichia coli* Ribonucleotide Reductase, *Proc. Natl. Acad. Sci. U. S. A. 108*, 21046-21051.
3. Ando, N., Li, H. R., Brignole, E. J., Thompson, S., McLaughlin, M. I., Page, J. E., Asturias, F. J., Stubbe, J., and Drennan, C. L. (2016) Allosteric Inhibition of Human Ribonucleotide Reductase by dATP Entails the Stabilization of a Hexamer, *Biochemistry 55*, 373-381.
4. Fairman, J. W., Wijerathna, S. R., Ahmad, M. F., Xu, H., Nakano, R., Jha, S., Prendergast, J., Welin, R. M., Flodin, S., Roos, A., Nordlund, P., Li, Z., Walz, T., and Dealwis, C. G. (2011) Structural Basis for Allosteric Regulation of Human Ribonucleotide Reductase by Nucleotide-Induced Oligomerization, *Nat. Struct. Mol. Biol. 18*, 316-322.

5. Zimanyi, C. M., Chen, P. Y. T., Kang, G., Funk, M. A., and Drennan, C. L. (2016) Molecular Basis for Allosteric Specificity Regulation in Class Ia Ribonucleotide Reductase from *Escherichia coli*, *eLife 5*, 23.

6. Lundin, D., Torrents, E., Poole, A. M., and Sjöberg, B. M. (2009) RNRdb, a Curated Database of the Universal Enzyme Family Ribonucleotide Reductase, Reveals a High Level of Misannotation in Sequences Deposited to Genbank, *BMC Genomics 10*, 8.

7. Cotruvo, J. A., and Stubbe, J. (2011) Class I Ribonucleotide Reductases: Metallocofactor Assembly and Repair *In Vitro* and *In Vivo*, In *Annual Review of Biochemistry* (Kornberg, R. D., Raetz, C. R. H., Rothman, J. E., and Thorner, J. W., Eds.), pp 733-767, Annual Reviews, Palo Alto.

8. Atkin, C. L., Thelander, L., Reichard, P., and Lang, G. (1973) Iron and Free Radical in Ribonucleotide Reductase. Exchange of Iron and Mössbauer Spectroscopy of Protein B2 Subunit of the *Escherichia coli* Enzyme, *J. Biol. Chem. 248*, 7464-7472.

9. Eliasson, R., Pontis, E., Jordan, A., and Reichard, P. (1996) Allosteric Regulation of the Third Ribonucleotide Reductase (NrdEF Enzyme) from Enterobacteriaceae, *J. Biol. Chem. 271*, 26582-26587.

10. Huque, Y., Fieschi, F., Torrents, E., Gibert, I., Eliasson, R., Reichard, P., Sahlin, M., and Sjöberg, B. M. (2000) The Active Form of the R2F Protein of Class Ib Ribonucleotide Reductase from *Corynebacterium ammoniagenes* is a Diferric Protein, *J. Biol. Chem. 275*, 25365-25371.

11. Jordan, A., Pontis, E., Atta, M., Krook, M., Gibert, I., Barbé, J., and Reichard, P. (1994) A Second Class I Ribonucleotide Reductase in *Enterobacteriaceae* - Characterization of the *Salmonella typhimurium* Enzyme, *Proc. Natl. Acad. Sci. U. S. A. 91*, 12892-12896.

12. Auling, G., and Follmann, H. (1994) Manganese-Dependent Ribonucleotide Reduction and Overproduction of Nucleotides in Coryneform Bacteria, In *Metalloenzymes Involving Amino Acid Residues and Related Radicals* (Sigel, H., and Sigel, A., Eds.), pp 131-164, Marcel Dekker Inc., New York, NY.

13. Schimpff-Weiland, G., Follmann, H., and Auling, G. (1981) A New Manganese-Activated Ribonucleotide Reductase Found in Gram-Positive Bacteria, *Biochem. Biophys. Res. Commun. 102*, 1276-1282.

14. Willing, A., Follmann, H., and Auling, G. (1988) Ribonucleotide Reductase of *Brevibacterium ammoniagenes* is a Manganese Enzyme, *Eur. J. Biochem. 170*, 603-611.

15. Cotruvo, J. A., and Stubbe, J. (2010) An Active Dimanganese(III)-Tyrosyl Radical Cofactor in *Escherichia coli* Class Ib Ribonucleotide Reductase, *Biochemistry 49*, 1297-1309.

16. Boal, A. K., Cotruvo, J. A., Stubbe, J., and Rosenzweig, A. C. (2010) Structural Basis for Activation of Class Ib Ribonucleotide Reductase, *Science 329*, 1526-1530.

17. Cotruvo, J. A., Stich, T. A., Britt, R. D., and Stubbe, J. (2013) Mechanism of Assembly of the Dimanganese-Tyrosyl Radical Cofactor of Class Ib Ribonucleotide Reductase:

Enzymatic Generation of Superoxide Is Required for Tyrosine Oxidation via a Mn(III)Mn(IV) Intermediate, *J. Am. Chem. Soc. 135*, 4027-4039.

18. Cotruvo, J. A., and Stubbe, J. (2011) *Escherichia coli* Class Ib Ribonucleotide Reductase Contains a Dimanganese(III)-Tyrosyl Radical Cofactor *in vivo*, *Biochemistry 50*, 1672-1681.

19. Cox, N., Ogata, H., Stolle, P., Reijerse, E., Auling, G., and Lubitz, W. (2010) A Tyrosyl-Dimanganese Coupled Spin System is the Native Metalloradical Cofactor of the R2F Subunit of the Ribonucleotide Reductase of *Corynebacterium ammoniagenes*, *J. Am. Chem. Soc. 132*, 11197-11213.

20. Stolle, P., Barckhausen, O., Oehlmann, W., Knobbe, N., Vogt, C., Pierik, A. J., Cox, N., Schmidt, P. P., Reijerse, E. J., Lubitz, W., and Auling, G. (2010) Homologous Expression of the *nrdF* Gene of *Corynebacterium ammoniagenes* Strain ATCC 6872 Generates a Manganese-metallocofactor (R2F) and a Stable Tyrosyl Radical (Y•) Involved in Ribonucleotide Reduction, *FEBS J. 277*, 4849-4862.

21. Zhang, Y., and Stubbe, J. (2011) *Bacillus subtilis* Class Ib Ribonucleotide Reductase is a Dimanganese(III)-Tyrosyl Radical Enzyme, *Biochemistry 50*, 5615-5623.

22. Uppsten, M., Färnegårdh, M., Jordan, A., Eliasson, R., Eklund, H., and Uhlin, U. (2003) Structure of the Large Subunit of Class Ib Ribonucleotide Reductase from *Salmonella typhimurium* and Its Complexes with Allosteric Effectors, *J. Mol. Biol. 330*, 87-97.

23. Eriksson, M., Jordan, A., and Eklund, H. (1998) Structure of *Salmonella typhimurium nrdF* Ribonucleotide Reductase in Its Oxidized and Reduced Forms, *Biochemistry 37*, 13359-13369.

24. Eriksson, M., Uhlin, U., Ramaswamy, S., Ekberg, M., Regnström, K., Sjöberg, B. M., and Eklund, H. (1997) Binding of Allosteric Effectors to Ribonucleotide Reductase Protein R1: Reduction of Active-Site Cysteines Promotes Substrate Binding, *Structure 5*, 1077-1092.

25. Nordlund, P., Sjöberg, B. M., and Eklund, H. (1990) Three-Dimensional Structure of the Free Radical Protein of Ribonucleotide Reductase, *Nature 345*, 593-598.

26. Uhlin, U., and Eklund, H. (1994) Structure of Ribonucleotide Reductase Protein R1, *Nature 370*, 533-539.

27. Aravind, L., Wolf, Y. I., and Koonin, E. V. (2000) The ATP-Cone: An Evolutionarily Mobile, ATP-Binding Regulatory Domain, *J. Mol. Microbiol. Biotechnol. 2*, 191-194.

28. Crona, M., Torrents, E., Røhr, A. K., Hofer, A., Furrer, E., Tomter, A. B., Andersson, K. K., Sahlin, M., and Sjöberg, B. M. (2011) NrdH-Redoxin Protein Mediates High Enzyme Activity in Manganese-Reconstituted Ribonucleotide Reductase from *Bacillus anthracis*, *J. Biol. Chem. 286*, 33053-33060.

29. Jordan, A., Pontis, E., Åslund, F., Hellman, U., Gibert, I., and Reichard, P. (1996) The Ribonucleotide Reductase System of *Lactococcus lactis* - Characterization of an NrdEF Enzyme and a New Electron Transport Protein, *J. Biol. Chem. 271*, 8779-8785.

30. Makhlynets, O., Boal, A. K., Rhodes, D. V., Kitten, T., Rosenzweig, A. C., and Stubbe, J. (2014) *Streptococcus sanguinis* Class Ib Ribonucleotide Reductase: High Activity with Both Iron and Manganese Cofactors and Structural Insights, *J. Biol. Chem. 289*, 6259-6272.

31. Yang, F. D., Curran, S. C., Li, L. S., Avarbock, D., Graf, J. D., Chua, M. M., Lui, G. Z., Salem, J., and Rubin, H. (1997) Characterization of Two Genes Encoding the *Mycobacterium tuberculosis* Ribonucleotide Reductase Small Subunit, *J. Bacteriol. 179*, 6408-6415.

32. Jordan, A., Aragall, E., Gibert, I., and Barbé, J. (1996) Promoter Identification and Expression Analysis of *Salmonella typhimurium* and *Escherichia coli* NrdEF Operons Encoding One of Two Class I Ribonucleotide Reductases Present in Both Bacteria, *Mol. Microbiol. 19*, 777-790.

33. Panosa, A., Roca, I., and Gibert, I. (2010) Ribonucleotide Reductases of *Salmonella* Typhimurium: Transcriptional Regulation and Differential Role in Pathogenesis, *PLoS One 5*, 11.

34. Fuchs, J. A., Warner, H. R., Karlström, H. O., and Reichard, P. (1972) Defective Gene Product in *dnaF* Mutant of *Escherichia coli*, *Nature New Biol. 238*, 69-71.

35. Martin, J. E., and Imlay, J. A. (2011) The Alternative Aerobic Ribonucleotide Reductase of *Escherichia coli*, NrdEF, is a Manganese-Dependent Enzyme that Enables Cell Replication During Periods of Iron Starvation, *Mol. Microbiol. 80*, 319-334.

36. Center for Disease Control and Prevention, Office of Infectious Disease. *Antibiotic Resistance Threats in the United States, 2013.* April, 2013. Available at: http://www.cdc.gov/drugresistance/threat-report-2013/. Accessed October 8, 2016.

37. Higgins, D., and Dworkin, J. (2012) Recent Progress in *Bacillus subtilis* Sporulation, *FEMS Microbiol. Rev. 36*, 131-148.

38. Dubnau, D. (1991) Genetic Competence in *Bacillus subtilis*, *Microbiol. Rev. 55*, 395-424.

39. Vlamakis, H., Chai, Y. R., Beauregard, P., Losick, R., and Kolter, R. (2013) Sticking Together: Building a Biofilm the *Bacillus subtilis* Way, *Nat. Rev. Microbiol. 11*, 157-168.

40. López, D., and Kolter, R. (2010) Extracellular Signals that Define Distinct and Coexisting Cell Fates in *Bacillus subtilis*, *FEMS Microbiol. Rev. 34*, 134-149.

41. Boal, A. K., Cotruvo, J. A., Stubbe, J., and Rosenzweig, A. C. (2012) The Dimanganese(II) Site of *Bacillus subtilis* Class Ib Ribonucleotide Reductase, *Biochemistry 51*, 3861-3871.

42. Parker, M. J., Zhu, X. L., and Stubbe, J. (2014) *Bacillus subtilis* Class Ib Ribonucleotide Reductase: High Activity and Dynamic Subunit Interactions, *Biochemistry 53*, 766-776.

43. Licht, S., and Stubbe, J. (1999) Mechanistic Investigations of Ribonucleotide Reductases, In *Comprehensive Natural Products Chemistry* (Barton, D. H. R., Nakanishi, K., and Meth-Cohn, O., Eds.), pp 163-203, Elsevier, New York.

44. Stubbe, J., and van der Donk, W. A. (1998) Protein Radicals in Enzyme Catalysis, *Chem. Rev. 98*, 705-762.

45. Licht, S., Gerfen, G. J., and Stubbe, J. A. (1996) Thiyl Radicals in Ribonucleotide Reductases, *Science 271*, 477-481.

46. Kolberg, M., Strand, K. R., Graff, P., and Andersson, K. K. (2004) Structure, Function, and Mechanism of Ribonucleotide Reductases, *BBA-Proteins Proteomics 1699*, 1-34.

47. Sun, X. Y., Eliasson, R., Pontis, E., Andersson, J., Buist, G., Sjöberg, B. M., and Reichard, P. (1995) Generation of the Glycyl Radical of the Anaerobic *Escherichia coli* Ribonucleotide Reductase Requires a Specific Activating Enzyme, *J. Biol. Chem. 270*, 2443-2446.

48. Andersson, J., Westman, M., Hofer, A., and Sjöberg, B. M. (2000) Allosteric Regulation of the Class III Anaerobic Ribonucleotide Reductase from Bacteriophage T4, *J. Biol. Chem. 275*, 19443-19448.

49. Eliasson, R., Pontis, E., Sun, X. Y., and Reichard, P. (1994) Allosteric Control of the Substrate-Specificity of the Anaerobic Ribonucleotide Reductase from *Escherichia coli*, *J. Biol. Chem. 269*, 26052-26057.

50. Torrents, E., Buist, G., Liu, A., Eliasson, R., Kok, J., Gibert, I., Gräslund, A., and Reichard, P. (2000) The Anaerobic (Class III) Ribonucleotide Reductase from *Lactococcus lactis* - Catalytic Properties and Allosteric Regulation of the Pure Enzyme System, *J. Biol. Chem. 275*, 2463-2471.

51. Wei, Y. F., Funk, M. A., Rosado, L. A., Baek, J., Drennan, C. L., and Stubbe, J. (2014) The Class III Ribonucleotide Reductase from *Neisseria bacilliformis* can Utilize Thioredoxin as a Reductant, *Proc. Natl. Acad. Sci. U. S. A. 111*, E3756-E3765.

52. Wei, Y. F., Li, B., Prakash, D., Ferry, J. G., Elliott, S. J., and Stubbe, J. (2015) A Ferredoxin Disulfide Reductase Delivers Electrons to the *Methanosarcina barkeri* Class III Ribonucleotide Reductase, *Biochemistry 54*, 7019-7028.

53. Fraser, C. M., Casjens, S., Huang, W. M., Sutton, G. G., Clayton, R., Lathigra, R., White, O., Ketchum, K. A., Dodson, R., Hickey, E. K., Gwinn, M., Dougherty, B., Tomb, J. F., Fleischmann, R. D., Richardson, D., Peterson, J., Kerlavage, A. R., Quackenbush, J., Salzberg, S., Hanson, M., vanVugt, R., Palmer, N., Adams, M. D., Gocayne, J., Weidman, J., Utterback, T., Watthey, L., McDonald, L., Artiach, P., Bowman, C., Garland, S., Fujii, C., Cotton, M. D., Horst, K., Roberts, K., Hatch, B., Smith, H. O., and Venter, J. C. (1997) Genomic Sequence of a Lyme Disease Spirochaete, *Borrelia burgdorferi*, *Nature 390*, 580-586.

54. Glass, J. I., Lefkowitz, E. J., Glass, J. S., Heiner, C. R., Chen, E. Y., and Cassell, G. H. (2000) The Complete Sequence of the Mucosal Pathogen *Ureaplasma urealyticum*, *Nature 407*, 757-762.

55. Dawes, S. S., Warner, D. F., Tsenova, L., Timm, J., McKinney, J. D., Kaplan, G., Rubin, H., and Mizrahi, V. (2003) Ribonucleotide Reduction in *Mycobacterium tuberculosis*: Function and Expression of Genes encoding Class Ib and Class II Ribonucleotide Reductases, *Infect. Immun. 71*, 6124-6131.

56. Sintchak, M. D., Arjara, G., Kellogg, B. A., Stubbe, J., and Drennan, C. L. (2002) The Crystal Structure of Class II Ribonucleotide Reductase Reveals How an Allosterically Regulated Monomer Mimics a Dimer, *Nat. Struct. Biol. 9*, 293-300.

57. Larsson, K. M., Andersson, J., Sjöberg, B. M., Nordlund, P., and Logan, D. T. (2001) Structural Basis for Allosteric Substrate Specificity Regulation in Anaerobic Ribonucleotide Reductases, *Structure 9*, 739-750.

58. Goulian, M., and Beck, W. S. (1966) Purification and Properties of Cobamide-Dependent Ribonucleotide Reductase from *Lactobacillus leichmannii*, *J. Biol. Chem. 241*, 4233-4242.

59. Torrents, E., Poplawski, A., and Sjöberg, B. M. (2005) Two Proteins Mediate Class II Ribonucleotide Reductase Activity in *Pseudomonas aeruginosa* - Expression and Transcriptional Analysis of the Aerobic Enzymes, *J. Biol. Chem. 280*, 16571-16578.

60. Eliasson, R., Pontis, E., Jordan, A., and Reichard, P. (1999) Allosteric Control of Three B_{12}-dependent (Class II) Ribonucleotide Reductases - Implications for the Evolution of Ribonucleotide Reduction, *J. Biol. Chem. 274*, 7182-7189.

61. Jiang, W., Xie, J. J., Nørgaard, H., Bollinger, J. M., and Krebs, C. (2008) Rapid and Quantitative Activation of *Chlamydia trachomatis* Ribonucleotide Reductase by Hydrogen Peroxide, *Biochemistry 47*, 4477-4483.

62. Rhodes, D. V., Crump, K. E., Makhlynets, O., Snyder, M., Ge, X. C., Xu, P., Stubbe, J., and Kitten, T. (2014) Genetic Characterization and Role in Virulence of the Ribonucleotide Reductases of *Streptococcus sanguinis*, *J. Biol. Chem. 289*, 6273-6287.

63. Sjöberg, B. M., and Torrents, E. (2011) Shift in Ribonucleotide Reductase Gene Expression in *Pseudomonas aeruginosa* during Infection, *Infect. Immun. 79*, 2663-2669.

64. Mao, S. S., Holler, T. P., Yu, G. X., Bollinger, J. M., Booker, S., Johnston, M. I., and Stubbe, J. (1992) A Model for the Role of Multiple Cysteine Residues Involved in Ribonucleotide Reduction - Amazing and Still Confusing, *Biochemistry 31*, 9733-9743.

65. Laurent, T. C., Moore, E. C., and Reichard, P. (1964) Enzymatic Synthesis of Deoxyribonucleotides. VI. Isolation and Characterization of Thioredoxin, the Hydrogen Donor from *Escherichia coli* B, *J. Biol. Chem. 239*, 3436-3444.

66. Meyer, Y., Buchanan, B. B., Vignols, F., and Reichheld, J. P. (2009) Thioredoxins and Glutaredoxins: Unifying Elements in Redox Biology, *Annu. Rev. Genet. 43*, 335-367.

67. Jordan, A., Åslund, F., Pontis, E., Reichard, P., and Holmgren, A. (1997) Characterization of *Escherichia coli* NrdH - A Glutaredoxin-like Protein with a Thioredoxin-like Activity Profile, *J. Biol. Chem. 272*, 18044-18050.

68. Stehr, M., Schneider, G., Åslund, F., Holmgren, A., and Lindqvist, Y. (2001) Structural Basis for the Thioredoxin-like Activity Profile of the Glutaredoxin-like NrdH-Redoxin from *Escherichia coli*, *J. Biol. Chem. 276*, 35836-35841.

69. Holmgren, A. (1976) Hydrogen Donor System for *Escherichia coli* Ribonucleoside-Diphosphate Reductase Dependent upon Glutathione, *Proc. Natl. Acad. Sci. U. S. A. 73*, 2275-2279.

70. Gustafsson, T. N., Osman, H., Werngren, J., Hoffner, S., Engman, L., and Holmgren, A. (2016) Ebselen and Analogs as Inhibitors of *Bacillus anthracis* Thioredoxin Reductase and Bactericidal Antibacterials Targeting *Bacillus* Species, *Staphylococcus aureus* and *Mycobacterium tuberculosis*, *Biochim. Biophys. Acta-Gen. Subj. 1860*, 1265-1271.

71. Lu, J., and Holmgren, A. (2014) The Thioredoxin Antioxidant System, *Free Radic. Biol. Med. 66*, 75-87.

72. Gustafsson, T. N., Sahlin, M., Lu, J., Sjöberg, B. M., and Holmgren, A. (2012) *Bacillus anthracis* Thioredoxin Systems, Characterization and Role as Electron Donors for Ribonucleotide Reductase, *J. Biol. Chem. 287*, 39686-39697.

73. Mulliez, E., Ollagnier, S., Fontecave, M., Eliasson, R., and Reichard, P. (1995) Formate is the Hydrogen Donor for the Anaerobic Ribonucleotide Reductase from *Escherichia coli*, *Proc. Natl. Acad. Sci. U. S. A. 92*, 8759-8762.

74. Nizet, V., and Johnson, R. S. (2009) Interdependence of Hypoxic and Innate Immune Responses, *Nat. Rev. Immunol. 9*, 609-617.

75. Wei, Y. F., Mathies, G., Yokoyama, K., Chen, J. H., Griffin, R. G., and Stubbe, J. (2014) A Chemically Competent Thiosulfuranyl Radical on the *Escherichia coli* Class III Ribonucleotide Reductase, *J. Am. Chem. Soc. 136*, 9001-9013.

76. Thelander, L. (1973) Physicochemical Characterization of Ribonucleoside Diphosphate Reductase from *Escherichia coli*, *J. Biol. Chem. 248*, 4591-4601.

77. Jonna, V. R., Crona, M., Rofougaran, R., Lundin, D., Johansson, S., Brännström, K., Sjöberg, B. M., and Hofer, A. (2015) Diversity in Overall Activity Regulation of Ribonucleotide Reductase, *J. Biol. Chem. 290*, 17339-17348.

78. Högbom, M., Galander, M., Andersson, M., Kolberg, M., Hofbauer, W., Lassmann, G., Nordlund, P., and Lendzian, F. (2003) Displacement of the Tyrosyl Radical Cofactor in Ribonucleotide Reductase Obtained by Single-Crystal High-Field EPR and 1.4-Angstrom X-Ray Data, *Proc. Natl. Acad. Sci. U. S. A. 100*, 3209-3214.

79. Petersson, L., Gräslund, A., Ehrenberg, A., Sjöberg, B. M., and Reichard, P. (1980) The Iron Center in Ribonucleotide Reductase from *Escherichia coli*, *J. Biol. Chem. 255*, 6706-6712.

80. Bender, C. J., Sahlin, M., Babcock, G. T., Barry, B. A., Chandrashekar, T. K., Salowe, S. P., Stubbe, J., Lindström, B., Petersson, L., Ehrenberg, A., and Sjöberg, B. M. (1989) An ENDOR Study of the Tyrosyl Free Radical in Ribonucleotide Reductase from *Escherichia coli*, *J. Am. Chem. Soc. 111*, 8076-8083.

81. Hoganson, C. W., Sahlin, M., Sjöberg, B. M., and Babcock, G. T. (1996) Electron Magnetic Resonance of the Tyrosyl Radical in Ribonucleotide Reductase From *Escherichia coli*, *J. Am. Chem. Soc. 118*, 4672-4679.

82. Brown, N. C., and Reichard, P. (1969) Ribonucleoside Diphosphate Reductase - Formation of Active and Inactive Complexes of Proteins B1 and B2, *J. Mol. Biol. 46*, 25-38.

83. Brown, N. C., Larsson, A., and Reichard, P. (1967) On Subunit Structure of Ribonucleoside Diphosphate Reductase, *J. Biol. Chem. 242*, 4272-4273.

84. Sjöberg, B. M., Karlsson, M., and Jörnvall, H. (1987) Half-site reactivity of the Tyrosyl Radical of Ribonucleotide Reductase from *Escherichia coli*, *J. Biol. Chem. 262*, 9736-9743.

85. Climent, I., Sjöberg, B. M., and Huang, C. Y. (1991) Carboxyl-Terminal Peptides as Probes for *Escherichia coli* Ribonucleotide Reductase Subunit Interaction - Kinetic Analysis of Inhibition Studies, *Biochemistry 30*, 5164-5171.

86. Hassan, A. Q., Wang, Y. T., Plate, L., and Stubbe, J. (2008) Methodology To Probe Subunit Interactions in Ribonucleotide Reductases, *Biochemistry 47*, 13046-13055.

87. Kasrayan, A., Birgander, P. L., Pappalardo, L., Regnström, K., Westman, M., Slaby, A., Gordon, E., and Sjöberg, B. M. (2004) Enhancement by Effectors and Substrate Nucleotides of R1-R2 Interactions in *Escherichia coli* Cass Ia Ribonucleotide Reductase, *J. Biol. Chem. 279*, 31050-31057.

88. Torrents, E., Westman, M., Sahlin, M., and Sjöberg, B. M. (2006) Ribonucleotide Reductase Modularity - Atypical Duplication of the ATP-cone Domain in *Pseudomonas aeruginosa*, *J. Biol. Chem. 281*, 25287-25296.

89. Ingemarson, R., and Thelander, L. (1996) A Kinetic Study on the Influence of Nucleoside Triphosphate Effectors on Subunit Interaction in Mouse Ribonucleotide Reductase, *Biochemistry 35*, 8603-8609.

90. Rofougaran, R., Crona, M., Vodnala, M., Sjöberg, B. M., and Hofer, A. (2008) Oligomerization Status Directs Overall Activity Regulation of the *Escherichia coli* Class Ia Ribonucleotide Reductase, *J. Biol. Chem. 283*, 35310-35318.

91. Bennati, M., Robblee, J. H., Mugnaini, V., Stubbe, J., Freed, J. H., and Borbat, P. (2005) EPR Distance Measurements Support a Model for Long-Range Radical Initiation in *E.coli* Ribonucleotide Reductase, *J. Am. Chem. Soc. 127*, 15014-15015.

92. Seyedsayamdost, M. R., Chan, C. T. Y., Mugnaini, V., Stubbe, J., and Bennati, M. (2007) PELDOR Spectroscopy with DOPA-b2 and NH_2Y-a2s: Distance Measurements between Residues Involved in the Radical Propagation Pathway of *E. coli* Ribonucleotide Reductase, *J. Am. Chem. Soc. 129*, 15748-15749.

93. Yokoyama, K., Smith, A. A., Corzilius, B., Griffin, R. G., and Stubbe, J. (2011) Equilibration of Tyrosyl Radicals (Y_{356}•, Y_{731}•, Y_{730}•) in the Radical Propagation Pathway of the *Escherichia coli* Class la Ribonucleotide Reductase, *J. Am. Chem. Soc. 133*, 18420-18432.

94. Minnihan, E. C., Ando, N., Brignole, E. J., Olshansky, L., Chittuluru, J., Asturias, F. J., Drennan, C. L., Nocera, D. G., and Stubbe, J. (2013) Generation of a Stable, Aminotyrosyl Radical-Induced a_2b_2 Complex of *Escherichia coli* Class Ia Ribonucleotide Reductase, *Proc. Natl. Acad. Sci. U. S. A. 110*, 3835-3840.

95. Zimanyi, C. M., Ando, N., Brignole, E. J., Asturias, F. J., Stubbe, J., and Drennan, C. L. (2012) Tangled Up in Knots: Structures of Inactivated Forms of *E. coli* Class Ia Ribonucleotide Reductase, *Structure 20*, 1374-1383.

96. Uppsten, M., Färnegårdh, M., Domkin, V., and Uhlin, U. (2006) The First Holocomplex Structure of Ribonucleotide Reductase Gives New Insight into Its Mechanism of Action, *J. Mol. Biol. 359*, 365-377.

97. Ge, J., Yu, G. X., Ator, M. A., and Stubbe, J. (2003) Pre-Steady-State and Steady-State Kinetic Analysis of *E. coli* Class I Ribonucleotide Reductase, *Biochemistry 42*, 10071-10083.

98. Minnihan, E. C., Nocera, D. G., and Stubbe, J. (2013) Reversible, Long-Range Radical Transfer in *E. coli* Class la Ribonucleotide Reductase, *Accounts Chem. Res. 46*, 2524-2535.

99. Climent, I., Sjöberg, B. M., and Huang, C. Y. (1992) Site-Directed Mutagenesis and Deletion of the Carboxyl Terminus of *Escherichia coli* Ribonucleotide Reductase Protein R2. Effects on Catalytic Activity and Subunit Interaction, *Biochemistry 31*, 4801-4807.

100. Ekberg, M., Sahlin, M., Eriksson, M., and Sjöberg, B. M. (1996) Two Conserved Tyrosine Residues in Protein R1 Participate in an Intermolecular Electron Transfer in Ribonucleotide Reductase, *J. Biol. Chem. 271*, 20655-20659.

101. Kasanmascheff, M., Lee, W., Nick, T. U., Stubbe, J., and Bennati, M. (2016) Radical Transfer in *E. coli* Ribonucleotide Reductase: a $NH_2Y_{731}/R_{411}A$-α Mutant Unmasks a New Conformation of the Pathway Residue 731, *Chem. Sci. 7*, 2170-2178.

102. Nick, T. U., Lee, W., Koßmann, S., Neese, F., Stubbe, J., and Bennati, M. (2015) Hydrogen Bond Network between Amino Acid Radical Intermediates on the Proton-Coupled Electron Transfer Pathway of *E. coli* α2 Ribonucleotide Reductase, *J. Am. Chem. Soc. 137*, 289-298.

103. Oyala, P. H., Ravichandran, K., Funk, M. A., Stucky, P. A., Stich, T. A., Drennan, C. L., Britt, R. D., and Stubbe, J. (2016) Biophysical Characterization of Fluorotyrosine Probes Site-specifically Incorporated into Enzymes: *E. coli* Ribonucleotide Reductase as an Example, *J. Am. Chem. Soc. 138*, 7951-7964.

104. Ravichandran, K., Minnihan, E. C., Wei, Y. F., Nocera, D. G., and Stubbe, J. (2015) Reverse Electron Transfer Completes the Catalytic Cycle in a 2,3,5-Trifluorotyrosine-Substituted Ribonucleotide Reductase, *J. Am. Chem. Soc. 137*, 14387-14395.

105. Wörsdorfer, B., Conner, D. A., Yokoyama, K., Livada, J., Seyedsayamdost, M., Jiang, W., Silakov, A., Stubbe, J., Bollinger, J. M., and Krebs, C. (2013) Function of the Diiron Cluster of *Escherichia coli* Class Ia Ribonucleotide Reductase in Proton-Coupled Electron Transfer, *J. Am. Chem. Soc. 135*, 8585-8593.

106. Baldwin, J., Krebs, C., Ley, B. A., Edmondson, D. E., Huynh, B. H., and Bollinger, J. H. (2000) Mechanism of Rapid Electron Transfer During Oxygen Activation in the R2 Subunit of *Escherichia coli* Ribonucleotide Reductase. 1. Evidence for a Transient Tryptophan Radical, *J. Am. Chem. Soc. 122*, 12195-12206.

107. Krebs, C., Chen, S. X., Baldwin, J., Ley, B. A., Patel, U., Edmondson, D. E., Huynh, B. H., and Bollinger, J. M. (2000) Mechanism of Rapid Electron Transfer During Oxygen Activation in the R2 Subunit of *Escherichia coli* Ribonucleotide Reductase. 2. Evidence for and Consequences of Blocked Electron Transfer in the W48F Variant, *J. Am. Chem. Soc. 122*, 12207-12219.

108. Ravichandran, K., (2016) *Mechanistic Investigations of the Radical Transport Pathway in Fluorotyrosine-Substituted Class Ia Ribonucleotide Reductase*. Ph.D. Thesis, Massachusets Institute of Technology.

109. Dassama, L. M. K., Jiang, W., Varano, P. T., Pandelia, M. E., Conner, D. A., Xie, J. J., Bollinger, J. M., and Krebs, C. (2012) Radical-Translocation Intermediates and Hurdling of Pathway Defects in "Super-oxidized" (Mn^{IV}/Fe^{IV}) *Chlamydia trachomatis* Ribonucleotide Reductase, *J. Am. Chem. Soc. 134*, 20498-20506.

110. Jiang, W., Xie, J. J., Varano, P. T., Krebs, C., and Bollinger, J. M. (2010) Two Distinct Mechanisms of Inactivation of the Class Ic Ribonucleotide Reductase from *Chlamydia trachomatis* by Hydroxyurea: Implications for the Protein Gating of Intersubunit Electron Transfer, *Biochemistry 49*, 5340-5349.

111. Livada, J., Martinie, R. J., Dassama, L. M. K., Krebs, C., Bollinger, J. M., and Silakov, A. (2015) Direct Measurement of the Radical Translocation Distance in the Class I Ribonucleotide Reductase from *Chlamydia trachomatis*, *J. Phys. Chem. B 119*, 13777-13784.

112. Mathews, C. K. (2006) DNA Precursor Metabolism and Genomic Stability, *FASEB J. 20*, 1300-1314.

113. Mathews, C. K. (2015) Deoxyribonucleotide Metabolism, Mutagenesis and Cancer, *Nat. Rev. Cancer 15*, 528-539.

114. Shao, J. M., Liu, X. Y., Zhu, L. J., and Yen, Y. (2013) Targeting Ribonucleotide Reductase for Cancer Therapy, *Expert Opin. Ther. Targets 17*, 1423-1437.

115. Wnuk, S. F., and Robins, M. J. (2006) Ribonucleotide Reductase Inhibitors as Anti-Herpes Agents, *Antiviral Res. 71*, 122-126.

116. Nordlund, N., and Reichard, P. (2006) Ribonucleotide Reductases, In *Annual Review of Biochemistry*, pp 681-706, Annual Reviews, Palo Alto.

117. Yao, R. J., Zhang, Z., An, X. X., Bucci, B., Perlstein, D. L., Stubbe, J., and Huang, M. X. (2003) Subcellular Localization of Yeast Ribonucleotide Reductase Regulated by the DNA Replication and Damage Checkpoint Pathways, *Proc. Natl. Acad. Sci. U. S. A. 100*, 6628-6633.

118. Guarino, E., Salguero, I., and Kearsey, S. E. (2014) Cellular Regulation of Ribonucleotide Reductase in Eukaryotes, *Semin. Cell Dev. Biol. 30*, 97-103.

119. Gerth, U., Kock, H., Kusters, I., Michalik, S., Switzer, R. L., and Hecker, M. (2008) Clp-Dependent Proteolysis Down-Regulates Central Metabolic Pathways in Glucose-Starved *Bacillus subtilis*, *J. Bacteriol. 190*, 321-331.

120. Arnaoutov, A., and Dasso, M. (2014) IRBIT is a Novel Regulator of Ribonucleotide Reductase in Higher Eukaryotes, *Science 345*, 1512-1515.

121. Kashlan, O. B., Scott, C. P., Lear, J. D., and Cooperman, B. S. (2002) A Comprehensive Model for the Allosteric Regulation of Mammalian Ribonucleotide Reductase. Functional Consequences of ATP- and dATP-Induced Oligomerization of the Large Subunit, *Biochemistry 41*, 462-474.

122. Brown, N. C., and Reichard, P. (1969) Role of Effector Binding in Allosteric Control of Ribonucleoside Diphosphate Reductase, *J. Mol. Biol. 46*, 39-55.

123. Holmgren, A., Reichard, P., and Thelander, L. (1965) Enzymatic Synthesis of Deoxyribonucleotides, VIII. Effects of ATP and dATP in the CDP Reductase System from *E. coli*, *Proc. Natl. Acad. Sci. U. S. A. 54*, 830-836.

124. Larsson, A., and Reichard, P. (1966) Enzymatic Synthesis of Deoxyribonucleotides. IX. Allosteric Effects in the Reduction of Pyrimidine Ribonucleotides by the Ribonucleoside Diphosphate Reductase System of *Escherichia coli*, *J. Biol. Chem. 241*, 2533-2539.

125. Larsson, A., and Reichard, P. (1966) Enzymatic Synthesis of Deoxyribonucleotides. X. Reduction of Purine Ribonucleotides; Allosteric Behavior and Substrate Specificity of the Enzyme System from *Escherichia coli* B, *J. Biol. Chem. 241*, 2540-2549.

126. Reichard, P. (1962) Enzymatic Synthesis of Deoxyribonucleotides. I. Formation of Deoxycytidine Diphosphate from Cytidine Diphosphate with Enzymes from *Escherichia coli*, *J. Biol. Chem. 237*, 3513-3519.

127. Reichard, P., Baldesten, A., and Rutberg, L. (1961) Formation of Deoxycytidine Phosphates from Cytidine Phosphates in Extracts from *Escherichia coli*, *J. Biol. Chem. 236*, 1150-1157.

128. Reichard, P., Canellakis, Z. N., and Canellakis, E. S. (1961) Studies on a Possible Regulatory Mechanism for Biosynthesis of Deoxyribonucleic Acid, *J. Biol. Chem. 236*, 2514-2519.

129. von Döbeln, U., and Reichard, P. (1976) Binding of Substrates to *Escherichia coli* Ribonucleotide Reductase, *J. Biol. Chem. 251*, 3616-3622.

130. Berglund, O. (1972) Ribonucleoside Diphosphate Reductase Induced by Bacteriophage T4. II. Allosteric Regulation of Substrate Specificity and Catalytic Activity, *J. Biol. Chem. 247*, 7276-7281.

131. Chen, A. K., Bhan, A., Hopper, S., Abrams, R., and Franzen, J. S. (1974) Substrate and Effector Binding to Ribonucleoside Triphosphate Reductase of *Lactobacillus leichmannii*, *Biochemistry 13*, 654-661.

132. Eriksson, S., Thelander, L., and Åkerman, M. (1979) Allosteric Regulation of Calf Thymus Ribonucleoside Diphosphate Reductase, *Biochemistry 18*, 2948-2952.

133. Hofer, A., Ekanem, J. T., and Thelander, L. (1998) Allosteric Regulation of *Trypanosoma brucei* Ribonucleotide Reductase Studied *in vitro* and *in vivo*, *J. Biol. Chem. 273*, 34098-34104.

134. Scott, C. P., Kashlan, O. B., Lear, J. D., and Cooperman, B. S. (2001) A Quantitative Model for Allosteric Control of Purine Reduction by Murine Ribonucleotide Reductase, *Biochemistry 40*, 1651-1661.

135. Hofer, A., Schmidt, P. P., Gräslund, A., and Thelander, L. (1997) Cloning and Characterization of the R1 and R2 Subunits of Ribonucleotide Reductase from *Trypanosoma brucei*, *Proc. Natl. Acad. Sci. U. S. A. 94*, 6959-6964.

136. Chimploy, K., and Mathews, C. K. (2001) Mouse Ribonucleotide Reductase Control - Influence of Substrate Binding Upon Interactions with Allosteric Effectors, *J. Biol. Chem. 276*, 7093-7100.

137. Hendricks, S. P., and Mathews, C. K. (1997) Regulation of T4 Phage Aerobic Ribonucleotide Reductase - Simultaneous Assay of the Four Activities, *J. Biol. Chem. 272*, 2861-2865.

138. Hendricks, S. P., and Mathews, C. K. (1998) Allosteric Regulation of Vaccinia Virus Ribonucleotide Reductase, Analyzed by Simultaneous Monitoring of its Four Activities, *J. Biol. Chem. 273*, 29512-29518.

139. Averett, D. R., Lubbers, C., Elion, G. B., and Spector, T. (1983) Ribonucleotide Reductase Induced by Herpes Simplex Type 1 Virus - Characterization of a Distinct Enzyme, *J. Biol. Chem. 258*, 9831-9838.

140. Friedrich, N. C., Torrents, E., Gibb, E. A., Sahlin, M., Sjöberg, B. M., and Edgell, D. R. (2007) Insertion of a Homing Endonuclease Creates a Genes-In-Pieces Ribonudeotide Reductase that Retains Function, *Proc. Natl. Acad. Sci. U. S. A. 104*, 6176-6181.

141. Huszar, D., and Bacchetti, S. (1981) Partial Purification and Characterization of the Ribonucleotide Reductase Induced by Herpes Simplex Virus Infection of Mammalian Cells, *J. Virol. 37*, 580-588.

142. Poncedeleon, M., Eisenberg, R. J., and Cohen, G. H. (1977) Ribonucleotide Reductase from Herpes Simplex Virus (Type 1 and Type 2) Infected and Uninfected KB Cells: Properties of the Partially Purified Enzymes, *J. Gen. Virol. 36*, 163-173.

143. Duffy, S., Shackelton, L. A., and Holmes, E. C. (2008) Rates of Evolutionary Change in Viruses: Patterns and Determinants, *Nat. Rev. Genet. 9*, 267-276.

144. Johansson, R., Jonna, V. R., Kumar, R., Nayeri, N., Lundin, D., Sjöberg, B. M., Hofer, A., and Logan, D. T. (2016) Structural Mechanism of Allosteric Activity Regulation in a Ribonucleotide Reductase with Double ATP Cones, *Structure 24*, 906-917.

145. Møllgaard, H., and Neuhard, J. (1978) Deoxycytidylate Deaminase from *Bacillus subtilis* - Purification, Characterization, and Physiological Function, *J. Biol. Chem. 253*, 3536-3542.

146. Weiss, B., and Wang, L. H. (1994) De Novo Synthesis of Thymidylate via Deoxycytidine in *dcd* (dCTP Deaminase) Mutants of *Escherichia coli*, *J. Bacteriol. 176*, 2194-2199.

147. Hofer, A., Steverding, D., Chabes, A., Brun, R., and Thelander, L. (2001) *Trypanosoma brucei* CTP Synthetase: A Target for the Treatment of African Sleeping Sickness, *Proc. Natl. Acad. Sci. U. S. A. 98*, 6412-6416.

148. Brignole, E. J., Ando, N., Zimanyi, C. M., and Drennan, C. L. (2012) The Prototypic Class Ia Ribonucleotide Reductase from *Escherichia coli*: Still Surprising After All These Years, *Biochem. Soc. Trans. 40*, 523-530.

149. Yaginuma, H., Kawai, S., Tabata, K. V., Tomiyama, K., Kakizuka, A., Komatsuzaki, T., Noji, H., and Imamura, H. (2014) Diversity in ATP Concentrations in a Single Bacterial Cell Population Revealed by Quantitative Single-Cell Imaging, *Sci Rep. 4*, 7.

150. Crona, M., Furrer, E., Torrents, E., Edgell, D. R., and Sjöberg, B. M. (2010) Subunit and Small-Molecule Interaction of Ribonucleotide Reductases Via Surface Plasmon Resonance Biosensor Analyses, *Protein Eng. Des. Sel. 23*, 633-641.

151. Ormö, M., and Sjöberg, B. M. (1990) An Ultrafiltration Assay for Nucleotide Binding to Ribonucleotide Reductase, *Anal. Biochem. 189*, 138-141.

152. Zimanyi, C. M., (2013) *Structural Studies of Allosteric Regulation in the Class Ia Ribonucleotide Reductase from Escherichia coli*. Ph.D. Thesis, Massachusetts Institute of Technology.

153. Grinberg, I., Shteinberg, T., Gorovitz, B., Aharonowitz, Y., Cohen, G., and Borovok, I. (2006) The *Streptomyces* NrdR Transcriptional Regulator is a Zn Ribbon/ATP Cone Protein that Binds to the Promoter Regions of Class Ia and Class II Ribonucleotide Reductase Operons, *J. Bacteriol. 188*, 7635-7644.

154. Roshick, C., Iliffe-Lee, E. R., and McClarty, G. (2000) Cloning and Characterization of Ribonucleotide Reductase from *Chlamydia trachomatis*, *J. Biol. Chem. 275*, 38111-38119.

155. Larsson, K. M., Jordan, A., Eliasson, R., Reichard, P., Logan, D. T., and Nordlund, P. (2004) Structural Mechanism of Allosteric Substrate Specificity Regulation in a Ribonucleotide Reductase, *Nat. Struct. Mol. Biol. 11*, 1142-1149.

156. Xu, H., Faber, C., Uchiki, T., Fairman, J. W., Racca, J., and Dealwis, C. (2006) Structures of Eukaryotic Ribonucleotide Reductase I Provide Insights into dNTP Regulation, *Proc. Natl. Acad. Sci. U. S. A. 103*, 4022-4027.

157. Nicholson, W. L., Munakata, N., Horneck, G., Melosh, H. J., and Setlow, P. (2000) Resistance of *Bacillus* Endospores to Extreme Terrestrial and Extraterrestrial Environments, *Microbiol. Mol. Biol. Rev. 64*, 548-572.

158. Mohamed, S. F., Gvozdiak, O. R., Stallmann, D., Griepenburg, U., Follmann, H., and Auling, G. (1998) Ribonucleotide Reductase in *Bacillus subtilis* - Evidence for a Mn-dependent Enzyme, *Biofactors 7*, 337-344.

159. Maaβ, S., Sievers, S., Zühlke, D., Kuzinski, J., Sappa, P. K., Muntel, J., Hessling, B., Bernhardt, J., Sietmann, R., Völker, U., Hecker, M., and Becher, D. (2011) Efficient, Global-Scale Quantification of Absolute Protein Amounts by Integration of Targeted Mass Spectrometry and Two-Dimensional Gel-Based Proteomics, *Anal. Chem. 83*, 2677-2684.

160. Maaβ, S., Wachlin, G., Bernhardt, J., Eymann, C., Fromion, V., Riedel, K., Becher, D., and Hecker, M. (2014) Highly Precise Quantification of Protein Molecules per Cell During Stress and Starvation Responses in *Bacillus subtilis*, *Mol. Cell. Proteomics 13*, 2260-2276.

161. Muntel, J., Fromion, V., Goelzer, A., Maaβ, S., Mäder, U., Büttner, K., Hecker, M., and Becher, D. (2014) Comprehensive Absolute Quantification of the Cytosolic Proteome of *Bacillus subtilis* by Data Independent, Parallel Fragmentation in Liquid Chromatography/Mass Spectrometry (LC/MS^E), *Mol. Cell. Proteomics 13*, 1008-1019.

162. Chen, J., Shen, J., Solem, C., and Jensen, P. R. (2015) A New Type of YumC-Like Ferredoxin (Flavodoxin) Reductase Is Involved in Ribonucleotide Reduction, *mBio 6*, 8.

163. Lofstad, M., Gudim, I., Hammerstad, M., Røhr, A. K., and Hersleth, H. P. (2016) Activation of the Class Ib Ribonucleotide Reductase by a Flavodoxin Reductase in *Bacillus cereus*, *Biochemistry 55*, 4998-5001.

164. Earl, A. M., Losick, R., and Kolter, R. (2008) Ecology and Genomics of *Bacillus subtilis*, *Trends Microbiol. 16*, 269-275.

165. Hong, H. A., Khaneja, R., Tam, N. M. K., Cazzato, A., Tan, S., Urdaci, M., Brisson, A., Gasbarrini, A., Barnes, I., and Cutting, S. M. (2009) *Bacillus subtilis* Isolated from the Human Gastrointestinal Tract, *Res. Microbiol. 160*, 134-143.

166. Schyns, G., Serra, C. R., Lapointe, T., Pereira-Leal, J., Potot, S., Fickers, P., Perkins, J. B., Wyss, M., and Henriques, A. O. (2013) Genome of a Gut Strain of *Bacillus subtilis*, *Genome Announc. 1*, e00184-00112.

167. Cazorla, F. M., Romero, D., Pérez-Garcia, A., Lugtenberg, B. J. J., de Vicente, A., and Bloemberg, G. (2007) Isolation and Characterization of Antagonistic *Bacillus subtilis* Strains from the Avocado Rhizoplane Displaying Biocontrol Activity, *J. Appl. Microbiol. 103*, 1950-1959.

168. Earl, A. M., Eppinger, M., Fricke, W. F., Rosovitz, M. J., Rasko, D. A., Daugherty, S., Losick, R., Kolter, R., and Ravel, J. (2012) Whole-Genome Sequences of *Bacillus subtilis* and Close Relatives, *J. Bacteriol. 194*, 2378-2379.

169. Kunst, F., and Ogasawara, N., and Moszer, I., and Albertini, A. M., and Alloni, G., and Azevedo, V., and Bertero, M. G., and Bessieres, P., and Bolotin, A., and Borchert, S., and Borriss, R., and Boursier, L., and Brans, A., and Braun, M., and Brignell, S. C., and Bron, S., and Brouillet, S., and Bruschi, C. V., and Caldwell, B., and Capuano, V., and Carter, N. M., and Choi, S. K., and Codani, J. J., and Connerton, I. F., and Cummings, N. J., and Daniel, R. A., and Denizot, F., and Devine, K. M., and Dusterhoft, A., and Ehrlich, S. D., and Emmerson, P. T., and Entian, K. D., and Errington, J., and Fabret, C., and Ferrari, E., and Foulger, D., and Fritz, C., and Fujita, M., and Fujita, Y., and Fuma, S., and Galizzi, A., and Galleron, N., and Ghim, S. Y., and Glaser, P. *et al.* (1997) The Complete Genome Sequence of the Gram-Positive Bacterium *Bacillus subtilis*, *Nature 390*, 249-256.

170. Härtig, E., and Jahn, D. (2012) Regulation of the Anaerobic Metabolism in Bacillus subtilis, In *Advances in Bacterial Respiratory Physiology* (Poole, R. K., Ed.), pp 195-216, Academic Press Ltd-Elsevier Science Ltd, London.

171. Nakano, M. M., and Zuber, P. (1998) Anaerobic Growth of a "Strict Aerobe" (*Bacillus subtilis*), *Annu. Rev. Microbiol. 52*, 165-190.

172. Härtig, E., Hartmann, A., Schätzle, M., Albertini, A. M., and Jahn, D. (2006) The *Bacillus subtilis nrdEF* Genes, Encoding a Class Ib Ribonucleotide Reductase, are Essential for Aerobic and Anaerobic Growth, *Appl. Environ. Microbiol. 72*, 5260-5265.

173. Karamata, D., and Gross, J. D. (1970) Isolation and Genetic Analysis of Temperature-Sensitive Mutants of *B. subtilis* Defective in DNA Synthesis, *Mol. Gen. Genet. 108*, 277-287.

174. Scotti, C., Valbuzzi, A., Perego, M., Galizzi, A., and Albertini, A. M. (1996) The *Bacillus subtilis* Genes for Ribonucleotide Reductase are Similar to the Genes for the Second Class I NrdE/NrdF Enzymes of *Enterobacteriaceae*, *Microbiology 142*, 2995-3004.

175. Nakano, M. M., Zuber, P., Glaser, P., Danchin, A., and Hulett, F. M. (1996) Two-Component Regulatory Proteins ResD-ResE are Required for Transcriptional Activation of *fnr* upon Oxygen Limitation in *Bacillus subtilis*, *J. Bacteriol. 178*, 3796-3802.

176. Sun, G. F., Sharkova, E., Chesnut, R., Birkey, S., Duggan, M. F., Sorokin, A., Pujic, P., Ehrlich, S. D., and Hulett, F. M. (1996) Regulators of Aerobic and Anaerobic Respiration in *Bacillus subtilis*, *J. Bacteriol. 178*, 1374-1385.

177. Ochi, K., Kandala, J., and Freese, E. (1982) Evidence that *Bacillus subtilis* Sporulation Induced by the Stringent Response is Caused by the Decrease in GTP or GDP, *J. Bacteriol. 151*, 1062-1065.

178. Kriel, A., Bittner, A. N., Kim, S. H., Liu, K. Q., Tehranchi, A. K., Zou, W. Y., Rendon, S., Chen, R., Tu, B. P., and Wang, J. D. (2012) Direct Regulation of GTP Homeostasis by (p)ppGpp: A Critical Component of Viability and Stress Resistance, *Mol. Cell 48*, 231-241.

179. Eymann, C., Homuth, G., Scharf, C., and Hecker, M. (2002) *Bacillus subtilis* Functional Genomics: Global Characterization of the Stringent Response by Proteome and Transcriptome Analysis, *J. Bacteriol. 184*, 2500-2520.

180. Westers, H., Dorenbos, R., van Dijl, J. M., Kabel, J., Flanagan, T., Devine, K. M., Jude, F., Séror, S. J., Beekman, A. C., Darmon, E., Eschevins, C., de Jong, A., Bron, S., Kuipers, O. P., Albertini, A. M., Antelmann, H., Hecker, M., Zamboni, N., Sauer, U., Bruand, C., Ehrlich, D. S., Alonso, J. C., Salas, M., and Quax, W. J. (2003) Genome Engineering Reveals Large Dispensable Regions in *Bacillus subtilis*, *Mol. Biol. Evol. 20*, 2076-2090.

181. Kobayashi, K., Ehrlich, S. D., Albertini, A., Amati, G., Andersen, K. K., Arnaud, M., Asai, K., Ashikaga, S., Aymerich, S., Bessieres, P., Boland, F., Brignell, S. C., Bron, S., Bunai, K., Chapuis, J., Christiansen, L. C., Danchin, A., Débarbouillé, M., Dervyn, E., Deuerling, E., Devine, K., Devine, S. K., Dreesen, O., Errington, J., Fillinger, S., Foster, S. J. *et al.* (2003) Essential *Bacillus subtilis* Genes, *Proc. Natl. Acad. Sci. U. S. A. 100*, 4678-4683.

182. Lazarevic, V., Soldo, B., Dusterhoft, A., Hilbert, H., Mauel, C., and Karamata, D. (1998) Introns and Intein Coding Sequence in the Ribonucleotide Reductase Genes of *Bacillus subtilis* Temperate Bacteriophage SPb, *Proc. Natl. Acad. Sci. U. S. A. 95*, 1692-1697.

183. Nicolas, P., Mäder, U., Dervyn, E., Rochat, T., Leduc, A., Pigeonneau, N., Bidnenko, E., Marchadier, E., Hoebeke, M., Aymerich, S., Becher, D., Bisicchia, P., Botella, E., Delumeau, O., Doherty, G., Denham, E. L., Fogg, M. J., Fromion, V., Goelzer, A., Hansen, A., Härtig, E., Harwood, C. R., Homuth, G., Jarmer, H., Jules, M., Klipp, E., Le Chat, L., Lecointe, F., Lewis, P., Liebermeister, W., March, A., Mars, R. A. T., Nannapaneni, P., Noone, D., Pohl, S., Rinn, B., Rügheimer, F., Sappa, P. K., Samson, F., Schaffer, M., Schwikowski, B., Steil, L., Stülke, J., Wiegert, T., Devine, K. M., Wilkinson, A. J., van Dijl, J. M., Hecker, M., Völker, U., Bessières, P., and Noirot, P. (2012) Condition-Dependent Transcriptome Reveals High-Level Regulatory Architecture in *Bacillus subtilis*, *Science 335*, 1103-1106.

184. Rasmussen, S., Nielsen, H. B., and Jarmer, H. (2009) The Transcriptionally Active Regions in the Genome of *Bacillus subtilis*, *Mol. Microbiol. 73*, 1043-1057.

185. Marolewski, A., Smith, J. M., and Benkovic, S. J. (1994) Cloning and Characterization of a New Purine Biosynthetic Enzyme: A Non-Folate Glycinamide Ribonucleotide Transformylase from *Escherichia coli*, *Biochemistry 33*, 2531-2537.

186. Saxild, H. H., Jacobsen, J. H., and Nygaard, P. (1994) Genetic and Physiological Characterization of a Formate-Dependent 5'-Phosphoribosyl-1-Glycinamide Transformylase Activity in *Bacillus subtilis*, *Molecular & General Genetics 242*, 415-420.

187. Ebbole, D. J., and Zalkin, H. (1987) Cloning and Characterization of a 12 Gene Cluster from *Bacillus subtilis* Encoding Nine Enzymes for *de novo* Purine Nucleotide Synthesis, *J. Biol. Chem. 262*, 8274-8287.

188. Ebbole, D. J., and Zalkin, H. (1989) Interaction of a Putative Repressor Protein with an Extended Control Region of the *Bacillus subtilis pur* Operon, *J. Biol. Chem. 264*, 3553-3561.

189. Mandal, M., Boese, B., Barrick, J. E., Winkler, W. C., and Breaker, R. R. (2003) Riboswitches Control Fundamental Biochemical Pathways in *Bacillus subtilis* and Other Bacteria, *Cell 113*, 577-586.

190. Weng, M. L., Nagy, P. L., and Zalkin, H. (1995) Identification of the *Bacillus subtilis pur* Operon Repressor, *Proc. Natl. Acad. Sci. U. S. A. 92*, 7455-7459.

191. Saxild, H. H., Brunstedt, K., Nielsen, K. I., Jarmer, H., and Nygaard, P. (2001) Definition of the *Bacillus subtilis* PurR Operator Using Genetic and Bioinformatic Tools and Expansion of the PurR Regulon with *glyA*, *guaC*, *pbuG*, *xpt-pbuX*, *yqhZ-folD*, and *pbuO*, *J. Bacteriol. 183*, 6175-6183.

192. Turner, R. J., Lu, Y., and Switzer, R. L. (1994) Regulation of the *Bacillus subtilis* Pyrimidine Biosynthetic (*pyr*) Gene Cluster by an Autogenous Transcriptional Attenuation Mechanism, *J. Bacteriol. 176*, 3708-3722.

193. Lu, Y., and Switzer, R. L. (1996) Transcriptional Attenuation of the *Bacillus subtilis pyr* Operon by the PyrR Regulatory Protein and Uridine Nucleotides *in vitro*, *J. Bacteriol. 178*, 7206-7211.

194. Kloudová, A., and Fučik, V. (1974) Transport of Nucleosides in *Bacillus subtilis*: Characteristics of Cytidine Uptake, *Nucleic Acids Res. 1*, 629-637.

195. Meng, Q., and Switzer, R. L. (2001) Regulation of Transcription of the *Bacillus subtilis pyrG* Gene, Encoding Cytidine Triphosphate Synthetase, *J. Bacteriol. 183*, 5513-5522.

196. Saxild, H. H., Andersen, L. N., and Hammer, K. (1996) *dra-nupC-pdp* Operon of *Bacillus subtilis*: Nucleotide Sequence, Induction by Deoxyribonucleosides, and Transcriptional Regulation by the *deoR*-Encoded DeoR Repressor Protein, *J. Bacteriol. 178*, 424-434.

197. Beaman, T. C., Hitchins, A. D., Ochi, K., Vasantha, N., Endo, T., and Freese, E. (1983) Specificity and Control of Uptake of Purines and Other Compounds in *Bacillus subtilis*, *J. Bacteriol. 156*, 1107-1117.

198. Ghim, S. Y., and Neuhard, J. (1994) Pyrimidine Biosynthesis Operon of the Thermophile *Bacillus caldolyticus* Includes Genes for Uracil Phosphoribosyltransferase and Uracil Permease, *J. Bacteriol. 176*, 3698-3707.

199. Belitsky, B. R., and Sonenshein, A. L. (2011) CodY-Mediated Regulation of Guanosine Uptake in *Bacillus subtilis*, *J. Bacteriol. 193*, 6276-6287.

200. Christiansen, L. C., Schou, S., Nygaard, P., and Saxild, H. H. (1997) Xanthine Metabolism in *Bacillus subtilis*: Characterization of the *xpt-pbuX* Operon and Evidence for Purine- and Nitrogen-Controlled Expression of Genes Involved in Xanthine Salvage and Catabolism, *J. Bacteriol. 179*, 2540-2550.

201. Johansen, L. E., Nygaard, P., Lassen, C., Agerso, Y., and Saxild, H. H. (2003) Definition of a Second *Bacillus subtilis pur* Regulon Comprising the *pur* and *xpt-pbuX* Operons Plus *pbuG*, *nupG* (*yxjA*), and *pbuE* (*ydhL*), *J. Bacteriol. 185*, 5200-5209.

202. Nygaard, P., Duckert, P., and Saxild, H. H. (1996) Role of Adenine Deaminase in Purine Salvage and Nitrogen Metabolism and Characterization of the *ade* Gene in *Bacillus subtilis*, *J. Bacteriol. 178*, 846-853.

203. Saxild, H. H., and Nygaard, P. (1987) Genetic and Physiological Characterization of *Bacillus subtilis* Mutants Resistant to Purine Analogs, *J. Bacteriol. 169*, 2977-2983.

204. de Giuseppe, P. O., Martins, N. H., Meza, A. N., dos Santos, C. R., Pereira, H. D., and Murakami, M. T. (2012) Insights into Phosphate Cooperativity and Influence of Substrate Modifications on Binding and Catalysis of Hexameric Purine Nucleoside Phosphorylases, *PLoS One 7*, 11.

205. Schuch, R., Garibian, A., Saxild, H. H., Piggot, P. J., and Nygaard, P. (1999) Nucleosides as a Carbon Source in *Bacillus subtilis*: Characterization of the *drm-pupG* Operon, *Microbiology-(UK) 145*, 2957-2966.

206. Endo, T., Uratani, B., and Freese, E. (1983) Purine Salvage Pathways of *Bacillus subtilis* and Effect of Guanine on Growth of GMP Reductase Mutants, *J. Bacteriol. 155*, 169-179.

207. Martinussen, J., Glaser, P., Andersen, P. S., and Saxild, H. H. (1995) Two Genes Encoding Uracil Phosphoribosyltransferase are Present in *Bacillus subtilis*, *J. Bacteriol. 177*, 271-274.

208. Andersen, R. B., and Neuhard, J. (2001) Deoxynucleoside Kinases Encoded by the *yaaG* and *yaaF* Genes of *Bacillus subtilis* - Substrate Specificity and Kinetic Analysis of Deoxyguanosine Kinase with UTP as the Preferred Phosphate Donor, *J. Biol. Chem. 276*, 5518-5524.

209. Møllgaard, H. (1980) Deoxyadenosine/Deoxycytidine Kinase from *Bacillus subtilis*: Purification, Characterization, and Physiological Function, *J. Biol. Chem. 255*, 8216-8220.

210. Coote, J. G., and Binnie, C. (1986) Tolerance to Bromodeoxyuridine in a Thymidine-Requiring Strain of *Bacillus subtilis*, *J. Gen. Microbiol. 132*, 481-492.

211. Johansson, E., Neuhard, J., Willemoes, M., and Larsen, S. (2004) Structural, Kinetic, and Mutational Studies of the Zinc Ion Environment in Tetrameric Cytidine Deaminase, *Biochemistry 43*, 6020-6029.

212. Jeon, Y. J., Park, S. C., Song, W. S., Kim, O. H., Oh, B. C., and Yoon, S. I. (2016) Structural and Biochemical Characterization of Bacterial YpgQ Protein Reveals a Metal-Dependent Nucleotide Pyrophosphohydrolase, *J. Struct. Biol. 195*, 113-122.

213. Terakawa, A., Natsume, A., Okada, A., Nishihata, S., Kuse, J., Tanaka, K., Takenaka, S., Ishikawa, S., and Yoshida, K. (2016) *Bacillus subtilis* 5 '-Nucleotidases with Various Functions and Substrate Specificities, *BMC Microbiol. 16*, 13.

214. Nygaard, P., and Saxild, H. H. (2005) The Purine Efflux Pump PbuE in *Bacillus subtilis* Modulates Expression of the PurR and G-box (XptR) Regulons by Adjusting the Purine Base Pool Size, *J. Bacteriol. 187*, 791-794.

215. Zakataeva, N. P., Gronskiy, S. V., Sheremet, A. S., Kutukova, E. A., Novikova, A. E., and Livshits, V. A. (2007) A New Function for the *Bacillus* PbuE Purine Base Efflux Pump: Efflux of Purine Nucleosides, *Res. Microbiol. 158*, 659-665.

216. Vogels, G. D., and van der Drift, C. (1976) Degradation of Purines and Pyrimidines by Microorganisms, *Bacteriological Reviews 40*, 403-468.

217. Ramazzina, I., Costa, R., Cendron, L., Berni, R., Peracchi, A., Zanotti, G., and Percudani, R. (2010) An Aminotransferase Branch Point Connects Purine Catabolism to Amino Acid Recycling, *Nat. Chem. Biol. 6*, 801-806.

218. Schultz, A. C., Nygaard, P., and Saxild, H. H. (2001) Functional Analysis of 14 Genes that Constitute the Purine Catabolic Pathway in *Bacillus subtilis* and Evidence for a Novel Regulon Controlled by the PucR Transcription Activator, *J. Bacteriol. 183*, 3293-3302.

219. Rodionov, D. A., and Gelfand, M. S. (2005) Identification of a Bacterial Regulatory System for Ribonucleotide Reductases by Phylogenetic Profiling, *Trends Genet. 21*, 385-389.

220. Grinberg, I., Shteinberg, T., Hassan, A. Q., Aharonowitz, Y., Borovok, I., and Cohen, G. (2009) Functional Analysis of the *Streptomyces coelicolor* NrdR ATP-Cone Domain: Role in Nucleotide Binding, Oligomerization, and DNA Interactions, *J. Bacteriol. 191*, 1169-1179.

221. McKethan, B. L., and Spiro, S. (2013) Cooperative and Allosterically Controlled Nucleotide Binding Regulates the DNA Binding Activity of NrdR, *Mol. Microbiol. 90*, 278-289.

222. McLennan, A. G. (2006) The Nudix Hydrolase Superfamily, *Cell. Mol. Life Sci. 63*, 123-143.

223. Mildvan, A. S., Xia, Z., Azurmendi, H. F., Saraswat, V., Legler, P. M., Massiah, M. A., Gabelli, S. B., Bianchet, M. A., Kang, L. W., and Amzel, L. M. (2005) Structures and Mechanisms of Nudix Hydrolases, *Arch. Biochem. Biophys. 433*, 129-143.

224. Bhatnagar, S. K., and Bessman, M. J. (1988) Studies on the Mutator Gene, *mutT* of *Escherichia coli* - Molecular Cloning of the Gene, Purification of the Gene Product, and Identification of a Novel Nucleoside Triphosphatase, *J. Biol. Chem. 263*, 8953-8957.

225. Hori, M., Fujikawa, K., Kasai, H., Harashima, H., and Kamiya, H. (2005) Dual Hydrolysis of Diphosphate and Triphosphate Derivatives of Oxidized Deoxyadenosine by Orf17 (NtpA), a MutT-type Enzyme, *DNA Repair 4*, 33-39.

226. O'Handley, S. F., Frick, D. N., Bullions, L. C., Mildvan, A. S., and Bessman, M. J. (1996) *Escherichia coli orf17* Codes for a Nucleoside Triphosphate Pyrophosphohydrolase Member of the MutT Family of Proteins - Cloning, Purification, and Characterization of the Enzyme, *J. Biol. Chem. 271*, 24649-24654.

227. Frick, D. N., and Bessman, M. J. (1995) Cloning, Purification, and Properties of a Novel NADH Pyrophosphatase - Evidence for a Nucleotide Pyrophosphatase Catalytic Domain in MutT-like Enzymes, *J. Biol. Chem. 270*, 1529-1534.

228. Gasmi, L., and McLennan, A. G. (2001) The Mouse Nudt7 Gene Encodes a Peroxisomal Nudix Hydrolase Specific for Coenzyme A and its Derivatives, *Biochem. J. 357*, 33-38.

229. Xu, W. L., Gauss, P., Shen, J. Y., Dunn, C. A., and Bessman, M. J. (2002) The Gene *e.1* (*nudE.1*) of T4 Bacteriophage Designates a New Member of the Nudix Hydrolase Superfamily Active on Flavin Adenine Dinucleotide, Adenosine 5 '-Triphospho-5 '-adenosine, and ADP-ribose, *J. Biol. Chem. 277*, 23181-23185.

230. Legler, P. M., Massiah, M. A., Bessman, M. J., and Mildvan, A. S. (2000) GDP-Mannose Mannosyl Hydrolase Catalyzes Nucleophilic Substitution at Carbon, Unlike All Other Nudix Hydrolases, *Biochemistry 39*, 8603-8608.

231. Duong-Ly, K. C., Gabelli, S. B., Xu, W. L., Dunn, C. A., Schoeffield, A. J., Bessman, M. J., and Amzel, L. M. (2011) The Nudix Hydrolase CDP-Chase, a CDP-Choline Pyrophosphatase, is an Asymmetric Dimer with Two Distinct Enzymatic Activities, *J. Bacteriol. 193*, 3175-3185.

232. Abdelghany, H. M., Gasmi, L., Cartwright, J. L., Bailey, S., Rafferty, J. B., and McLennan, A. G. (2001) Cloning, Characterisation and Crystallisation of a Diadenosine 5 ',5 '''-P-1,P-4-tetraphosphate Pyrophosphohydrolase from *Caenorhabditis elegans*, *Biochim. Biophys. Acta-Protein Struct. Molec. Enzym. 1550*, 27-36.

233. Guranowski, A. (2000) Specific and Nonspecific Enzymes Involved in the Catabolism of Mononucleoside and Dinucleoside Polyphosphates, *Pharmacol. Ther. 87*, 117-139.

234. Iwai, T., Kuramitsu, S., and Masui, R. (2004) The Nudix Hydrolase Ndx1 from *Thermus thermophilus* HB8 is a Diadenosine Hexaphosphate Hydrolase with a Novel Activity, *J. Biol. Chem. 279*, 21732-21739.

235. Deana, A., Celesnik, H., and Belasco, J. G. (2008) The Bacterial Enzyme RppH Triggers Messenger RNA Degradation by 5' Pyrophosphate Removal, *Nature 451*, 355-358.

236. Richards, J., Liu, Q. S., Pellegrini, O., Celesnik, H., Yao, S. Y., Bechhofer, D. H., Condon, C., and Belasco, J. G. (2011) An RNA Pyrophosphohydrolase Triggers 5 '-Exonucleolytic Degradation of mRNA in *Bacillus subtilis*, *Mol. Cell 43*, 940-949.

237. Lawhorn, B. G., Gerdes, S. Y., and Begley, T. P. (2004) A Genetic Screen for the Identification of Thiamin Metabolic Genes, *J. Biol. Chem. 279*, 43555-43559.

238. Fisher, D. I., Safrany, S. T., Strike, P., McLennan, A. G., and Cartwright, J. L. (2002) Nudix Hydrolases that Degrade Dinucleoside and Diphosphoinositol Polyphosphates also have 5-Phosphoribosyl 1-Pyrophosphate (PRPP) Pyrophosphatase Activity that Generates the Glycolytic Activator Ribose 1,5-Bisphosphate, *J. Biol. Chem. 277*, 47313-47317.

239. Gabelli, S. B., Bianchet, M. A., Xu, W. L., Dunn, C., Niu, Z. D., Amzel, L. M., and Bessman, M. J. (2007) Structure and Function of the *E. coli* Dihydroneopterin Triphosphate Pyrophosphatase: A Nudix Enzyme Involved in Folate Biosynthesis, *Structure 15*, 1014-1022.

240. Klaus, S. M. J., Wegkamp, A., Sybesma, W., Hugenholtz, J., Gregory, J. F., and Hanson, A. D. (2005) A Nudix Enzyme Removes Pyrophosphate from Dihydroneopterin Triphosphate in the Folate Synthesis Pathway of Bacteria and Plants, *J. Biol. Chem. 280*, 5274-5280.

241. Dunn, C. A., O'Handley, S. F., Frick, D. N., and Bessman, M. J. (1999) Studies on the ADP-ribose Pyrophosphatase Subfamily of the Nudix Hydrolases and Tentative Identification of *trgB*, a Gene Associated with Tellurite Resistance, *J. Biol. Chem. 274*, 32318-32324.

242. Kang, L. W., Gabelli, S. B., Bianchet, M. A., Xu, W. L., Bessman, M. J., and Amzel, L. M. (2003) Structure of a Coenzyme A Pyrophosphatase from *Deinococcus radiodurans*: a Member of the Nudix Family, *J. Bacteriol. 185*, 4110-4118.

243. Xu, W. L., Shen, J. Y., Dunn, C. A., and Bessman, M. J. (2003) A New Subfamily of the Nudix Hydrolase Superfamily Active on 5-Methyl-UTP (ribo-TTP) and UTP, *J. Biol. Chem. 278*, 37492-37496.

244. Nakamura, T., Meshitsuka, S., Kitagawa, S., Abe, N., Yamada, J., Ishino, T., Nakano, H., Tsuzuki, T., Doi, T., Kobayashi, Y., Fujii, S., Sekiguchi, M., and Yamagata, Y. (2010) Structural and Dynamic Features of the MutT Protein in the Recognition of Nucleotides with the Mutagenic 8-Oxoguanine Base, *J. Biol. Chem. 285*, 444-452.

245. Wang, S. S., Mura, C., Sawaya, M. R., Cascio, D., and Eisenberg, D. (2002) Structure of a Nudix Protein from *Pyrobaculum aerophilum* Reveals a Dimer with Two Intersubunit β-sheets, *Acta Crystallogr. Sect. D-Biol. Crystallogr. 58*, 571-578.

246. Bessman, M. J., Frick, D. N., and O'Handley, S. F. (1996) The MutT Proteins or "Nudix" Hydrolases, a Family of Versatile, Widely Distributed, "Housecleaning" Enzymes, *J. Biol. Chem. 271*, 25059-25062.

247. Maki, H., and Sekiguchi, M. (1992) MutT Protein Specifically Hydrolyzes a Potent Mutagenic Substrate for DNA Synthesis, *Nature 355*, 273-275.

248. Fowler, R. G., White, S. J., Koyama, C., Moore, S. C., Dunn, R. L., and Schaaper, R. M. (2003) Interactions Among the *Escherichia coli mutT*, *mutM*, and *mutY* Damage Prevention Pathways, *DNA Repair 2*, 159-173.

249. Yang, F. D., Lu, G. Z., and Rubin, H. (1994) Isolation of Ribonucleotide Reductase from *Mycobacterium tuberculosis* and Cloning, Expression, and Purification of the Large Subunit, *J. Bacteriol. 176*, 6738-6743.

Chapter 2

Bacillus subtilis Class Ib Ribonucleotide Reductase: High Activity and Dynamic Subunit Interactions

Adapted with permission from Parker, M.J.; Zhu, X.; Stubbe, J. (**2014**) *Biochemistry* **53**: 766-776.
 DOI 10.1021/bi401056e

2.1. INTRODUCTION

Ribonucleotide reductases (RNRs, **Figure 2.1**) catalyze the conversion of nucleoside 5′-diphosphates (NDPs) to 2′-deoxynucleotides (dNDPs) and play an essential role in supplying balanced dNTP pools for DNA replication and repair.[1] The RNRs have been classified based on the metallo-cofactors used to generate a transient cysteine thiyl radial that is essential for catalysis.[2] The class I RNRs contain two subunits (α and β) with subclassification (Ia, Ib and Ic, **Figure 2.1**) based on the dimetallo-cofactor located in β. Class Ib RNRs, as with class Ia enzymes, were long thought to use a diferric-tyrosyl radical ($Fe(III)_2$-Y•) cofactor because the cluster can self-assemble and has catalytic activity.[3-5] Recently, however, the class Ib RNRs have been isolated and characterized from their endogenous sources (*Corynebacterium ammoniagenes*, *Escherichia coli*, and *Bacillus subtilis*) in sufficient amounts to establish that the cofactor is a dimanganic-tyrosyl radical ($Mn(III)_2$-Y•).[6-9] While most prokaryotes have multiple RNRs, *B. subtilis* has a single essential class Ib RNR found in the *nrdI-nrdE-nrdF-ymaB* operon.[10-12] The apparent tight binding of the subunits of the *B. subtilis* RNR[9] and the low reported specific activities of this and other class Ib enzymes (Appendix 2, **Table A2.1**), provided the impetus to maximize its catalytic activity for future studies and for structural determination.

In all class I RNRs, α (designated NrdE for class Ib and NrdA for class Ia, **Figure 2.1**) contains the active site for NDP reduction and binding sites for allosteric effectors that control substrate specificity.[1] The class Ia α contains a second allosteric site in its N-terminal ATP-cone domain that controls the overall rate of nucleotide reduction and is absent in the class Ib α subunits. The β subunit (NrdF for class Ib and NrdB for class Ia) are all dimeric (β_2) with the class Ia housing a $Fe(III)_2$-Y• cofactor (typically with 3.6 Fe/β_2 and 1 Y•/β_2) and the class Ib, housing a $Mn(III)_2$-Y•, whose stoichiometry remains to be optimized in all systems.[2]

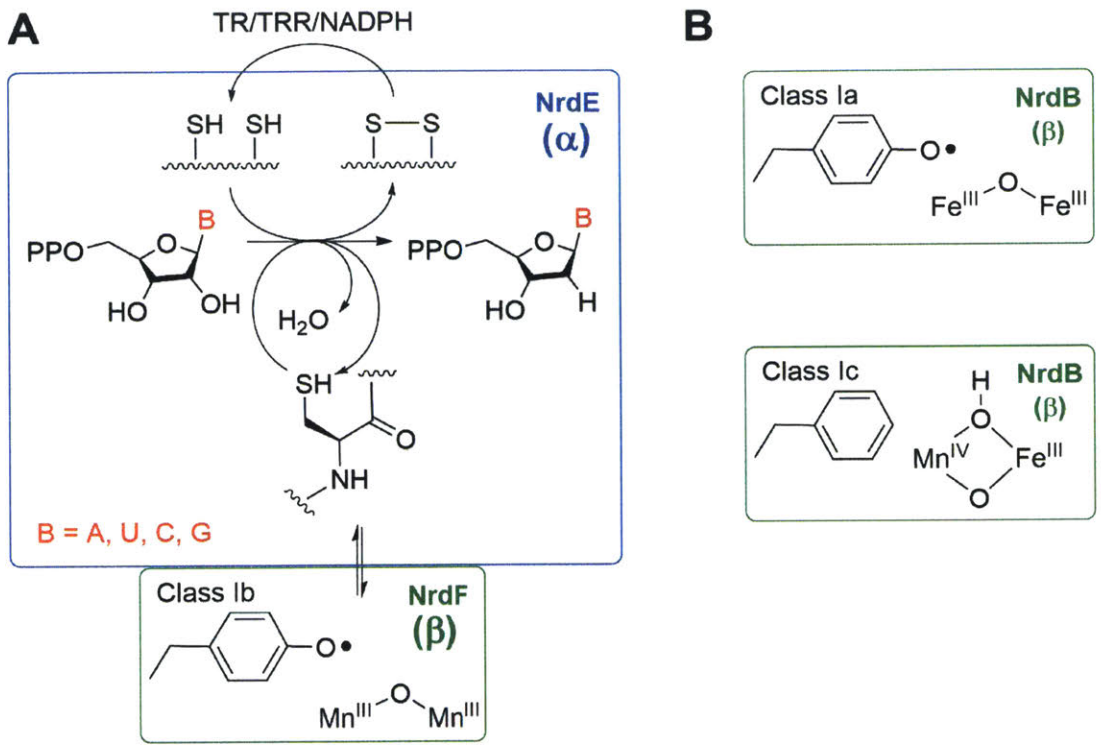


Figure 2.1. Ribonucleotide reduction catalyzed by class I RNRs. (**A**) The class Ib RNRs are composed of NrdE (α) and NrdF (β) which are proposed to form an $\alpha_2\beta_2$ complex during substrate turnover (only one α/β pair depicted). The stable dimanganic-Y• housed in β (1 Y•/β_2) oxidizes an essential Cys residue in the active site of α to a thiyl radical through a long range radical transfer process involving redox-active amino acids in both subunits. Reduction of NDPs to dNDPs occurs concomitant with oxidation of two Cys residues to a disulfide. The disulfide is re-reduced by a thioredoxin (TR) or NrdH (a glutaredoxin-like thioredoxin), thioredoxin reductase (TRR), and NADPH. (**B**) The dimetallo-cofactors of the class Ia and Ic β subunits.

Recent data suggest that in addition to allosteric effectors, quaternary structure(s) of the class I RNRs, including subunit affinity and its dependence on nucleotide binding, are all important in regulating RNR activity.[1] The subunit affinity in the *E. coli* and human class Ia RNRs have reported K_ds of ~ 0.2 μM in the presence of substrate (CDP) and effector (ATP).[13-15] While there is now a consensus that the active quaternary structure of *E. coli* class Ia RNR is $\alpha_2\beta_2$,[16-18] the eukaryotic Ia structure is open to debate [$(\alpha_2)_m(\beta_2)_n$ (m = 1, 3 and n = 1, 3)].[19-21] Recent studies have further revealed that the active complexes of both the prokaryotic and eukaryotic RNRs are

dynamic.[16, 17, 19, 22] The quaternary structures of the class Ib RNRs, especially with the $Mn(III)_2$-Y• cofactor, have been much less extensively studied and remain to be determined.[3, 23, 24]

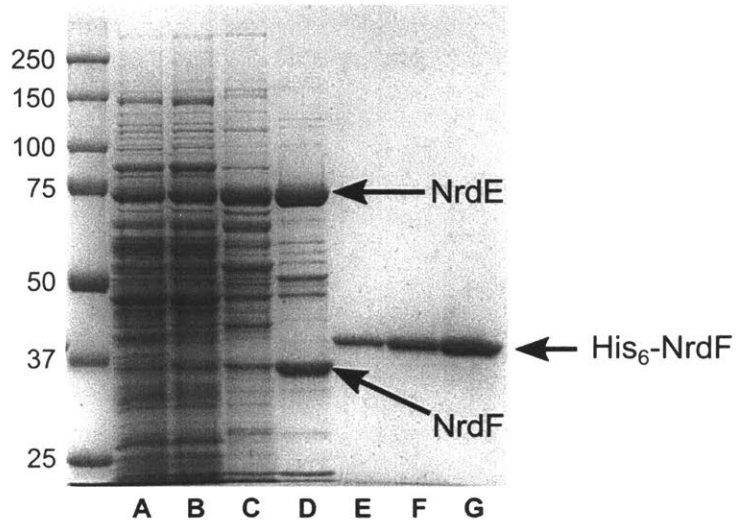


Figure 2.2. Co-purification of endogenous NrdE and NrdF from a strain (1B-UP) of *B. subtilis* engineered to overexpress the class Ib RNR operon. Samples from each purification step were loaded and run on a 10% (w/v) SDS-PAGE gel: *lane A* – 30 μg crude cell extract, *lane B* – 30 μg of resolubilized 40% (w/v) NH_4SO_4 pellet, *lane C* – 15 μg of protein recovered from Butyl-Sepharose hydrophobic affinity chromatography, *lane D* – 10 μg of protein recovered from Poros HQ/20 FPLC anion exchange chromatography, *lanes E–G* – 1, 2, and 4 μg of purified recombinant His_6-NrdF, respectively. Adapted with permission from reference 9. Copyright 2011 American Chemical Society.

Since the establishment that class Ib RNRs possess a $Mn(III)_2$-Y• cofactor, many efforts have been made to isolate and characterize these enzymes.[6-9, 23-25] With one exception,[1] the reported activities of purified class Ib RNRs, whether endogenously or recombinantly sourced, are substantially lower (Appendix 2, **Table A2.1**) than their class Ia counterparts (5000 – 8000 nmol min^{-1} mg^{-1}).[26] Recently, our lab isolated the class Ib RNR from *B. subtilis* with NrdE and NrdF co-purifying in a 1:1 ratio through three chromatographic steps in the absence of nucleotides and Mg^{2+}, suggesting "tight" subunit affinity (**Figure 2.2**).[9] However, while the RNR was 80% pure,

[1] The reported activity of the *C. ammoniagenes* class Ib RNR[8] is inconsistent with activities of all known RNRs and is likely incorrect.

its catalytic activity was low (~160 nmol min^{-1} mg^{-1}). Efforts to reconstitute active RNR using recombinantly expressed NrdE and reconstituted $Mn(III)_2$-Y• NrdF also resulted in low catalytic activity.[9, 27]

The work reported in this chapter describes the optimization of the catalytic activity of the *B. subtilis* RNR by separation of active $Me(III)_2$-Y•-loaded NrdF (Me = Fe or Mn) from the apo- and mismetallated forms and reveals that the active form has 1 Y•/β_2 and 4 Mn/β_2. Also reported is the identification, cloning, expression, and purification of candidate endogenous reductants for *B. subtilis* RNR: TrxA, YosR, and YtnI, a thioredoxin and two glutaredoxin-like thioredoxins, respectively, and TrxB, a thioredoxin reductase. Both TrxA and YosR supported NDP reduction in the presence of TrxB and NADPH, whereas YtnI was not reduced by TrxB and, thus, was eliminated from consideration as a reductant of RNR. Activity assays with a 1:1 ratio of $Mn(III)_2$-Y• NrdF:NrdE using either TrxA or YosR with TrxB produced specific activities of ~1250 nmol min^{-1} mg^{-1}, 7-fold higher activity than the $Fe(III)_2$-Y• cofactor, and substantially higher than other characterized class Ib RNRs (Appendix 2, **Table A2.1**). A comparison of the ability of reductants (DTT or TrxA/TrxB/NADPH) to support RNR activity with purified or unpurified $Mn(III)_2$-Y• NrdF revealed that the use of the endogenous reductants contributed significantly more to the activity enhancement relative to the cluster loading of NrdF. Activity assays with CDP/ATP and increasing concentrations of NrdE:NrdF maintained in a 1:1 ratio gave an apparent K_m of 25 nM, 10-fold lower than the apparent K_d for the Ia RNRs with the same substrate/effector pair.[13, 14] The quaternary structure of each subunit and a 1:1 mixture of NrdE:NrdF at physiological concentrations (1 μM)[9, 28-30] was examined by size exclusion chromatography (SEC) and sedimentation velocity analytical ultracentrifugation (SV-AUC) in the absence of nucleotides. These studies reveal that NrdE is a monomer, $Me(III)_2$-Y• NrdF is a dimer, and the 1:1 mixture of

NrdE and NrdF is a mixture of subunits and higher molecular weight species that are interconverting on the timescale of the techniques. These studies set the stage for determination of nucleotide binding and how it effects subunit interactions and quaternary structure in a class Ib RNR.

2.2. EXPERIMENTAL

2.2.1. Materials. General chemicals were from the following vendors, purchased at the highest quality available, and were used as received: BDH Chemicals, EMD Millipore, Mallinckrodt, and Sigma Aldrich. EMD Millipore also supplied Amicon Stirred cells, Ultracel PL regenerated cellulose membranes (3, 5, 10, and 30 kDa MWCO), Biomax 50 kDa MWCO polyethersulfone membranes, and Amicon Ultra-15 and 0.5 YM regenerated cellulose centrifugal filters (3, 10, 30, and 50 kDa MWCO). Chemically competent BL21 (DE3) *E. coli* cells were from Agilent. Bacto™ Agar was from Beckton Dickinson. Gravity columns, sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) supplies, and AG® 1-X8 resin were from BioRad. The resin was converted to the borate form by washing 500 g sequentially with 4 L NaOH, 7 L water, and 4 L saturated potassium tetraborate solution. GE Healthcare Lifesciences supplied directly the pre-packed MonoQ™ 10/100 GL and Superdex™ 200 10/300 GL columns, and indirectly via Sigma Q-Sepharose, SP-Sepharose, and Sephadex resins. Primers were from either Integrated DNA Technologies or Invitrogen. BL21 (DE3) Codon Plus RIL and RIL-X and BL21 (DE3) pLysS chemically competent *E. coli* cells were from Invitrogen. Luria-Bertani (LB) medium was from MO BIO Laboratories. Restriction enzymes were from either New England Biolabs or Promega. pET14b empty vector was from Novagen. DL-dithiothreitol (DTT), isopropyl β-D-1-thiogalactopyranoside (IPTG), and T4 DNA ligase were from Promega. Ni-charged nitrilotriacetic acid (Ni-NTA) agarose resin was from Qiagen. Calf intestinal phosphatase

and Complete ethylenediaminetetraacetic acid (EDTA)-free protease inhibitor cocktail tablets were from Roche. The vectors pET14b-*nrdI*, pET14b-*nrdF*, and pET14b-*nrdE*, prepared by Dr. Yan Zhang, have been described previously.[9]

2.2.2. General methodology. Sequencing by the MIT Biopolymers Laboratory confirmed the successful cloning of all genes. Optical density measurements at 600 nm (OD_{600}) during cell cultivation were made using a SmartSpec^{TM} 3000 spectrophotometer (BioRad) and disposable semi-micro polystyrene cuvettes (VWR). All gravity column dimensions reported for protein purification procedures are given as height x internal diameter. Solvent flow rates through columns were maintained during protein purifications using a P1 peristaltic pump (GE Healthcare Lifesciences). Fractions eluted from gravity columns during protein purifications were screened for A_{260}, A_{280}, and (in the cases of His_6-$NrdI_{ox}$ and His_6-TrxB) $A_{449\text{-}450}$ using an Epoch microplate spectrophotometer (BioTek) and 175 μL of each fraction loaded into a UV-transparent, 96-well plate (Coring). UV-visible spectra were acquired using an Agilent 8453 diode array or Varian Cary 3 spectrophotometer interfaced with a personal computer and run with ChemStation or WinUV software, respectively. For protein concentration determination, the following ε_{280}, calculated using the ProtParam tool of the ExPASy proteomics server,[31] were used: His_6-NrdE = 79100 M^{-1} cm^{-1}, apo-His_6-NrdF = 54800 M^{-1} cm^{-1}, His_6-TrxA = 12700 M^{-1} cm^{-1}, His_6-YosR = 1500 M^{-1} cm^{-1}, and His_6-YtnI = 11600 M^{-1} cm^{-1}. His_6-NrdI concentrations were determined based on bound flavin using ε_{449} = 12300 M^{-1} cm^{-1}.[27] Unless noted otherwise, concentrations of His_6-NrdE are reported as monomer and His_6-NrdF as dimer. Anaerobic procedures and manipulations were conducted in a custom-built LabMaster 130 glove box (MBraun) under an atmosphere of N_2 at ~4 °C. Protein solutions and buffer used for anaerobic work were rendered anoxic on a Schlenk line using 5 – 6 cycles (protein) or 3 cycles (buffers) of vacuum evacuation followed by backfilling

with Ar gas (ultra-high purity, Airgas) before being introduced into the glove box. Protein solutions were centrifuged in the glove box (14100 x g, 1 min, ~4 °C) prior to use to remove the small amounts of precipitate resulting from the Schlenk line procedure.

2.2.3. Buffers. Similar buffers were used in the purification of all proteins. **Buffer A** consisted of 50 mM sodium phosphate, pH 7.6, 10 mM imidazole, 5% (w/v) glycerol. **Buffer B** was 50 mM sodium phosphate, pH 7.6, 250 mM imidazole, 5% (w/v) glycerol. **Buffer C** was 50 mM Tris(hydroxymethyl)aminomethane (Tris), pH 7.6, 5% (w/v) glycerol. **Buffer D** was 50 mM 4-(2-hydroxyethyl)-1-piperazineethanesuflonic acid (HEPES), pH 7.6, 5% (w/v) glycerol). **Edelhoch buffer** was 6 M guanidinium HCl, 20 mM sodium phosphate, pH 6.5.[32, 33] O_2-staturated Buffer D was prepared immediately before use by sparging with O_2 gas (zero grade, Airgas) either at room temperature (for $Mn(III)_2$-Y• cofactor reconstitution) or on ice (for $Fe(III)_2$-Y• cofactor reconstitution) for at least 45 min. For analytical ultracentrifugation experiments, **Buffer 1** consisted of 50 mM Tris, pH 7.6, 5% (w/v) glycerol, 150 mM NaCl, 15 mM $MgCl_2$, 1 mM DTT and **Buffer 2** consisted of 50 mM sodium phosphate, pH 7.6, 150 mM NaCl, 5% (w/v) glycerol, 1 mM tris(2-carboxyethyl)phosphine (TCEP).

Table 2.1. Sequences of primers used to clone *B. subtilis* redoxins.

Gene	Primer direction	Sequence[a]
trxA	Forward	5′-AGTCAGCATATGATGGCTATCGTAAAAGCAACTGAT-3′
	Reverse	5′-AGTCAGCTCGAGTCAAAGATGTTTGTTTACAAGCTC-3′
trxB	Forward	5′-AGTCAGCATATGATGAGATTAATTAAATTAGAGCAG-3′
	Reverse	5′-AGTCAGCTCGAGTCATCGTAATTCCTTTAGTAACTC-3′
yosR	Forward	5′-AGTCAGCATATGGTGTCAGAAGAAAAAATTTATGAC-3′
	Reverse	5′-AGTCAGCTCGAGTCATTTTAAGGTTTTCAGCGTTTC-3′
ytnI	Forward	5′-GGGACAGACATATGAGTGATGTTGTGAACATTGTGGTTTGG-3′
	Reverse	5′-AATCTGCTCGAGTCATCTTTTTGCCTCCTTTATCTGTGC-3′

[a] Restriction sites are underlined, and start and stop codons are indicated in boldface font.

2.2.4. Cloning of N-terminally His_6-tagged redoxins. Cloning of *trxA*, *trxB*, and *yosR* into pET14b was carried out by Dr. Xuling Zhu. Each gene was amplified from *B. subtilis* JH624 genomic DNA (a gift from Dr. Alan Grossman, Department of Biology, Massachusetts Institute of Technology) using the primers listed in **Table 2.1** to introduce a 5′ *NdeI* and 3′ *XhoI* restriction site into the amplicons. The amplicons were digested with NdeI and XhoI and then ligated into similarly digested pET14b using T4 DNA ligase.

2.2.5. General growth and expression procedure. The growth and expression of all proteins followed a similar general procedure with slight modifications for each protein. Vectors were transformed into the following *E. coli* strains: BL21 (DE3) = pET14b-*ytnI*; (BL21-DE3) Codon Plus RIL = pET14b-*nrdE*, -*nrdF*, and -*nrdI*; BL21 (DE3) Codon Plus RIL-X = pET14b-*trxB*; and BL21 (DE3) pLysS = pET14b-*trxA* and -*yosR*. Transformants were selected for on LB agar plates (1% (w/v)) containing 100 μg mL^{-1} ampicillin (Amp) that were incubated at 37 °C overnight. A single colony was inoculated into 100 mL of LB supplemented with 100 μg mL^{-1} Amp and, when expressing His_6-NrdE, -NrdF, -NrdI, and -TrxB, 50 μg mL^{-1} chloramphenicol (Cm) contained in a 250 mL Erlenmeyer flask and grown to saturation (13 – 15 h) at 37 °C with 200 rpm shaking. For growth scale up, saturated starter culture was diluted 200-fold in four 1 L volumes of antibiotic-supplemented LB contained in 6 L Erlenmeyer flasks. Cultures were incubated at 37 °C with 200 rpm shaking to an OD_{600} of 0.6 – 0.8 (typically 4 h post-inoculation). IPTG was then added to a final concentration of 0.2 mM, and cultures continued to incubate at 37 °C (200 rpm) for 2 h prior to harvesting by centrifugation (3500 x g, 15 min, 4 °C). Cell pellets were collected in pre-weighed 50 mL falcon tubes, frozen (liq. N_2), and stored at -80 °C.

Modifications of the general cultivation procedure were introduced for production of His_6-NrdF, -TrxB, -YtnI, and, occasionally, -NrdI. For His_6-NrdF cultivation, large scale cultures were

2 L of LB in 6 L Erlenmeyer flasks. Cultures were grown at 37 °C (200 rpm) to an OD_{600} of 0.8 – 1.0 at which point the chelator 1,10-phenanthroline (0.1 M in 0.1 M HCl) was added to a final concentration of 100 μM. After 20 min, ITPG was then added to a final concentration of 0.4 mM and the cultures were incubated for 4 h prior to harvest. For His_6-TrxB production, cultures (1 L) were grown in 2.8 L Fernbach flasks and the post-induction incubation period was increased to 3 h. For His_6-YtnI production, the post-induction incubation temperature was lowered to 30 °C. Finally, His_6-NrdI was occasionally grown in large scale cultures of 2 L, induced with 0.4 mM IPTG, and incubated post-induction for 4 h.

Typical yields were in the range of 1.6 – 2.2 g cell paste L^{-1} for all strains except for those expressing His_6-NrdI (4.8 ± 1.1 g L^{-1}) and His_6-TrxB (3.0 g L^{-1}, strain grown once).

2.2.6. Purification of the His_6-tagged RNR proteins, TrxA, TrxB, and YosR. Purification of His_6-tagged NrdE, NrdF, NrdI, TrxA, TrxB, and YosR followed the same general procedure up through the immobilized metal affinity chromatography (IMAC) step with slight modifications for each protein. Per preparation, the starting amounts of cell paste were typically as follows: *His_6-NrdE* – 11.5 g, *His_6-NrdF* – 18 g, *His_6-NrdI* – 22 g, *His_6-TrxA* – 8.5 g, and *His_6-TrxB* – 12 g. All steps of the procedure were carried out at 4 °C. Cell pellets were thawed on ice and resuspended to ~0.2 g cell paste mL^{-1} in Buffer A supplemented with Complete EDTA-free protease inhibitor cocktail tablets (one per 50 mL suspension), 250 μM phenylmethylsulfonyl fluoride (PMSF), and > 1 mg mL^{-1} lysozyme (from chicken egg whites, Sigma). Lysis buffers were additionally supplemented for the following proteins: *His_6-NrdE* – 150 mM NaCl and 10 mM β-mercaptoethanol (BME); *His_6-NrdF* – 100 μM 1,10-phenanthroline; *His_6-TrxA and -YosR* – 300 mM NaCl; and *His_6-TrxB* – 2 mM flavin adenine dinucleotide (FAD), 2 U mL^{-1} DNase I, and 1 U

mL^{-1} RNase A. Suspensions were homogenized using a tissue homogenizer prior to lysing cells by one passage through a French pressure cell at 14000 psi.

After removal of the cell debris by centrifugation (25000 x g, 30 min), streptomycin sulfate solution was added dropwise over 10 min to the stirring cell extract to a final concentration of 1.3% (w/v). The treated extracts were stirred for an additional 15 min before removing precipitated material by centrifugation (25000 x g, 30 min). This step was omitted in the purification of His_6-TrxB as it caused the protein to precipitate.

The clarified extracts were loaded onto Ni-NTA columns containing ~5-6 mL resin per gram cell paste for His_6-NrdE and His_6-NrdF, or ~2 mL resin per gram cell paste for His_6-NrdI and the redoxin proteins. Prior to loading, columns were equilibrated in Buffer A containing additional supplements for the following proteins: *His_6-NrdE* – 150 mM NaCl and 10 mM BME; *His_6-TrxA*, *-TrxB, and -YosR* – 300 mM NaCl. The columns were washed with equilibration buffer until the A_{280} of the eluate decreased below 0.1 (~25 column volumes (CVs)) and then were eluted with 5 CVs of Buffer B containing additional supplements for the following proteins: *His_6-NrdE* – 150 mM NaCl and 10 mM BME; *His_6-TrxA*, *-TrxB*, *and -YosR* – 300 mM NaCl. For the purification of His_6-NrdI, an additional 5 CV wash with 50 mM sodium phosphate, pH 7.6, 30 mM imidazole, 5% (w/v) glycerol was included prior to elution.

His_6-NrdE and His_6-NrdF were further purified by anion exchange chromatography. Protein-containing fractions from the Ni-NTA column were pooled (typically 285 mg His_6-NrdE and 120 mg His_6-NrdF total), diluted 5-fold in Buffer C supplemented with NaCl (*His_6-NrdE* = 100 mM, *His_6-NrdF* = 150 mM) and 10 mM DTT (*His_6-NrdE only*), and loaded onto a Q-Sepharose column: *His_6-NrdE* – 5.0 x 2.5 cm, 25 mL; *His_6-NrdF* – 6.3 x 2.5 cm, 30 mL). The column was washed with 2 CVs of buffer before development with the following linear gradients:

His$_6$-NrdE – 300 mL from 100 – 500 mM NaCl in Buffer C supplemented with 10 mM DTT; *His$_6$-NrdF* – 200 mL from 150 – 450 mM NaCl in Buffer C. Protein-containing fractions were pooled, exchanged into storage buffer (*His$_6$-NrdE*: 50 mM sodium phosphate, pH 7.6, 150 mM NaCl, 5% (w/v) glycerol, 10 mM DTT; *His$_6$-NrdF*: 50 mM HEPES, pH 7.6, 5% (w/v) glycerol), and concentrated (*His$_6$-NrdE* = 50 kDa MWCO filters, *His$_6$-NrdF* = 30 kDa MWCO filters). Typical yields of 25 mg His$_6$-NrdE and 5.4 mg His$_6$-NrdF g^{-1} cell paste (> 90% pure as judged by SDS-PAGE) were obtained.

His$_6$-NrdI was further purified using cation exchange chromatography. Protein-containing fractions from the Ni-NTA column were pooled (typically 25 mg protein total), diluted 5-fold in Buffer D, and loaded onto a SP-Sepharose column (3 x 2.5 cm, 10 mL). The column was washed with 2 CVs of Buffer D and then eluted with 5 CVs of Buffer D supplemented with 200 mM NaCl. Protein-containing fractions were pooled, desalted, and concentrated (10 kDa MWCO filters). Typical yields of 185 μg His$_6$-NrdI$_{ox}$ g^{-1} cell paste (~85% pure as judged by SDS-PAGE) were obtained.

His$_6$-TrxA was further purified using SEC. Protein-containing fractions from the Ni-NTA column were pooled (typically 230 mg protein total), concentrated to 1 – 1.5 mL (3 kDa MWCO filters), and chromatographed on a Sephadex G75 column (27 x 2.5 cm, 130 mL) using 50 mM sodium phosphate, pH 7.6, 150 mM NaCl. Fractions were screened by A_{280} and SDS-PAGE using 17% (w/v) gels, and those containing pure His$_6$-TrxA were pooled and concentrated. Typical yields of 12 mg His$_6$-TrxA g^{-1} cell paste (≥ 99% pure by SDS-PAGE) were obtained.

His$_6$-TrxB and His$_6$-YosR were further purified by anion exchange chromatography using a BioCAD SPRINT FPLC system (PerSeptive Biosystems). Protein (~90 mg) was loaded onto a Poros HQ/20 column (Applied Biosystems, 10 x 1.6 cm, 20 mL) equilibrated in Buffer C

supplemented with 100 mM NaCl. The column was washed with 0.5 CV of equilibration buffer before development with a linear gradient from 100 mM to 700 mM NaCl in Buffer C. His_6-TrxB eluted in the range of 230 – 280 mM NaCl (1 mL min^{-1}) and His_6-YosR in the range of 350 – 450 mM NaCl (2 mL min^{-1}). Fractions containing His_6-TrxB were screened by the 5,5′-dithiobis-(2-nitrobenzoic acid) (DTNB) assay (**Section 2.2.9**) using 1 – 10 μL to initiate turnover. The most active fractions were pooled, exchanged into storage buffer (50 mM sodium phosphate, pH 7.6, 150 mM NaCl), and concentrated (30 kDa MWCO filters). Fractions containing His_6-YosR, as assessed by SDS-PAGE on 15% (w/v) gels, were pooled and concentrated (3 kDa MWCO filter) prior to further purification by Sephadex G75 SEC as described for His_6-TrxA. Typical yields of 0.4 mg His_6-TrxB and 3 mg His_6-YosR g^{-1} cell paste (≥ 95% pure as judged by SDS-PAGE) were obtained. The FAD content of His_6-TrxB was determined as described previously.[34, 35]

2.2.7. Purification of His_6-YtnI. The cell pellet (1.6 g) was thawed and subsequently resuspended in 7.9 mL Buffer A supplemented with 300 mM NaCl, 0.25 mM PMSF, a small amount of lysozyme (> 1 mg), 13 U mL^{-1} DNase I, and 1 U mL^{-1} RNase A. The suspension was placed in an ice water bath while cells were lysed with six cycles of 1 min of sonication (~1 s bursts, ~19 – 20 W per burst) followed by 2 min rest. Cell debris was removed by centrifugation (23224 x g, 25 min) and supernatant (14 mg mL^{-1}) was frozen and stored at -20 °C. Later, the extract was thawed and loaded onto a Ni-NTA column (3.5 x 0.8 cm, 1.8 mL). The column was washed with Buffer A supplemented with 300 mM NaCl until the A_{280} reached 0.04 and was subsequently eluted with 6 CVs of Buffer B supplemented with 300 mM NaCl. Protein-containing fractions were pooled, exchanged into storage buffer (50 mM sodium phosphate, pH 7.6, 150 mM NaCl), and concentrated (5 kDa MWCO filter) to give 25 mg of His_6-YtnI (> 95% pure as judged by SDS-PAGE).

2.2.8. Determination of ε_{455} for FAD bound to His_6-TrxB. The ε_{455} for FAD bound to His_6-TrxB was determined by the method of Aliverti.[34] Briefly, 100 μL samples of 35 μM His_6-TrxB were incubated for 10 min in a boiling water bath prior to being placed on ice to chill for ~1 min. Samples were then centrifuged (21130 x g, 10 min, 4 °C) and the free FAD in the supernatants was quantitated using ε_{450} = 11300 M^{-1} cm^{-1}. The protein pellet was dissolved in 80 μL of Edelhoch buffer and quantitated by UV-vis using ε_{280} = 22330 M^{-1} cm^{-1}, calculated using equation 2.1[32, 33] where n_W, n_Y, and n_{CSSC} are the number of Trp, Tyr, and Cys disulfides, respectively, in the protein. His_6-TrxB contains one W and 13 Y residues, and all C residues were assumed to be reduced.

$$\varepsilon_{280} = (5690\ M^{-1}\ cm^{-1} \times n_W) + (1280\ M^{-1}\ cm^{-1} \times n_Y) + (120\ M^{-1}\ cm^{-1} \times n_{CSSC}) \qquad (2.1)$$

An ε_{455} = 11300 M^{-1} cm^{-1} was calculated using equation 2.2 and was subsequently used to determine His_6-TrxB protein concentrations.

$$\frac{A_{450\,free}}{\varepsilon_{450\,free} \times \ell} = c_{FAD} = \frac{A_{455bound}}{\varepsilon_{455bound} \times \ell} \qquad (2.2)$$

$A_{450free}$ and $\varepsilon_{450free}$ are the absorption and extinction coefficient at 450 nm for free FAD, $A_{455bound}$ and $\varepsilon_{455bound}$ are the absorption and extinction coefficient at 455 nm for His_6-TrxB-bound FAD, c_{FAD} is the concentration of FAD, and ℓ is the path length of the cell (1 cm).

2.2.9. Redoxin activity assays. The activities of the redoxins were assayed using DTNB as a substrate at 25 °C in a total volume of 300 μL as described previously.[36, 37] DTNB cleavage was monitored at 412 nm, where the product 2-nitro-5-sulfidobenzoate maximally absorbs (ε_{412} = 13600 M^{-1} cm^{-1}). DTNB stock solutions were prepared fresh in ethanol prior to running the assays. When determining His_6-TrxA, His_6-YosR, and His_6-YtnI activities, assay solutions consisted of 100 mM sodium phosphate, pH 7.6, 10 mM EDTA, 0.02 mg mL^{-1} bovine serum albumin (BSA), 0.1 mM NADPH, 0.1 mM DTNB, 0.2 μM His_6-TrxB, and 5 – 25 μM His_6-TrxA (His_6-YosR/His_6-

YtnI). When determining His_6-TrxB activity, assay solutions consisted of 50 mM sodium phosphate, pH 7.6, 0.3 mM EDTA, 0.1 mM NADPH, 0.15 mM DTNB, 25 μM His_6-TrxA, and 0.1 μM His_6-TrxB. Reactions were initiated by addition of mixtures of His_6-TrxA (His_6-YosR/His_6-YtnI) and His_6-TrxB. Controls omitting the assayed enzyme were run to account for background levels of DTNB cleavage, which were subtracted from the rates measured in the complete assays. One unit of activity corresponds to the amount of enzyme required to cleave 1 nmol DTNB in 1 min.

2.2.10. Reconstitution of $Fe(III)_2$-Y• His_6-NrdF. Reconstitution of the $Fe(III)_2$-Y• cofactor followed the previously described procedure for *E. coli* His_6-NrdF.[38] Anaerobic solutions of apo-His_6-NrdF were diluted with anaerobic Buffer D to ~200 μM β_2 prior to adding $(NH_4)_2Fe(SO_4)_2$ (~20 mM freshly prepared stock solutions in Buffer D) to give 5 equiv Fe^{2+} per β_2. The $Fe(II)_2$-His_6-NrdF sample was incubated at ~4 °C for 20 min before being brought out of the glove box and placed on ice. Ice cold O_2-saturated Buffer D (~1.9 mM O_2) was added to the sample to give 3.5 equivalents O_2/β_2. Samples were incubated at RT for 8 – 10 min before being subjected to either immediate purification by MonoQ anion exchange chromatography (**Section 2.2.12**) or freezing and storage at -80 °C for future purification.

2.2.11. Reconstitution of $Mn(III)_2$-Y• His_6-NrdF. Reconstitution of the $Mn(III)_2$-Y• cofactor followed the previously published procedure.[9] The hydroquinone form of His_6-NrdI (His_6-$NrdI_{hq}$) was prepared by incubating anaerobic His_6-$NrdI_{ox}$ solutions (~300 – 400 μM) with 1 equiv sodium dithionite ($Na_2S_2O_4$, ~6 mM stock solutions) for 20 min in the glove box. Stock solutions of $Na_2S_2O_4$ were prepared fresh in anaerobic Buffer D and standardized using 1 mM solutions of $K_3Fe(CN)_6$ (ε_{420} = 1020 M^{-1} cm^{-1}). Concomitantly, $MnCl_2$•$4H_2O$ (10 – 20 mM solutions in Buffer D) was added to samples of apo-His_6-NrdF (200 – 400 μM β_2) to give 4 equiv Mn^{2+}, and the

protein was incubated for 20 min. Samples (500 μL) containing 60 – 100 μM $Mn(II)_4$-His_6-β_2 and 120 – 200 μM His_6-$NrdI_{hq}$ (2:1 ratio of NrdI:β_2) in Buffer D were prepared, brought out of the glove box, and warmed to RT. Five hundred microliters of room temperature O_2-saturated Buffer D (~260 μM O_2) was added to the protein samples to give 2.6 – 4.3 equivalents O_2/β_2. Samples were incubated at RT in the dark for 8 – 10 min before being subjected to either immediate purification by MonoQ anion exchange chromatography (**Section 2.2.12**) or freezing and storage at -80 °C for future purification.

2.2.12. Purification of holo-$Me(III)_2$-Y• His_6-NrdF subsequent to cofactor assembly. EDTA (pH 7.6) was added to samples of reconstituted $Me(III)_2$-Y• NrdF (30 – 50 μM Mn-loaded β_2 or ~300 μM Fe-loaded β_2, ~0.6 Y•/β_2) to a final concentration of 5 mM. Samples were mixed, centrifuged (20817 x g, 2 min, 4 °C), and loaded onto a MonoQ 10/100GL anion exchange column (10 x 1 cm, 8 mL) equilibrated with Buffer C supplemented with 100 mM NaCl. The column was washed with 1.3 CVs of equilibration buffer before development with a 100 mL linear gradient from 100 – 350 mM NaCl in Buffer C at a flow rate of 1 mL min^{-1}. Apo/mismetallated NrdF eluted at 250 – 280 mM NaCl and holo-protein at 280 – 310 mM (**Figure 2.3**). Protein-containing fractions from each peak were pooled separately, exchanged into Buffer D, and concentrated. Typical yields were 35 – 45% of $Mn(III)_2$-Y• His_6-β_2 and 53 – 63% of $Fe(III)_2$-Y• His_6-β_2. Protein concentrations were determined using ε_{280} = 119000 M^{-1} cm^{-1} and 109600 M^{-1} cm^{-1} for $Fe(III)_2$-Y•- and $Mn(III)_2$-Y•-loaded protein, respectively.

2.2.13. Metal quantification. Iron bound to $Fe(III)_2$-Y• His_6-NrdF was determined by the ferrozine assay.[39] A set of standards with final concentrations of 10, 20, 30, 40, and 50 μM Fe were prepared from a Fe standard stock solution (Fluka, 1000 ± 4 mg L^{-1}). Samples of $Fe(III)_2$-Y• His_6-NrdF were prepared to final concentrations in the range of 5 – 12 μM β_2. Two hundred

microliters of Fe standard or protein sample was mixed with 100 μL Reagent A (2.25% (w/v) $KMnO_4$ dissolved in 0.6 M HCl) and incubated at 65 °C for 2 h. After cooling to RT, 20 μL of Reagent B (5 M ammonium acetate, 2 M ascorbic acid, 13.1 mM 2,9-dimethyl-1,10-phenanthroline (neocuprione), 6.5 mM ferrozine) was added and mixed in by vortexing. Samples were then incubated in the dark at RT for 30 min, centrifuged (14100 x g, 1 min), and the A_{562} recorded.

Manganese bound to $Mn(III)_2$-Y• His_6-NrdF was quantified with a Perkin-Elmer Analyst 600 atomic absorption spectrometer in the laboratory of Prof. Stephen Lippard using a standard curve (0, 1.25, 2.5, 3.75, and 5 μg L^{-1} Mn) prepared from a Mn standard solution (Fluka, 1000 ± 4 mg L^{-1}) serially diluted to 5 μg L^{-1} in volumetric flasks. Protein samples were serially diluted to appropriate concentrations for analysis in distilled-deionized water. Each analysis was done in triplicate and the results averaged.

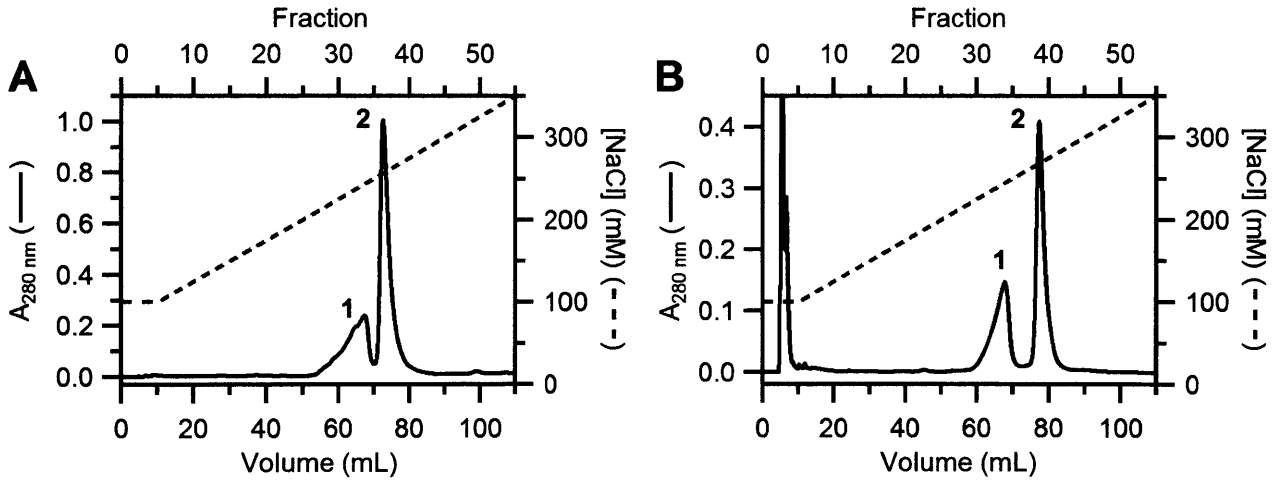


Figure 2.3. Separation of (**A**) holo-$Fe(III)_2$-Y• His_6-NrdF and (**B**) holo-$Mn(III)_2$-Y• His_6-NrdF from apo-/mismetallated protein by anion exchange chromatography on a MonoQ column. Peak 1 corresponds to apo/mismetallated NrdF and peak 2 to holo-$Me(III)_2$-Y• His_6-β_2. The material eluting in the void volume in (**B**) is His_6-$NrdI_{ox}$.

2.2.14. Y• quantitation by EPR spectroscopy. Spectra were acquired at 77 K on a Bruker EMX X band spectrometer using a quartz finger dewar. All spectra were acquired at 9.3 GHz with 100 kHz modulation amplitude. $Mn(III)_2$-Y• cofactor-specific acquisition parameters were as follows: power = 1 mW, modulation amplitude = 1.5 G, gain = 2.52 x 10^4, time constant = 10.24 ms. $Fe(III)_2$-Y• cofactor-specific parameters were as follows: power = 50 μW, modulation amplitude = 1.5 G, gain = 2.52 x 10^4 or 10^3, time constant = 10.24 ms. Spin quantification was carried out by comparing the double integral of the signal intensity to that of $CuSO_4$ standard[40] or a reference sample of *E. coli* NrdB (417 μM Y•, 1.2 Y•/β_2). The Y• content of the NrdB reference was determined by the dropline method[41] and at 77 K by EPR with comparison to a $CuSO_4$ standard.[40]

2.2.15. RNR radioactive activity assays. Routine assays were carried out at 37 °C in a final volume of 200 μL containing: 0.5 to 1 μM $Me(III)_2$-Y• His_6-β_2 (Me = Fe or Mn), 2 equiv of His_6-α, 3 mM ATP, 1 mM [5-^{3}H]-CDP (ViTrax, specific activity ≈ 1100 cpm $nmol^{-1}$), reductant [8 – 20 mM DTT or 40 μM redoxin (His_6-TrxA, His_6-YosR, or His_6-YtnI), 0.4 μM TrxB, and 1 mM NADPH], 50 mM HEPES, pH 7.6, 15 mM $MgCl_2$, and 1 mM EDTA. The reaction was initiated by addition of [5-^{3}H]-CDP. Aliquots (35 μL) were removed over 4 min and quenched by incubation at 100 °C for 2 min. The method of Steeper and Steuart was used to quantitate dCDP.[42] One unit of activity is the amount of enzyme required to catalyze the formation of 1 nmol dCDP min^{-1}.

Concentration dependent effects of ATP or dATP (Sigma, ε_{260} = 15400 M^{-1} cm^{-1})[43] on CDP reduction activity were determined with final concentrations ranging from 50 μM – 4 mM and 0.4 μM – 4 mM, respectively. The concentration dependence of RNR activity using a 1:1 ratio of α_2:β_2 was measured with final subunit concentrations varying from 0.01 – 1 μM. At concentrations below 0.1 μM, BSA (Thermo Fisher Scientific) was included in the assay solution to a final

concentration of 0.2 mg mL^{-1}. These data were fit to the Michaelis-Menten equation (equation 2.3) with Igor Pro (Wavemetrics, Lake Oswego, OR).

$$v = \frac{V_{max} \times [subunit]}{K_m + [subunit]} \tag{2.3}$$

2.2.16. Quaternary structural analysis of 1:1 mixtures of α:β using biophysical methods.

2.2.16.A. Anion exchange chromatography. His_6-NrdE and $Me(III)_2$-Y• His_6-NrdF (Me = Fe or Mn, 1 Y•/β_2, ~3.6 Me^{3+}/β_2, 1450 U/mg [$Mn(III)_2$-Y•] or 125 U/mg [$Fe(III)_2$-Y•]) were mixed in a 1:1 ratio at 5 μM in Buffer D in a total volume of 0.5 – 1 mL. Samples were incubated on ice for 20 min and at room temperature for 10 min, followed by centrifugation (20817 x g, 1 min, 4 °C). All subsequent steps were conducted at 4 °C. Samples were injected onto a MonoQ anion exchange column equilibrated in Buffer C supplemented with 100 mM NaCl. The column was washed with 1 CV of equilibration buffer and then developed with a 160 mL linear gradient from 100 – 500 mM NaCl in Buffer C at a flow rate of 1 mL min^{-1}. Protein-containing fractions were pooled, exchanged into Buffer D, and concentrated. The metal content, radical content, and activity of the isolated complex was measured as described in **Sections 2.2.13 – 2.2.15**. The ratio of α:β was estimated by SDS-PAGE densitometry as described in **Section 2.2.17**.

2.2.16.B. Size exclusion chromatography. Samples of His_6-α_2 (2.7 and 27 μM), holo-$Fe(III)_2$-Y• His_6-β_2 (23 μM), holo-$Mn(III)_2$-Y• His_6-β_2 (18 μM), and 1:1 mixtures of α_2:β_2 (1.0, 4.1, 8.1, and 23 μM) in a total volume of 200 μL were centrifuged (20817 x g, 10 min, 4 °C) and then injected onto a Superdex 200 10/300 GL column (10 x 300 mm, ~24 mL bed volume) connected to an ÄKTA Purifier FPLC system (GE Healthcare). Protein was eluted at 4 °C at a flow rate of 0.25 mL min^{-1} using 50 mM HEPES, pH 7.6, 100 mM NaCl, 15 mM $MgCl_2$, 1 mM EDTA. Molecular weights were calculated as previously described[44] using the experimentally measured $s_{20,w}$ (**Section 2.2.16.C**), the Stokes radius (R_s) estimated from the SEC retention times,

and the correlation function of Laurent and Killander.[45] The column was calibrated using a High Molecular Weight Gel Filtration kit (GE Healthcare) containing the following standards: Blue Dextran 2000 (void volume determination), thyroglobulin (669 kDa, R_s = 85.0 Å), ferritin (440 kDa, R_s = 61.0 Å), aldolase (158 kDa, R_s = 48.1 Å), conalbumin (75 kDa, R_s = 36.4 Å[46]), and ovalbumin (44 kDa, R_s = 30.5 Å).

2.2.16.C. Sedimentation velocity analytical ultracentrifugation (SV-AUC). Concentration-dependent SV-AUC experiments were performed using a Beckman XL-I analytical ultracentrifuge at the MIT Biophysical Instrumentation Facility with His_6-Fe-β_2 (1.2 Y•/β_2, 4.0 Fe/β_2, 130 U/mg), His_6-Mn-β_2 (0.9 Y•/β_2, 4.0 Mn/β_2, 1545 U/mg), and His_6-α. Individual subunits were analyzed in Buffer 1 whereas 1:1 mixtures of α and β were analyzed in Buffer 2 to allow experiments to be monitored by A_{230}. Protein was exchanged into buffer using three cycles of concentration-dilution with an Amicon Ultra-15 YM30 centrifugal filter. The final filtrate served as reference buffer for the SV-AUC experiments. Samples spanning 0.1 – 1.5 AU at A_{230}, A_{250}, or A_{280} were prepared in a total volume of 500 µL by mixing concentrated protein solution with reference buffer to the desired final concentration for each experiment (**Table 2.2**). Cells, assembled with Epon charcoal double sector centerpieces and quartz windows, were loaded with ~440 µL reference solution and ~430 µL sample solution, radially calibrated, and thermally equilibrated at 20 °C for 1.5 – 2.0 h prior to initiating the experiment. Sedimentation was measured over 18 – 19 h with absorption detection at 20 °C and an angular velocity of 35000 rpm. Scans were collected every 1.2 min using ProteomeLab XL-I Graphical User Interface, version 4.5b (Beckman).

Individual data sets were loaded into the program Sedfit, time-stamp corrected,[47] and fit to the c(s) model using the Simplex optimization algorithm with a resolution = 200 – 250 and a

regularization factor = 0.95.[48, 49] Distributions were converted to standard state ($s_{20,w}$) using a solvent density (ρ) = 1.018 g mL^{-1} (Buffer 1) or 1.023 g mL^{-1} (Buffer 2), solvent viscosity (η) = 1.164 cP (Buffer 1) or 1.170 cP (Buffer 2), and the partial specific volume ($\bar{v}$) of the species analyzed (**Table 2.3**). These solvent and protein hydrodynamic parameters were calculated and temperature corrected using the program Sednterp.[50] Weight-averaged $s_{20,w}$ values were determined by integrating the distributions over the $s_{20,w}$ range containing all peaks, and error in this value was assessed using 500 rounds of Monte-Carlo simulations. All $Me(III)_2$-Y• His_6-NrdF data sets and the His_6-NrdE and His_6-NrdE:$Fe(III)_2$-Y• His_6-NrdF data collected at physiological concentrations (~1 μM) were imported into and fit with the program Sedphat using the "hybrid local continuous/global discrete species model" for a more rigorous determination of $s_{20,w}$ and M_W.[48] The best fit $s_{20,w}$ for each species was used to determine R_s and the frictional ratio (f/f_o) with Sednterp by turning off the temperature corrections, setting ρ = 0.998 g mL^{-1} and η = 1.002 cP, and calculating the molecular weights using the amino acid sequences of His_6-NrdE and His_6-NrdF.

Table 2.2. SV-AUC sample concentration ranges and monitoring wavelengths.

Protein	λ (nm)	Sample concentrations (μM)
His_6-NrdE	280	1.0, 2.4, 4.3, 6.7, 8.9, 11.4
	250	12.3, 14.7, 16.6, 18.9, 26.7, 32.3
holo-$Fe(III)_2$-Y• His_6-β_2	280	0.9, 1.9, 4.1, 6.8, 10.2, 11.7
holo-$Mn(III)_2$-Y• His_6-β_2	280	0.8, 1.6, 3.4, 5.2, 8.1, 10.9
1:1 His_6-α_2 + holo-$Fe(III)_2$-Y• His_6-β_2	230	0.025, 0.04, 0.06, 0.23, 0.56, 0.67
	280	0.48, 1.2, 2.0, 2.7, 3.6
	250	4.8, 5.3, 6.2, 6.8

Table 2.3. Partial specific volumes and molecular weights of His-tagged RNR components.

Subunit/complex	His_6-α	His_6-β	His_6-α_2	His_6-β_2	His_6-$\alpha_2\beta_2$
M_W (kDa)	83	41	166	81	247
$\bar{v}$ (cm^3 g^{-1})	0.7315	0.7352	0.7315	0.7352	0.7327

2.2.16.D. Hydrodynamic modeling of the His$_6$-tagged *B. subtilis* RNR subunits. Hydrodynamic parameters for each subunit and the $\alpha_2\beta_2$ complex were predicted using the program HYDROPRO[51] and structural models of NrdF (PDB 4DRO[52]) and NrdE. As a structure of *B. subtilis* NrdE was not available, the web server PHYRE[53] was used to build a homology model based on the structure of *Salmonella enterica* Serovar Typhimurium (*S.* typhimurium) NrdE (PDB 1PEM[54]). A pairwise BLAST sequence alignment of *B. subtilis* NrdE (accession # NP_389620) and *S.* Typhimurium NrdE (accession # AAL21692) exhibited 45% identity (65% with positive substitutions) between the two proteins after introduction of 10 gaps (1% of total sequence). Two models of the *B. subtilis* $\alpha_2\beta_2$ complex were prepared and analyzed using UCSF Chimera.[55] In one model, two NrdE monomers and the $Mn(II)_2$-NrdF dimer were aligned with the symmetric docking model of the *E. coli* class Ia RNR (RMSD = 1.1 Å over 400 atom pairs).[56] In the second model, they were aligned with the asymmetric structure of the class Ib RNR holo-complex from *S.* Typhimurium (PDB 2BQ1, RMSD = 1.2 Å over 474 atom pairs).[57]

Predictions of $s_{20,w}$ for α, β, α_2, β_2, and the $\alpha_2\beta_2$ complex were calculated from atomic-level shell models using the program HYDROPRO under default settings using ρ = 0.998 g mL^{-1}, η = 0.01 poise, and the M_W and $\bar{v}$ values listed in **Table 2.3**. Predictions of R_s were calculated using equation 2.4, where N_A is Avogadro's number.

$$R_s = \frac{M_W(1-\bar{v}\rho)}{6\pi\eta N_A s_{20,w}} \tag{2.4}$$

The maximum expected sedimentation coefficient (s_{max}) is the value expected for a perfect sphere with M_W equal to that of the protein of interest. The s_{max} values for the subunits and complex at 20 °C in water were calculated using equation 2.4 rearranged for $s_{20,w}$ and using the Stokes radius for a perfect sphere (R_o) as defined in equation 2.5.[58]

$$R_o = \sqrt[3]{\frac{3M_W}{4\pi N_A}} \tag{2.5}$$

The predicted frictional ratios (f/f_o) were calculated from R_s and R_o for each component using equation 2.6.

$$R_x = \frac{f_x}{6\pi\eta} \tag{2.6}$$

2.2.17. SDS-PAGE densitometry. Samples of 1:1 mixtures of α and β (~4 μM, 10 – 20 μL) were diluted with an equivalent volume of 2X Laemmli buffer supplemented with 3% (v/v) BME and incubated at ~100 °C for 5 min. The samples were centrifuged (14100 x g, 1 min, RT) and 10 – 20 μL were loaded per well on an 18 well, 10% (w/v) Tris-HCl Criterion denaturing gel (BioRad). NrdE and NrdF standards (10, 20, 30, 40, and 50 pmol in 25 μL) were loaded onto the same gel. Gels were run at 4 °C and 180 V, stained with Coomassie Brilliant Blue R-250, and imaged on a Gel Doc™ 2000 Imager (BioRad). Band intensities were quantified using Quality One software (BioRad), and the relative amounts of NrdE and NrdF in samples of complex were determined using the corresponding linear fits generated from the respective protein standard curves.

2.3. RESULTS

2.3.1. Characterization of as-isolated His_6-tagged *B. subtilis* RNR proteins. The following results and observations regarding the purified *B. subtilis* RNR proteins were not described in the previous studies[9, 27, 52] of this enzyme and warrant brief mention here. Optimization of the protocol for purifying His_6-NrdE[9] by inclusion of at least 150 mM NaCl in all buffers for protein stabilization and an additional anion exchange chromatographic step after Ni-NTA IMAC resulted in high yields of ≥ 95% homogenous protein (**Figure 2.4A**). As observed previously[9], the molecular weight of His_6-NrdE appears smaller than expected based on sequence by SDS-PAGE (~70 – 74 kDa), but the following observations demonstrate that full-length protein is present: (1)

the entire *nrdE* gene is accounted for by repeated sequencing of pET14b-*nrdE*, (2) the ability to reduce CDP in the presence of His_6-TrxA (or -YosR) and His_6-TrxB, as described later, indicates that the C-terminal Cys residues of NrdE are intact,[59] and (3) the ability of the protein to be retained on a Ni-NTA column indicates the N-terminus is intact. Recombinant His_6-tagged *B. subtilis* NrdF and NrdI were purified to > 90% and > 85% homogeneity, respectively, using optimized protocols reported previously (**Figure 2.4B – C**).[9, 27] Yields of His_6-NrdF were good and contained < 0.1 Fe atoms per β_2 and no radical observable by UV-vis and EPR spectroscopic methods. His_6-NrdI was isolated fully loaded with one FMN per peptide, but the total protein yields were very low (usually < 1 mg from 4 L of cells), thus providing the impetus to explore the SUMO-Pro expression system described in Chapter 3.

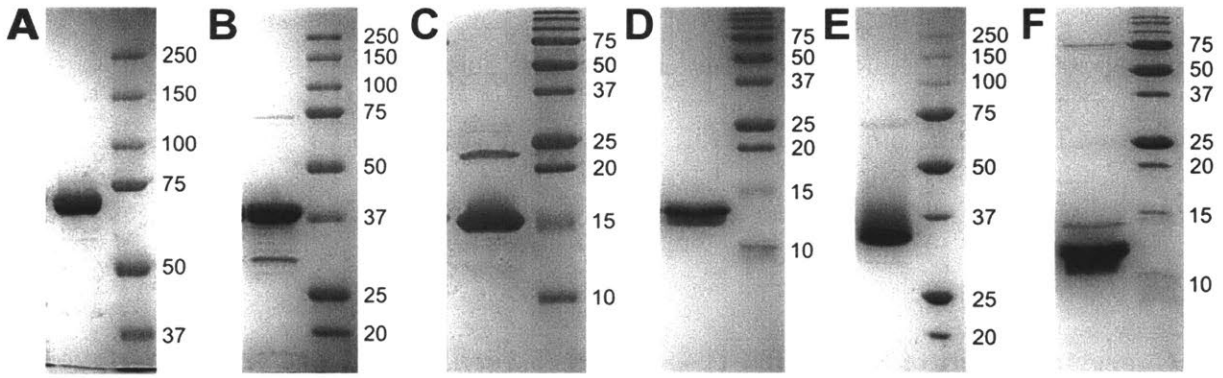


Figure 2.4. SDS-PAGE analysis of representative examples of purified *B. subtilis* (**A**) His_6-NrdE [83 kDa, 7%], (**B**) His_6-NrdF [41 kDa, 10%], (**C**) His_6-NrdI [16.9 kDa, 17%], (**D**) His_6-TrxA [13.7 kDa, 17%], (**E**) His_6-TrxB [36.8 kDa, 10%], and (**F**) His_6-YtnI [12.8 kDa, 17%]. Expected M_W based on sequence and percent acrylamide (w/v) gel used in the analysis is indicated in brackets for each protein. The double bands observed for His_6-TrxA and His_6-YtnI are artifacts due to their low M_W.

A thrombin-cleavable 20-residue His_6-tag with the sequence $MGSSH_6SSG\underline{LVPRGS}H$ (thrombin site underlined) was fused to the N-terminus of each protein. The tag was not removed prior to use of the protein. In addition to the His-tag, an extra Leu (His_6-NrdE) or Val (His_6-NrdF

and His_6-NrdI) was inadvertently inserted between the natural N-terminal Met and second residue of each protein during cloning.[9] Neither of these additions had any effect on activity or structure when compared with results obtained with the tagless protein constructs reported in Chapter 3.

2.3.2. Holo-Me(III)$_2$-Y• His_6-NrdF (Me = Fe or Mn) can be purified by anion exchange on a MonoQ column. Our initial results with endogenous and His_6-tagged recombinant NrdF reconstituted *in vitro* with Mn(III)$_2$-Y• revealed the protein was loaded with substoichiometirc amounts of Mn and Y• (endogenous: 0.5 Y•/β_2, 2.1 Mn/β_2; recombinant: 0.6 Y•/β_2, 1.9 Mn/β_2),[9] indicating the samples were heterogenous mixtures of holo- and apo-/mismetallated protein. Based on the report of partial separation of *B. anthracis* holo- and apo-NrdF,[24] reconstituted samples of *B. subtilis* His_6-NrdF were examined on different anion exchange resins for the ability to resolve apo- from holo-protein. A MonoQ column ultimately effected nearly complete separation of holo-His_6-Feβ_2 and His_6-Mnβ_2 from apo-/mismetallated protein (**Figure 2.3**). Electronic spectra of the pooled protein from peak 2 (**Figure 2.5A** and **C**) exhibited features consistent with the presence of Y• (λ = 390 nm (shoulder), 410 nm).[60] In the case of holo-Mn(III)$_2$-Y• His_6-NrdF, the electronic spectrum also exhibited an absorption feature from 450 – 700 nm that was consistent with Mn(III)$_2$ as reported previously (**Figure 2.5A**),[61] and quantitation of the metal and radical content by atomic absorption and EPR spectroscopy (**Figure 2.5B**) revealed 1.0 ± 0.1 Y•/β_2 and 4.1 ± 0.3 Mn/β_2, respectively. Similar analyses (**Figure 2.5C** and **D**) of holo-Fe(III)$_2$-Y• His_6-NrdF revealed absorption features consistent with Fe(III)$_2$ (λ = 325 nm, 370 nm, 500 nm, and 600 nm),[62] a radical content of 1.1 ± 0.1 Y•/β_2, and a metal content of 4.0 ± 0.1 Fe/β_2 by the ferrozine assay. In contrast, pooled protein from peak 1 (**Figure 2.3**) lacked characteristic electronic absorption features of the dinuclear clusters and Y•, and typically had 0.1 Y•/β_2, ~1.0

metal ion per β_2, and no detectable activity. Holo-His_6-NrdF from peak 2 was used in all subsequent experiments.

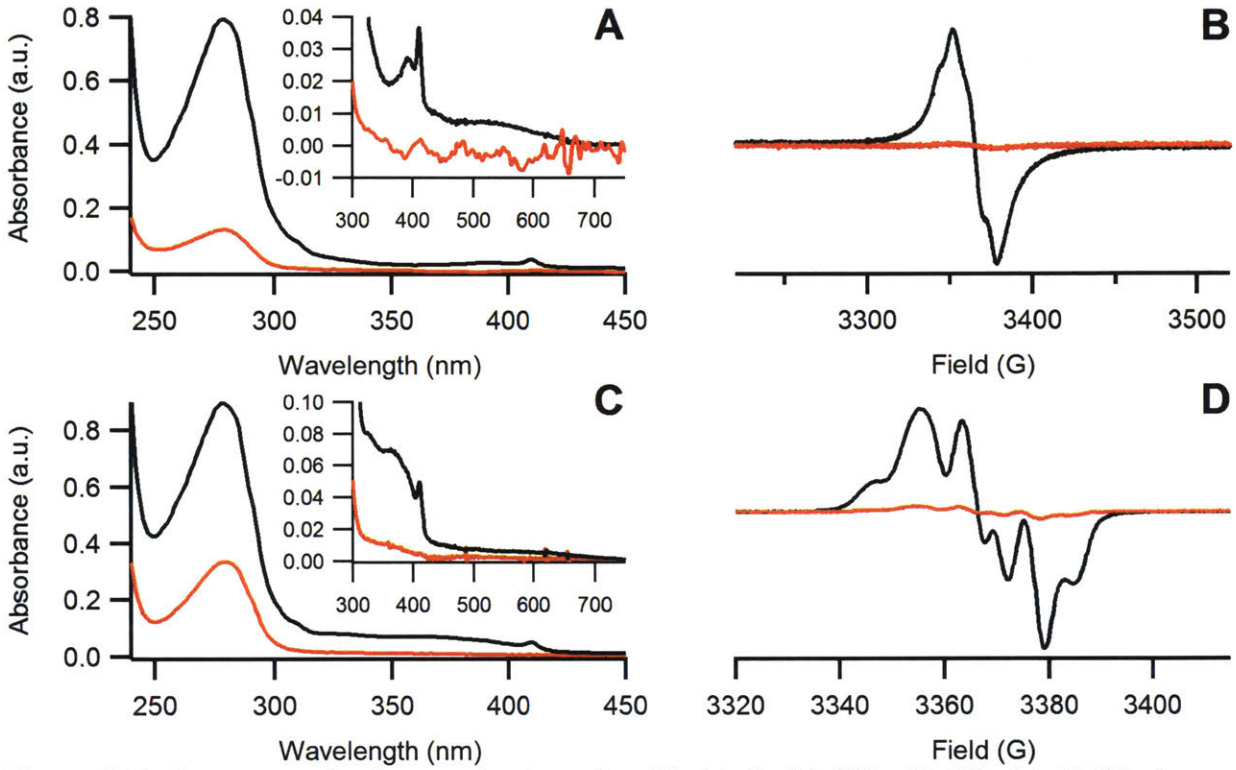


Figure 2.5. Spectroscopic characterization of purified holo-$Me(III)_2$-Y• His_6-NrdF (black traces) and apo-/mismetallated protein (red traces). (**A**) UV-visible (*inset*: blow up of 300 – 750 nm) and (**B**) 77 K EPR spectra of purified protein from $Mn(III)_2$-Y• reconstitution. (**C**) UV-visible (*inset*: blow up of 300 – 750 nm) and (**D**) 77 K EPR spectra of purified protein from $Fe(III)_2$-Y• reconstitution.

2.3.3. Purification and characterization of the *B. subtilis* redoxins. Thioredoxin (*trxA)* and thioredoxin reductase (*trxB)* were cloned, over-expressed in *E. coli*, and purified by Ni-affinity chromatography. The purity of the proteins was not satisfactory after IMAC, so each protein was subjected to additional purification steps. For His_6-TrxA, SEC on a Sephadex G-75 column yielded ≥ 99% pure protein (**Figure 2.4D**) with a specific activity of 210 ± 50 nmol DTNB min^{-1} mg^{-1} in the presence of NADPH and TrxB. Nearly homogenous His_6-TrxB (≥ 95% pure, **Figure 2.4E**) with one equivalent of FAD bound per peptide was obtained with anion exchange

chromatography using a Poros HQ/20 semi-preparative FPLC column and had a specific activity of 17.6 ± 2.7 μmol DTNB min^{-1} mg^{-1} in the presence of TrxA. A full complement of FAD bound to His_6-TrxB is associated with an A_{280}:A_{455} ratio of ~5.5, similar to that reported for fully loaded *E. coli* TrxB (A_{271}:A_{456} = 5.8).[63] In addition to the N-terminal His_6-tag, both proteins were cloned by Dr. Zhu with an extra residue (M for His_6-TrxA and V for His_6-TrxB) encoded after the natural start codon. Both additions appeared to have no detrimental effect on the activity of either protein.

Two potential NrdH-like genes, *yosR* and *ytnI*, that could serve as electron donors for RNR were annotated in the genome.[64] These genes were cloned into pET14b, expressed in *E. coli*, and the resulting protein purified using standard IMAC procedures. His_6-YtnI was ≥ 95% pure after this procedure (**Figure 2.4F**), whereas His_6-YosR was purified by Dr. Zhu to ≥ 95% homogeneity using Poros HQ/20 FPLC anion exchange chromatography followed by Sephadex G-75 SEC. His_6-YosR was competent to reduction by His_6-TrxB and had a specific activity of 440 – 580 nmol DTNB min^{-1} mg^{-1}. Like His_6-TrxA and -TrxB, an extra Met was inserted between the tag and native protein, but did not affect the activity of the protein. In contrast, incubation of His_6-YtnI with His_6-TrxB, NADPH, and DTNB showed an initial burst of 2-nitro-5-sulfidobenzoate formation, indicating the thiols on His_6-YtnI were mostly in the reduced state prior to addition of DTNB. Following the burst, however, background levels of substrate cleavage were observed, indicating that His_6-YtnI was inefficiently reduced by His_6-TrxB and, therefore, was not a NrdH that supported RNR turnover.

2.3.4. The endogenous reducing system for *B. subtilis* RNR is TrxA, TrxB, and NADPH. Radioactivity-based assays modeled after those reported for the *E. coli* class Ia RNR[26] revealed both His_6-TrxA and His_6-YosR were capable of reducing RNR during turnover, whereas His_6-YtnI yielded no detectable CDP reduction activity, as expected from the results of the DTNB assays.

Titrations of RNR (0.7 μM His_6-$Mn\beta_2$ (0.2 Y•/β_2), 7 μM His_6-α_2) with His_6-TrxA (1 – 80 μM) or YosR (0.5 – 40 μM) and His_6-TrxB held at 0.25 μM revealed that the effects of either redoxin were saturable with apparent K_m values of ~1 μM, indicating both could interact similarly with His_6-NrdE (Xuling Zhu, unpublished). Under optimal conditions (**Table 2.4**), holo-$Mn(III)_2$-Y• and $Fe(III)_2$-Y• His_6-NrdF gave, respectively, 1480 ± 130 and 125 ± 20 U/mg with His_6-TrxA and 1000 ± 10 U/mg and 110 ± 10 U/mg with His_6-YosR. Therefore, both proteins can serve as the possible endogenous reductants for the *B. subtilis* RNR. However, genetic[10, 65] and bioengineering[12] evidence indicate *yosR* expression is low and that the gene is dispensable for cell viability, thus leading to the conclusion that TrxA is the physiologically relevant reductant.

Table 2.4. Optimized CDP reduction activity of *B. subtilis* His_6-NrdF.

Metal	M/H[a]	Reductant[b]	Effector	Y•/β_2	Me^{3+}/β_2	SA (U mg^{-1})[c]	Reference
Fe	M	DTT	1 mM ATP	0.18	0.9	22	[9]
	M	DTT	0.3 mM dATP	0.18	0.9	15	[9]
	M	TR	3 mM ATP	0.5	2.8	90 ± 5	This work
	H	DTT	3 mM ATP	1.1	3.6	50 ± 3	This work
	H	TR	3 mM ATP	1.1	3.6	125 ± 20	This work
	H	YosR	3 mM ATP	1.1	3.6	110 ± 10	This work
	H	TR	4 μM dATP	1.1	3.6	170 ± 20	This work
Mn	M	DTT	1 mM ATP	0.44	2.1	160	[9]
	M	DTT	0.3 mM dATP	0.44	2.1	52	[9]
	M	TR	3 mM ATP	0.63	2.6	720 ± 170	This work
	H	DTT	3 mM ATP	1.0	3.2	150 ± 15	This work
	H	TR	3 mM ATP	1.0	3.2	1480 ± 130	This work
	H	YosR	3 mM ATP	1.0	3.2	1000 ± 10	This work
	H	TR	7 μM dATP	1.0	3.2	1260 ± 130	This work
	H	YtnI	3 mM ATP	0.94	nd[d]	None detected	This work

[a] M = mixture of holo- and apo-protein. H = MonoQ-purified holo-protein.
[b] DTT final concentration = 20 mM. TrxA/YosR/YtnI final concentration = 40 μM with 0.4 μM TRR and 1 mM NADPH.
[c] One unit of enzyme catalyzes the reduction of 1 nmol CDP min^{-1}.
[d] Not determined.

2.3.5. Activity assay optimization. Historically, the activity of each subunit for class Ia RNRs has been assayed independently in the presence of an excess of the second subunit due to weak subunit interactions.[26, 66] Similar assays of Ib RNRs have been carried out using this protocol.[3, 4, 8, 9, 24, 61, 67-69] Our previous results, however, suggested the *B. subtilis* RNR could be assayed as a holo-enzyme by using equivalent amounts of each subunit.[9] To validate the use of a 1:1 ratio of subunits, assays were initially carried out using 0.5 μM $Me(III)_2$-Y• His_6-β_2, increasing amounts of His_6-α_2 (1, 2.5, 5, or 10 equiv.), 3 mM ATP, and the endogenous reductants His_6-TrxA/His_6-TrxB/NADPH. The results shown in **Table 2.5** revealed both Fe- and Mn-loaded His_6-NrdF in a 1:1 ratio with His_6-NrdE had ~80% of the maximum activity observed at higher α:β ratios. Therefore, subsequent assay optimization was continued using a 1:1 ratio of the subunits.

Table 2.5. *B. subtilis* RNR activity as a function of the α:β ratio.[a]

α:β ratio	$Fe(III)_2$-Y• His_6-NrdF		$Mn(III)_2$-Y• His_6-NrdF	
	SA (U mg^{-1})[b]	% Maximum	SA (U mg^{-1})[b]	% Maximum
1	140 ± 10	80.4	1630 ± 20	82.4
2.5	170 ± 6	97.9	1980 ± 20	100
5	173 ± 15	100	1890 ± 40	95.3
10	171 ± 17	99.0	1900 ± 20	96.1

[a] Activity was measured under optimal assay conditions.
[b] One unit is the amount of enzyme required to catalyze the reduction of 1 nmol CDP min^{-1}.

Reduction of NDPs by RNRs shows an absolute requirement for Mg^{2+}.[70] The chelator EDTA is also routinely included in assay buffers to prevent non-specific oxidation of α cysteine residues by redox-active metal ions. The dependence of Fe- and Mn-loaded *B. subtilis* Ib RNR CDP reduction activity on Mg^{2+} and EDTA concentration was examined in the presence of 3 mM ATP and the endogenous reducing system. Magnesium ion concentrations in the range of 10 – 15 mM maximized activity with 0.5 μM Mn-loaded enzyme whereas 20 – 40 mM maximized activity with 1 μM Fe-loaded RNR (**Figure 2.6A**). Concentrations of EDTA up to 3 mM had no effect on

the activity (data not shown). In all subsequent assays, 15 mM Mg^{2+} (in the form of $MgCl_2$) and 1 mM EDTA were included in the assay buffer.

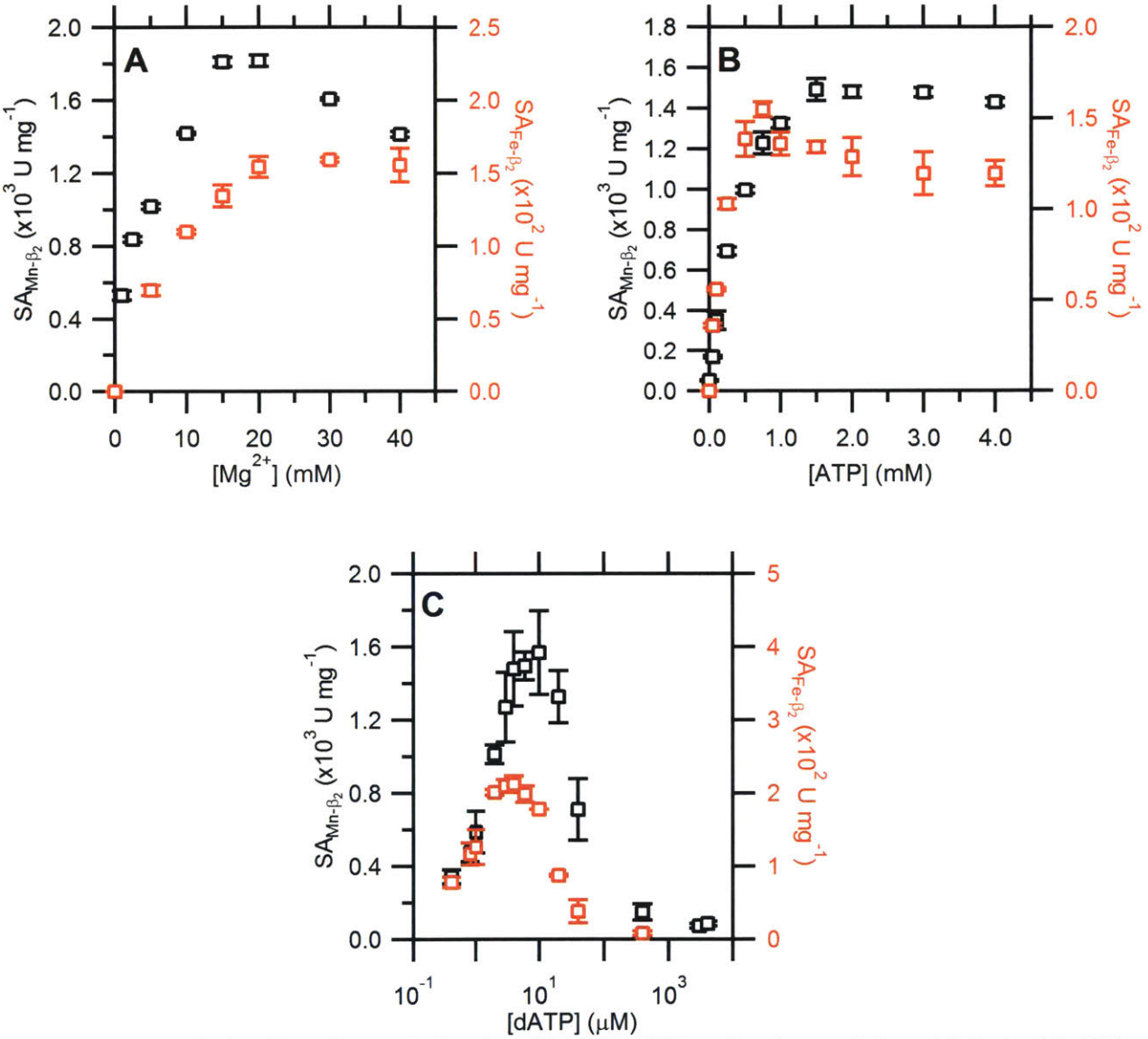


Figure 2.6. Optimization of *B. subtilis* class Ib RNR CDP reduction activity with holo-$Mn(III)_2$-Y• His_6-NrdF (black squares) and holo-$Fe(III)_2$-Y• His_6-NrdF (red squares). In all experiments, [5-^{3}H]-CDP (~1000 cpm $nmol^{-1}$) was held at 1 mM. (**A**) CDP reduction activity dependence on Mg^{2+}. (**B**) CDP reduction dependence on ATP (15 mM Mg^{2+}). (**C**) CDP reduction dependence on dATP (15 mM Mg^{2+}). Assays were run in triplicate and the results averaged. In several cases, the error bars (± 1 standard deviation) are smaller than the marker size.

Previous studies of the Fe-loaded *S.* Typhimurium, *Lactococcus lactis*, and *Mycobacterium tuberculosis* and Mn-loaded *Streptococcus sanguinis* class Ib RNRs reported that dATP strongly stimulated CDP reduction while ATP had only a marginal stimulatory affect.[3, 5, 68, 69, 71] To validate this observation with the *B. subtilis* Ib RNR, we examined the dependence of CDP reduction on both ATP and dATP concentration. The data shown in **Figure 2.6** revealed an altogether different and unexpected result. ATP showed saturation kinetics (K_m = 100 ± 40 μM using holo-Fe(III)$_2$-Y• His_6-NrdF and 320 ± 50 μM using holo-Mn(III)$_2$-Y• His_6-NrdF) with maximum activity observed at 1.5 – 2 mM with CDP held at 1 mM (**Figure 2.6B**). In contrast, dATP showed maximal stimulation at 8 – 10 μM followed by potent inhibition at higher concentrations (**Figure 2.6C**). This behavior is reminiscent of *E. coli* and other class Ia RNRs, but distinct from the previously characterized class Ib RNRs. Inhibition of the *B. subtilis* class Ib RNR represents a new allosteric regulatory mechanism and warrants further investigation.

The specific activity of *B. subtilis* RNR under optimized conditions with CDP/(d)ATP is among the highest reported for any member of the Ib subclass (Appendix 2, **Table A2.1**). Given that cluster loading and reductant identity are essential for RNR activity, the activities of purified and unpurified Me(III)$_2$-Y• NrdF using both DTT and His_6-TrxA/His_6-TrxB/NADPH were compared to establish the relative contributions of each factor to the overall activity enhancement observed with *B. subtilis* RNR. Surprisingly, the use of the endogenous reducing system had a much greater effect on activity (5 – 10-fold enhancement with Mn(III)$_2$-Y•) than did cluster assembly (0 – 2-fold enhancement, **Table 2.4**). This starkly contrasts with the *B. anthracis* class Ib RNR, where similar activities were observed with DTT and Trx1.[24]

2.3.6. Reconstitution of holo-RNR from His_6-NrdE and reconstituted His_6-NrdFs. The most distinct aspect of the co-purification of endogenous NrdE and NrdF in the preliminary study of the

B. subtilis Ib RNR was that the complex was maintained in the absence of substrates, effectors, and Mg^{2+} (**Figure 2.2**).[9] This observation was recapitulated using recombinant His_6-NrdE and holo-$Fe(III)_2$-Y• His_6-NrdF in a 1:1 mixture (5 μM of each subunit) on a MonoQ FPLC column (**Figure 2.7A**). A single peak eluted between 380 – 410 mM NaCl, behavior distinct from His_6-NrdE alone (two peaks that eluted between 265 – 300 mM NaCl and 325 – 340 mM NaCl, Appendix 2, **Figure A2.4**) and reconstituted His_6-NrdF (two peaks that eluted in the range of 250 – 280 mM NaCl and 280 – 310 mM NaCl (**Figure 2.3**)) using the same NaCl gradient. SDS-PAGE densitometry analysis of the fractions showed an approximately 1:1 subunit ratio (**Figure 2.7B**). Both the recovered Fe- and Mn-loaded complexes were active in CDP reduction (120 ± 5 U/mg and 890 ± 6 U/mg, respectively) and contained a nearly full complement of metal (Fe/β_2 = 3.8 ± 0.2; Mn/β_2 = 3.8 ± 1.6) and Y• ($Fe\text{-}\beta_2$ = 1.2 Y•/β_2, $Mn\text{-}\beta_2$ = 0.7 Y•/β_2). The lower activities measured likely stem from several factors, including radical loss and His_6-NrdE inactivation (the chromatography buffer did not contain DTT or TCEP).

2.3.7. Estimation of α/β interaction strength. The affinity of His_6-NrdE and His_6-NrdF for one another under turnover conditions was assessed using with a 1:1 subunit ratio at concentrations ranging from 0.01 μM to 1 μM in the presence of ATP (3 mM), [5-^{3}H]-CDP (1 mM), and optimized concentrations of His_6-TrxA/His_6-TrxB/NADPH. The results are shown in **Figure 2.8**. Fitting the data to equation 2.3 gave V_{max} = 120 ± 5 U/mg and K_m = 0.06 ± 0.01 μM for assays using holo-$Fe(III)_2$-Y• His_6-NrdF and V_{max} = 1080 ± 40 U/mg and K_m = 0.03 ± 0.003 μM for assays using holo-$Mn(III)_2$-Y• His_6-NrdF. Though not a true K_d, the apparent K_ms for the affinity between the subunits is still ~ 10 x greater than the K_d reported for the *E. coli* class Ia RNR in steady-state turnover (0.2 μM) and the *B. anthracis* class Ib RNR (0.2 – 0.5 μM).[13, 23, 24] Recent quantitative western blot results of *B. subtilis* JH624[9] and 2D gel electrophoresis and targeted mass

spectrometric results from Dörte Becher's group[28-30] revealed that the two proteins are present in crude cell extracts in approximately a 1:1 subunit ratio in the concentration range of 0.1 – 1 μM (see Appendix 2, **Tables A2.2 – A2.5**). The apparent affinity reported here suggests that NrdE and NrdF will be mostly complexed *in vivo*.

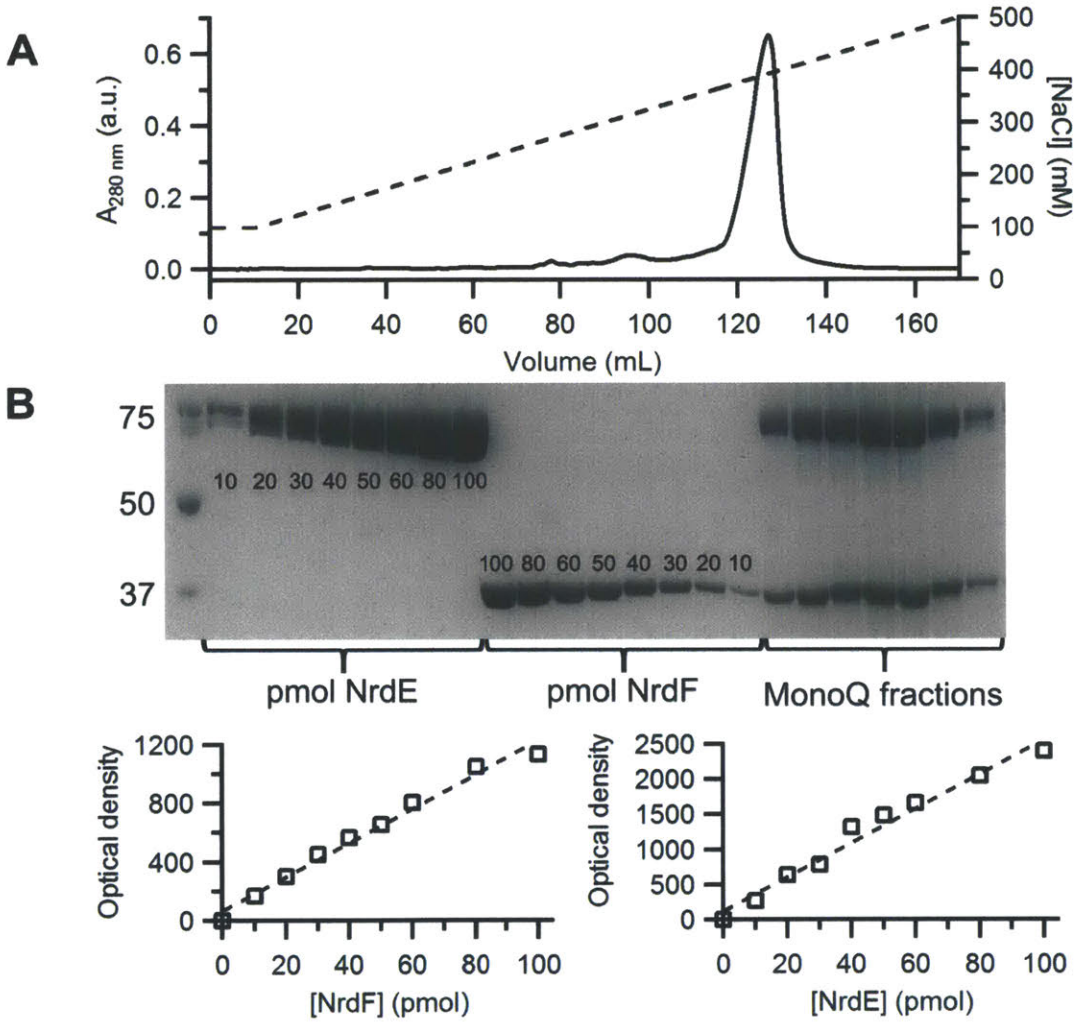


Figure 2.7. Reconstitution of *B. subtilis* holo-RNR. (**A**) MonoQ anion exchange chromatography of a 1:1 mixture of holo-Fe(III)$_2$-Y• His_6-NrdF:His_6-NrdE. (**B**) (*top panel*) SDS-PAGE densitometry analysis of His_6-NrdE standards (left), His_6-NrdF standards (middle), and of fractions from the anion exchange purification of a 1:1 mixture of holo-Fe(III)$_2$-Y• His_6-NrdF:His_6-NrdE (right). The number of pmol of standard loaded per lane is indicated below (His_6-NrdE) or above (His_6-NrdF) the bands. Linear fits of the NrdE and NrdF standard curves (*bottom panel*: left graph = NrdF, right graph = NrdE) yielded acceptable correlation coefficients (r^2 = 0.97 and 0.98, respectively).

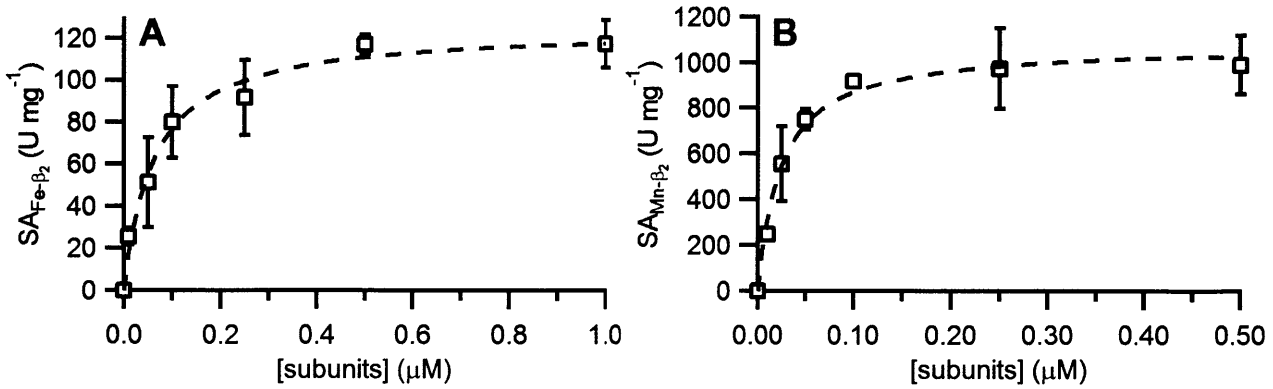


Figure 2.8. Subunit concentration dependence of the activity of a 1:1 mixture of (**A**) His_6-NrdE and holo-$Fe(III)_2$-Y• His_6-NrdF (1.0 Y•/β_2), and (**B**) His_6-NrdE and holo-$Mn(III)_2$-Y• His_6-NrdF (1.1 Y•/β_2). Data points are the average of three replicates. Error bars represent ± 1 standard deviation from the mean. The dotted curves show the fit of the data to eq 2.3.

2.3.8. Analysis of the quaternary structure of the *B. subtilis* class Ib RNR by multiple methods reveals a dynamic system.

2.3.8A. Analysis of the quaternary structure by SEC. SEC was used initially to estimate the molecular weights of the subunits and quaternary structure(s) of holo-RNR. His_6-NrdF (Fe- or Mn-loaded) eluted predominantly as a single peak (**Figure 2.9A**) with a molecular weight of 80 kDa (**Table 2.6**), consistent with a dimer. A small amount of monomer was also observed (50 kDa, **Table 2.6**). Near physiologically relevant concentrations (~2.7 μM), His_6-NrdE also eluted as a single peak (**Figure 2.9B**) with a molecular weight of 83 kDa (**Table 2.6**), consistent with a monomer. When the concentration of His_6-NrdE was increased 10-fold (27 μM), the protein forms several higher molecular weight oligomers (**Figure 2.9B**), including dimer (155 kDa) and species with molecular weights in the range of pentamers to heptamers (for a full list of retention volumes, see Appendix 2, **Table A2.6**). Studies with 1:1 mixtures of His_6-NrdE:His_6-NrdF at concentrations from 1 – 23 μM are shown in **Figure 2.9C**. Even at physiological concentrations (1 μM), complexity in the *B. subtilis* Ib RNR quaternary structure was observed. Broad peaks with retention volumes ranging from 9.8 – 14.5 mL are indicative of interconverting species. At 1 μM

complex, the highest percentage of protein eluted at V_e = 11.8 mL, regardless of the metal loading. This retention volume corresponded to a molecular weight of 205 kDa (**Table 2.6**), suggesting, in conjunction with the other hydrodynamic properties, the presence of an $\alpha_2\beta_2$-like complex with a structure similar to the elongated *S.* Typhimurium class Ib holo-RNR (**Table 2.6**).[57] As the concentration increased, a species eluting at $V_e \approx 10$ mL accumulates, corresponding to a molecular weight of ~400 kDa (see Appendix 2, **Table A2.6**). The peak, however, is broad, indicating that it is composed of multiple species with molecular weights ranging from 370 – 450 kDa. These results clearly indicate that the quaternary structure(s) of the *B. subtilis* RNR are not stable in the absence of nucleotides, resulting in an ensemble of different interconverting species, despite the low experimentally measured K_m for the subunit interaction (**Figure 2.8**).

Table 2.6. Biophysical characteristics of *B. subtilis* class Ib RNR in the absence of effectors.[a]

	NrdE (1 μM)		**$Mn(III)_2$-Y• NrdF 1 (μM)**				**1:1 NrdE:$Fe(III)_2$-Y• NrdF (1 μM)**			
Parameter	**Meas.**	**Pred.**	**Meas.**	**Pred.**	**Meas.**	**Pred.**	**Meas.**	**Pred.**	**Meas.**	**Pred.**
M_w (kDa)	71 *83*	83.0	45 *50*	40.6	90 *80*	81.3	31	40.6	199 *205*	247.2
$s_{20,w}$ (S)	5.6	5.4	3.7	3.5	4.8	5.5	3.6	3.5	9.6	10.8[b] 11.4[c]
f/f_o	1.23 *1.22*	1.27	1.15 *1.40*	1.20	1.40 *1.38*	1.20	1.15	1.20	1.46 *1.21*	1.30[b] 1.23[c]
R_s (Å)	35 *35*	36.4	26 *32*	27.2	40 *40*	34.6	26	27.2	61 *50*	54.0[b] 51.2[c]
State(s)	α		β		β_2		β		$\alpha_m(\beta_2)_n$	

[a] Values derived from fits of the SV-AUC in Sedphat (regular font) and from the SEC data (*italicized font*) using the $s_{20,w}$ values determined by AUC. [b] Predicted values from HYDROPRO based on alignment of the *B. subtilis* class Ib RNR with the *S.* Typhimurium class Ib holo-RNR.[57] [c] Predicted values from HYDROPRO based on alignment of the *B. subtilis* class Ib RNR with the *E. coli* class Ia docking model.[56]

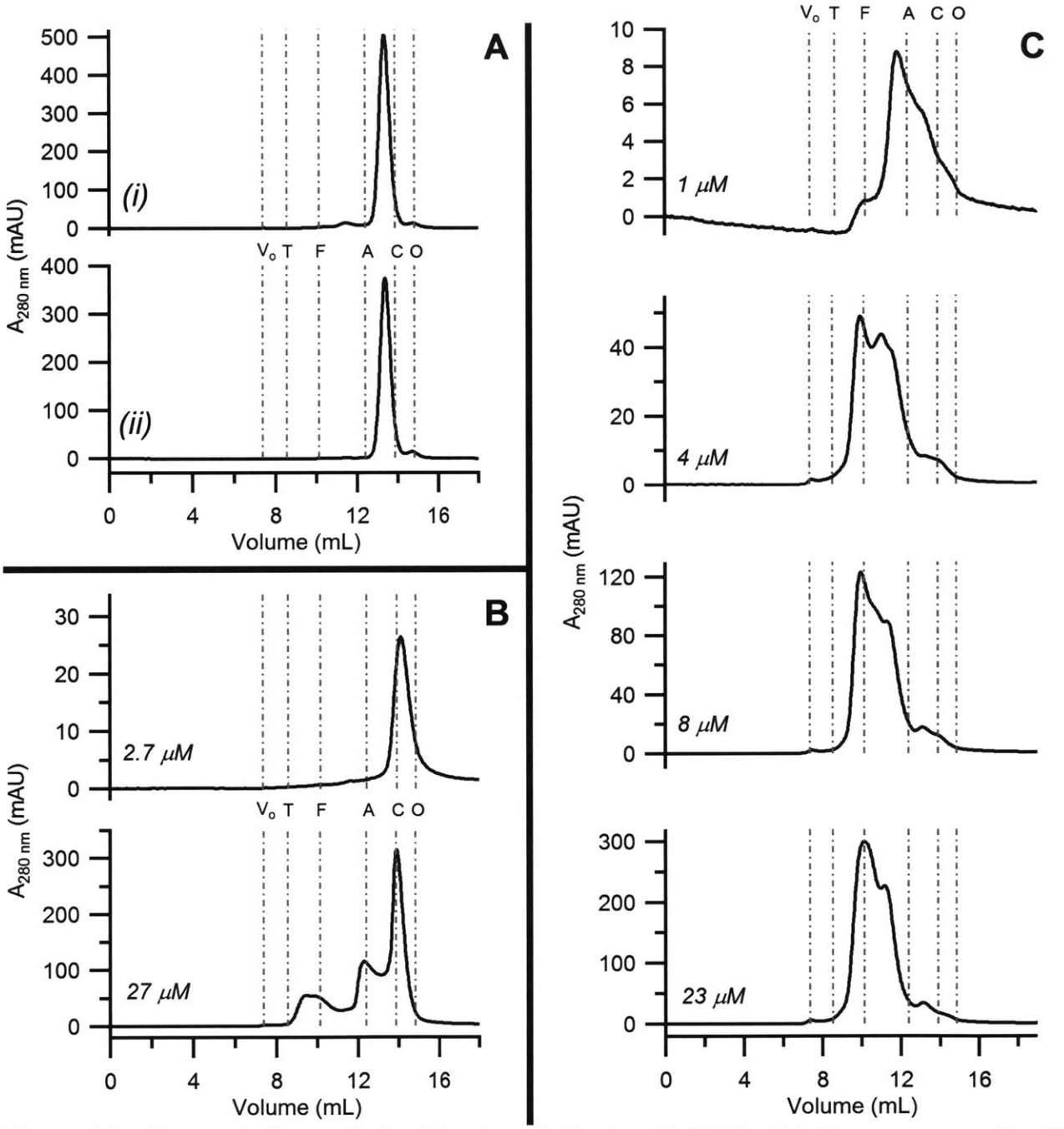


Figure 2.9. Size exclusion analysis of the *B. subtilis* class Ib RNR. (**A**) Chromatograms for *(i)* holo-Fe(III)$_2$-Y• His$_6$-NrdF and *(ii)* holo-Mn(III)$_2$-Y• His$_6$-NrdF. (**B**) Chromatograms for His$_6$-NrdE samples at loading concentrations of 2.7 μM and 27 μM dimer. (**C**) Chromatograms for 1:1 mixtures of His$_6$-α_2:holo-Me(III)$_2$-Y• His$_6$-β_2 at loading concentrations of 1 μM, 4 μM, 8 μM, and 23 μM. The vertical grey lines indicate the retention volumes of the calibrants: V_o = Blue Dextran 2000 (void volume), T = thyroglobulin, F = ferritin, A = aldolase, C = conalbumin, and O = ovalbumin.

2.3.8B. Analysis of the quaternary structure by SV-AUC. SV experiments with each subunit and with a 1:1 mixture of subunits were carried out at different concentrations (**Table 2.2**) to gain further insight into their quaternary states. To aid in the analysis of our experimental data, HYDROPRO[51] was used to predict hydrodynamic properties for His_6-NrdF and His_6-NrdE from the NrdF crystal structure (PDB 4DRO[52]) and a threading model of NrdE generated using the *S.* Typhimurium NrdE structure as a template (PDB 1PEM[54]). Predictions were also made for docking models of the *B. subtilis* $\alpha_2\beta_2$ complex prepared by *in silico* alignment of the subunit models with the *E. coli* class Ia RNR docking model[56] (symmetric $\alpha_2\beta_2$ composed of PDB 4R1R[72] and 1MXR[73]) and the elongated X-ray structure of the *S.* Typhimurium class Ib holo-RNR (asymmetric $\alpha_2\beta_2$, PDB 2BQ1[57]). The results of the HYDROPRO calculations are summarized in **Table 2.7**.

Table 2.7. Predicted hydrodynamic properties of the His_6-tagged *B. subtilis* Ib RNR subunits and $\alpha_2\beta_2$ complex.

Protein	R_s (Å)	$s_{20,w}$ (S)[a]	f x 10^{-7} (gs^{-1})[a]	s_{max} (S)[b]	f_o x 10^{-7} (gs^{-1})[b]	f/f_o
α	36.4	5.4	0.692	6.8	0.545	1.27
β	27.2	3.5	0.518	4.5	0.431	1.20
α_2	48.0	8.2	0.903	10.8	0.687	1.31
β_2	34.6	5.5	0.652	6.6	0.542	1.20
Symmetric $\alpha_2\beta_2$[c]	51.2	11.4	0.967	14.1	0.785	1.23
Asymmetric $\alpha_2\beta_2$[d]	54.0	10.8	1.021	14.1	0.785	1.30

[a] Predicted from crystal structures in HYDROPRO using ρ = 0.998 g cm^{-3} and η = 0.01 poise.
[b] Maximum sedimentation and frictional coefficients expected for a perfect sphere of the same molecular weight (**Table 2.3**).
[c] Based on alignment of *B. subtilis* Ib RNR with the *E. coli* Ia docking model.
[d] Based on alignment of *B. subtilis* Ib RNR with the *S.* Typhimurium holo-RNR complex.

The sedimentation behavior of $Me(III)_2$-Y• His_6-β_2 (1 - 11 μM), His_6-NrdE (1 - 30 μM), and His_6-α_2:$Fe(III)_2$-Y• His_6-β_2 (25 nM to 7 μM) were determined and analyzed using the c(s) model in Sedfit and, due to the detection by the c(s) fits of interactions occurring on the timescale

of sedimentation, with Sedphat. Further details about these analyses are presented in Appendix 2. The results for 1 μM $Mn(III)_2$-Y• His_6-β_2 are shown in **Figure 2.10A** and reveal two peaks representing stable, non-interacting entities at all concentrations examined (Appendix 2, **Figure A2.1**). Globally fitting the data sets in Sedphat yielded $s_{20,w}$ = 3.7 S and 4.8 S for each entity, along with the corresponding M_Ws listed in **Table 2.6**. These results, in comparison with HYDROPRO predictions (**Table 2.7**) and SEC results (**Table 2.6**) suggest that the predominant species is a dimer and the minor species is a monomer. Similar results were observed with $Fe(III)_2$-Y• His_6-β_2 (**Appendix 2**, Figure A2.1). These results are consistent with the general reports that all class I small subunits are dimers.

The results of similar experiments with His_6-NrdE are shown in **Figure 2.10A** and in Appendix 2 (**Figure A2.2**). At 1 μM His_6-NrdE, a single peak with an $s_{20,w}$ = 5.6 S is observed. However, analysis of samples of increasing protein concentrations up to 9 μM reveal a second peak at $s_{20,w}$ = 7.4 S (Appendix 2, **Figure A2.2A**) and at still higher concentrations (30 μM), larger oligomeric states (**Figure A2.2B**). The 1 μM His_6-NrdE data set was analyzed in Sedphat, and when compared to the HYDROPRO predictions (**Table 2.7**) and the SEC data (**Table 2.6**), suggest that the peak at low concentrations is His_6-NrdE monomer. As the protein concentration increases, His_6-NrdE appears to form a temporally unstable dimer which can further oligomerize (Appendix 2, **Figure A2.2B**). It is also possible that the larger species correspond to aggregates since it has been noted that α from prokaryotic and eukaryotic class Ia RNRs in general exhibit low solubility and aggregate, as has been reported previously by SAXS and electron microscopy analyses.[16]

The analysis of a 1:1 mixture His_6-α_2:$Fe(III)_2$-Y• His_6-β_2 using c(s) is shown in **Figure 2.10B** (also see Appendix 2, **Figure A2.3**). Even at 1 μM, dynamic behavior of the subunits is indicated by the reaction boundary that has apparent peaks at 4.4 S and 9.7 S. A plot of the best fit

f/f_o from the c(s) models versus loading concentration (**Figure 2.10C**) reveals a sharp decrease as the protein concentration increases. As described in more detail in the Supporting Information, this observation indicates the presence of interconverting species. The data was imported into Sedphat and the peaks were treated as discrete species with individual f/f_o.[48] The calculated hydrodynamic properties of the smaller species ($s_{20,w}$ = 3.6 S) are consistent with a His_6-NrdF monomer. For the larger species ($s_{20,w}$ = 9.6 S) the analysis suggests the presence of oligomeric structures that are similar to, but smaller than that predicted for an $\alpha_2\beta_2$ complex (**Table 2.7**). As noted however, this is clearly not a discrete species as this peak shifts to larger $s_{20,w}$ and larger complex(es) of unknown composition become evident as the protein concentration increases (**Figure 2.10D**, see also Appendix 2, **Figure A2.3**). Thus, we currently do not have a model for the behavior of the 1:1 His_6-NrdE:His_6-NrdF complex. Further analysis requires the presence of different substrate and effector pairs after their binding stoichiometries are established, and requires development of a method to specifically reduce the Y• in NrdF to prevent chemistry during the analysis.[61]

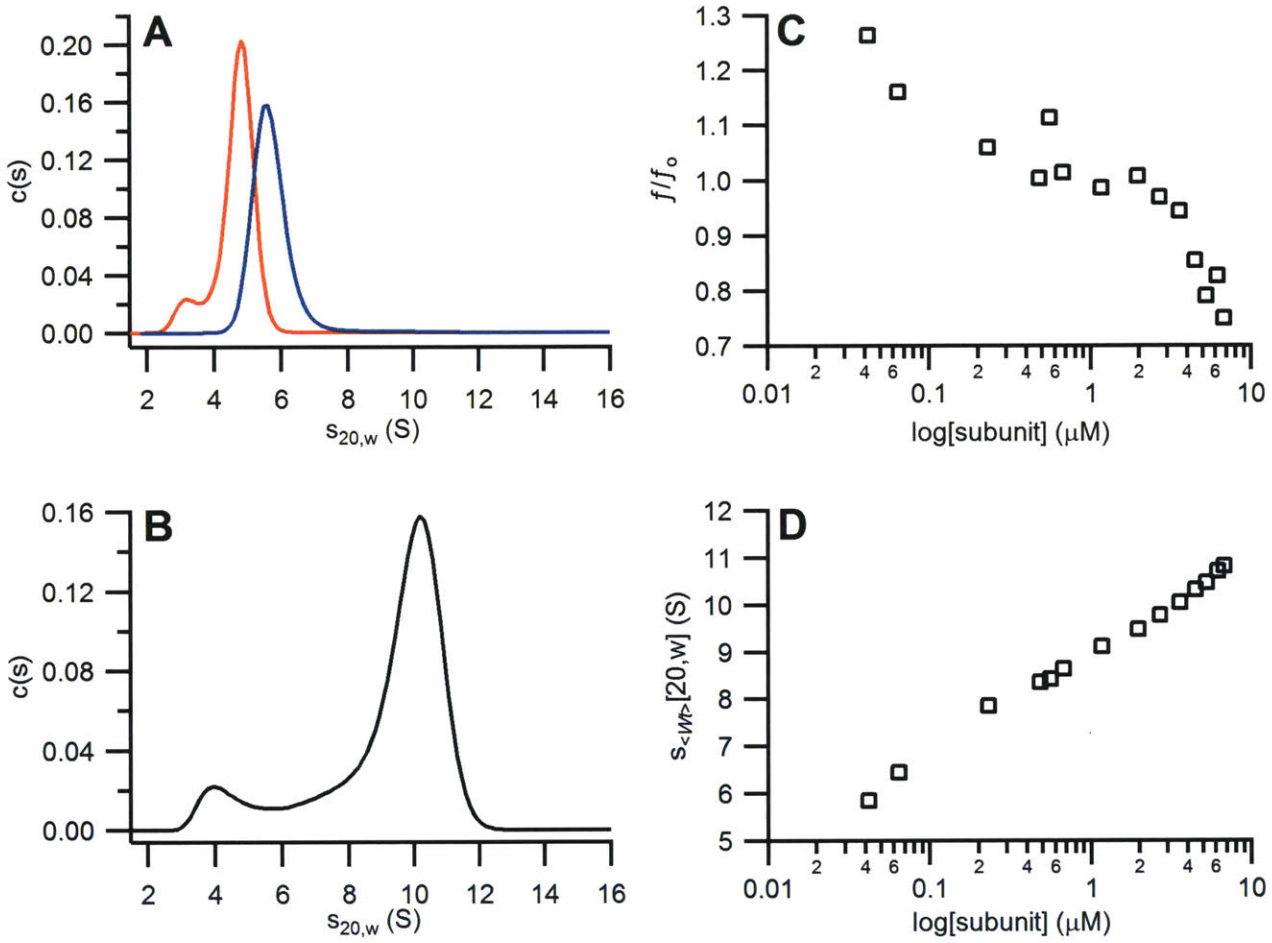


Figure 2.10. Characterization of the *B. subtilis* class Ib RNR by SV-AUC (all samples were run at 1 μM final concentration). (**A**) Overlaid $s_{20,w}$ distributions of holo-$Mn(III)_2$-Y• His_6-β_2 (red trace) and His_6-NrdE (blue trace). (**B**) The $s_{20,w}$ distribution of a 1:1 mixture of $Fe(III)_2$-Y• His_6-β_2:His_6-α_2. (**C**) An isotherm of the weight averaged frictional ratios versus loading concentration of 1:1 mixtures of $Fe(III)_2$-Y• His_6-β_2:His_6-α_2 shows a steep decrease as subunit concentrations increase. (**D**) An isotherm plotting the weight averaged $s_{20,w}$ against loading concentration of 1:1 mixtures of $Fe(III)_2$-Y• His_6-β_2:His_6-α_2 indicates the *B. subtilis* class Ib RNR exhibits complicated quaternary structural dynamics.

2.4. DISCUSSION

In 2010, it was shown that class Ib RNRs can assemble a $Mn(III)_2$-Y• cofactor with the aid of the unusual flavodoxin NrdI and O_2 *in vitro*[61] and that enzymes isolated from three endogenous sources have a similar cluster.[6, 7, 9] More recently, *in vitro* assembly of a $Mn(III)_2$-Y• cofactor using species-specific NrdIs have been carried out in *B. subtilis*,[9, 27] *S. sanguinis*,[71] *B. anthracis*, and *B. cereus*[24] with similar results. The interesting observation that a $Fe(III)_2$-Y• cofactor can self-assemble in these enzymes *in vitro* and maintain catalytic activity, although demonstrably lower than the Mn-loaded cofactor (Appendix 2, **Table A2.1**),[3-5, 54, 57] raises important biological questions as to the nature of the active cofactor *in vivo* and whether it changes with the growth conditions of the organism. Given that many pathogenic organisms[74] require a class Ib RNR for aerobic growth and that humans use a $Fe(III)_2$-Y• class Ia RNR, the distinctions in the mechanisms of cofactor biosynthesis, allosteric regulation, and quaternary structure may offer new opportunities for therapeutic intervention. Thus, understanding the basic properties of the Ib RNRs, as well as their similarities and differences to one another and to class Ia RNRs, is an important first step.

Recent studies on the class Ib RNRs (with $Mn(III)_2$-Y• NrdF) from *B. anthracis* and *B. cereus* have identified and characterized the endogenous reductants required to make deoxynucleotides at optimum rates.[24, 25] Despite this important contribution, these and other class Ib RNRs (Appendix 2, **Table A2.1**) still have low catalytic activity. Reconstitution of the $Mn(III)_2$-Y• cofactor in *B. anthracis* NrdF using NrdI gave 0.3 – 0.4 Y•/β_2, but the reported activity was only 44 – 70 U/mg.[24, 25] Similar activities were observed using DTT (40 – 63 U/mg) alone or with either TR or NrdH.[24, 25] The *B. anthracis* $Fe(III)_2$-Y• NrdF, reconstituted with 0.6 Y•/β_2, was reported to have activity of only 8 U/mg.[24] This value is very low given the high Y• content, which

in the case of the Ia RNRs, is correlated with activity.[75] The initial report on the activity of endogenous *B. subtilis* RNR using DTT as reductant was also low, 160 U/mg.[9] To study the class Ib RNRs, it was thus important to better understand the basis for the low catalytic activity, which was believed to lie with cluster assembly in NrdF and the reductant used to support NDP reduction. The studies presented in this chapter show that apo/mismetallated forms of NrdF can be separated from holo-protein, yielding $Mn(III)_2$-Y• NrdF with 1 Y•/β_2 and 4 Mn/β_2. Contrary to expectations, the purity of $Me(III)_2$-Y• NrdF was only a minor factor in activity enhancement (**Table 2.4**). Rather, the identity of the reductant used to support turnover was the major limiting factor in obtaining highly active preparations of recombinant *B. subtilis* RNR. The optimum endogenous reducing system was TrxA/TrxB/NADPH, which supported catalytic activity of ~1450 U/mg. However, the major contributor to high enzymatic activity may vary from species to species, given that with *B. anthracis*, Trx1 and DTT yielded similar activities using poorly loaded $Mn(III)_2$-Y• NrdF.[25] The basis for the low activity in the *B. anthracis* system is likely associated with mismetallated NrdFs which interfere with formation of an active RNR complex.

The availability of highly active *B. subtilis* NrdF and NrdE enabled the assembly of an active RNR using a 1:1 ratio of subunits at physiologically relevant concentrations (**Figure 2.8**). The specific activity of the enzyme measured under these conditions is one of the highest reported for any class Ib RNR (Appendix 2, **Table A2.1**). Furthermore, the apparent K_m of 0.03 μM measured for *B. subtilis* RNR contrasts with a K_d of ~0.2 μM measured for the *E. coli* Ia RNR subunits, which was determined with each subunit assayed independently in the presence of a 5- to 10-fold excess of the second subunit to generate the "active" complex.[13, 14] The latter conditions are routinely used in the study of class I RNRs, and all the assays reported for the *B. anthracis* enzyme were carried out using this method.[24, 25] While assaying one subunit in excess of the

second often yields high and reproducible specific activities, these conditions are problematic in that they could alter quaternary structures (both α and α/β) and result in inhibition under these assay conditions (see discussion below). With bacterial systems, it now seems reasonable that all assays of enzymatic activity should be carried out with a 1:1 ratio of the two subunits, accompanied by a study to establish the optimum concentration range for maximum active complex formation with each RNR.

Two issues warrant further discussion of the results with the *B. subtilis* Ib RNR in comparison to those with *B. anthracis* (and *B. cereus*) enzymes: the nature of the endogenous reductant(s),[24, 25] and the nature and dynamics of their quaternary structure(s). In the studies of Gustafsson *et al.*, interrogation of the *B. anthracis* genome identified three thioredoxins, two potential thioredoxin reductases, and three potential glutaredoxin-like proteins, all of which were cloned, expressed, and screened for their ability to support RNR activity.[25] The turnover numbers for the *B. anthracis* RNR were highest with Trx1/TR1 and 7-fold greater than with NrdH/TR1 (Appendix 2, **Table A2.1**). Furthermore, under their growth conditions, western analysis revealed that Trx1 was the predominant RNR reductant in the cell.[25]

The *B. anthracis* genes[25] were used to query the *B. subtilis* genome for similar proteins, resulting in the identification of six thioredoxin-like proteins (TrxA, Ydbp, YtpP, YdfQ, YusE, and YusI), one thioredoxin reductase (TrxB), and two thioredoxin-like glutaredoxins (YosR and BdbA). TrxA and TrxB are homologs of *B. anthracis* Trx1 (75% identity) and TR1 (87% identity), suggesting they likely function as the reducing system for the *B. subtilis* RNR. Three of the other potential reductants for RNR were eliminated from consideration based on previous studies: BdbA participates in the biosynthesis of the S-linked glycopeptide sublancin 168,[76, 77] YusE is under the control of the sporulation transcription factor Spo0A and, thus, likely functions during *B. subtilis*

sporulation,[78] and YdfO is induced when *B. subtilis* is exposed to aromatic compounds and, thus, likely participates in their catabolism.[79, 80] Moreover, previous gene knock out experiments of all of these proteins and the growth of the resulting deletion mutants in rich or minimal medium, revealed that only TrxA and TrxB are essential.[10, 81] Using these purified proteins, it was established that endogenous reductants effect a ~10-fold increase on RNR activity relative to DTT (**Table 2.4**) in contrast with the *B. anthracis* RNR.[25] It is also noteworthy that studies by other Stubbe group members have reported a 20-fold difference between endogenous reductant and DTT with the *S. sanguinis* class Ib RNR.[71]

It was pointed out previously by Sjöberg *et al.* that *B. subtilis* has a second class Ib RNR, located in the *bnrdEFI* operon, which could potentially be physiologically important.[24] The gene *yosR*, recently reannotated as *nrdH*, is adjacent to this operon,[25, 74] prompting an examination of this protein in assays as a potential reductant for RNR as well. As was observed with the *B. anthracis* RNR, YosR(NrdH) exhibits activity similar to TrxA in ribonucleotide reduction (**Table 2.4**). In *B anthracis*, as noted above, Western blot analysis suggested that TrxA was the endogenous reductant. Previous studies in *B. subtilis* have established that *bnrdEFI* operon is dispensable for its survival under numerous growth conditions,[11, 12, 82] supported by transcriptional analysis under rich and minimal growth conditions in which both *yosR* and the *bnrdEFIH* were expressed at very low levels.[65] Therefore, YosR, although able to support RNR activity, does not appear to be the physiologically relevant reductant.

One gene not returned as a hit to the BLAST searches using *B. anthracis* redoxins[25] was *ytnI*, encoding a glutaredoxin-like protein annotated as a NrdH by the Conserved Domain Database tool of NCBI.[83] In addition to the annotation, *ytnI* was also interesting because it was found to be the most upregulated in expression (~450-fold) in a *trxA* knockout mutant of *B. subtilis* that was

kept viable by complementation with an IPTG-inducible vector encoding the intact gene.[81] By changing the concentration of IPTG in the growth medium, the cells were depleted of TrxA to study the transcriptional response of *B. subtilis* to TrxA limitation, and this condition resulted in significant upregulation of *ytnI* expression.[81] A hallmark of NrdH reported for other organisms is that it is reduced by thioredoxin reductase despite having a glutaredoxin-like tertiary structure.[25, 64, 84, 85] This was found not to be the case with YtnI, which led to the conclusion that YtnI was not the endogenous RNR reductant and, thus, further substantiates the claim that TrxA is the most likely the physiological reductant for RNR in *B. subtilis*.

The most interesting observation from the studies reported in this chapter is associated with the analysis of quaternary structure. Tight subunit association was expected in the absence of nucleotides (dNTPs), based on the isolation of endogenous *B. subtilis* RNR[9] and the concentration dependence of activity (**Figure 2.8**), in contrast to the class Ia RNRs (K_d of 0.4 μM).[14, 86] The results of the SEC (**Figure 2.9**) and SV-AUC (**Figure 2.10**) analyses, however, revealed that this is not the case. Under physiological concentrations (1 μM) with no nucleotides, both methods revealed mixtures of subunits and complexes of subunits. GEMMA analysis of the *B. anthracis* class Ib RNR in the absence of nucleotides also revealed mixtures of subunits and an $\alpha_2\beta_2$ complex.[24] However, a thorough study of the subunit affinities and the rates of association and dissociation has not been carried out with either the *B. subtilis* or *B. anthracis* RNRs. In order to initiate such studies, it is necessary to find a way to specifically reduce Y• to prevent turnover form occurring during the analyses.

Finally, the observation of potent inhibition of RNR activity by dATP given the absence of an ATP cone domain deserves comment. The results are distinct from all class Ib RNRs reported to date, which show no inhibition, even at concentrations of 1 mM dATP.[3, 5, 24, 69, 71, 87]

Furthermore, with dATP and the *B. anthracis* RNR, the GEMMA analysis revealed individual subunits and $\alpha_2\beta_2$, but no large quaternary states were reported.[24] Thus, the cause of the potent inhibition with the *B. subtilis* RNR requires further study.

The results of the studies reported in this chapter suggest that the paradigm of RNR as $\alpha_2\beta_2$ in the literature needs to be re-examined.[16, 17, 22] The difficulties associated with obtaining structures of prokaryotic RNRs with both subunits present are likely associated with the dynamics of their interactions. Undoubtedly these dynamics will be affected by the presence of ATP, dNTPs, and Mg^{2+}, and studies of these effects under physiologically relevant conditions are currently in progress. In order to carry out these types of studies, highly active protein with fully loaded metallo-cofactor is essential, and the studies reported herein provide the foundation to investigate quaternary structure and its connection to the enzymatic activity and allosteric regulation of RNRs.

2.5. REFERENCES

1. Hofer, A., Crona, M., Logan, D. T., and Sjöberg, B. M. (2012) DNA Building Blocks: Keeping Control of Manufacture, *Crit. Rev. Biochem. Mol. Biol. 47*, 50-63.
2. Cotruvo, J. A., and Stubbe, J. (2011) Class I Ribonucleotide Reductases: Metallocofactor Assembly and Repair *In Vitro* and *In Vivo*, In *Annual Review of Biochemistry* (Kornberg, R. D., Raetz, C. R. H., Rothman, J. E., and Thorner, J. W., Eds.), pp 733-767, Annual Reviews, Palo Alto.
3. Eliasson, R., Pontis, E., Jordan, A., and Reichard, P. (1996) Allosteric Regulation of the Third Ribonucleotide Reductase (NrdEF Enzyme) from Enterobacteriaceae, *J. Biol. Chem. 271*, 26582-26587.
4. Huque, Y., Fieschi, F., Torrents, E., Gibert, I., Eliasson, R., Reichard, P., Sahlin, M., and Sjöberg, B. M. (2000) The Active Form of the R2F Protein of Class Ib Ribonucleotide Reductase from *Corynebacterium ammoniagenes* is a Diferric Protein, *J. Biol. Chem. 275*, 25365-25371.
5. Jordan, A., Pontis, E., Atta, M., Krook, M., Gibert, I., Barbé, J., and Reichard, P. (1994) A Second Class I Ribonucleotide Reductase in *Enterobacteriaceae* - Characterization of the *Salmonella typhimurium* Enzyme, *Proc. Natl. Acad. Sci. U. S. A. 91*, 12892-12896.
6. Cotruvo, J. A., and Stubbe, J. (2011) *Escherichia coli* Class Ib Ribonucleotide Reductase Contains a Dimanganese(III)-Tyrosyl Radical Cofactor *in vivo*, *Biochemistry 50*, 1672-1681.

7. Cox, N., Ogata, H., Stolle, P., Reijerse, E., Auling, G., and Lubitz, W. (2010) A Tyrosyl-Dimanganese Coupled Spin System is the Native Metalloradical Cofactor of the R2F Subunit of the Ribonucleotide Reductase of *Corynebacterium ammoniagenes*, *J. Am. Chem. Soc. 132*, 11197-11213.

8. Stolle, P., Barckhausen, O., Oehlmann, W., Knobbe, N., Vogt, C., Pierik, A. J., Cox, N., Schmidt, P. P., Reijerse, E. J., Lubitz, W., and Auling, G. (2010) Homologous Expression of the *nrdF* Gene of *Corynebacterium ammoniagenes* Strain ATCC 6872 Generates a Manganese-metallocofactor (R2F) and a Stable Tyrosyl Radical (Y•) Involved in Ribonucleotide Reduction, *FEBS J. 277*, 4849-4862.

9. Zhang, Y., and Stubbe, J. (2011) *Bacillus subtilis* Class Ib Ribonucleotide Reductase is a Dimanganese(III)-Tyrosyl Radical Enzyme, *Biochemistry 50*, 5615-5623.

10. Kobayashi, K., Ehrlich, S. D., Albertini, A., Amati, G., Andersen, K. K., Arnaud, M., Asai, K., Ashikaga, S., Aymerich, S., Bessieres, P., Boland, F., Brignell, S. C., Bron, S., Bunai, K., Chapuis, J., Christiansen, L. C., Danchin, A., Débarbouillé, M., Dervyn, E., Deuerling, E., Devine, K., Devine, S. K., Dreesen, O., Errington, J., Fillinger, S., Foster, S. J., Fujita, Y., Galizzi, A., Gardan, R., Eschevins, C., Fukushima, T., Haga, K., Harwood, C. R., Hecker, M., Hosoya, D., Hullo, M. F., Kakeshita, H., Karamata, D., Kasahara, Y., Kawamura, F., Koga, K., Koski, P., Kuwana, R., Imamura, D., Ishimaru, M., Ishikawa, S., Ishio, I., Le Coq, D., Masson, A., Mauël, C., Meima, R., Mellado, R. P., Moir, A., Moriya, S., Nagakawa, E., Nanamiya, H., Nakai, S., Nygaard, P., Ogura, M., Ohanan, T., O'Reilly, M., O'Rourke, M., Pragai, Z., Pooley, H. M., Rapoport, G., Rawlins, J. P., Rivas, L. A., Rivolta, C., Sadaie, A., Sadaie, Y., Sarvas, M., Sato, T., Saxild, H. H., Scanlan, E., Schumann, W., Seegers, J., Sekiguchi, J., Sekowska, A., Séror, S. J., Simon, M., Stragier, P., Studer, R., Takamatsu, H., Tanaka, T., Takeuchi, M., Thomaides, H. B., Vagner, V., van Dijl, J. M., Watabe, K., Wipat, A., Yamamoto, H., Yamamoto, M., Yamamoto, Y., Yamane, K., Yata, K., Yoshida, K., Yoshikawa, H., Zuber, U., and Ogasawara, N. (2003) Essential *Bacillus subtilis* Genes, *Proc. Natl. Acad. Sci. U. S. A. 100*, 4678-4683.

11. Lazarevic, V., Soldo, B., Dusterhoft, A., Hilbert, H., Mauel, C., and Karamata, D. (1998) Introns and Intein Coding Sequence in the Ribonucleotide Reductase Genes of *Bacillus subtilis* Temperate Bacteriophage SPβ, *Proc. Natl. Acad. Sci. U. S. A. 95*, 1692-1697.

12. Westers, H., Dorenbos, R., van Dijl, J. M., Kabel, J., Flanagan, T., Devine, K. M., Jude, F., Séror, S. J., Beekman, A. C., Darmon, E., Eschevins, C., de Jong, A., Bron, S., Kuipers, O. P., Albertini, A. M., Antelmann, H., Hecker, M., Zamboni, N., Sauer, U., Bruand, C., Ehrlich, D. S., Alonso, J. C., Salas, M., and Quax, W. J. (2003) Genome Engineering Reveals Large Dispensable Regions in *Bacillus subtilis*, *Mol. Biol. Evol. 20*, 2076-2090.

13. Climent, I., Sjöberg, B. M., and Huang, C. Y. (1991) Carboxyl-Terminal Peptides as Probes for *Escherichia coli* Ribonucleotide Reductase Subunit Interaction - Kinetic Analysis of Inhibition Studies, *Biochemistry 30*, 5164-5171.

14. Hassan, A. Q., Wang, Y. T., Plate, L., and Stubbe, J. (2008) Methodology To Probe Subunit Interactions in Ribonucleotide Reductases, *Biochemistry 47*, 13046-13055.

15. Ingemarson, R., and Thelander, L. (1996) A Kinetic Study on the Influence of Nucleoside Triphosphate Effectors on Subunit Interaction in Mouse Ribonucleotide Reductase, *Biochemistry 35*, 8603-8609.

16. Ando, N., Brignole, E. J., Zimanyi, C. M., Funk, M. A., Yokoyama, K., Asturias, F. J., Stubbe, J., and Drennan, C. L. (2011) Structural Interconversions Modulate Activity of *Escherichia coli* Ribonucleotide Reductase, *Proc. Natl. Acad. Sci. U. S. A. 108*, 21046-21051.

17. Minnihan, E. C., Ando, N., Brignole, E. J., Olshansky, L., Chittuluru, J., Asturias, F. J., Drennan, C. L., Nocera, D. G., and Stubbe, J. (2013) Generation of a Stable, Aminotyrosyl Radical-Induced $\alpha_2\beta_2$ Complex of *Escherichia coli* Class Ia Ribonucleotide Reductase, *Proc. Natl. Acad. Sci. U. S. A. 110*, 3835-3840.

18. Rofougaran, R., Crona, M., Vodnala, M., Sjöberg, B. M., and Hofer, A. (2008) Oligomerization Status Directs Overall Activity Regulation of the *Escherichia coli* Class Ia Ribonucleotide Reductase, *J. Biol. Chem. 283*, 35310-35318.

19. Fairman, J. W., Wijerathna, S. R., Ahmad, M. F., Xu, H., Nakano, R., Jha, S., Prendergast, J., Welin, R. M., Flodin, S., Roos, A., Nordlund, P., Li, Z., Walz, T., and Dealwis, C. G. (2011) Structural Basis for Allosteric Regulation of Human Ribonucleotide Reductase by Nucleotide-Induced Oligomerization, *Nat. Struct. Mol. Biol. 18*, 316-322.

20. Kashlan, O. B., Scott, C. P., Lear, J. D., and Cooperman, B. S. (2002) A Comprehensive Model for the Allosteric Regulation of Mammalian Ribonucleotide Reductase. Functional Consequences of ATP- and dATP-Induced Oligomerization of the Large Subunit, *Biochemistry 41*, 462-474.

21. Rofougaran, R., Vodnala, M., and Hofer, A. (2006) Enzymatically Active Mammalian Ribonucleotide Reductase Exists Primarily as an $\alpha_6\beta_2$ Octamer, *J. Biol. Chem. 281*, 27705-27711.

22. Zimanyi, C. M., Ando, N., Brignole, E. J., Asturias, F. J., Stubbe, J., and Drennan, C. L. (2012) Tangled Up in Knots: Structures of Inactivated Forms of *E. coli* Class Ia Ribonucleotide Reductase, *Structure 20*, 1374-1383.

23. Crona, M., Furrer, E., Torrents, E., Edgell, D. R., and Sjöberg, B. M. (2010) Subunit and Small-Molecule Interaction of Ribonucleotide Reductases Via Surface Plasmon Resonance Biosensor Analyses, *Protein Eng. Des. Sel. 23*, 633-641.

24. Crona, M., Torrents, E., Røhr, A. K., Hofer, A., Furrer, E., Tomter, A. B., Andersson, K. K., Sahlin, M., and Sjöberg, B. M. (2011) NrdH-Redoxin Protein Mediates High Enzyme Activity in Manganese-Reconstituted Ribonucleotide Reductase from *Bacillus anthracis*, *J. Biol. Chem. 286*, 33053-33060.

25. Gustafsson, T. N., Sahlin, M., Lu, J., Sjöberg, B. M., and Holmgren, A. (2012) *Bacillus anthracis* Thioredoxin Systems, Characterization and Role as Electron Donors for Ribonucleotide Reductase, *J. Biol. Chem. 287*, 39686-39697.

26. Ge, J., Yu, G. X., Ator, M. A., and Stubbe, J. (2003) Pre-Steady-State and Steady-state Kinetic Analysis of *E. coli* Class I Ribonucleotide Reductase, *Biochemistry 42*, 10071-10083.

27. Cotruvo, J. A., Stich, T. A., Britt, R. D., and Stubbe, J. (2013) Mechanism of Assembly of the Dimanganese-Tyrosyl Radical Cofactor of Class Ib Ribonucleotide Reductase: Enzymatic Generation of Superoxide Is Required for Tyrosine Oxidation via a Mn(III)Mn(IV) Intermediate, *J. Am. Chem. Soc. 135*, 4027-4039.

28. Maaβ, S., Sievers, S., Zühlke, D., Kuzinski, J., Sappa, P. K., Muntel, J., Hessling, B., Bernhardt, J., Sietmann, R., Völker, U., Hecker, M., and Becher, D. (2011) Efficient, Global-Scale Quantification of Absolute Protein Amounts by Integration of Targeted Mass Spectrometry and Two-Dimensional Gel-Based Proteomics, *Anal. Chem. 83*, 2677-2684.

29. Maaβ, S., Wachlin, G., Bernhardt, J., Eymann, C., Fromion, V., Riedel, K., Becher, D., and Hecker, M. (2014) Highly Precise Quantification of Protein Molecules per Cell During Stress and Starvation Responses in *Bacillus subtilis*, *Mol. Cell. Proteomics 13*, 2260-2276.

30. Muntel, J., Fromion, V., Goelzer, A., Maaβ, S., Mäder, U., Büttner, K., Hecker, M., and Becher, D. (2014) Comprehensive Absolute Quantification of the Cytosolic Proteome of *Bacillus subtilis* by Data Independent, Parallel Fragmentation in Liquid Chromatography/Mass Spectrometry (LC/MS^E), *Mol. Cell. Proteomics 13*, 1008-1019.

31. Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D., and Bairoch, A. (2005) Protein Identification and Analysis Tools on the ExPASy Server., In *The Proteomics Protocols Handbook* (Walker, J. M., Ed.), pp 571-607, Humana Press, Totowa, NJ.

32. Edelhoch, H. (1967) Spectroscopic Determination of Tryptophan and Tyrosine in Proteins, *Biochemistry 6*, 1948-1954.

33. Gill, S. C., and von Hippel, P. H. (1989) Calculation of Protein Extinction Coefficients from Amino Acid Sequence Data, *Anal. Biochem. 182*, 319-326.

34. Aliverti, A., Curti, B., and Vanoni, M. A. (1999) Identifying and Quantitating FAD and FMN in Simple and in Iron-Sulfur-Containing Flavoproteins, In *Flavoprotein Protocols* (Chapman, S. K., and Reid, G. A., Eds.), pp 9-23, Humana Press.

35. Prongay, A. J., Engelke, D. R., and Williams, C. H. (1989) Characterization of Two Active-Site Mutations of Thioredoxin Reductase from *Escherichia coli*, *J. Biol. Chem. 264*, 2656-2664.

36. Moore, E. C., Reichard, P., and Thelander, L. (1964) Enzymatic Synthesis of Deoxyribonucleotides. V. Purification and Properties of Thioredoxin Reductase from *Escherichia coli* B, *J. Biol. Chem. 239*, 3445-3452.

37. Thelander, L. (1967) Thioredoxin Reductase - Characterization of a Homogeneous Preparation from *Escherichia coli* B, *J. Biol. Chem. 242*, 852-859.

38. Cotruvo, J. A., and Stubbe, J. (2008) NrdI, A Flavodoxin Involved in Maintenance of the Diferric-Tyrosyl Radical Cofactor in *Escherichia coli* Class Ib Ribonucleotide Reductase, *Proc. Natl. Acad. Sci. U. S. A. 105*, 14383-14388.

39. Fish, W. W. (1988) Rapid Colorimetric Micromethod for the Quantitation of Complexed Iron in Biological Samples, *Method Enzymol. 158*, 357-364.

40. Malmström, B. G., Reinhammar, B., and Vanngard, T. (1970) The State of Copper in Stellacyanin and Laccase from the Lacquer Tree *Rhus vernicifera*, *Biochim. Biophys. Acta 205*, 48-57.

41. Bollinger, J. M., Jr., Tong, W. H., Ravi, N., Huynh, B. H., Edmondson, D. E., and Stubbe, J. (1995) Use of Rapid Kinetics Methods to Study the Assembly of the Diferric-Tyrosyl Radical Cofactor of *Escherichia coli* Ribonucleotide Reductase, *Methods Enzymol. 258*, 278-303.

42. Steeper, J. R., and Steuart, C. D. (1970) A Rapid Assay for CDP Reductase Activity in Mammalian Cell Extracts, *Anal. Biochem. 34*, 123-130.

43. (1996) Characteristics of Nucleic Acids, In *Current Protocols in Molecular Biology*, pp A.1D.1-A.1D.11, John Wiley & Sons, Inc.

44. Siegel, L. M., and Monty, K. J. (1966) Determination of Molecular Weights and Frictional Ratios of Proteins in Impure Systems by Use of Gel Filtration and Density Gradient Centrifugation. Application to Crude Preparations of Sulfite and Hydroxylamine Reductases, *Biochim. Biophys. Acta 112*, 346-362.

45. Laurent, T. C., and Killander, J. (1964) Theory of Gel Filtration and its Experimental Verification, *J. Chromatogr. 14*, 317-330.

46. Meirelles, G. V., Silva, J. C., Mendonca, Y. D., Ramos, C. H. I., Torriani, I. L., and Kobarg, J. (2011) Human Nek6 is a Monomeric Mostly Globular Kinase with an Unfolded Short N-terminal Domain, *BMC Struct. Biol. 11:12*.

47. Zhao, H. Y., Ghirlando, R., Piszczek, G., Curth, U., Brautigam, C. A., and Schuck, P. (2013) Recorded Scan Times can Limit the Accuracy of Sedimentation Coefficients in Analytical Ultracentrifugation, *Anal. Biochem. 437*, 104-108.

48. Balbo, A., and Schuck, P. (2005) Analytical Ultracentrifugation in the Study of Protein Self-Association and Heterogeneous Protein-Protein Interactions, In *Protein-Protein Interactions: A Molecular Cloning Manual* (Golemis, E. A., and Adams, P. D., Eds.), pp 253-277, Cold Spring Harbor Laboratory Press, New York.

49. Schuck, P. (2003) On the Analysis of Protein Self-Association by Sedimentation Velocity Analytical Ultracentrifugation, *Anal. Biochem. 320*, 104-124.

50. Laue, T. M., Shah, B. D., Ridgeway, T. M., and Pelletier, S. L. (1992) Computer-aided Interpretation of Analytical Sedimentation Data for Proteins, In *Analytical Ultracentrifugation in Biochemistry and Polymer Science* (Harding, S. E., Rowe, A. J., and Horton, J. C., Eds.), pp 90-125, Royal Society of Chemistry, Cambridge.

51. Ortega, A., Amorós, D., and de la Torre, J. G. (2011) Prediction of Hydrodynamic and Other Solution Properties of Rigid Proteins from Atomic- and Residue-Level Models, *Biophys. J. 101*, 892-898.

52. Boal, A. K., Cotruvo, J. A., Stubbe, J., and Rosenzweig, A. C. (2012) The Dimanganese(II) Site of *Bacillus subtilis* Class Ib Ribonucleotide Reductase, *Biochemistry 51*, 3861-3871.

53. Kelley, L. A., and Sternberg, M. J. E. (2009) Protein Structure Prediction on the Web: A Case Study Using the Phyre Server, *Nat. Protoc. 4*, 363-371.

54. Uppsten, M., Färnegårdh, M., Jordan, A., Eliasson, R., Eklund, H., and Uhlin, U. (2003) Structure of the Large Subunit of Class Ib Ribonucleotide Reductase from *Salmonella typhimurium* and Its Complexes with Allosteric Effectors, *J. Mol. Biol. 330*, 87-97.

55. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004) UCSF Chimera - A Visualization System for Exploratory Research and Analysis, *J. Comput. Chem. 25*, 1605-1612.

56. Uhlin, U., and Eklund, H. (1994) Structure of Ribonucleotide Reductase Protein R1, *Nature 370*, 533-539.

57. Uppsten, M., Färnegårdh, M., Domkin, V., and Uhlin, U. (2006) The First Holocomplex Structure of Ribonucleotide Reductase Gives New Insight into Its Mechanism of Action, *J. Mol. Biol. 359*, 365-377.

58. Lebowitz, J., Lewis, M. S., and Schuck, P. (2002) Modern Analytical Ultracentrifugation in Protein Science: A Tutorial Review, *Protein Sci. 11*, 2067-2079.

59. Mao, S. S., Holler, T. P., Yu, G. X., Bollinger, J. M., Booker, S., Johnston, M. I., and Stubbe, J. (1992) A Model for the Role of Multiple Cysteine Residues Involved in Ribonucleotide Reduction - Amazing and Still Confusing, *Biochemistry 31*, 9733-9743.

60. Atkin, C. L., Thelander, L., Reichard, P., and Lang, G. (1973) Iron and Free Radical in Ribonucleotide Reductase. Exchange of Iron and Mössbauer Spectroscopy of Protein B2 Subunit of the *Escherichia coli* Enzyme, *J. Biol. Chem. 248*, 7464-7472.

61. Cotruvo, J. A., and Stubbe, J. (2010) An Active Dimanganese(III)-Tyrosyl Radical Cofactor in *Escherichia coli* Class Ib Ribonucleotide Reductase, *Biochemistry 49*, 1297-1309.

62. Petersson, L., Gräslund, A., Ehrenberg, A., Sjöberg, B. M., and Reichard, P. (1980) The Iron Center in Ribonucleotide Reductase from *Escherichia coli*, *J. Biol. Chem. 255*, 6706-6712.

63. Pigiet, V. P., and Conley, R. R. (1977) Purification of Thioredoxin, Thioredoxin Reductase, and Glutathione Reductase by Affinity Chromatography, *J. Biol. Chem. 252*, 6367-6372.

64. Jordan, A., Åslund, F., Pontis, E., Reichard, P., and Holmgren, A. (1997) Characterization of *Escherichia coli* NrdH - A Glutaredoxin-like Protein with a Thioredoxin-like Activity Profile, *J. Biol. Chem. 272*, 18044-18050.

65. Rasmussen, S., Nielsen, H. B., and Jarmer, H. (2009) The Transcriptionally Active Regions in the Genome of *Bacillus subtilis*, *Mol. Microbiol. 73*, 1043-1057.

66. Thelander, L., Sjöberg, B. R., and Eriksson, S. (1978) Ribonucleoside Diphosphate Reductase (*Escherichia coli*), *Method Enzymol. 51*, 227-237.

67. Jordan, A., Gibert, I., and Barbé, J. (1994) Cloning and Sequencing of the Genes from *Salmonella typhimurium* Encoding a New Bacterial Ribonucleotide Reductase, *J. Bacteriol. 176*, 3420-3427.

68. Jordan, A., Pontis, E., Åslund, F., Hellman, U., Gibert, I., and Reichard, P. (1996) The Ribonucleotide Reductase System of *Lactococcus lactis* - Characterization of an NrdEF Enzyme and a New Electron Transport Protein, *J. Biol. Chem. 271*, 8779-8785.

69. Yang, F. D., Curran, S. C., Li, L. S., Avarbock, D., Graf, J. D., Chua, M. M., Lui, G. Z., Salem, J., and Rubin, H. (1997) Characterization of Two Genes Encoding the *Mycobacterium tuberculosis* Ribonucleotide Reductase Small Subunit, *J. Bacteriol. 179*, 6408-6415.

70. Reichard, P., Baldesten, A., and Rutberg, L. (1961) Formation of Deoxycytidine Phosphates from Cytidine Phosphates in Extracts from *Escherichia coli*, *J. Biol. Chem. 236*, 1150-1157.

71. Makhlynets, O., Boal, A. K., Rhodes, D. V., Kitten, T., Rosenzweig, A. C., and Stubbe, J. (2014) *Streptococcus sanguinis* Class Ib Ribonucleotide Reductase: High Activity with Both Iron and Manganese Cofactors and Structural Insights, *J. Biol. Chem. 289*, 6259-6272.

72. Eriksson, M., Uhlin, U., Ramaswamy, S., Ekberg, M., Regnström, K., Sjöberg, B. M., and Eklund, H. (1997) Binding of Allosteric Effectors to Ribonucleotide Reductase Protein R1: Reduction of Active-Site Cysteines Promotes Substrate Binding, *Structure 5*, 1077-1092.

73. Högbom, M., Galander, M., Andersson, M., Kolberg, M., Hofbauer, W., Lassmann, G., Nordlund, P., and Lendzian, F. (2003) Displacement of the Tyrosyl Radical Cofactor in Ribonucleotide Reductase Obtained by Single-Crystal High-Field EPR and 1.4-Angstrom X-Ray Data, *Proc. Natl. Acad. Sci. U. S. A. 100*, 3209-3214.

74. Lundin, D., Torrents, E., Poole, A. M., and Sjöberg, B. M. (2009) RNRdb, a Curated Database of the Universal Enzyme Family Ribonucleotide Reductase, Reveals a High Level of Misannotation in Sequences Deposited to Genbank, *BMC Genomics 10*, 8.

75. Bollinger, J. M., Jr., (1993) *On the Chemical Mechanism of Assembly of the Tyrosyl Radical-Dinuclear Iron Cluster Cofactor of E. coli Ribonucleotide Reductase*. PhD Thesis, Massachusetts Institute of Technology.

76. Dorenbos, R., Stein, T., Kabel, J., Bruand, C., Bolhuis, A., Bron, S., Quax, W. J., and van Dijl, J. M. (2002) Thiol-Disulfide Oxidoreductases are Essential for the Production of the Lantibiotic Sublancin 168, *J. Biol. Chem. 277*, 16682-16688.

77. Oman, T. J., Boettcher, J. M., Wang, H. A., Okalibe, X. N., and van der Donk, W. A. (2011) Sublancin is Not a Lantibiotic but an S-Linked Glycopeptide, *Nat. Chem. Biol. 7*, 78-80.

78. Molle, V., Fujita, M., Jensen, S. T., Eichenberger, P., González-Pastor, J. E., Liu, J. S., and Losick, R. (2003) The Spo0A Regulon of *Bacillus subtilis*, *Mol. Microbiol. 50*, 1683-1701.

79. Töwe, S., Leelakriangsak, M., Kobayashi, K., van Duy, N., Hecker, M., Zuber, P., and Antelmann, H. (2007) The MarR-Type Repressor MhqR (YkvE) Regulates Multiple Dioxyenases/Glyoxalases and an Azoreductase which Confer Resistance to 2-Methylhydroquinone and Catechol in *Bacillus subtilis*, *Mol. Microbiol. 66*, 40-54.

80. van Duy, N., Wolf, C., Mäder, U., Lalk, M., Langer, P., Lindequist, U., Hecker, M., and Antelmann, H. (2007) Transcriptome and Proteome Analyses in Response to 2-Methylhydroquinone and 6-Brom-2-Vinyl-Chroman-4-on Reveal Different Degradation Systems Involved in the Catabolism of Aromatic Compounds in *Bacillus subtilis*, *Proteomics 7*, 1391-1408.

81. Smits, W. K., Dubois, J. Y. F., Bron, S., van Dijl, J. M., and Kuipers, O. P. (2005) Tricksy Business: Transcriptome Analysis Reveals the Involvement of Thioredoxin A in Redox Homeostasis, Oxidative Stress, Sulfur Metabolism, and Cellular Differentiation in *Bacillus subtilis*, *J. Bacteriol. 187*, 3921-3930.

82. Härtig, E., Hartmann, A., Schätzle, M., Albertini, A. M., and Jahn, D. (2006) The *Bacillus subtilis nrdEF* Genes, Encoding a Class Ib Ribonucleotide Reductase, are Essential for Aerobic and Anaerobic Growth, *Appl. Environ. Microbiol. 72*, 5260-5265.

83. Marchler-Bauer, A., Derbyshire, M. K., Gonzales, N. R., Lu, S. N., Chitsaz, F., Geer, L. Y., Geer, R. C., He, J., Gwadz, M., Hurwitz, D. I., Lanczycki, C. J., Lu, F., Marchler, G. H., Song, J. S., Thanki, N., Wang, Z. X., Yamashita, R. A., Zhang, D. C., Zheng, C. J., and Bryant, S. H. (2015) CDD: NCBI's Conserved Domain Database, *Nucleic Acids Res. 43*, D222-D226.

84. Rabinovitch, I., Yanku, M., Yeheskel, A., Cohen, G., Borovok, I., and Aharonowitz, Y. (2010) *Staphylococcus aureus* NrdH Redoxin Is a Reductant of the Class Ib Ribonucleotide Reductase, *J. Bacteriol. 192*, 4963-4972.

85. Stehr, M., Schneider, G., Åslund, F., Holmgren, A., and Lindqvist, Y. (2001) Structural Basis for the Thioredoxin-like Activity Profile of the Glutaredoxin-like NrdH-Redoxin from *Escherichia coli*, *J. Biol. Chem. 276*, 35836-35841.

86. Brown, N. C., and Reichard, P. (1969) Ribonucleoside Diphosphate Reductase - Formation of Active and Inactive Complexes of Proteins B1 and B2, *J. Mol. Biol. 46*, 25-38.

87. Roca, I., Torrents, E., Sahlin, M., Gibert, I., and Sjöberg, B. M. (2008) NrdI Essentiality for Class Ib Ribonucleotide Reduction in *Streptococcus pyogenes*, *J. Bacteriol. 190*, 4849-4858.

Chapter 3

Characterization of the Allosteric Regulation of the Class Ib RNR of *Bacillus subtilis* and its Links to Quaternary Structure

3.1. INTRODUCTION

The purification and utilization of the endogenous reducing system, consisting of thioredoxin (TrxA), thioredoxin reductase (TrxB), and NADPH, in assays of the *Bacillus subtilis* class Ib RNR, as reported in Chapter 2, produced some of the highest CDP reduction activities ever measured for a class Ib system (Appendix 2, **Table A2.1**). An additional benefit of using the endogenous reducing system is the capability to spectrophotometrically measure the rates of reduction of any NDP by monitoring NADPH oxidation at 340 nm.[1] In this chapter, the spectrophotometric assay is utilized to study the allosteric regulation of the Mn-loaded *B. subtilis* RNR in order to gain a better understanding of this atypical class Ib system. The information gained from these experiments was applied in the design of SV-AUC studies that probed for nucleotide-based changes in quaternary structure of NrdE which, as described in Chapter 2, behaves as a monomer at physiological concentrations. While a comparison of the allosteric regulation of the Fe- and Mn-forms of *B. subtilis* RNR could be informative, low signal-to-noise and non-correlation of NADPH oxidation with CDP reduction prevented accurate rate measurements with the Fe-loaded enzyme, and given the physiological irrelevance of this cofactor in *B. subtilis*,[2] studies of this form of the enzyme were not pursued any further.

The allosteric regulation of class Ib RNRs is a relatively unexplored area in the field as previous studies primarily focused on identifying the physiologically relevant $Me(III)_2$-Y• cluster (Me = Fe or Mn) assembled in NrdF[2-8] and determining the function of the two accessory proteins, NrdI and NrdH, that were associated with class Ib operons.[9-16] The only case in which experiments were reported that examined the reduction of substrates other than CDP and utilized effectors other than ATP and dATP was a study of the Fe-loaded form of the *S. enterica* Serovar Typhimurium class Ib RNR by the Reichard group.[17] They demonstrated that the ability of the enzyme to reduce

all four NDPs to the corresponding dNDPs utilized essentially the same allosteric control of specificity as observed in nearly all previously characterized RNRs:[18] ATP/dATP maximally stimulated the reduction of the pyrimidine ribonucleotides, TTP stimulated GDP reduction, and dGTP stimulated ADP reduction. Nucleotide binding and competition experiments revealed the existence of only two allosteric binding sites per α_2 which, based on their ability to bind all three effector dNTPs and ATP, corresponded to the specificity sites for this RNR.[17] Subsequent structural characterization of *S.* Typhimurium NrdE crystals soaked with dATP (PDB 1PEU, resolution: 3.2 Å), dCTP (1PEO, 3.0 Å), and TTP (1PEQ, 2.8 Å) confirmed that the location of the specificity sites at the α_2 dimer interface was the same as in *E. coli* class Ia structures (**Figure 3.1**).[19, 20] Furthermore, the loops responsible for binding effector (Loop 1) and mediating communication between the catalytic and specificity sites (Loop 2) were positioned similarly in the two structures and were ordered in *S.* Typhimurium NrdE only when effector had been soaked into the specificity sites (**Figure 3.1**).

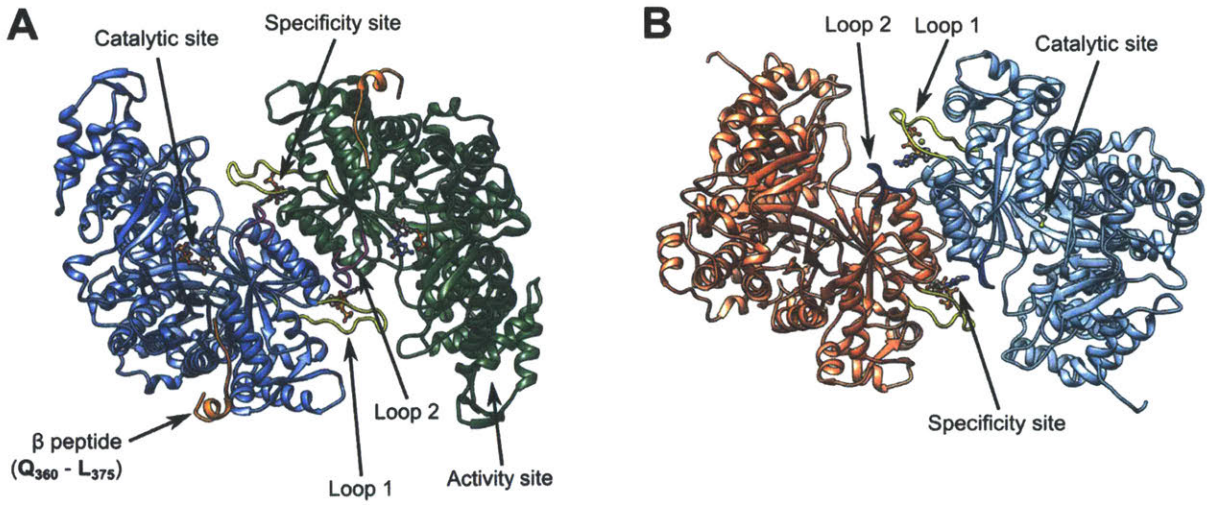


Figure 3.1. Comparison of *E. coli* Ia (**A**) and *S.* Typhimurium Ib (**B**) α_2 crystal structures soaked with nucleotides (PDB 4R1R[19] and 1PEU,[20] respectively). TTP and GDP occupy the specificity and catalytic sites in (**A**), respectively, while dATP occupies the specificity site in (**B**). Loop 1 in both panels is colored yellow. Loop 2 is magenta in (**A**) and blue in (**B**).

The similarities of the allosteric regulation of the *S.* Typhimurium class Ib RNR to other characterized systems has led to the assumption that other Ib RNRs will exhibit the same allosteric regulatory pattern as well.[18] However, the *S.* Typhimurium Ib RNR, as noted in Chapter 1, is not a prototype for this subclass given that the enzyme is expressed only under stressful conditions (Fe limitation, oxidative stress) and cannot otherwise support growth of the bacterium.[21, 22] In contrast to the *Enterobacteriaceae*, many other bacterial families, including many human pathogens, utilize a Ib RNR as their sole aerobic source for dNTPs.[23] It can be expected that the Ib RNRs from these bacteria will likely exhibit differences in their allosteric regulation so that the cell can maintain its genome integrity and the high fidelity of DNA replication and repair through controlling dNTP pool sizes and ratios via RNR activity. Therefore, a study of the allosteric regulation of a Ib RNR isolated from a bacterium that is entirely dependent on this enzyme for dNTPs under aerobic conditions will likely be more pertinent to the Ib systems in general as compared to the *S.* Typhimurium Ib RNR.

By far the most distinct characteristic of the class Ib RNRs in terms of allosteric regulation is the lack of activity inhibition by dATP, as reported for the enzymes from *S.* Typhimurium,[7] *Lactococcus lactis*,[13] *Mycobacterium tuberculosis*,[24] *Bacillus anthracis/Bacillus cereus*,[10] and *Streptococcus sanguinis*.[25] This trait is attributed to the absence on NrdE proteins of an N-terminal ATP-cone domain[26] and, therefore, the second class of allosteric binding sites that regulate the overall enzymatic activity in response to ATP (active state) and dATP (inhibited state) via formation of alternate quaternary structures.[27-30] As reported in Chapter 2, CDP reduction catalyzed by the *B. subtilis* class Ib RNR was shown to be inhibited by physiologically relevant dATP concentrations despite not having an ATP-cone domain. This unusual observation, while interesting in itself, raises the question as to whether or not the *B. subtilis* enzyme might exhibit

other allosteric regulatory properties that are distinct from the more generally accepted rules of RNR regulation.[18] As mounting evidence strongly supports $Mn(III)_2$-Y• as the physiologically relevant cofactor of class Ib RNRs,[2, 4, 5, 31] an added incentive for studying the allosteric regulation of the *B. subtilis* enzyme was that it would be the first characterization of an RNR loaded with this cofactor and, thus, should provide valuable insight for subsequent allosteric studies of Ib RNRs in general.

The allosteric regulation of class I RNRs is also directly linked with their quaternary structural dynamics;[18] therefore, a study of one of these aspects requires some knowledge about the other. One of the goals of working with the *B. subtilis* RNR is to see if it might be the system that finally yields a complete structure of the $\alpha_2\beta_2$ complex proposed to be active in NDP reduction. In Chapter 2, it was demonstrated that the subunits of the *B. subtilis* enzyme, while forming a qualitatively "tight" interaction, did not form temporally stable complexes on the timescales of SEC or SV-AUC (**Figure 2.9C** and **2.10B**). However, in those studies no substrates or effectors were included in the samples. Since RNRs are always in the presence of nucleotides *in vivo*, these molecules are likely to play an important role in stabilizing complexes of these enzymes. Therefore, a knowledge of concentrations of nucleotides *in vivo* (see Appendix 3) and some fundamentals of the allosteric regulation of the *B. subtilis* RNR should greatly help in identifying conditions under which stable complexes can be generated.

Finally, an important issue requiring further investigation is the quaternary structure of NrdE. The results of SEC and SV-AUC analysis (Chapter 2) showed that at physiological concentrations in the absence of nucleotides, NrdE exists as a monomer alone (**Figure 2.9B** and **2.10A**) and as an equilibrating system of monomers and dimers in the presence of an equivalent amount of β_2 (**Figure 2.9C** and **2.10B**). Several other class I RNRs have also been reported to

have monomeric α subunits,[8, 10, 32-35] and in most cases, dimer formation was observed in the presence of effectors, substrates, and/or β_2. In an extreme case, nearly complete dimerization of the *Pseudomonas aeruginosa* Ia α subunit was observed by GEMMA in the presence of TTP and GDP.[33] These observations suggest that one or a combination of three factors (substrate, effector, β_2) stabilize α_2, and as *B. subtilis* NrdF alone was unable to do this, it was important to establish if NrdE dimerization could be driven to completion by substrate and/or effector nucleotides. If a stable NrdE dimer can be generated, it could increase the chance of finally obtaining the elusive structure of a class I RNR active complex with the *B. subtilis* system.

The results of the spectrophotometric assays reported in this chapter show that, despite generally obeying the previously established pattern of allosteric specificity control widely observed in all RNR classes,[18] the Mn-loaded *B. subtilis* enzyme does exhibit peculiar attributes that are distinct from other prokaryotic class I RNRs. For a variety of reasons, including the low yields of His_6-tagged NrdI (Chapter 2) and the recently reported aggregation issues of the His_6-tagged human RNR in the presence of ATP,[28] tagless constructs of the *B. subtilis* Ib RNR were first prepared and tested to show that results reported here and in Chapter 2 were not an artifact of using His_6-tagged protein. The inhibition of the enzyme by dATP was confirmed with the spectrophotometric assay using both CDP and UDP as substrates. Other notable features of the allosteric regulation of the *B. subtilis* Ib RNR include: (1) positive and negative homotropic cooperative effects on activity by varying GDP and ADP concentrations, respectively, (2) stimulation of maximum GDP reduction by the combination of ATP and TTP, (3) no activity inhibition by ATP in combination with another dNTP effector, and (4) apparent product inhibition by dADP that cannot be alleviated by the presence of ATP. These results led to the examination of NrdE quaternary structure in the presence of GDP, TTP, and dATP using SV-AUC. The results

show that within the range of nucleotide concentrations tested, TTP alone or in combination with GDP did not affect the structure of NrdE whereas, counterintuitively, GDP alone appeared to cause the small amount of NrdE dimer detected to monomerize. On the other hand, dATP appeared to cause NrdE to form dimers at physiologically relevant concentrations. Interestingly, dATP also caused formation of larger oligomers when examined with 1:1 mixtures of NrdE and NrdF, suggesting that the mechanism of dATP inhibition of B. subtilis RNR involves formation of alternate inhibitory quaternary structures.

3.2. EXPERIMENTAL

3.2.1. Materials and methods. Chemicals, equipment, and routine methods were generally the same as described in Chapter 2. Nucleotides were obtained from Sigma or MP Biomedicals and used as received. BL21 (DE3) Codon Plus RPIL cells were from Agilent. The vector pTB145, encoding His_6-tagged SUMO protease, was a gift from Bradley Pentelute (Department of Chemistry, Massachusetts Institute of Technology). His_6-tagged NrdE (600 nmol min^{-1} mg^{-1} α_2), NrdF, NrdI, TrxA (200 nmol min^{-1} mg^{-1}), and TrxB (18 μmol min^{-1} mg^{-1}) were prepared as described in Chapter 2. His_6-NrdF was reconstituted with $Mn(III)_2$-Y• and the holo-protein (~1300 nmol min^{-1} mg^{-1} β_2) purified as described in Chapter 2. Unless noted otherwise, NrdE concentrations are reported relative to monomer and NrdF relative to dimer. Given the results described in Chapter 4, it should also be noted that the NrdE samples used in these experiments were mixtures of protein with and without an equivalent of tightly bound dAMP.

3.2.2. Expression and purification of SUMO protease. The vector pTB145 was transformed into BL21 (DE3) Codon Plus RPIL cells and selected for on an LB agar plate containing 100 μg mL^{-1} Amp. A single colony was grown to saturation at 37 °C with 200 rpm shaking overnight (~14 h) in 100 mL LB supplemented with 100 μg mL^{-1} Amp, 50 μg mL^{-1} Cm, and 75 μg mL^{-1}

streptomycin (Strep). Large scale growth was initiated by adding 5 mL of saturated starter culture to each of four 1 L volumes of Amp/Cm/Strep-supplemented LB medium and incubating at 37 °C with 200 rpm shaking. When the OD_{600} reached 0.6, the cultures were removed to a cold room (4 °C) while the shaker was cooled to 30 °C (~10 min). IPTG was then added to each liter of culture to a final concentration of 0.5 mM before replacing the flasks in the shaker and incubating the cultures at 30 °C for 4 h. Cells were harvested by centrifugation (3500 x g, 15 min, 4 °C), frozen (liq N_2), and stored at -80 °C. A total of 9.5 g cell paste (2.4 g L^{-1}) was recovered.

All steps of the purification were completed at 4 °C. The cell pellet was placed on ice and thawed for 1 h prior to resuspension in 48 mL SUMO buffer (25 mM Tris, pH 8.0, 5% (w/v) glycerol) supplemented with 100 mM NaCl, 20 mM imidazole, 10 mM BME, 2.6 U mL^{-1} DNase I, 1.3 U mL^{-1} RNase A, and ~1 mg lysozyme (from chicken egg white, Sigma). After homogenization, cells were lysed by one passage through a chilled French pressure cell at 14000 psi and the cell debris pelleted by centrifugation (25000 x g, 30 min). The extract was loaded onto a Ni-NTA Fast flow column (4.7 x 1.0 cm, ~4 mL bed volume) equilibrated in SUMO buffer supplemented with 100 mM NaCl, 20 mM imidazole, and 10 mM BME. The column was washed with equilibration buffer until the A_{280} stabilized at 0.16 a.u. (~400 mL) and then eluted using SUMO buffer supplemented with 100 mM NaCl, 400 mM imidazole, and 10 mM BME. Protein-containing fractions were pooled, diluted 5-fold with SUMO buffer supplemented with 100 mM NaCl and 0.5 mM DTT, and loaded onto a Q-Sepharose column (3.5 x 2.5 cm, ~17 mL bed volume) equilibrated in diluent. The column was washed with three CVs of equilibration buffer and the flow through collected and concentrated. The concentrated protein (40 mL) was loaded into two 10000 MWCO, 12 – 30 mL capacity Slide-a-Lyzer cassettes (Thermo Fisher Scientific) and dialyzed against 4 L of SUMO buffer supplemented with 250 mM NaCl and 0.5 mM DTT for

18 h with one buffer exchange. A total of 3.6 mg SUMO protease g^{-1} cell paste was recovered (> 97% pure by SDS-PAGE). Protein concentration was estimated using $\varepsilon_{280} = 29910\ M^{-1}\ cm^{-1}$ (Dr. Amy Rabideau, Pentelute group, personal communication).

Table 3.1. Primers used to clone *nrdE*, *nrdF*, and *nrdI* into pE-SUMO.

Gene	Primer direction	Sequence[a]
nrdE	Forward	5′-AGGAGGTCTCAAGGT**ATG**TCACAAAATCAAGTGCCAAAATGG-3′
	Reverse	5′-TCTCGAGC**TCA**AACAACACAAGAAAGGCAGC-3′
nrdI	Forward	5′-AGGTGGTCTCAAGGT**ATG**GTACAAATCATATTTGATTCG-3′
	Reverse	5′-CATGCTCGAG**TTA**TTTAACTGGATCCATTTTGGC-3′
nrdF	Primer#1	5′-AGAGCTCGGTCTCCTTGCTCAGGAAATTTATAATAAACAG-3′
	Primer#2	5′-ATATATCTCGAG**TTA**TATCTGCTCTTTTTCATCTTCAAAA-3′
	Primer#3	5′-ATATATGGTCTCCAGGT**ATG**ACAAAAATTTATGACGCAGC-3′
	Primer#4	5′-CCTGAGCAAGGAGACCGACATACACGCCGTGAATCGC-3′

[a] Restriction sites are underlined, and start and stop codons are indicated in boldface font.

3.2.3. Cloning the *B. subtilis* RNR genes into pE-SUMO. All cloning was completed using the touchdown PCR technique.[36] *nrdE* and *nrdI* were amplified from *B. subtilis* JH624 genomic DNA using the primers listed in **Table 3.1** to introduce 5′ *BsaI* and 3′ *SacI* (*nrdE*) or *XhoI* (*nrdI*) restriction sites. The amplicons were digested and ligated into similarly treated pE-SUMO (LifeSensors, Kan^r) using T4 DNA ligase. *nrdF* has an internal *BsaI* restriction site, and attempts to remove it with site-directed mutagenesis failed, thus the gene was cloned in two steps. In the first step, the last 375 bp of *nrdF* was amplified using Primer#1 and Primer#2 (**Table 3.1**) to introduce 5′ *SacI* and 3′ *XhoI* restriction sites. The digested amplicon was ligated into pE-SUMO to generate the vector pE-SUMO-$nrdF_{Cterm}$. In the second step, the first 631 bp of *nrdF* was amplified using Primer#3 and Primer#4 (**Table 3.1**) to introduce *BsaI* sites at the 5′ and 3′ ends of the amplicon. The primers were designed such that mutually exclusive sticky ends were generated at the amplicon ends so that it could be ligated into pE-SUMO-$nrdF_{Cterm}$ in only one orientation.

The resulting amplicon was digested with BsaI and ligated into a sample of similarly digested pE-SUMO-*nrdF*$_{Cterm}$, thus generating an intact copy of *nrdF* in pE-SUMO. Sequencing of the vectors (MIT Biopolymers Laboratory) confirmed the successful cloning of all three genes.

3.2.4. Expression and purification of the His$_6$-Smt3-tagged *B. subtilis* RNR proteins. His$_6$-Smt3-NrdE, -NrdF, and -NrdI were overproduced following the procedure described in Chapter 2 with the modification that the vectors were transformed into standard BL21 (DE3) cells. Typical yields (in g cell paste L^{-1}) were 2.8 and 1.7 for His$_6$-Smt3-NrdE and -NrdF, respectively. His$_6$-Smt3-NrdI was grown only once with a yield of 2.6 g L^{-1}.

His$_6$-Smt3-NrdE and -NrdF were purified following the procedures described in Chapter 2 with no modifications. The purification of His$_6$-Smt3-NrdI was the same as described in Chapter 2 through the IMAC step. Protein-containing fractions from the Ni-NTA column were pooled, diluted 5-fold in Tris buffer (50 mM Tris, pH 8.5, 5% (w/v) glycerol) supplemented with 100 mM NaCl, and loaded onto a Q-Sepharose column (5.7 x 2.5 cm, ~28 mL bed volume) equilibrated in diluent. The column was washed with 2 CVs equilibration buffer before development with a 300 mL linear gradient from 100 – 500 mM NaCl in Tris buffer; His$_6$-Smt3-NrdI eluted at ~230 mM NaCl. Protein-containing fractions were pooled, concentrated, frozen, and stored at -80 °C. All proteins were carried on to the tag removal step without an intervening buffer exchange step; thus, the buffer composition of the purified tagged protein stocks was 50 mM Tris, pH 7.6, ~350 mM NaCl, 5% (w/v) glycerol, and, in the case of His$_6$-Smt3-NrdE, 10 mM DTT. Concentrations of His$_6$-Smt3-NrdE and His$_6$-Smt3-NrdF were estimated using ε_{280} = 80600 and 56300 M^{-1} cm^{-1}, respectively, while the concentration of His$_6$-Smt3-NrdI was estimated using the bicinchoninic acid (BCA) assay using BSA as a standard.[37, 38]

3.2.5. Removal of the His_6-Smt3 tag to produce tagless *B. subtilis* RNR subunits. All steps for tag removal and purification of the tagless proteins were conducted at 4 °C. The His_6-Smt3 tag was cleaved from NrdE and NrdF using a 500:1 mole ratio of tagged protein:SUMO protease. A typical reaction mixture contained 100 – 300 μM His_6-Smt3-tagged protein in 1 – 1.5 mL. Samples of His_6-Smt3-NrdF were supplemented with DTT to 5 mM prior to the addition of SUMO protease. The samples were then incubated on a nutating platform for 2 h before being frozen and stored at -80 °C. At a later date, the digest products were thawed and loaded onto a Ni-NTA column (4.0 x 1.5 cm, 7 mL) equilibrated in 50 mM sodium phosphate, pH 7.6, 150 mM NaCl, 5% (w/v) glycerol and either 40 mM imidazole and 10 mM BME (NrdE) or 20 mM imidazole (NrdF). The column was washed with 9 CVs of equilibration buffer and twenty 3 mL fractions were collected. Protein-containing fractions were pooled, concentrated, and exchanged into storage buffer (*NrdE*: 50 mM sodium phosphate, pH 7.6, 150 mM NaCl, 5% (w/v) glycerol, 10 mM DTT; *NrdF*: 50 mM HEPES, pH 7.6, 5% (w/v) glycerol) by gel filtration on a Sephadex G25 column (28 x 1.5 cm, ~49 mL). Average yields, relative to the starting amounts of His_6-Smt3-tagged protein, of 90 ± 10% tagless NrdE and 89 ± 13% tagless NrdF were recovered (> 97% pure as judged by SDS-PAGE). Protein concentrations were estimated using the same ε_{280} reported for the His_6-tagged proteins in Chapter 2 (79100 M^{-1} cm^{-1} and 54800 M^{-1} cm^{-1} for NrdE and NrdF, respectively).

Tag removal from NrdI was performed using one 1.2 mL aliquot of ~550 μM His_6-Smt3-NrdI per preparation. The sample was supplemented sequentially with 5 mM DTT and 20 μM SUMO protease (~25 mol His_6-Smt3-NrdI per mol protease), bringing the total volume of the aliquot to ~1.5 mL. The digestion incubated for 24 – 48 h on a nutating platform at 4 °C before being frozen and stored at -80 °C. At a later date, the sample was thawed and all of it was loaded onto a Ni-NTA column (4.0 x 1.5 cm, 7 mL) equilibrated in 50 mM sodium phosphate, pH 7.6,

150 mM NaCl, 20 mM imidazole, 5% (w/v) glycerol. The column was washed with 7 CVs equilibration buffer and twenty 2.5 mL fractions were collected. Tagless NrdI was found to wash completely out of the column in fractions 1 – 5 (12.5 mL total), and it was assumed that complete recovery was achieved relative to the starting amount of tagged protein. These fractions were pooled and transferred to a 50 mL beaker. Solid FMN (10 mg, Sigma Aldrich, 73 – 79%) and 7.5 mL 50 mM sodium phosphate (pH 7.6) were added to the contents of the beaker to bring the total volume to 20 mL, yielding final concentrations of 50 mM sodium phosphate, pH 7.6, 94 mM NaCl, 12.5 mM imidazole, 3% (w/v) glycerol, 1 mM FMN, and ~35 μM NrdI. The beaker was wrapped in aluminum foil and the contents stirred at 4 °C for 4 – 6 h. The solution was then concentrated to ~1 – 1.5 mL using an Amicon stirred cell pressurized with N_2 and outfitted with a 10000 MWCO regenerated cellulose membrane. The entirety of the concentrate was loaded onto and eluted from a Sephadex G25 column (28 x 1.5 cm, 49 mL) equilibrated in 50 mM HEPES, pH 7.6, 5% (w/v) glycerol. Protein-containing fractions were pooled and concentrated to ~1 mL, giving a final NrdI concentration of 500 – 600 μM. This procedure gave an average NrdI recovery of 84 ± 13% (> 97% pure by SDS-PAGE) relative to the starting amount of His_6-Smt3-NrdI.

3.2.6. Determination of NrdI FMN loading and ε_{449} for protein-bound flavin. The procedure described by Mayhew and Massey[39] was used to determine the FMN load and ε_{449} for NrdI. In order to estimate the flavin load, the protein pellets recovered from the denaturation procedure were dissolved in 200 μL Edelhoch buffer[40, 41] and the A_{280} measured. The protein concentration of denatured NrdI (one W and three Y) was determined using ε_{280} = 9530 M^{-1} cm^{-1}, which was calculated according to eq 2.1 (**Section 2.2.8**).[40, 41] The ε_{449} for FMN bound to NrdI was calculated according to equation 3.1.

$$\frac{A_{445\,free}}{\varepsilon_{445\,free} \times \ell} = c_{FMN} = \frac{A_{449\,bound}}{\varepsilon_{449\,bound} \times \ell} \tag{3.1}$$

$A_{445free}$ and $\varepsilon_{445free}$ are the absorption and extinction coefficient at 445 nm for free FMN, $A_{449bound}$ and $\varepsilon_{449bound}$ are the absorption and extinction coefficient at 449 nm for NrdI-bound FMN, c_{FMN} is the concentration of FMN, and ℓ is the path length of the cell (1 cm). The ε_{449} determined from this procedure (11900 M^{-1} cm^{-1}) was used to estimate NrdI concentrations.

3.2.7. Reconstitution and purification of holo-Mn(III)$_2$-Y• tagless NrdF. Holo-Mn(III)$_2$-Y• NrdF (1300 nmol min^{-1} mg^{-1} β_2) was reconstituted, purified, and characterized following the procedures described in Chapter 2.

3.2.8. Measuring RNR activity using the spectrophotometric assay. Routine assays (500 μL) monitoring the consumption of NADPH (ε_{340} = 6220 M^{-1} cm^{-1}) were conducted at 37 °C and contained 50 mM HEPES, pH 7.6, 15 mM $MgCl_2$, 1 mM EDTA, 40 μM TrxA, 0.4 μM TrxB, 0.2 mM NADPH, 0.5 μM Mn-β_2 (0.8 – 1.0 Y•/β_2), 1 μM NrdE, 3 mM ATP, and 1 mM CDP. Assay solutions were thermally equilibrated for 2 min prior to adding either substrate or Mn-β_2 to initiate turnover.

3.2.9. Correlating NADPH consumption with [5-^{3}H]-dCDP production. Assay mixtures were prepared as described above using [5-^{3}H]-CDP (940 cpm $nmol^{-1}$). The reaction was monitored continuously at 340 nm and 30 μL aliquots were withdrawn every minute over 10 min and quenched with 30 μL 2% (v/v) $HClO_4$. After pH neutralization (26 μL 0.4 M KOH), the samples were frozen, thawed, and worked up using the method of Steeper and Steuart.[42]

3.2.10. Measuring steady-state kinetics of *B. subtilis* RNR with different substrate/effector pairs. Measurement of the steady-state kinetic parameters of *B. subtilis* RNR followed the same procedure as described above except that the desired substrate/effector pair was used in place of CDP/ATP. To measure kinetic parameters for effectors, the appropriate substrate was present at 1 mM while the concentration of the effector varied: ATP (25 μM – 4 mM); dGTP (250 nM – 4

mM); TTP (100 nM – 4 mM). To measure the kinetic parameters for substrates, the appropriate effector was present at ~10x its apparent K_m while the substrate concentration was varied: ADP (25 μM – 8 mM); CDP (10 μM – 2 mM); GDP (10 μM – 1 mM); UDP (25 μM – 4 mM). Data were fit to the Michaelis-Menten and Hill equations[43] (equation 3.2) in the graphing program Igor Pro (Wavemetrics, Lake Oswego, OR) using non-linear regression. For fits to the Michaelis-Menten equation, which assumes multiple ligand binding events are non-cooperative, n_H in eq 3.2 was fixed at 1.

$$v = \frac{V_{max}[nucleotide]^{n_H}}{K_m^{n_H} + [nucleotide]^{n_H}} \tag{3.2}$$

The statistical significance of better fits to the Hill equation versus the Michaelis-Menten equation were evaluated using an extra sum-of-squares F test.[44] A threshold of $p \leq 0.01$ was chosen as the cutoff for statistical significance.

3.2.11. Characterization of *B. subtilis* RNR subunit quaternary structure(s) in the presence of nucleotides by SV-AUC.

3.2.11A. Sample preparation. Experiments examining the sedimentation of NrdE in the absence of nucleotides and 1:1 mixtures of His_6-α_2:$Mn(III)_2$-Y• His_6-β_2 in the presence of dATP were run in Buffer 1 (50 mM Tris, pH 7.6, 5% (w/v) glycerol, 150 mM NaCl, 15 mM $MgCl_2$, 1 mM DTT). All other experiments were run in Buffer 2 (50 mM HEPES, pH 7.6, 150 mM NaCl, 15 mM $MgCl_2$, and 1 mM TCEP for experiments with NrdE). Proteins were buffer exchanged by eluting samples of NrdE (400 μL, 240 μM) or $Mn(III)_2$-Y• β_2 (790 μL, 30 μM) from a Sephadex G25 column (28 x 1.5 cm, ~49 mL) equilibrated with > 20 CVs of the desired buffer. Nucleotide stocks were prepared fresh in sample buffer and the concentrations determined by UV-vis.[45] Samples were prepared in a total volume of 500 μL by mixing the proper amounts of concentrated protein stock, concentrated nucleotide stock, and buffer to give the desired final concentrations.

For experiments run in the absence of nucleotides, NrdE or NrdF were present at concentrations (NrdE: 2 – 14 μM, β_2: 0.5 – 10 μM) that spanned A_{280} = 0.05 – 1.5 a.u. For experiments involving sedimentation of individual subunits in the presence of nucleotides, NrdE and holo-Mn(III)$_2$-Y• β_2 concentrations were 2 – 2.5 μM and 1 μM, respectively, while TTP and GDP concentrations spanned 0.1x to 10x their apparent K_m values and dATP concentrations spanned 0.5 to 50 μM. For experiments involving 1:1 mixtures of His$_6$-α_2:Mn(III)$_2$-Y• His$_6$-β_2, a 1.2 mL sample consisting of 27 μM of each subunit was exchanged into Buffer 1 supplemented with 40 μM dATP using three rounds of dilution-concentration with an Amicon Ultra-15 YM30 centrifugal filter (30000 MWCO, regenerated cellulose membranes). Samples were prepared by diluting the appropriate amount of concentrated 1:1 α_2:β_2 mixture with dATP supplemented Buffer 1 to yield the desired final protein concentration (0.4 – 3.5 μM) in a volume of 500 μL.

3.2.11B. Instrument setup and experiment execution. SV-AUC experiments were performed using the Beckman XL-I analytical ultracentrifuge at the MIT Biophysical Instrumentation Facility. Cells assembled with Epon charcoal double sector centerpieces (1.2 cm) and either quartz (A_{280} optical monitoring in the absence of nucleotides) or sapphire (interference optical monitoring in the presence of nucleotides) windows were carefully loaded with reference buffer and sample so that as close to matching menisci as possible were obtained (~440 μL in each sector). Laser delay setup and radial calibration of the samples were completed with the rotor spinning at 3000 rpm; no sedimentation was detected during these procedures. The rotor was then brought to rest and the system thermally equilibrated to 20 °C for 1.5 – 2.0 h before sedimenting the samples over 17 – 19 h at 20 °C with an angular velocity of 42000 rpm. Absorbance scans were collected every 1.2 min and interference scans every 1.5 min using the ProteomeLab XL-I graphical user interface (version 4.5b).

3.2.11C. Data analysis. To accelerate the data analysis, all scans of a set were first loaded into Sedfit and manually inspected to determine the subset that covered the entire sedimentation process only; typically, these were the first 90 – 120 scans of samples monitored by A_{280} or the first 250 scans of samples monitored by interference. These scans were then reloaded, time-stamp corrected,[46] and fit to the c(s) model with a resolution of 100 – 130 and a regularization factor of 0.68 using alternating rounds of Simplex and Marquardt-Levenberg optimization algorithms until the fits converged. With interference data sets, the recently incorporated buffer sedimentation correction module[47] was implemented to account for any menisci and/or buffer mismatches between the reference and sample sector (see discussion in Appendix 3 for further details). For the samples of 1:1 mixtures of α:β in the presence of dATP, the presence of extremely large oligomers necessitated the use of the ls-g*(s) model in order to obtain interpretable fits. Solvent parameters (ρ, η) for corrections to standard state are as follows: *Buffer 1*: ρ = 1.018 g mL^{-1}, η = 1.164 cP; *Buffer 2*: ρ = 1.009 g mL^{-1}, η = 1.054 cP. These values were calculated and temperature corrected with the program Sednterp.[48] The partial specific volumes and molecular weights of the tagless proteins are listed in **Table 3.2**. Weight averaged $s_{20,w}$ values and their associated error were determined as described in Chapter 2 for all results except those for the sedimentation of 1:1 mixtures of His_6-NrdE:His_6-NrdF in the presence of dATP. The error in the $s_{<Wt>}[20,w]$ values was determined from 5000 rounds of Monte Carlo simulation.

Table 3.2. Partial specific volumes and molecular weights of tagless RNR components.

Subunit/Complex	α	β	α_2	β_2	$\alpha_2\beta_2$
M_W (kDa)	81	38	161	77	238
$\bar{v}$ (cm^3 g^{-1})	0.7324	0.7375	0.7324	0.7375	0.7340

3.2.11D. Hydrodynamic modeling of the tagless *B. subtilis* RNR subunits. Predictions of the hydrodynamic parameters of the tagless *B. subtilis* Ib RNR subunits were carried out with the same structural models as described in Chapter 2 using the Mws listed in **Table 3.2**.

3.3. RESULTS

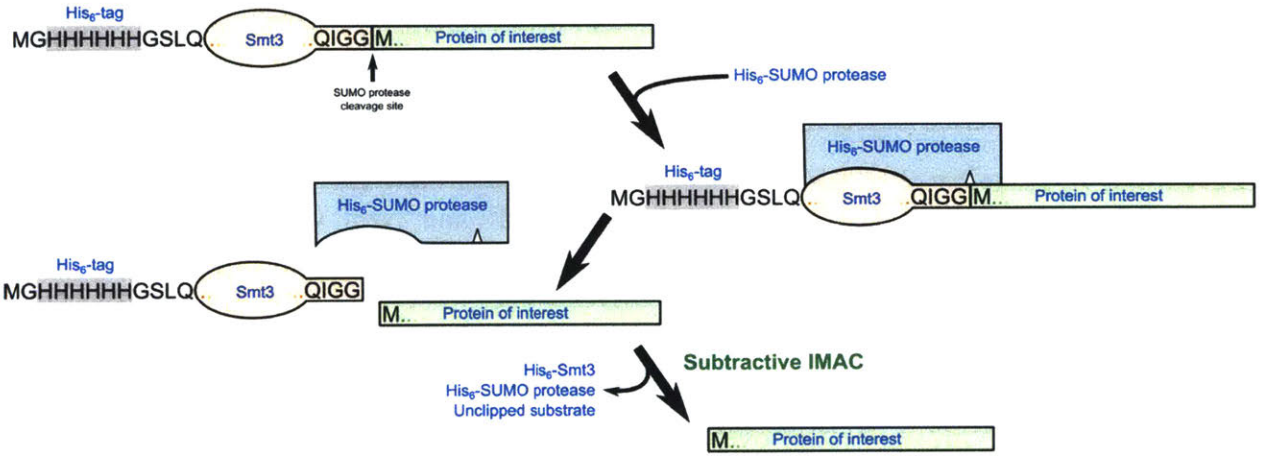


Figure 3.2. Illustration of a His_6-Smt3 tagged protein of interest and the tag removal process using SUMO protease. The Smt3 domain (12.4 kDa) has an N-terminal His_6-affinity tag with a four-residue linker. The C-terminal end of the domain is fused directly to the N-terminus of the protein of interest. SUMO protease recognizes the fold of the Smt3 domain rather than a specific cleavage sequence and cleaves the peptide bond after the Gly-Gly sequence separating the domain from the protein of interest. This cleavage is traceless, thus the natural N-terminal residue of the protein of interest can be exposed (a protein of interest with its N-terminal Met residue intact is illustrated in the figure). The *B. subtilis* RNR proteins have all been cloned to include their N-terminal Met residue in the constructs reported in this chapter.

3.3.1. Fusion of a cleavable His_6-Smt3 tag allows for relatively facile production of tagless *B. subtilis* RNR proteins. In order to facilitate rapid and facile purification of tagless constructs of the *B. subtilis* Ib RNR proteins, the *nrdE*, *nrdF*, and *nrdI* genes were cloned into the vector pE-SUMO. Expression of the proteins from this vector fuses a His_6-tagged Smt3 protein (a ubiquitin-like protein from *Saccharomyces cerevisiae*) to the N-terminus (**Figure 3.2**), which has been shown in a number of cases to improve protein expression and solubility.[49, 50] After purification of the proteins by standard IMAC methods, the tag can be removed using N-terminally

His_6-tagged SUMO protease (the catalytically active domain (residues 403-621) of the *S. cerevisiae* Ulp1 protein[51]), which recognizes the Smt3 tertiary structure instead of a specific amino acid sequence (**Figure 3.2**). The cleavage reaction is traceless, thus revealing the natural N-terminus of the desired protein. The *B. subtilis* RNR genes were cloned with their start codon included, thus the resulting tagless recombinant proteins have an N-terminal Met residue revealed after cleavage of the tag. A final subtractive IMAC step removes the His_6-Smt3 tag, His_6-SUMO protease, and unclipped substrate to yield pure tagless protein (**Figure 3.2**).

Overexpression and purification of His_6-Smt3-NrdE and -NrdF was completed following the procedures described in Chapter 2 with minimal changes. This was true also for His_6-Smt3-NrdI up through the IMAC step. As a consequence of fusing a His_6-Smt3 tag to its N-terminus, the pI of His_6-Smt3-NrdI was 6.12,[37] over two units lower as compared to His_6-NrdI (8.90). This difference necessitated the use of anion rather than cation exchange chromatography in the final step of purifying His_6-Smt3-NrdI. Yields (in mg protein g^{-1} cell paste) of 13.4, 11.0, and 6.2 were obtained for His_6-Smt3-NrdE, -NrdF, and -NrdI, respectively (all > 97% pure as judged by SDS-PAGE). The amount of His_6-Smt3-NrdI recovered was 35-fold higher in comparison to His_6-NrdI (Chapter 2), but the protein was found to be substoichiometrically loaded with FMN (35%).

Removal of the His_6-Smt3 tag from NrdE and NrdF was robust and required very little His_6-SUMO protease (1 mol protease per 500 mol His_6-Smt3-tagged substrate). Incubation at 4 °C for 2 h was sufficient to cleave ≥ 95% of the total protein, and nearly quantitative recoveries of the tagless product were achieved after subtractive IMAC (**Figure 3.3**). In contrast, only 30% of His_6-Smt3-NrdI was cleaved after 2 h at 4 °C, which correlates very well with the observed substoichiometric FMN loading of this protein. These observations suggest His_6-Smt3-NrdI lacking FMN is more inert to cleavage, possibly due to sequestration of the Smt3 domain from

recognition by SUMO protease. Using larger amounts of SUMO protease and incubating the digestion reactions for up to 48 h improved the yield of tagless NrdI (≥ 80%, **Figure 3.4A**) without any apparent protein precipitation.

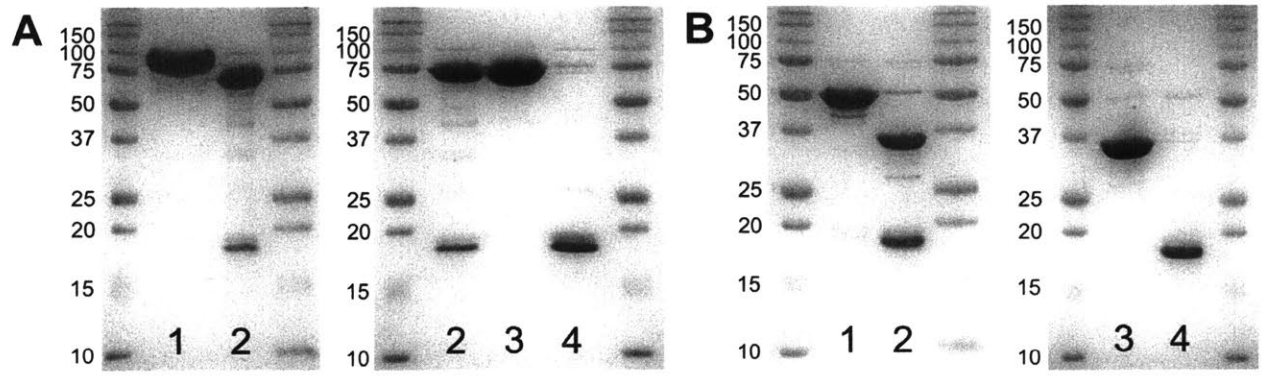


Figure 3.3. Preparation of (**A**) tagless NrdE (80.7 kDa) and (**B**) tagless NrdF (38.4 kDa). Shown are representative SDS-PAGE gels (12.5% (w/v)) illustrating the His_6-Smt3 tag removal procedure for both proteins. *Lane 1*: His_6-Smt3-tagged substrate, *lane 2*: digest product after 2 h incubation with 500:1 mol ratio of substrate:SUMO protease at 4 °C, *lane 3*: pure tagless protein after subtractive IMAC to remove the tag, protease, and unclipped substrate, and *lane 4*: His_6-tagged components removed by subtractive IMAC. The ~18 kDa species in *lanes 2* and *4* is His_6-Smt3. His_6-SUMO protease (27 kDa) is faintly visible in *lane 4* in panel (**A**) above the 25 kDa molecular weight standard.

One of the goals of moving *nrdI* into the pE-SUMO vector was to increase the amount of FMN-loaded protein that was recovered per preparation. Therefore, a reconstitution procedure was developed in order to increase the FMN content of NrdI after digestion. Precedent in the literature indicated that phosphate ions exhibit cooperative effects on the insertion of FMN into *Desulfovibrio vulgaris* flavodoxin.[52] Since NrdIs and flavodoxins are structurally similar, it seemed feasible that a similar effect could aid in the insertion of FMN into NrdI. Therefore, after removal of His_6-Smt3, SUMO protease, and unclipped substrate by subtractive IMAC, NrdI was incubated with 1 mM FMN in 50 mM sodium phosphate buffer for 4 h prior to removal of unbound flavin cofactor by gel filtration. Quantitation of the FMN load of the recovered protein revealed a full equivalent of cofactor was bound, confirming the utility of the procedure in loading *B. subtilis*

NrdI. With optimized digestion conditions and FMN reconstitution, nearly quantitative recovery of fully loaded NrdI (**Figure 3.4B**) was achieved. This protein was associated with an A_{280}:A_{449} of 2.5, which is a useful diagnostic for the FMN load of NrdI.

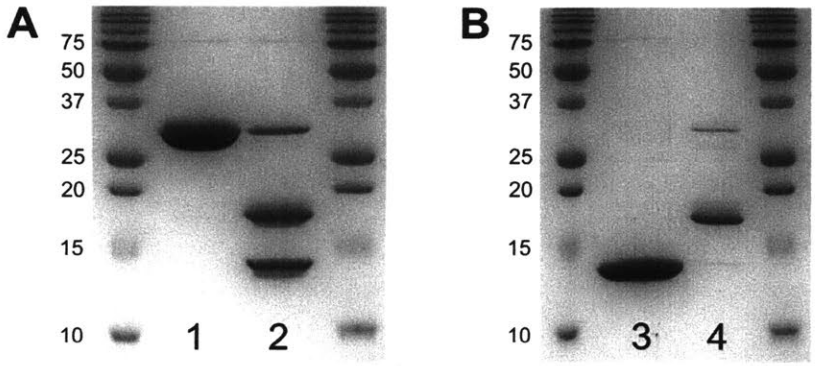


Figure 3.4. Preparation of tagless NrdI. Shown are representative SDS-PAGE gels (17% (w/v)) of protein from each stage of the procedure. (**A**) Incubation of His_6-Smt3-NrdI (*lane 1*) at 4 °C for 24 h with a 25:1 mol ratio of substrate:SUMO protease resulted in > 90% cleavage of the tag from NrdI (*lane 2*). (**B**) Pure tagless NrdI after subtractive IMAC and FMN reconstitution (*lane 3*) and the His_6-tagged components retained on the Ni-NTA column during the subtractive IMAC step (*lane 4*).

3.3.2. Reconstitution of tagless $Mn(III)_2$-Y• NrdF is similar to the His_6-tagged counterpart.

NrdI and NrdF behaved similarly to their His_6-tagged counterparts during the reconstitution of the $Mn(III)_2$-Y• cofactor and subsequent purification of holo-β_2 using MonoQ anion exchange chromatography (Chapter 2). For example, chromatograms of reconstituted tagless protein were identical to that shown in **Figure 2.3B**, including the concentration of NaCl at which the apo-/mis-metallated (~280 mM) and holo-protein (~310 mM) eluted. Yields of holo-$Mn(III)_2$-Y• NrdF (0.83 ± 0.03 Y•/β_2) were similar to the His_6-tagged counterpart (average = 33%, range = 27 – 40%). Finally, the CDP reduction activity using the radioactive assay with NrdE under routine conditions demonstrated comparable specific activities of Mn-loaded NrdF and His_6-NrdF (1300 ± 50 nmol min^{-1} mg^{-1} β_2 versus 1480 ± 130 nmol min^{-1} mg^{-1} β_2, respectively).

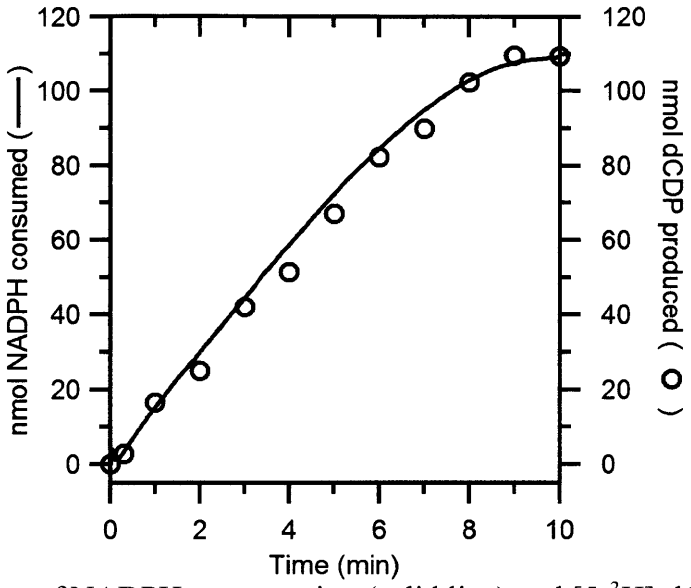


Figure 3.5. Correlation of NADPH consumption (solid line) and [5-^{3}H]-dCDP production (open circles) in activity assays of Mn-loaded *B. subtilis* Ib RNR. The assay was conducted at 37 °C in a total volume of 500 μL. The composition of the assay was a 1:1 ratio of 0.5 μM subunits, 3 mM ATP, 1 mM [5-^{3}H]-CDP (~940 cpm nmol^{-1}), and the endogenous reducing system.

3.3.3. NADPH oxidation and [5-^{3}H]-dCDP formation by Mn-loaded *B. subtilis* RNR are correlated. As mentioned in Chapter 2, class I RNR activities have typically been assessed using an excess (2 – 5x) of one subunit over the other. However, it was shown by the radioactivity-based assay that a 1:1 subunit ratio of the *B. subtilis* Ib RNR at physiologically relevant concentrations (0.5 μM) yielded > 80% the maximal activity (**Table 2.5**) observed with higher α_2:β_2 ratios. Such conditions were desirable for measuring the steady-state kinetics of the enzyme, but it was unclear whether or not the spectrophotometric assay would yield accurate data under these conditions given the complexity of the coupling reactions. Therefore, an experiment was carried out to determine if NADPH consumption and dCDP generation were correlated. The former was monitored by A_{340} and the latter by [5-^{3}H]-dCDP formation.[42] The data, as shown in **Figure 3.5**, revealed excellent correlation between nmol NADPH consumed and nmol dCDP formed. The reaction slows and then ceases after the total amount of NADPH (~100 nmol) has been oxidized, and the specific activity calculated from the linear portions of the curves (460 U

mg^{-1} $\alpha_2\beta_2$ by spectrophotometric and 470 U mg^{-1} $\alpha_2\beta_2$ by radioactive) agrees reasonably well. The spectrophotometric assay, therefore, was suitable for measuring Mn-loaded *B. subtilis* RNR activity.

3.3.4. The catalytic efficiency of *B. subtilis* RNR is highest when bound with dNTP effectors. Evidence accumulated from many studies of different classes of RNRs indicates there is apparent universality in the allosteric control of substrate specificity.[18] However, similar studies with the Mn-loaded form of a class Ib RNR from a bacterium that uses the enzyme as its sole aerobic source of dNTPs have not been reported. Therefore, an examination of the steady-state kinetics of the *B. subtilis* Ib RNR with the canonical substrate/effector pairs (ADP/dGTP, CDP/(d)ATP, GDP/TTP, and UDP/(d)ATP) was undertaken to further understand this unique enzyme as well as complete the knowledgebase of the allosteric regulation of RNRs in general. The steady-state kinetics were first studied in the presence of the effectors ATP, dGTP, and TTP to check for any abnormalities that might be induced by these nucleotides, given the unusual inhibitory effect of dATP on activity observed previously (Chapter 2). For this reason, a large range of effector concentrations (sub-micromolar to millimolar) were tested. A substrate concentration of 1 mM was chosen to ensure complete saturation of RNR, as it was expected that the affinity of the enzyme for these nucleotides was going to be much tighter (on the order of 5- to 10-fold). Furthermore, RNRs do not exhibit product inhibition with the exception of one case,[53] thus the buildup of product during assays of the *B. subtilis* enzyme was expected to have no effects on activity.

The results of the steady-state analysis of the *B. subtilis* RNR with varying effector nucleotide concentrations are shown in **Table 3.3**. The data were additionally fit to the Hill equation to check for cooperative effects on activity, which could indicate that the effectors were capable of causing NrdE to dimerize. However, these fits were statistically no better than those to

the Michaelis-Menten equation; thus, the latter model was accepted as the best. The steady-state parameters measured for *B. subtilis* RNR, when compared to similar kinetic data (Appendix 3, **Table A3.1**) and measured K_ds (**Table A3.2** and **A3.3**) for the *E. coli* Ia and other class I RNRs, reveals a few common characteristics. Turnover numbers (k_{cat}) of *B. subtilis* RNR fall in the range of 1 – 2 s^{-1}, lower but still comparable to what has been reported for the *E. coli* class Ia RNR (2 – 14 s^{-1}).[54] The apparent K_m values show that *B. subtilis* RNR bound ATP ~190-fold more weakly than dNTPs, similar to the trend in the K_ds measured for the *E. coli* RNR (**Table A3.2**), the apparent K_ms with all effectors for the T4 phage Ia (**Table A3.1**) and Fe-loaded *S.* Typhimurium Ib RNRs (**Table A3.3**), and K_ms with ATP and dATP for other Fe-loaded Ib RNRs (**Table A3.1**). The *E. coli*, *S.* Typhimurium, and T4 enzymes showed slightly higher affinities for the purine dNTP effectors relative to TTP, whereas nearly identical affinities for dGTP and TTP were observed with *B. subtilis* RNR. Lastly, the k_{cat} K_m^{-1} values calculated for *B. subtilis* RNR indicate that under conditions of substrate saturation, the dNTPs are much more potent effectors of RNR activity than ATP. A possible reason for this result, at least in the class Ib RNRs, is that a highly-conserved Tyr residue in the specificity site (Y_{236} in *B. subtilis* NrdE) occludes the space where the 2′-OH of ATP would sit, thus reducing the affinity of the protein for this nucleotide.[20]

Since assays were run with a 1:1 ratio of α:β, V_{max} can be calculated in terms of either subunit (α_2 or β_2) or of the active complex ($\alpha_2\beta_2$). Under the conditions of the assay reported here, the differences in maximal activity are due to the different molecular weights of the components (β_2 = 81 kDa, α_2 = 166 kDa, $\alpha_2\beta_2$ = 247 kDa). **Table 3.3** and **3.4** list V_{max} relative to both subunits and $\alpha_2\beta_2$ to facilitate comparison of the *B. subtilis* enzyme with other class Ib RNRs (Appendix 2, **Table A2.1**). In the remainder of this chapter and in Chapter 4, however, specific activities are only reported relative to $\alpha_2\beta_2$ since this is the species active in NDP reduction.

Table 3.3. *B. subtilis* class Ib RNR steady-state kinetic parameters with 1 mM substrate and variable effector concentrations.[a]

E	S	[E] range (μM)	V_{max} (nmol min^{-1} mg^{-1})			K_m (μM)	k_{cat} (s^{-1})	k_{cat} K_m^{-1} (M^{-1} s^{-1})
			β_2	α_2	$\alpha_2\beta_2$			
ATP	CDP	50 - 4000	1200 ± 40	590 ± 20	400 ± 10	200 ± 30	1.6 ± 0.06	[8.3 ± 1.3] x 10^3
ATP	UDP	25 – 4000	1600 ± 40	770 ± 20	520 ± 10	350 ± 40	2.1 ± 0.06	[6.0 ± 0.6] x 10^3
dGTP	ADP	0.25 - 4000	1300 ± 10	640 ± 9	430 ± 6	1.7 ± 0.2	1.8 ± 0.03	[1.0 ± 0.1] x 10^6
TTP	GDP	0.10 - 4000	680 ± 8	330 ± 4	220 ± 3	1.2 ± 0.1	0.9 ± 0.01	[7.8 ± 0.6] x 10^5

[a] Measurements made with His_6-tagged proteins. His_6-NrdE was a mixture of apo- and dAMP-loaded protein.

Table 3.4. *B. subtilis* class Ib RNR steady-state kinetic parameters with constant effector and variable substrate concentrations.[a]

S	E[b]	[S] range (μM)	Fit[c]	V_{max} (nmol min^{-1} mg^{-1})			K_m (μM)	k_{cat} (s^{-1})	k_{cat} K_m^{-1} (M^{-1} s^{-1})	n_H	p^d
				β_2	α_2	$\alpha_2\beta_2$					
ADP	dGTP (20 μM)	25 – 8000	MM	1500 ± 40	750 ± 20	500 ± 10	340 ± 40	2.1 ± 0.05	[6.1 ± 0.7] x 10^3	1	0.0003
			H	1700 ± 70	840 ± 40	570 ± 20	490 ± 80	2.3 ± 0.10	[4.8 ± 0.8] x 10^3	0.7 ± 0.1	
GDP	TTP (15 μM)	10 – 1000	MM	710 ± 30	320 ± 10	210 ± 10	26 ± 6	0.9 ± 0.03	[3.4 ± 0.7] x 10^4	1	0.002
			H	610 ± 20	300 ± 8	200 ± 5	25 ± 2	0.8 ± 0.02	[3.3 ± 0.3] x 10^4	1.8 ± 0.3	
CDP	ATP (2.9 mM)	10 – 2000	MM	930 ± 30	460 ± 10	310 ± 9	71 ± 9	1.3 ± 0.04	[1.7 ± 0.2] x 10^4	1	0.03
			H	880 ± 30	430 ± 10	290 ± 9	66 ± 6	1.2 ± 0.03	[1.8 ± 0.2] x 10^4	1.3 ± 0.2	
CDP	dATP (2.5 μM)	25 – 2000	MM	790 ± 20	390 ± 8	260 ± 6	150 ± 10	1.1 ± 0.02	[7.1 ± 0.5] x 10^3	1	0.09
			H	750 ± 20	370 ± 10	250 ± 7	130 ± 10	1.0 ± 0.03	[7.6 ± 0.6] x 10^3	1.2 ± 0.1	
UDP	ATP (2.9 mM)	25 – 4000	MM	1200 ± 40	600 ± 20	400 ± 10	160 ± 20	1.7 ± 0.05	[1.1 ± 0.1] x 10^4	1	0.05
			H	1200 ± 40	570 ± 20	380 ± 10	140 ± 10	1.6 ± 0.05	[1.1 ± 0.1] x 10^4	1.3 ± 0.2	
UDP	dATP (5 μM)	25 – 3000	MM	800 ± 20	390 ± 8	270 ± 5	160 ± 10	1.1 ± 0.02	[6.8 ± 0.6] x 10^3	1	0.06
			H	770 ± 20	380 ± 9	250 ± 6	140 ± 10	1.0 ± 0.03	[7.3 ± 0.6] x 10^3	1.2 ± 0.1	

[a] Measurements made with His_6-tagged proteins. His_6-NrdE was a mixture of apo- and dAMP-loaded protein.

[b] Numbers in brackets are the final concentrations of effector nucleotide used in the assays (~10x the apparent K_m (**Table 3.3**)). The final dATP concentrations used were those giving maximal activity (**Figure 3.8**).

[c] MM = Michaelis-Menten equation, H = Hill equation.

[d] Nested model F-ratio. A $p \leq 0.01$ indicates a statistically better fit to the Hill equation as opposed to the Michaelis-Menten equation.

3.3.5. Cooperative effects on steady-state activity are observed with the purine substrates. **Table 3.4** shows the steady-state kinetic parameters for the *B. subtilis* RNR with effector concentrations held at ~10x their apparent K_m while substrate concentrations were varied. Assays using dATP as an effector were also run for comparison to the results obtained with ATP; concentrations of dATP that gave optimal CDP (2.5 μM dATP) and UDP (5 μM dATP) reduction activities were used in place of an apparent K_m (**Section 3.3.7**). Notably, the data collected with ADP and GDP were fit significantly better with the Hill equation (**Figure 3.6**), suggesting the possibility of cooperative binding of these nucleotides to the RNR. Strangely, the effects of each purine were opposing, as GDP exhibited strong positive cooperativity (n_H = 1.8 ± 0.3) whereas ADP exhibited moderate negative cooperativity (n_H = 0.7 ± 0.1). The latter may explain the strikingly weak apparent K_m (490 μM) of the *B. subtilis* enzyme for ADP in the presence of dGTP when compared to similar measurements (Appendix 3, **Table A3.4**) or K_ds (**Table A3.5**) reported for other class I systems. *B. subtilis* RNR also exhibited somewhat weaker affinity for CDP in comparison to most other characterized systems (exception: *Streptococcus pyogenes* Ib RNR exhibited an apparent K_m = 340 μM[15]), whereas the K_ms for GDP and UDP were more similar. All the turnover numbers were in the range of 1 – 2 s^{-1}, resulting in catalytic efficiencies (k_{cat} K_m^{-1}) that fall within the middle of the range reported for other class I RNRs (Appendix 3, **Table A3.6**). For the *B. subtilis* Ib RNR, GDP was the most efficient substrate and ADP the least, a similar scheme compared to the mouse and *Trypanosoma brucei* Ia enzymes[34, 55] but different from *E. coli*,[53, 56, 57] calf thymus,[58] and other RNRs,[17, 33, 59] for which UDP was reported to be least efficient substrate.

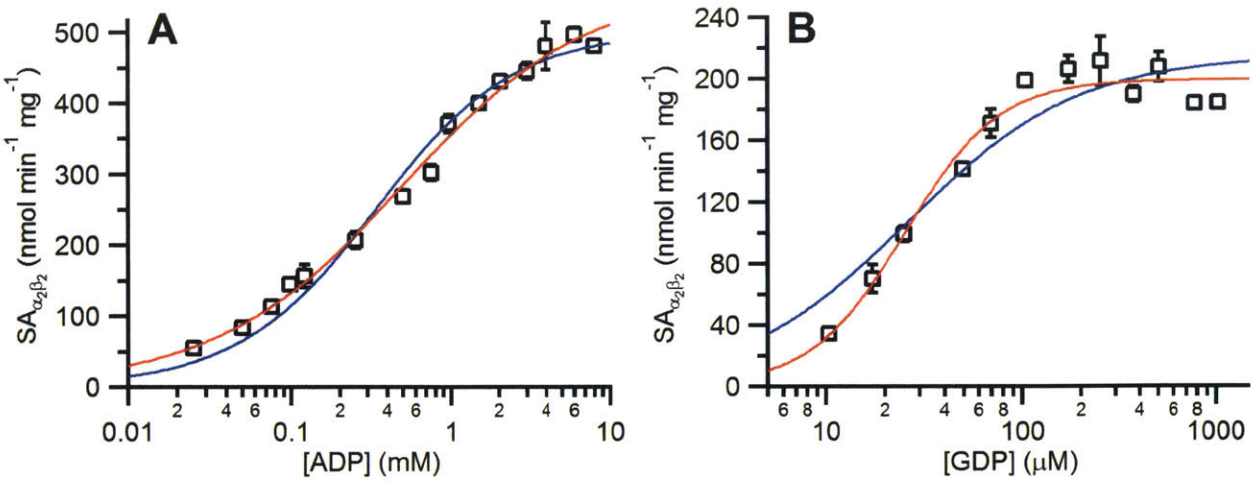


Figure 3.6. Comparison of fits of the Michaelis-Menten equation (blue line) versus Hill equation (red line) to the steady-state kinetic data (black boxes) obtained for (**A**) ADP (dGTP = 20 μM), and (**B**) GDP (TTP = 15 μM) with the *B. subtilis* Ib RNR. Assays were conducted at 37 °C in a total volume of 500 μL and contained a 1:1 mixture of 0.5 μM subunits, the endogenous reducing system, and effectors at ~10x their apparent K_m (**Table 3.3**). Each data point is the average of three measurements. In many cases, the error bars (± 1 standard deviation) are smaller than the marker size.

3.3.6. Substrate reduction is best stimulated by the expected prime effector. The steady-state kinetic measurements made with *B. subtilis* RNR were done under the assumption that the prime effector for a given substrate was the same as expected from the general scheme of RNR allosteric regulation.[18] However, all substrate and effector nucleotides are present simultaneously *in vivo*, and previous studies have noted the ability of individual effectors to stimulate low levels of reduction for all four substrates by, for example, the *E. coli* Ia[53, 57] and the Fe-loaded *S.* Typhimurium Ib RNRs.[17] It was important, given the inhibitory effects of dATP, to determine if the substrate specificity of *B. subtilis* RNR for NDPs was dictated any differently than expected from the general allosteric scheme.[18] Furthermore, since the *S.* Typhimurium Ib RNR is not a good model system for this subclass, the examination of the substrate specificity of a Ib RNR that is used as the primary aerobic source of dNTPs will be valuable in future studies. Therefore, assays examining the ability of each effector to stimulate the reduction of each substrate were carried out.

Also included in the experiments for completeness were assays testing the effector capabilities of dCTP, which has been reported to stimulate varying levels of reduction of all four substrates by the T4 phage Ia and *Thermoplasma acidophilum* class II RNRs.[59-61] Since no steady-state measurements were made with this nucleotide, a final concentration of 50 μM dCTP was arbitrarily chosen and later found to be consistent with pools sizes measured in *B. subtilis* cells (Appendix 3, **Table A3.11**). All other nucleotides were used at concentrations that were ~5 – 10x the apparent K_m measured in the steady-state assays described above (**Tables 3.3** and **3.4**).

The results of examining the substrate specificity determinants of *B. subtilis* RNR are shown in **Figure 3.7**, and demonstrate that the allosteric regulation of substrate specificity of this enzyme is similar to the general scheme for all RNRs.[18] ATP and low concentrations of dATP maximally stimulated the reduction of pyrimidines, TTP maximally stimulated GDP reduction, dGTP maximally stimulated ADP reduction, and dCTP was not an effector of RNR activity. The largest difference between the general allosteric scheme[18] and the results reported here concerns TTP, which appears to be equally effective in stimulating GDP, CDP, and UDP reduction, although it was not as potent as ATP for the latter two substrates (**Figure 3.7**). Similar results have been observed with the *E. coli* Ia[57] and *S.* Typhimurium Ib RNRs.[17] Further comparison of the two Ib RNRs reveals differences in substrate specificity regulation by dGTP. With the *B. subtilis* enzyme, this effector only stimulated ADP reduction, whereas with the *S.* Typhimurium enzyme, substantial levels of reduction of CDP, GDP, and UDP were also stimulated by this effector (30 – 45% relative to the levels stimulated by the prime effectors of these nucleotides). Finally, in the absence of effectors, only the pyrimidine substrates showed low rates (3 – 4% of that observed with ATP as the effector) of reduction by the enzyme (**Figure 3.7**). Here again, the

B. subtilis and *S.* Typhimurium RNRs are different in that the latter exhibits a much faster rate of pyrimidine reduction (10 – 20% the maximum) in the absence of effectors.[17]

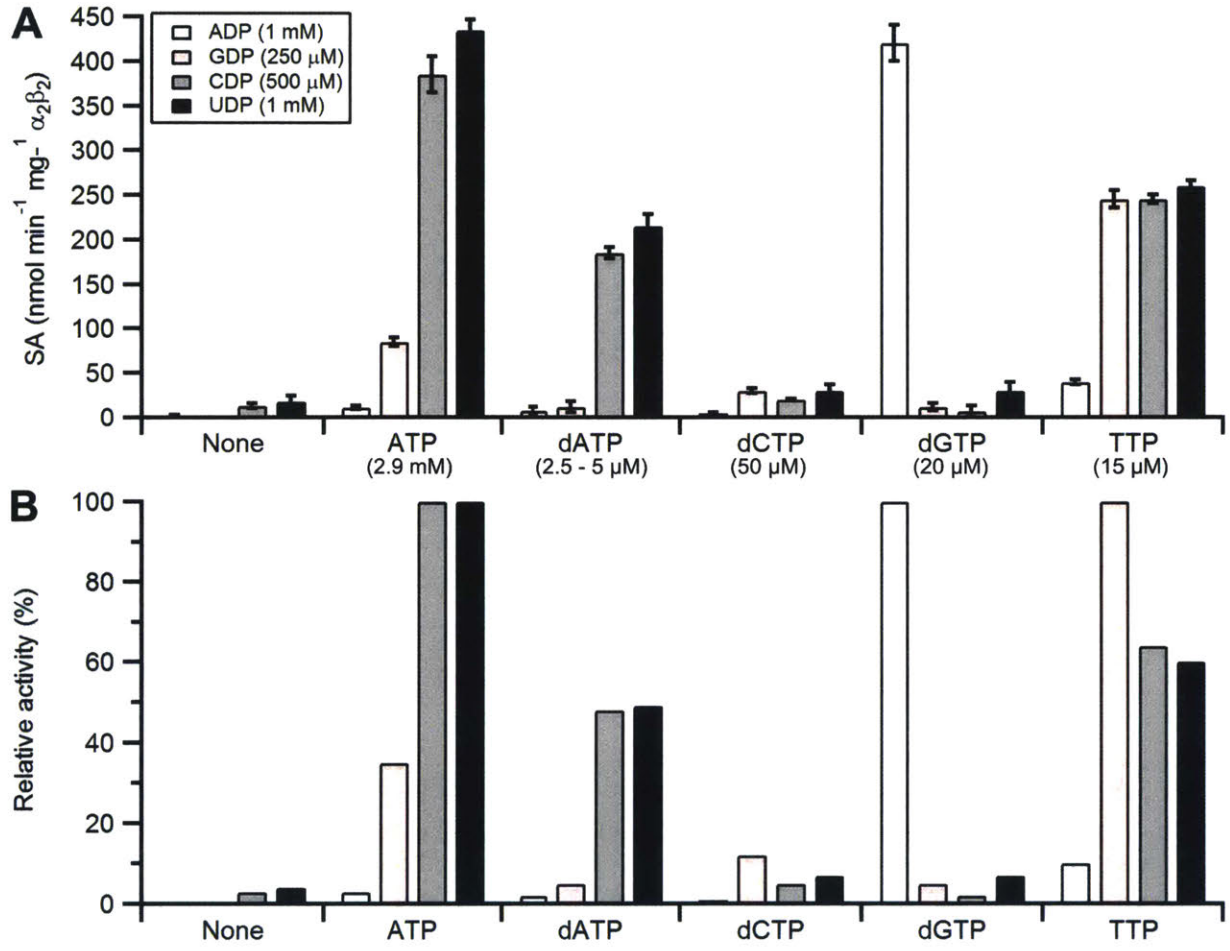


Figure 3.7. Allosteric regulation of the substrate specificity of Mn-loaded *B. subtilis* Ib RNR. (**A**) Specific activities of substrate reduction stimulated by single effectors. Assays (500 μL) were conducted at 37 °C and consisted of 0.5 μM Mn-β_2, 1 μM NrdE, the endogenous reducing system, and substrate/effector at the indicated final concentrations in 50 mM HEPES, pH 7.6, 15 mM $MgCl_2$, 1 mM EDTA. (**B**) Specific activities expressed as a percentage of the maximum for a given substrate. These data are tabulated in Appendix 3, **Table A3.13**.

3.3.7. Inhibition of *B. subtilis* RNR by dATP is an inherent property of the enzyme. The unexpected inhibitory effect of dATP on CDP reduction catalyzed by the *B. subtilis* Ib RNR described in Chapter 2 was confirmed by the spectrophotometric assay and extended to include UDP reduction as well (**Figure 3.8**). Inhibition was observed regardless of the presence versus

absence of a N-terminal His_6-tag, thus eliminating the possibility of reduced activity due to enhanced aggregation in the presence of dATP, a scenario analogous to that observed with tagged human RNR in the presence of ATP.[28] Substrate reduction was stimulated up to dATP concentrations of 2.5 μM (CDP) to 5 μM (UDP) before inhibitory effects began to set in, similar to what was observed by the radioactive assay (**Figure 2.6C**).

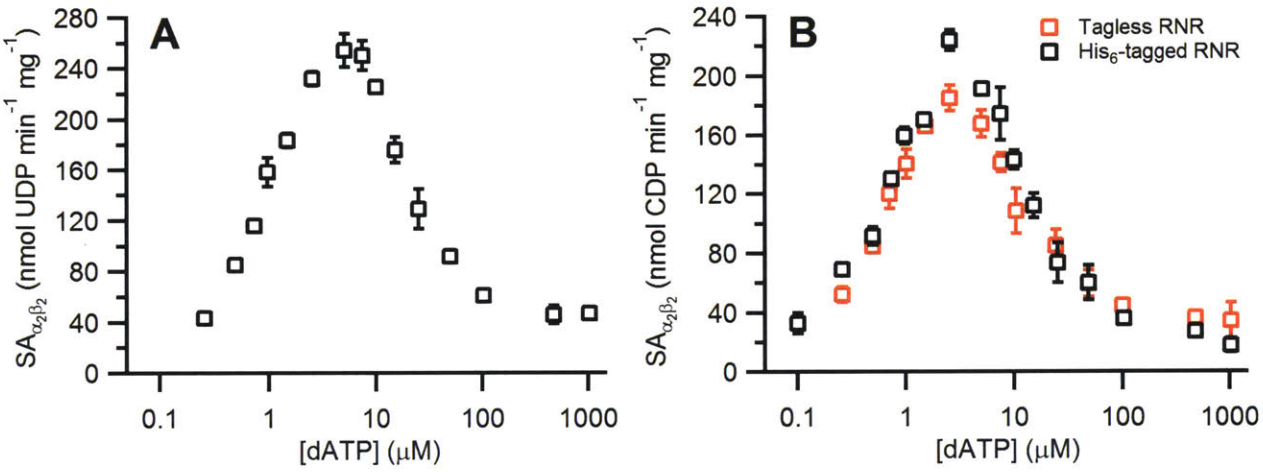


Figure 3.8. Inhibition of (**A**) UDP reduction and (**B**) CDP reduction catalyzed by the *B. subtilis* Ib RNR by dATP as measured by the spectrophotometric assay. Assays were conducted at 37 °C in 500 μL using a 1:1 ratio of 0.5 μM subunits. Black and red squares correspond to measured activity with the His_6-tagged and tagless proteins, respectively. Assays were run in triplicate and the results averaged. In most cases, the error bars (± 1 standard deviation) are smaller than the marker size.

3.3.8. Maximum GDP reduction rates require the simultaneous presence of TTP and ATP.

The confirmation of the inhibitory effects of dATP on *B. subtilis* RNR activity prompted a further examination of the allosteric regulation of the overall activity of this enzyme. In the class Ia RNRs, ATP plays an important role in the overall activity by binding to the ATP-cone domain and serving as a general activator of NDP reduction.[18] In the case of several eukaryotic RNRs, ATP in combination with the prime effector is required in order to stimulate the maximum rates of ADP

and GDP reduction, presumably with ATP occupying the ATP-cone and the dNTP occupying the specificity site.[33, 34, 55, 58] In contrast, it has been reported that the *E. coli* Ia RNR is inhibited by the combination of ATP and TTP,[29, 30] which is thought to be the result of a vaguely defined "cross-talk" between the specificity and overall activity sites.[18, 30] However, recent data from the Drennan group indicates that two effector nucleotides can bind in the ATP-cone domain of the *E. coli* enzyme, and recently published structures of the $\alpha_4\beta_4$ inhibited complex showed that TTP could be one of the two nucleotides bound at this site (Appendix 3, **Figure A3.1**).[62, 63] Therefore, it is possible that the inhibitory effects from the combination of ATP and TTP results from both nucleotides binding to the ATP-cone domain.

Given the observed inhibition of *B. subtilis* RNR by dATP despite not having an ATP-cone, it was important to determine if ATP had effects on enzymatic activity (besides specificity dictation for pyrimidine NDPs) that could be indicative of a more general allosteric mechanism regulating overall activity. If such a mechanism were to exist, then it may be possible that other class Ib RNRs can also exhibit overall activity regulation in the absence of an ATP-cone domain. Therefore, two sets of experiments were carried out: (1) the reduction of ADP and GDP was assessed with the prime effector (dGTP and TTP, respectively) in combination with ATP to see if the latter could act as a general activator, and (2) the reduction of CDP and UDP was assessed with the combination of ATP and TTP to see if this combination could inhibit the enzyme. The nucleotide concentrations used in these assays were again ~10x their apparent K_m values (**Table 3.3** and **3.4**).

The presence of saturating amounts of both ATP and TTP had little effect on the rate of CDP (360 ± 15 nmol min^{-1} mg^{-1} $\alpha_2\beta_2$) and UDP (380 ± 10 nmol min^{-1} mg^{-1} $\alpha_2\beta_2$) reduction catalyzed by *B. subtilis* RNR in comparison with the activities measured in the presence of ATP

only (**Table 3.4**). These results indicate that, at least under the conditions of these assays, the combination of ATP and TTP does not inhibit the *B. subtilis* enzyme. In contrast, the activities measured for GDP reduction stimulated by TTP or ATP alone (**Figure 3.7**) combined synergistically (280 ± 15 nmol min^{-1} mg^{-1} $\alpha_2\beta_2$) when both effectors were simultaneously included in the assay solution. The augmentation effect was saturable (**Figure 3.9**), with a rough apparent K_m of 70 ± 37 μM determined from fitting the data to the Michaelis-Menten equation expanded to include a constant, *C*, to account for the non-zero activity in the absence of ATP (equation 3.3).

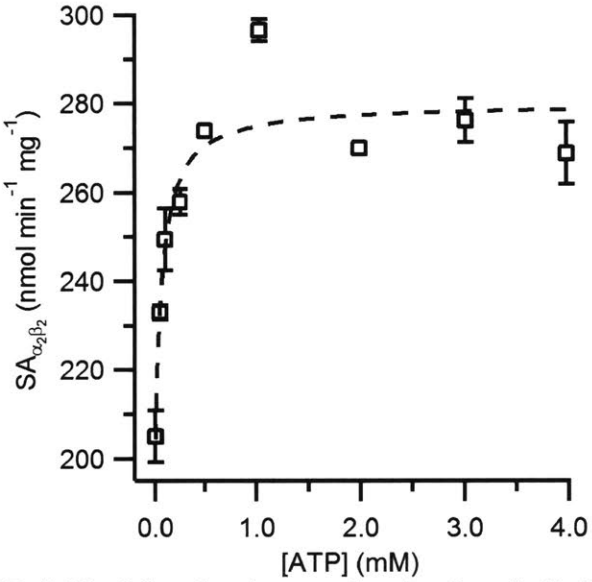


Figure 3.9. Increased GDP (250 μM) reduction rate by the *B. subtilis* Ib RNR in the presence of TTP (15 μM) and ATP. Assays were conducted at 37 °C in 500 μL using a 1:1 ratio of 0.5 μM subunits. Data points are the average of two replicates, and in a few cases the error bars (± 1 standard deviation) are smaller than the markers. The dotted line is the fit of the data to equation 3.3: V_{max} = 80 ± 10 U mg^{-1} $\alpha_2\beta_2$, K_m = 0.07 ± 0.04 mM, *C* = 204 ± 10 U mg^{-1} $\alpha_2\beta_2$.

$$v = \left(\frac{V_{max} \times [ATP]}{K_m + [ATP]} \right) + C \tag{3.3}$$

The value is smaller in magnitude than the apparent K_ms measured for ATP when it was the variable component in the steady-state kinetic assays described previously (**Table. 3.3**). Furthermore, ATP does not appear to be successfully competing with TTP for the specificity sites,

otherwise the activity should decrease due to the lower potency of the former to stimulate GDP reduction (**Figure 3.7**). The higher affinity for ATP and lack of competition for binding to the specificity site suggests that ATP may be binding to a distinct and uncharacterized allosteric site on the *B. subtilis* RNR, which supports the idea of a more generalized overall activity regulation of the enzyme. Notably, the cumulative effects of ATP and TTP bring the maximal GDP reduction rate catalyzed by *B. subtilis* RNR to within 90% of that observed for CDP with ATP as the sole effector (310 ± 9 nmol min^{-1} mg^{-1} $\alpha_2\beta_2$) and within 50 – 70% of that observed with other substrates using saturating concentrations of the prime effector (**Table 3.4**). The effects of ATP on dGTP-stimulated ADP reduction are more complex and are described in the next section.

3.3.9. *B. subtilis* RNR exhibits product inhibition by dADP that is not reversed by ATP. As mentioned earlier in this chapter, there was one case in which a RNR was reported to exhibit product inhibition. The *E. coli* Ia RNR, when examined with ADP and dGTP, showed curvature of kinetic traces when enzyme activity was plotted as a function of time, indicating that enzyme activity was slowly decreasing.[53] Control assays demonstrated that this gradual activity loss was attributable to the build-up of dADP rather than inactivation of the enzyme.[53] Further experimentation revealed that inhibition by dADP could be reversed and/or prevented by addition of ATP to the assays, which led to the conclusion that the inhibitory effects of dADP resulted from it binding to the same inhibitory allosteric site as dATP, albeit with a lower affinity.[53] The fact that the *B. subtilis* RNR is inhibited by dATP (**Figure 3.8**), exhibits an unusually high K_m for ADP (**Table 3.4**), and has negative cooperative effects on activity that are dependent on ADP concentration (**Figure 3.6A**), raised the question of whether or not dADP product inhibition occurs with this enzyme as well. Furthermore, if product inhibition was observed, it seemed logical that ATP could prevent and or reverse it, as was observed with the *E. coli* Ia RNR.

To test this hypothesis, the standard spectrophotometric assay conditions using dGTP and ADP were modified to allow NADPH oxidation to be monitored over 10 min. A 1:1 ratio of α_2:β_2 was maintained with the subunit concentrations reduced to 0.25 μM, and the total ADP concentration was lowered to 500 μM which, given the apparent K_m for the substrate (**Table 3.4**), makes the rate of ADP reduction more sensitive to the presence of inhibitors and activators. The concentrations of all other components remained the same, as well as the procedure for assay setup and execution. In these experiments, ATP was either included from the beginning of the assay or was added to the assay mixture during the course of the reaction to determine what affects it would have on the rate of ADP reduction.

In the presence of only dGTP, the rate of ADP reduction catalyzed by *B. subtilis* RNR was initially linear (SA = 340 nmol min^{-1} mg^{-1} $\alpha_2\beta_2$), but gradually began to decrease within 1 min of reaction initiation (**Figure 3.10A**, SA = ~170 U mg^{-1} $\alpha_2\beta_2$ prior to ATP addition), a phenomenon that was distinct from other substrates and is indicative of enzyme inhibition by increasing concentrations of dADP. Addition of ATP to 500 μM (**Figure 3.10A**) restored linearity to the rate of NADPH consumption for ~1 – 2 min, but did not increase it. This observation sharply contrasts with the results observed with the *E. coli* Ia RNR (**Figure 3.10C**), where ATP addition restored linearity and significantly increased the rate of ADP reduction. The ability of ATP to prolong the activity of *B. subtilis* RNR before the apparent product inhibition set in was tested by including increasing concentrations of ATP in the initial assay solutions. As shown in **Figure 3.10B**, increasing amounts of ATP caused modest increases in the initial rate of ADP reduction (320 U mg^{-1} $\alpha_2\beta_2$ in the absence of ATP to 400 U mg^{-1} $\alpha_2\beta_2$ with 2 mM ATP) and delayed the onset of curvature of the kinetic traces. These results indicate that the *B. subtilis* Ib RNR is susceptible to inhibition by dADP that can be delayed, but not prevented by the presence of ATP.

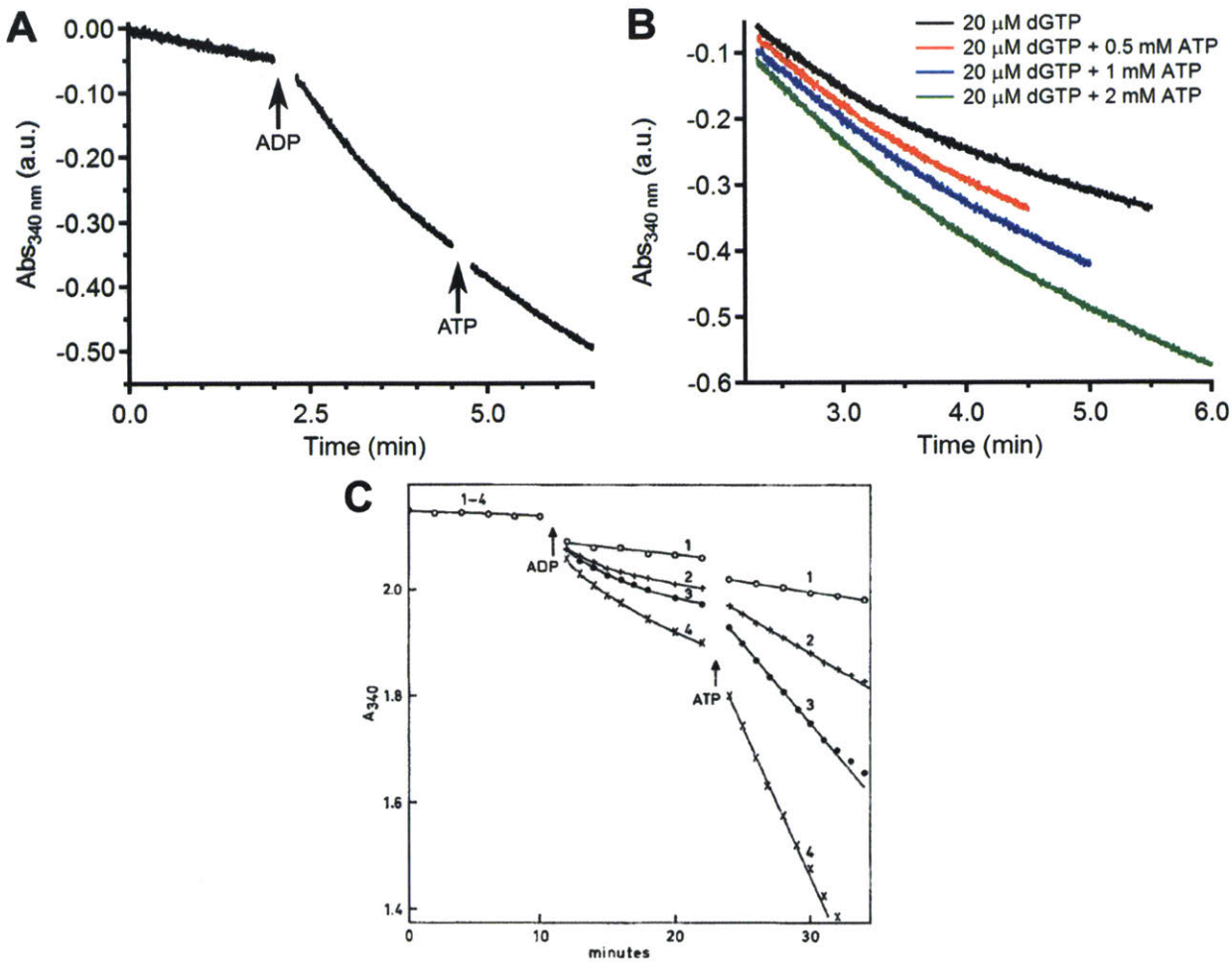


Figure 3.10. (**A**) *B. subtilis* RNR exhibits product inhibition by dADP that is not reversed by ATP. The assay solution (500 μL) was incubated at 37 °C and initially consisted of 0.5 μM NrdE, 0.25 μM $Mn(III)_2$-Y•-β_2, 20 μM dGTP, and the endogenous reducing system. At the *first arrow*, 270 nmol ADP (final concentration = 500 μM) was added to initiate turnover. At the *second arrow*, 240 nmol ATP was added to the assay to yield a final concentration of ~ 500 μM. (**B**) Increasing concentrations of ATP prolong turnover before product inhibition sets in. Assays (500 μL) prepared as described in (**A**) were supplemented with the indicated ATP concentrations and incubated for 2 min at 37 °C prior to ADP addition to initiate turnover. (**C**) Previously reported results for dADP inhibition of the *E. coli* Ia RNR. Assays (130 μL) were conducted at RT and consisted of 4 mM Tris, pH 8.0, 8 – 15 mM $MgCl_2$, 80 μM EDTA, 4 – 8 mM dithioerythritol, 0.46 mM NADPH, 1.3 – 2 μM thioredoxin, 0.7 – 1.8 μM thioredoxin reductase, 1.8 μM α_2, 1.5 μM β_2, and the following effectors: *curve 1*: no effector, *curve 2*: 100 μM dCTP, *curve 3*: 100 μM TTP, and *curve 4*: 100 μM dGTP. At the *first arrow*, 50 nmol ADP was added to initiate turnover (final concentration = 385 μM), and at the *second arrow*, 50 nmol ATP was added. This figure was originally published in the Journal of Biological Chemistry. Larsson, A.; Reichard, P. Enzymatic Synthesis of Deoxyribonucleotides. X. Reduction of Purine Ribonucleotides; Allosteric Behavior and Substrate Specificity of the Enzyme System from *Escherichia coli* B. *J. Biol. Chem.* 1966; 241: 2540-2549. © the American Society for Biochemistry and Molecular Biology.

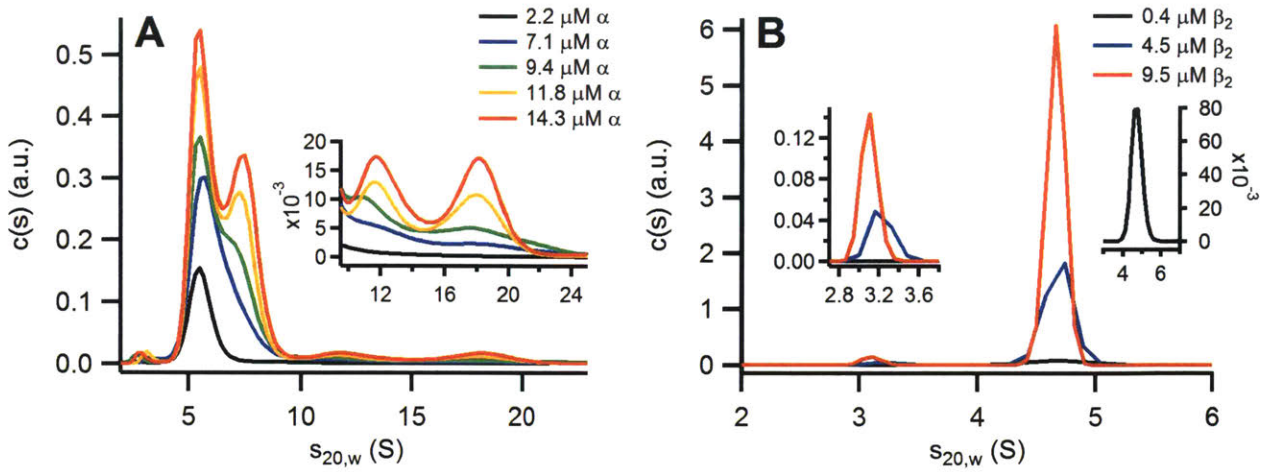


Figure 3.11. Concentration dependent SV-AUC analysis of tagless *B. subtilis* NrdE (**A**) and holo-Mn(III)$_2$-Y• NrdF (**B**) in the absence of nucleotides. The *inset* in (**A**) is a blow up of the 9 – 25 S region of the NrdE sample distributions to show the large oligomers that form with increasing protein concentration. The *left inset* in (**B**) is a blow up of the 2.7 – 3.8 S region of the NrdF sample distributions to show the small amounts of monomer. The *right inset* is a blow up of the 0.4 μM Mn-β_2 distribution to show the dimer peak which is obscured in the main image.

3.3.10. The quaternary structure of the *B. subtilis* RNR subunits is not significantly affected by tag removal. Concentration-dependent SV-AUC experiments of the individual tagless *B. subtilis* RNR subunits in the absence of nucleotides were run first to assess if the removal of the N-terminal affinity tag significantly affected their sedimentation behavior. As was done with the His$_6$-tagged constructs, the program HYDROPRO[64] was used to predict the hydrodynamic properties of the tagless subunits and symmetric and asymmetric $\alpha_2\beta_2$ complexes to use as a guide for interpreting the results of SV-AUC experiments (Appendix 3, **Table A3.12**). As expected, the tagless proteins behaved essentially the same as the His$_6$-tagged constructs (**Figure 3.11**). NrdE exists as a monomer ($s_{<Wt>}$[20,w] = 5.7 S, M_W = 75.4 kDa) near physiological concentrations (~2.2 μM) and showed concentration-dependent formation of dimer ($s_{<Wt>}$[20,w] = 7.5 S at 14.3 μM) and larger species at ~12.5 S and ~18 S (**Figure 3.11A**). Compared to the results with His$_6$-NrdE (Appendix 2, **Figure A2.2**) run under similar conditions (Buffer 1, 20 °C, 35000 rpm), the tagless protein formed larger oligomers at ~2-fold lower concentrations (7.1 μM versus ~16 – 19 μM for

His_6-NrdE), suggesting NrdE had a higher propensity to oligomerize. NrdF exists almost exclusively as a dimer (**Figure 3.11B**), with global fits to the three data sets yielding $s_{20,w}$ = 4.65 S, 68.3% confidence interval (CI) [4.59, 4.72] and M_W = 73.1 kDa, 68.3% CI [66.0, 84.3]. The amount of NrdF monomer was reduced compared to that observed with His_6-NrdF (Appendix 2, **Figure A2.1**).

3.3.11. General considerations for SV-AUC analysis of *B. subtilis* RNR in the presence of nucleotides. TTP and GDP were chosen for the initial experiments due to their stability over the timeframe required for data acquisition (20 °C, 17 – 19 h). More importantly, however, the positive cooperativity observed during steady-state assays of GDP raised the question of whether or not this effect manifested as quaternary structural changes of NrdE and/or holo-RNR. Along the same lines, given that dATP inhibition of Ia RNRs results from formation of inhibitory complexes,[27, 28, 30, 65] it seemed feasible that the mechanism of inhibition of *B. subtilis* RNR by dATP might also involve alterations in the quaternary structure of NrdE and/or holo-RNR that are, albeit, distinctly different from the Ia systems. Therefore, additional experiments probing the sedimentation of the *B. subtilis* RNR subunits and complexes were carried out in the presence dATP. In these experiments, protein concentrations were 2 – 2.5 μM, TTP and GDP concentrations varied from 0.1x – 10x their apparent K_m (**Table 3.3** and **3.4**), and dATP concentrations varied from 0.5 – 50 μM, spanning both stimulatory and inhibitory ranges as judged by the activity assays (**Figure 3.8**).

The nucleotides used in these experiments all absorb to a certain extent at 280 nm, thus their inclusion in the samples would contribute additional signal to the A_{280} which, at the higher end of the concentration ranges used, might overwhelm the signal from NrdE. This issue necessitated the use of the interference detection system, which monitors changes in the refractive

index and, therefore, detects solution components independently from their light absorption properties. The downside to this detection method is that anything that changes the refractive index of the sample solution relative to the reference, including buffer molecules and salts, will contribute to the signal. Therefore, it is important to match both the composition and geometry of the reference and sample sectors in order to minimize the changes in refractive index in the sample signal that are not due to sedimenting protein. As it is often difficult to get perfectly matched reference and sample sectors, a buffer mismatch model implementable in Sedfit during data fitting was developed by Peter Schuck's group (National Institute of Health, Bethesda, MD) to help overcome this issue.[47] A more detailed discussion of the model is presented in Appendix 3. Pertinent for the following results is that this model was used during data fitting. In the case of the NrdE data sets, it resulted in artifactual sharpening of the peaks in the $s_{20,w}$ distributions. This observation is counterintuitive as interconversion of NrdE between different oligomeric states was expected to result in peak broadening such as that observed in **Figure 3.11A**. While the cause for the odd appearance is unknown, it did not affect the data as the total signal accounted for all the protein loaded, and the best fit parameters are consistent with HYDROPRO predictions (Appendix 3, **Table A3.12**) and the previously obtained data for both tagless (**Figure 3.11A**) and His_6-tagged NrdE (**Table 2.6** and Appendix 2, **Figure A2.2**).

3.3.12. The quaternary structure of NrdF is unaffected by nucleotides. The sedimentation of $Mn(III)_2$-Y• β_2 (~1 μM) was unaffected by the presence of nucleotides as expected based on previous experiments with other class I RNR β_2 subunits (data not shown).[66-68] In all experiments, a single major peak was observed with a sedimentation coefficient ($s_{<Wt>}[20,w] = 4.6$ S), consistent with a dimer.

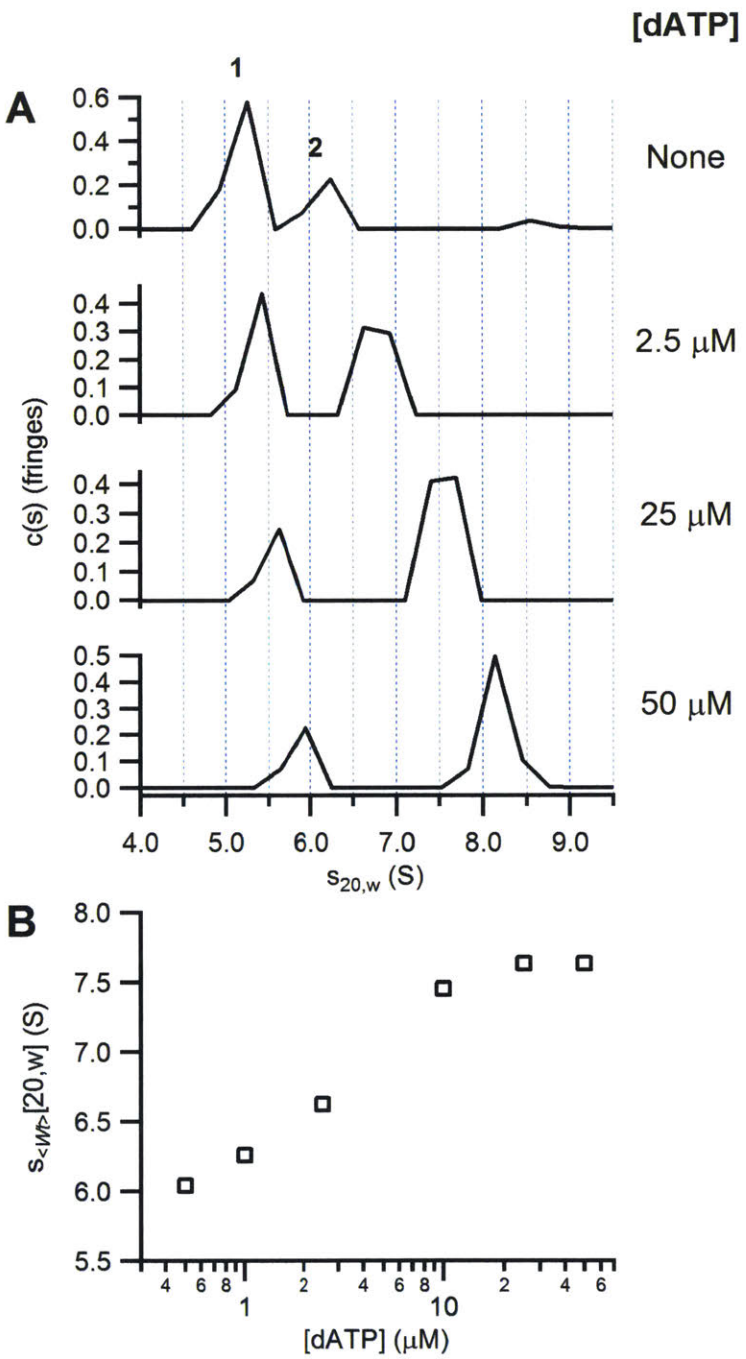


Figure 3.12. SV-AUC analysis of *B. subtilis* NrdE in the presence of dATP. (**A**) $s_{20,w}$ distributions for samples of 2.1 μM NrdE sedimented in buffer containing the final dATP concentrations indicated to the right. Peak **1** corresponds to NrdE monomer and Peak **2** to dimer. Note that the peak at ~8.5 S in the top distribution (no dATP) is an artifact from data fitting. (**B**) Weight averaged $s_{20,w}$ isotherm as a function of dATP concentration illustrating an overall shift of the protein quaternary structure to a form that is consistent with NrdE dimer. Note that samples with final dATP concentrations of 0.5, 1, and 10 μM were also run, but their $s_{20,w}$ distributions are not shown in (**A**).

3.3.13. The binding of dATP stabilizes NrdE dimers. The results of sedimenting ~2 μM NrdE in the presence of increasing concentrations of dATP are shown in **Figure 3.12**. As a control, one sample was run without nucleotides to see how the appearance of the $s_{20,w}$ distribution (**Figure 3.12A**, top distribution) differed from that obtained using A_{280} to monitor sedimentation (**Figure 3.11A**, black trace). As elaborated upon further in Appendix 3, technical factors in the acquisition of A_{280} versus interference data results in the collection of higher resolution scans by the latter method. Consequently, two major peaks (**1** and **2**) are observed in results obtained with interference detection (**Figure 3.12A**, top distribution), whereas only one is observed by A_{280} (**Figure 3.11A**, black trace), despite both samples having similar NrdE concentrations (2.1 – 2.2 μM). Given that NrdE interconversion between monomer and dimer is an equilibrium reaction, as indicated by the concentration-dependent sedimentation results shown in **Figure 3.11A**, the identity of Peak 2 in **Figure 3.12** and subsequent figures corresponds to a small amount of NrdE dimer. The reason that α_2 does not sediment as predicted (Appendix 3, **Table A3.12**) is likely due to fast rates of NrdE association and dissociation relative to the timescale of sedimentation.[69] Further discussion of this topic is presented in Appendix 3.

In the absence of dATP, most of the NrdE (**Figure 3.12A**, peak **1**, ~73%) sediments with $s_{<Wt>}[20,w]$ = 5.2 S (M_W = 72.3 kDa) and nearly all the rest (peak **2**, ~24%) at $s_{<Wt>}[20,w]$ = 6.2 S (M_W = 93.0 kDa). As the concentration of dATP increases, both peaks migrate to higher $s_{20,w}$ values concomitant with a shift in the amount of protein from the smaller to larger species. Ultimately, ~70% of the total protein is found in peak **2** ($s_{<Wt>}[20,w]$ = 8.2 S) and ~27% in peak **1** ($s_{<Wt>}[20,w]$ = 5.9 S) when the dATP concentration reaches 50 μM dATP (**Figure 3.12A**). The mass transfer is more readily evident in the weight averaged $s_{20,w}$ isotherm (**Figure 3.12B**) for the entire distribution (1 – 30 S). These observations indicate both major species correspond to a

different oligomeric form of NrdE and that dATP induces a shift in quaternary structure to the larger one. The results are most consistent with a NrdE monomer-dimer system that, in the absence of dATP, favors monomer, but in the presence of dATP, favors dimer. In support of this interpretation, the $s_{<Wt>}[20,w]$ for peak **1** in the sample lacking nucleotides (5.2 S) is identical to that predicted for NrdE monomer by HYDROPRO whereas the $s_{<Wt>}[20,w]$ for peak **2** in the 50 μM dATP sample (8.1 S) is very similar to the value predicted for NrdE dimer (8.0 S, Appendix 3, **Table A3.12**). The concentration of dATP at which the total amount of NrdE is equally distributed between monomer and dimer is roughly 5 μM.

The $s_{<Wt>}[20,w]$ isotherm (**Figure 3.12B**) appears to level off when the dATP concentration reaches 25 – 50 μM, suggesting that complete conversion of NrdE to dimer had taken place. However, the $s_{20,w}$ distributions (**Figure 3.12A**) show that close to 30% of the protein still remained in the monomeric state, indicating that either higher dATP concentrations were required to get complete NrdE dimerization or that the remaining monomer was incompetent in binding dATP and/or forming dimers. Additional SV-AUC experiments using dATP concentrations of 100 μM and 500 μM were inconclusive regarding its ability to completely convert NrdE to dimer due to the formation of large aggregates (data not shown). In regards to generation of incompetent protein, it has been noted previously with the *E. coli* Ia RNR that oxidation of Cys residues on the α_2 subunit caused protein monomerization and abolished activity, and the process was irreversible if the protein was left in the oxidized state for a few hours.[70] Whether this type of protein inactivation is relevant to *B. subtilis* NrdE has not been established, but it should be noted that the NrdE preparation used in these experiments was kept in a reducing environment (5 – 10 mM DTT or 1 mM TCEP), therefore protein oxidation should not have been an issue. Thus, the reason for the incomplete conversion of NrdE to dimer in the presence of dATP is currently unknown.

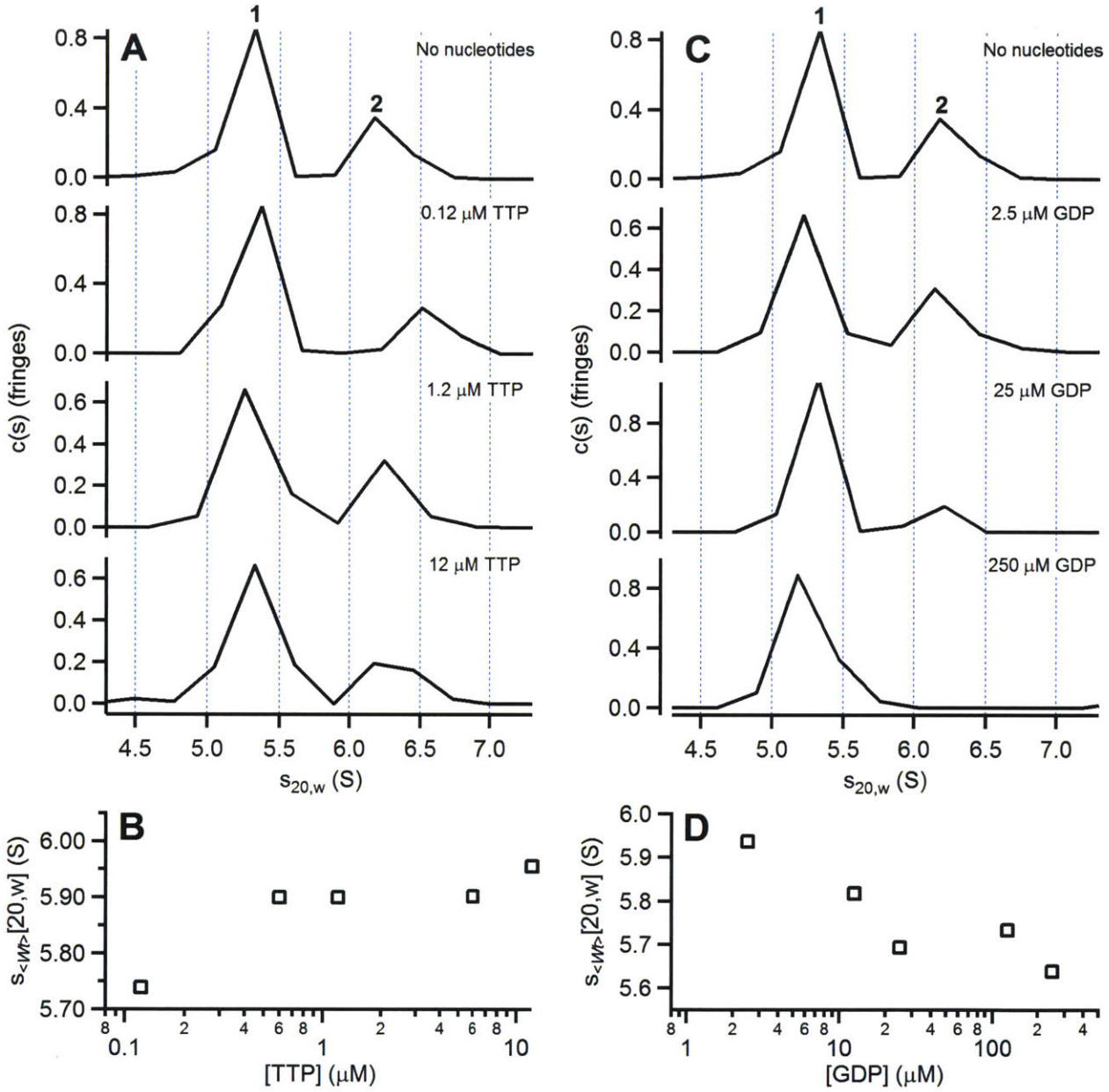


Figure 3.13. SV-AUC analysis of *B. subtilis* NrdE in the presence of TTP and GDP. (**A**) $s_{20,w}$ distributions for samples of 2.3 μM NrdE sedimented in buffer containing the final TTP concentrations indicated in the upper right corner of each graph. Peak **1** corresponds to NrdE monomer and Peak **2** to dimer. (**B**) The full distribution weight averaged $s_{20,w}$ isotherm as a function of TTP concentration. Note that two additional samples were run (0.6 and 6 μM TTP), but their $s_{20,w}$ distributions are not shown in (**A**). (**C**) $s_{20,w}$ distributions for samples of 2.3 μM NrdE sedimented in the presence of GDP at the final concentrations indicated in the upper right corner of each graph. (**D**) The full distribution weight averaged $s_{20,w}$ isotherm as a function of GDP concentration. Again note that two additional samples (12.5 and 125 μM GDP) were run, but their $s_{20,w}$ distributions are not shown in **(C)**.

3.3.14. NrdE dimer is not stabilized by TTP. Figure 3.13A shows the results of sedimenting ~2 μM NrdE in the presence of TTP at concentrations 0.1x, 1x, and 10x its apparent K_m. The $s_{20,w}$ distribution from the control run without nucleotides again showed two major peaks, with ~74% of the protein sedimenting as a monomer (**Figure 3.13A**, peak **1**, $s_{<Wt>}[20,w]$ = 5.3 S, M_W = 77.7 kDa) and the rest as dimer (peak **2**, $s_{<Wt>}[20,w]$ = 6.5 S, M_W = 106 kDa). In contrast to the effects observed with dATP, the quaternary structure of NrdE appears to be unaltered by TTP, a conclusion that is supported by the relatively constant $s_{<Wt>}[20,w]$ value for the full distribution (1 – 30 S) observed as a function of TTP concentration in the isotherm shown in **Figure 3.13B**. In all samples, most of the NrdE (~70%) sedimented at $s_{<Wt>}[20,w]$ = 5.4 S (M_W = ~ 73 kDa) and the rest at $s_{<Wt>}[20,w]$ = 6.4 S (M_W = ~ 96 kDa), consistent with monomer and dimer, respectively. Thus, TTP appears to be unable to stabilize NrdE dimer in the range of concentrations tested.

3.3.15. GDP binding causes NrdE dimers to dissociate. The results of sedimenting ~2 μM NrdE in the presence of GDP at concentrations spanning 0.1 – 10x its apparent K_m are shown in **Figure 3.13C**. Unexpectedly, increasing amounts of GDP causes the small amount of NrdE dimer observed in the absence of nucleotides (**Figure 3.13C**, peak **2**, $s_{<Wt>}[20,w]$ = ~6.6 S, M_W = ~99 kDa) to dissociate into monomers. This observation is supported by the gradual disappearance of peak **2**, the displacement of the total amount of protein into the monomer peak (peak **1**, $s_{<Wt>}[20,w]$ = ~5.3 S, M_W = ~74 kDa in all experiments), and by the decreasing $s_{<Wt>}[20,w]$ values for the full distribution (1 – 30 S) in the isothermal plot (**Figure 3.13D**). The dissociation of α_2 in the presence of substrate is unprecedented in the studies of class I RNRs. These results with the *B. subtilis* enzyme have been confirmed in preliminary studies of the protein using SAXS by Nozomi Ando's group at Princeton University (Will Thomas, N. Ando, unpublished), thus providing support that this phenomenon is an actual attribute of the enzyme. Further studies are warranted to understand

the mechanism by which the dissociation occurs and to determine whether the phenomenon has any physiological relevance.

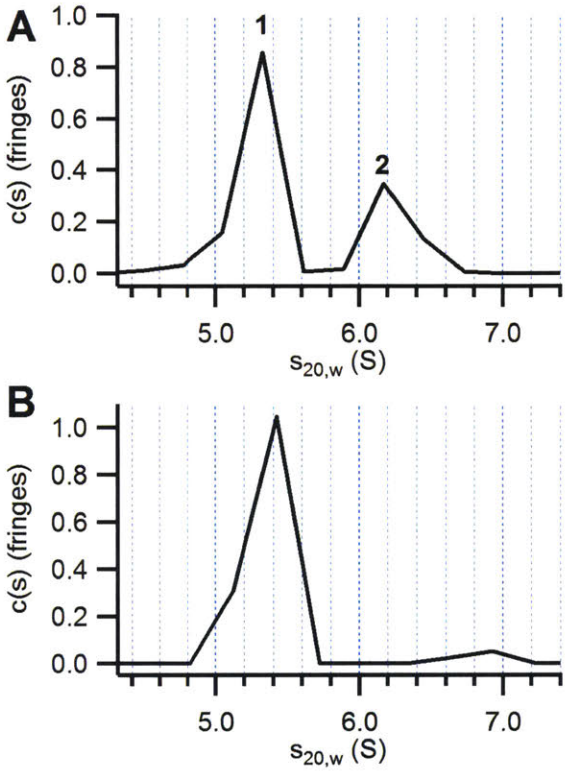


Figure 3.14. SV-AUC analysis of 2.3 μM NrdE in (**A**) the absence of nucleotides and (**B**) in the presence of 12 μM TTP and 250 μM GDP. Peak **1** corresponds to NrdE monomer and peak **2** to dimer.

3.3.16. NrdE dimerization is not induced by the presence of saturating amounts of both substrate and effector. The last SV-AUC experiment run with NrdE alone was sedimenting ~2 μM of the protein in the presence of saturating amounts of both TTP (12 μM) and GDP (250 μM) to see if NrdE dimers might be stabilized by having both the active and specificity sites occupied. The results of this experiment are shown in **Figure 3.14** and look very similar to those observed with 250 μM GDP. Most of the protein (~90%) sediments at $s_{<WT>}[20,w]$ = 5.4 S (peak **1**, Mw = 72.6 kDa), which is consistent with monomer. Unlike the case with 250 μM GDP, however, a

small amount of dimer (peak **2**, ~5%, $s_{<Wt>}[20,w]$ = 6.8 S, M_W = 105 kDa) is clearly present, indicating that TTP might slightly mitigate the monomerizing effects of GDP. Nevertheless, the results indicate that the combination of substrate and effector (at least in the case of GDP and TTP) does not stabilize NrdE dimer.

3.3.17. Large NrdE/NrdF complexes form in the presence of dATP. The effects of dATP on the quaternary structure of 1:1 mixtures of His_6-α_2:$Mn(III)_2$-Y• His_6-β_2 were also examined by SV-AUC. The design of this study differed from the previously discussed experiments in that the concentration of dATP was held constant at 40 μM while the subunit concentrations were varied from 0.4 – 3.5 μM. Nevertheless, the results obtained still unambiguously demonstrate that major changes in the quaternary structure of the *B. subtilis* RNR occur in the presence of dATP. As shown in **Figure 3.15**, a series of oligomers of increasing size had formed, with the largest species sedimenting at $s_{<Wt>}[20,w]$ = ~32 – 35 S and having an apparent M_W of over 1 MDa. The large range of $s_{20,w}$ values observed necessitated the use of the ls-g*(s) model in data fitting as opposed to the c(s) model used for the previous data sets. The major difference between the two models is that ls-g*(s) model does not attempt to deconvolute the diffusion in the $s_{20,w}$ distributions, which results in lower resolution of the different species present.[71, 72] This is the reason for a single broad peak in the range of 0 – 7 S (**Figure 3.15**) rather than individual peaks representing α, α_2, and β_2. However, the large macromolecules in the range of 15 – 35 S should show negligible diffusion on the timescale of sedimentation, thus the use of the ls-g*(s) model is acceptable. Unfortunately, other than the $s_{<Wt>}[20,w]$ values, the hydrodynamic properties of the different species are ill-defined and, therefore, makes it very difficult to attempt to identify the subunit composition and stoichiometry of each without more information. The major conclusion that can be drawn from the data is that dATP does induce changes in the quaternary structure of the *B. subtilis* RNR that

likely represent the mechanism by which the nucleotide inhibits enzymatic activity. On a final note, the formation of large oligomers of the *B. subtilis* RNR in the presence of dATP has been confirmed in preliminary SAXS experiments by the Ando group (Thomas, Ando, unpublished).

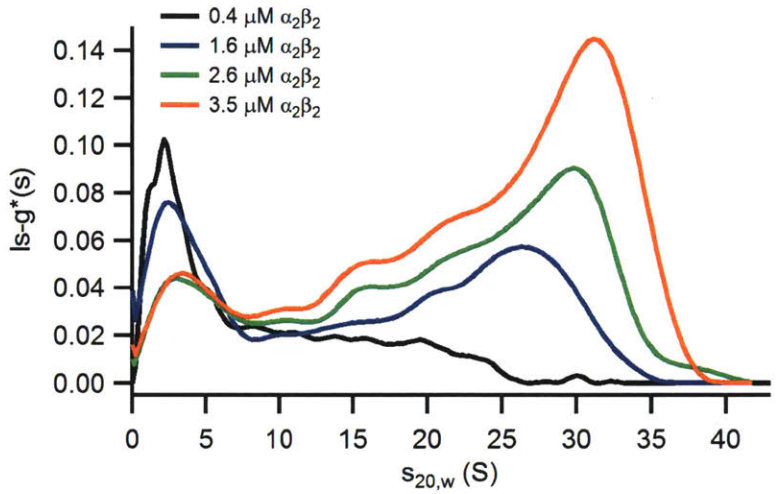


Figure 3.15. SV-AUC analysis of 1:1 mixtures of His_6-NrdE:$Mn(III)_2$-Y• His_6-NrdF at the indicated concentrations in the presence of 40 μM dATP.

3.4. DISCUSSION

The allosteric regulation of RNR is a key component in the cell's ability to maintain the proper amounts and ratios of the dNTP pools required for DNA replication and repair.[18] Allosteric effects observed with RNRs are induced by the binding of the effector nucleotides ATP, dATP, dGTP, and TTP to as many as two different classes of binding sites, depending on the enzyme, on the large subunit of the RNR. One class of sites affects the specificity of the enzyme for its four substrates while the other tunes the overall activity of the enzyme to the intracellular dATP/ATP ratio. Allosterically regulated enzymes are often found to be multimeric complexes, and the translation of a ligand binding event to a measurable effect on the protein's activity is the result of protein conformational changes that alter the quaternary structures of these complexes. The class

Ia RNRs offer a dramatic example of the types of quaternary structural changes that can occur in allosteric enzymes. Under conditions in which dNTPs are in demand, substrate and effector binding induces the formation of a globular $\alpha_2\beta_2$ complex that is active in NDP reduction.[73, 74] Once the demands of the cell are met, the binding of dATP induces conformational changes in α that cause the striking transformation of the globular enzyme into inhibited $\alpha_4\beta_4$ and/or α_6 ring quaternary structures.[27, 28, 65]

In contrast to class Ia systems, the allosteric regulation and quaternary structure of class Ib RNRs has remained a largely unexplored area,[10, 17, 75] and such studies with the physiologically relevant $Mn(III)_2$-Y• cofactor[2, 4, 5, 31] are virtually non-existent.[10] As reported in this chapter, the ability to produce highly active Mn-loaded *B. subtilis* Ib RNR in conjunction with the utilization of the endogenous reducing system has allowed, for the first time, an in-depth examination of these attributes for a Mn-loaded class Ib system. As *B. subtilis* is a model organism for Gram-positive and sporulating bacteria, substantial amounts of information about its genome, transcriptional/translational regulation, metabolism, and *in vivo* protein (see Appendix 2) and nucleotide concentrations (see Appendix 3) are available. By comparison of the results of this chapter with the published information in the literature, a better understanding of what may or may not be relevant to RNR function *in vivo* can be obtained.

3.4.1. Steady- state kinetic parameters are consistent with estimated in vivo concentrations of nucleotides in B. subtilis. It can be concluded from the steady-state kinetic analyses of the *B. subtilis* Ib RNR that the enzyme obeys the general rules of allosteric regulation of all RNRs[18] that were first laid down over 50 years ago with the *E. coli* class Ia system.[53, 56, 57] The presence of ATP or dATP (at low concentrations) cause the *B. subtilis* RNR to preferentially reduce pyrimidine NDPs whereas TTP and dGTP stimulate the reduction of GDP and ADP, respectively. The

apparent K_ms measured by the activity assays (**Table 3.3**) are reasonable relative to the estimated *in vivo* concentrations of these nucleotides (Appendix 3, **Table A3.11**).[76-83] The measurements suggest that the specificity sites of NrdE may often be occupied by TTP as the estimated concentration *in vivo* appears to be ≥ 8x its apparent K (Appendix 3, **Table A3.11**). If this scenario is true, it suggests that the ability of TTP to act as a general stimulator of NDP reduction (**Figure 3.7**) may be physiologically relevant. Further measurements of the *in vivo* levels of the dNTPs in *B. subtilis* using modern quantitative mass spectrometry techniques as well as the measurement of the affinity of the specificity sites for dATP will allow stronger conclusions to be drawn about what effectors will generally be occupying the specificity sites on the *B. subtilis* RNR *in vivo*. It should be noted that data from other class I systems supports the idea of full specificity site occupancy *in vivo* as similar conclusions were drawn or inferred for the Ia RNRs from mouse[34] and *T. brucei*,[55, 84] two organisms with lifestyles distinct from *B. subtilis*.

A comparison of the substrate apparent K_m values for the *B. subtilis* Ib RNR (**Table 3.4**) and the estimates of the *in vivo* concentrations of these nucleotides (Appendix 3) indicate that the enzyme will often be less than half saturated with substrate. This observation agrees with the generally held belief that the K_ms of enzymes often reflect the *in vivo* concentration of their substrates.[85] Indeed, a theoretical treatment by Athel Cornish-Bowden on evolutionary selective pressures on cells to maximize metabolic flux suggested that the K_ms of enzymes will often not differ from *in vivo* substrate concentrations by more than an order of magnitude.[86] For all four substrates, the ratio of K_m:$[NDP]_{in\ vivo}$ is generally consistent with this prediction (K_m:[ADP] = 0.12 – 1.12, K_m:[CDP] = 0.07 – 0.99, K_m:[GDP] = 0.08 – 2.80, K_m:[UDP] = 0.13 – 0.45). From the perspective of regulating the *in vivo* dNTP pools, non-saturating substrate concentrations also mean that the *B. subtilis* RNR will be more sensitive to changes in NDP concentrations (i.e. the

enzyme is operating under conditions where k_{cat} K_m^{-1} is important), thus allowing the enzyme specificity to adapt accordingly so that balanced pools of dNTPs are maintained.

3.4.2. The role of substrate competition in specificity regulation of class I RNRs. The specific activities reported with the *B. subtilis* RNR for each substrate with a particular effector (**Figure 3.7**) were measured by the conventional procedure of examining substrate/effector pairs one at a time. However, it has become increasingly apparent that competition of substrates for the catalytic site on class Ia RNRs also plays an important role in determining the prime effector for a given substrate.[18] This view has emerged from the development of an assay procedure, as pioneered by Christopher Mathews' group, capable of simultaneously monitoring the reduction of all four substrates.[87-89] In general, the results from such assays support the universal model of allosteric regulation of RNR substrate specificity.[30, 33, 87-89] Furthermore, the presence of all four substrates has revealed that most deviations from the general allosteric scheme,[18] such as TTP being a general activator of NDP reduction for the *E. coli* Ia RNR,[30, 53, 57] were artifacts likely arising from the use of single substrate/effector pairs in assays.

A comparison of the results obtained with Mn-loaded *B. subtilis* RNR using single substrate/effector pairs (**Figure 3.7**) with those of four substrate assays conducted with the *E. coli*[30] and *P. aeruginosa*[33] Ia RNRs reveals that, for the most part, the same trends are exhibited by the three systems. There are only two major differences observed: (1) with the *B. subtilis* RNR, TTP stimulated significant levels of reduction for CDP and UDP in addition to GDP, whereas with the Ia enzymes, the effector only stimulated GDP reduction; and (2) high levels of UDP reduction are attainable with *B. subtilis* RNR, whereas for the other enzymes, UDP was the poorest substrate. Given that the ability of TTP to act as a general stimulator of the *E. coli* Ia RNR was debunked in the four substrate assays,[30] it is probable that TTP-stimulated CDP and UDP reduction by the *B.*

subtilis enzyme is also not physiologically relevant. The highly efficient utilization of UDP will be further discussed below. Ultimately, the significance of the differences in allosteric regulation of the three systems can be determined if the *B. subtilis* RNR is examined with four substrate assays to see how substrate competition affects the specificity of the enzyme.

3.4.3. UDP is efficiently utilized by B. subtilis RNR. Probably the most distinct and surprising characteristic of the *B. subtilis* Ib RNR relative to most characterized Ia[34, 35, 53, 55-58] and the *S.* Typhimurium Ib[17] systems is the relatively high efficiency by which the enzyme can reduce UDP. UDP reduction is one of two ways utilized by cells to generate dUMP, the substrate for thymidylate synthase, in the *de novo* synthesis of TTP. The second route to dUMP is the deamination of deoxycytidine and/or its 5′-phosphorylated analogs acquired either exogenously or by the action of RNR on CDP. For many RNRs, UDP is often the least efficiently utilized substrate (see Appendix 3, **Table A3.6**) and suggests that deamination of dC analogs will likely be the major source of dUMP for TTP production in many organisms. Indeed, 75 – 80% of the TTP pool in *E. coli* has been reported to be derived from dCTP deamination.[90] In *B. subtilis*, both routes to dUMP are present (**Figure 3.16**) and enzymes that deaminate dC and dCMP provide two different ways in which deoxycytidine analogs can enter the dU pool for TTP synthesis. However, quantitation of the TTP pools of pyrimidine auxotrophs of *B. subtilis* fed [6-^{3}H]-uracil and unlabeled cytidine revealed that 55 – 65% of the total TTP had incorporated radiolabel,[78, 79] indicating that a small majority of the TTP synthesized by this bacterium is derived from UDP reduction by RNR. Furthermore, when deaminase activities for both cytidine/deoxycytidine analogs (Dcd and Cdd in **Figure 3.16**) were knocked out in these auxotrophs, the mutants grew just as well as wildtype and all of the TTP pool was found to be labeled with [^{3}H].[78] These results indicate that UDP reduction by *B. subtilis* RNR plays a significant role in production and maintenance of the *in vivo* TTP pools.

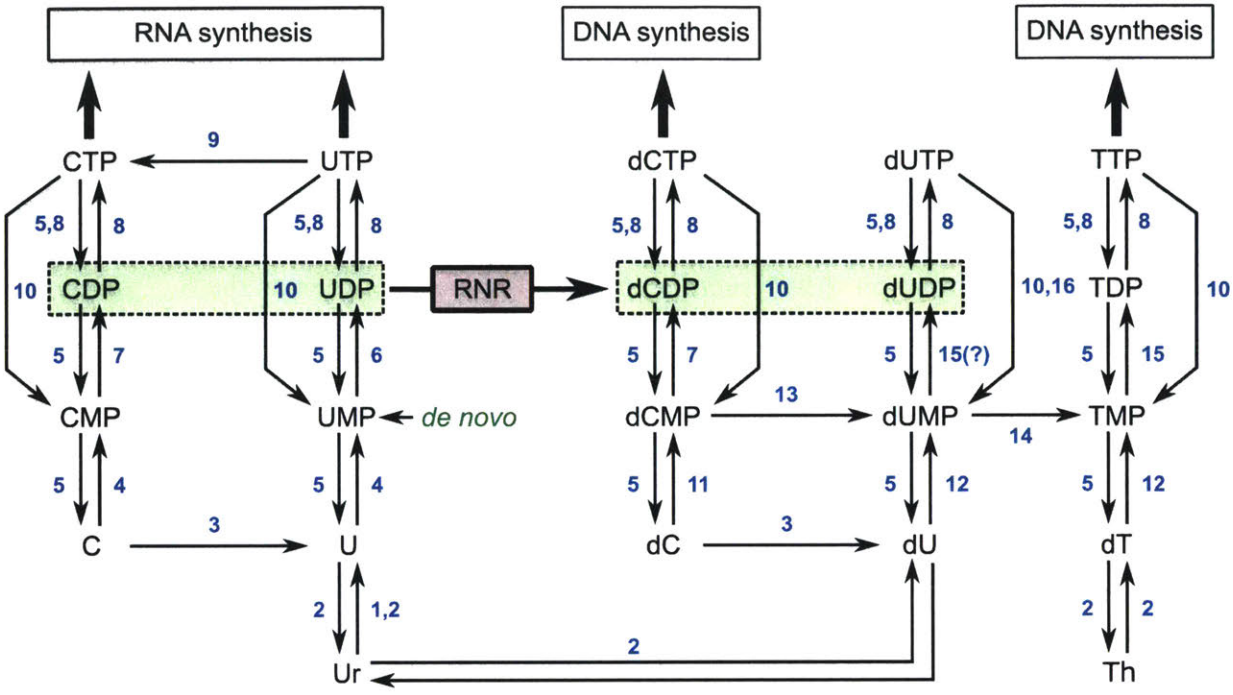


Figure 3.16. Pyrimidine nucleotide biosynthetic and salvage pathways in *B. subtilis*. With the exception of RNR (red box, middle of diagram), numbers are used to indicate the enzymes of these pathways as follows: 1 – uracil phosphoribosyltransferase (Upp, PyrR), 2 – pyrimidine nucleoside phosphorylase (Pdp), 3 – cytidine/deoxycytidine deaminase (Ccd), 4 – uridine kinase (Udk), 5 – specific or non-specific phosphatases, 6 – uridylate kinase (PyrH), 7 – cytidylate/deoxycytidylate kinase (Cmk), 8 – Nucleoside diphosphate kinase (Ndk), 9 – CTP synthetase (PyrG), 10 – Mn^{2+}-dependent pyrophosphohydrolase (YpgQ), 11 – deoxyadenosine/deoxycytidine kinase (Dck), 12 – thymidine kinase (Tdk), 13 – dCMP deaminase (ComEB, also known as Dcd), 14 – thymidylate synthase (ThyA or ThyB), 15 – thymidylate kinase (Tmk), and 16 – dUTP diphosphatase (YcnF or YosS).

3.4.4. The S. Typhimurium and B. subtilis Ib RNRs differ in response to ATP versus dATP.

As mentioned in the introduction to this chapter, the Fe-loaded *S.* Typhimurium Ib RNR is the only other case in which the allosteric regulation of a Ib system has been studied in any detail.[17] A comparison of the ability of single effectors to stimulate the reduction of each substrate with the Mn-loaded *B. subtilis* (**Figure 3.7**) and Fe-loaded *S.* Typhimurium Ib RNRs reveals that these enzymes behave quite similarly. The largest difference between the two systems was the stimulation of pyrimidine reduction by ATP versus dATP. In the case of the *S.* Typhimurium

RNR, dATP was reported as the best effector of pyrimidine reduction, whereas ATP was 90% and 60% as effective at stimulating UDP and CDP reduction, respectively.[17] The data in **Figure 3.7** show the reverse trend with the *B. subtilis* RNR: ATP was the best effector of pyrimidine reduction whereas dATP was ~50% as effective as ATP for both substrates. This difference is likely due to the ability of dATP to inhibit the *B. subtilis* enzyme (**Figure 3.8**) whereas the *S.* Typhimurium RNR was reported to exhibit no such inhibition up to a final concentration of 1 mM dATP.[7] However, it may also be possible that different Ib RNRs are inherently more sensitive to one effector versus the other, as was the case with the Fe-loaded *M. tuberculosis* enzyme. ATP was found to be a better stimulator of CDP reduction despite the fact that dATP up to 1 mM concentration was not inhibitory.[24] Characterization of the ability of these two effectors to stimulate pyrimidine reduction in class Ib RNRs from other organisms should help clarify this issue.

3.4.5. The ability of ATP to augment TTP-stimulated GDP reduction is similar to eukaryotic class Ia RNRs. Comparison of the steady-state parameters of the *B. subtilis* Ib RNR with characterized class Ia systems[34, 35, 53, 55-58] also reveals high levels of similarity. The catalytic efficiency (k_{cat} K_m^{-1}) of the *B. subtilis* enzyme for its substrates falls on the lower end of the spectrum observed with the Ia RNRs. However, it is noteworthy that the efficiencies measured with the *B. subtilis* system were made using a 1:1 ratio of subunits at protein concentrations that are relevant *in vivo* based on previous measurements (see Appendix 2).[2, 91-93] Catalytic efficiencies in other systems were determined using an excess of one subunit over the other and with nucleotide concentrations that likely fall outside of the range of physiological relevancy. The ability of ATP to increase the rate of TTP-stimulated GDP reduction by the *B. subtilis* RNR (**Figure 3.9**) is very reminiscent of the effects this nucleotide has on GDP reduction activity by eukaryotic RNRs.[33, 34,

[55, 58] In these cases, it is proposed that ATP occupies the overall activity sites on these RNRs and TTP the specificity sites in order to for such effects to occur. However, as *B. subtilis* NrdE does not contain an ATP-cone domain, the basis for the augmentative effect of ATP is going to be distinct from the eukaryotic cases. As discussed in Chapter 4, NrdE does appear to have more than one allosteric site to which effectors can bind to tune the overall activity of the enzyme, so it can be speculated that ATP might bind there or to an as yet uncharacterized site on NrdE.

3.4.6. Cooperative effects on RNR activity by purine substrates is very unusual. The observation that the purine substrates, of all the nucleotides tested, were the only ones that exhibited opposing cooperative effects on the activity of the *B. subtilis* RNR is an unusual result. Given that NrdE, in the absence of nucleotides, behaves as a monomer (**Figure 2.9B** and **2.10A**), one expectation going into the studies reported in this chapter was that the NDP reduction activities of the *B. subtilis* Ib RNR would exhibit positive cooperative effects with increasing effector, rather than substrate, concentrations. The rationale for this expectation was based on the location of the specificity site at the α_2 interface in class I RNRs.[19, 20] It was thought that the first effector nucleotide binding event would stabilize loosely-associated NrdE dimers somewhat, the result of which would be to "pre-form" the second specificity site for binding of a second equivalent of effector. However, this expectation was incorrect as the results of the steady-state assays demonstrated that the effectors bound non-cooperatively, and the SV-AUC results with NrdE in the presence of TTP alone (**Figure 3.13A**) or in combination with GDP (**Figure 3.14**) showed that in the effector was unable to induce and/or stabilize further dimerization of the protein in the concentration range tested. On the other hand, as dATP was found to have a stabilizing effect on NrdE dimers by SV-AUC (**Figure 3.12**), cooperative effects may still be relevant for this effector nucleotide. However, they are currently masked due to the inhibitory nature of the nucleotide at

concentrations above ~5 μM. If a mutation in NrdE can be found that prevents dATP inhibition of the *B. subtilis* RNR, activity assays can be conducted to test this hypothesis.

Cooperativity, in any case, implies that NrdE is a dimer at some point during the catalytic cycle. As suggested by the SV-AUC data presented in Chapter 2 for 1:1 mixtures of NrdE:NrdF (**Figure 2.10B**), it is probable that binding of β_2 induces α to dimerize in order to form the catalytically relevant $\alpha_2\beta_2$ complex. If this is the case, then the observed negative cooperativity in ADP reduction activity can likely be explained by the ability of the RNR to be inhibited by the product dADP (**Figure 3.10**). As the concentration of the product builds up, it is better able to bind to its inhibitory site on NrdE and, thus, induce a conformational change that causes NrdE to re-monomerize or prevents the second protomer from binding substrate. The fact that ATP can delay this effect, but not reverse it (**Figure 3.10A**), indicates that the mechanism by which dADP (and dATP, for that matter) is distinct from that observed in the class Ia RNRs. Whether dADP occupies the active site or binds with weaker affinity to the dATP site in order to cause RNR inhibition is unknown. Further investigations can be made if a mutation can be identified that prevents the *B. subtilis* RNR from becoming inhibited by dATP. On the other hand, the positive cooperativity observed on RNR activity for GDP reduction (**Table 3.4**) and the observed monomerization of NrdE by SV-AUC (**Figure 3.13C**) is an extremely puzzling result for which an explanation is not available. Further studies are required in order to gain an understanding of these unusual attributes of the *B. subtilis* Ib RNR.

3.4.7. Quaternary structural alterations are still the mechanism by which dATP inhibits the B. subtilis RNR. The results of sedimenting 1:1 mixtures of His_6-NrdE:His_6-NrdF in the presence of dATP (**Figure 3.15**) indicate that the mechanism of inhibition of the *B. subtilis* Ib RNR by this nucleotide follows the same principle as observed for the class Ia enzymes, that

principle being formation of alternate inhibitory quaternary structures. Since *B. subtilis* NrdE does not have an ATP cone domain, the structural basis for formation of the inhibitory complexes is going to be distinctly different. Preliminary cryo-electron microscopy results (Ed Brignole, Yan Zhang, Cathy Drennan, JoAnne Stubbe, unpublished) indicated that *B. subtilis* RNR was able to form fibril-like structures in the presence of dATP, thus suggesting that such structures may be the inhibited form of the enzyme. Fibril formation has been reported previously for many different enzymes as a potential mechanism for activity regulation.[94] Further studies of this phenomenon as a possible mechanism by which dATP inhibits the *B. subtilis* Ib RNR are warranted.

Despite the ability to generate a fibril with the *B. subtilis* Ib RNR and the fascinating ring structures of the class Ia enzymes *in vitro*, the issue remains as to whether or not they have any relevance *in vivo*. Knowledge of the different quaternary structures that can be formed, the kinetics of substrate turn over, the strength and concentration dependence of binding affinities, and the *in vivo* concentrations of the components are all required in order to connect *in vitro* observations with *in vivo* functionality and relevance. The only case in which such a connection has been tentatively established is for the formation of an α_6 quaternary structure of the human Ia RNR in response to the antineoplastic therapeutic clofarabine.[95, 96] In other cases, the different quaternary structures reported for class Ia RNRs[27-30, 32, 34, 65] have often been observed at both protein and nucleotide concentrations that are considered to be non-physiological. One of the major difficulties in establishing the physiological relevance of the different RNR quaternary structures is the lack of minimally perturbative methods that allow visualization of individual proteins inside the cell. As scientific technology improves, new method development should allow this obstacle to be overcome.

3.5. REFERENCES

1. Thelander, L., Sjöberg, B. R., and Eriksson, S. (1978) Ribonucleoside Diphosphate Reductase (*Escherichia coli*), *Method Enzymol. 51*, 227-237.

2. Zhang, Y., and Stubbe, J. (2011) *Bacillus subtilis* Class Ib Ribonucleotide Reductase is a Dimanganese(III)-Tyrosyl Radical Enzyme, *Biochemistry 50*, 5615-5623.

3. Cotruvo, J. A., and Stubbe, J. (2010) An Active Dimanganese(III)-Tyrosyl Radical Cofactor in *Escherichia coli* Class Ib Ribonucleotide Reductase, *Biochemistry 49*, 1297-1309.

4. Cotruvo, J. A., and Stubbe, J. (2011) *Escherichia coli* Class Ib Ribonucleotide Reductase Contains a Dimanganese(III)-Tyrosyl Radical Cofactor *in vivo*, *Biochemistry 50*, 1672-1681.

5. Cox, N., Ogata, H., Stolle, P., Reijerse, E., Auling, G., and Lubitz, W. (2010) A Tyrosyl-Dimanganese Coupled Spin System is the Native Metalloradical Cofactor of the R2F Subunit of the Ribonucleotide Reductase of *Corynebacterium ammoniagenes*, *J. Am. Chem. Soc. 132*, 11197-11213.

6. Huque, Y., Fieschi, F., Torrents, E., Gibert, I., Eliasson, R., Reichard, P., Sahlin, M., and Sjöberg, B. M. (2000) The Active Form of the R2F Protein of Class Ib Ribonucleotide Reductase from *Corynebacterium ammoniagenes* is a Diferric Protein, *J. Biol. Chem. 275*, 25365-25371.

7. Jordan, A., Pontis, E., Atta, M., Krook, M., Gibert, I., Barbé, J., and Reichard, P. (1994) A Second Class I Ribonucleotide Reductase in *Enterobacteriaceae* - Characterization of the *Salmonella typhimurium* Enzyme, *Proc. Natl. Acad. Sci. U. S. A. 91*, 12892-12896.

8. Willing, A., Follmann, H., and Auling, G. (1988) Ribonucleotide Reductase of *Brevibacterium ammoniagenes* is a Manganese Enzyme, *Eur. J. Biochem. 170*, 603-611.

9. Cotruvo, J. A., and Stubbe, J. (2008) NrdI, A Flavodoxin Involved in Maintenance of the Diferric-Tyrosyl Radical Cofactor in *Escherichia coli* Class Ib Ribonucleotide Reductase, *Proc. Natl. Acad. Sci. U. S. A. 105*, 14383-14388.

10. Crona, M., Torrents, E., Røhr, A. K., Hofer, A., Furrer, E., Tomter, A. B., Andersson, K. K., Sahlin, M., and Sjöberg, B. M. (2011) NrdH-Redoxin Protein Mediates High Enzyme Activity in Manganese-Reconstituted Ribonucleotide Reductase from *Bacillus anthracis*, *J. Biol. Chem. 286*, 33053-33060.

11. Gustafsson, T. N., Sahlin, M., Lu, J., Sjöberg, B. M., and Holmgren, A. (2012) *Bacillus anthracis* Thioredoxin Systems, Characterization and Role as Electron Donors for Ribonucleotide Reductase, *J. Biol. Chem. 287*, 39686-39697.

12. Jordan, A., Åslund, F., Pontis, E., Reichard, P., and Holmgren, A. (1997) Characterization of *Escherichia coli* NrdH - A Glutaredoxin-like Protein with a Thioredoxin-like Activity Profile, *J. Biol. Chem. 272*, 18044-18050.

13. Jordan, A., Pontis, E., Åslund, F., Hellman, U., Gibert, I., and Reichard, P. (1996) The Ribonucleotide Reductase System of *Lactococcus lactis* - Characterization of an NrdEF Enzyme and a New Electron Transport Protein, *J. Biol. Chem. 271*, 8779-8785.

14. Rabinovitch, I., Yanku, M., Yeheskel, A., Cohen, G., Borovok, I., and Aharonowitz, Y. (2010) *Staphylococcus aureus* NrdH Redoxin Is a Reductant of the Class Ib Ribonucleotide Reductase, *J. Bacteriol. 192*, 4963-4972.

15. Roca, I., Torrents, E., Sahlin, M., Gibert, I., and Sjöberg, B. M. (2008) NrdI Essentiality for Class Ib Ribonucleotide Reduction in *Streptococcus pyogenes*, *J. Bacteriol. 190*, 4849-4858.

16. Stehr, M., Schneider, G., Åslund, F., Holmgren, A., and Lindqvist, Y. (2001) Structural Basis for the Thioredoxin-like Activity Profile of the Glutaredoxin-like NrdH-Redoxin from *Escherichia coli*, *J. Biol. Chem. 276*, 35836-35841.

17. Eliasson, R., Pontis, E., Jordan, A., and Reichard, P. (1996) Allosteric Regulation of the Third Ribonucleotide Reductase (NrdEF Enzyme) from Enterobacteriaceae, *J. Biol. Chem. 271*, 26582-26587.

18. Hofer, A., Crona, M., Logan, D. T., and Sjöberg, B. M. (2012) DNA Building Blocks: Keeping Control of Manufacture, *Crit. Rev. Biochem. Mol. Biol. 47*, 50-63.

19. Eriksson, M., Uhlin, U., Ramaswamy, S., Ekberg, M., Regnström, K., Sjöberg, B. M., and Eklund, H. (1997) Binding of Allosteric Effectors to Ribonucleotide Reductase Protein R1: Reduction of Active-Site Cysteines Promotes Substrate Binding, *Structure 5*, 1077-1092.

20. Uppsten, M., Färnegårdh, M., Jordan, A., Eliasson, R., Eklund, H., and Uhlin, U. (2003) Structure of the Large Subunit of Class Ib Ribonucleotide Reductase from *Salmonella typhimurium* and Its Complexes with Allosteric Effectors, *J. Mol. Biol. 330*, 87-97.

21. Jordan, A., Aragall, E., Gibert, I., and Barbé, J. (1996) Promoter Identification and Expression Analysis of *Salmonella typhimurium* and *Escherichia coli* NrdEF Operons Encoding One of Two Class I Ribonucleotide Reductases Present in Both Bacteria, *Mol. Microbiol. 19*, 777-790.

22. Panosa, A., Roca, I., and Gibert, I. (2010) Ribonucleotide Reductases of *Salmonella* Typhimurium: Transcriptional Regulation and Differential Role in Pathogenesis, *PLoS One 5*, 11.

23. Lundin, D., Torrents, E., Poole, A. M., and Sjöberg, B. M. (2009) RNRdb, a Curated Database of the Universal Enzyme Family Ribonucleotide Reductase, Reveals a High Level of Misannotation in Sequences Deposited to Genbank, *BMC Genomics 10*, 8.

24. Yang, F. D., Curran, S. C., Li, L. S., Avarbock, D., Graf, J. D., Chua, M. M., Lui, G. Z., Salem, J., and Rubin, H. (1997) Characterization of Two Genes Encoding the *Mycobacterium tuberculosis* Ribonucleotide Reductase Small Subunit, *J. Bacteriol. 179*, 6408-6415.

25. Makhlynets, O., Boal, A. K., Rhodes, D. V., Kitten, T., Rosenzweig, A. C., and Stubbe, J. (2014) *Streptococcus sanguinis* Class Ib Ribonucleotide Reductase: High Activity with Both Iron and Manganese Cofactors and Structural Insights, *J. Biol. Chem. 289*, 6259-6272.

26. Aravind, L., Wolf, Y. I., and Koonin, E. V. (2000) The ATP-Cone: An Evolutionarily Mobile, ATP-Binding Regulatory Domain, *J. Mol. Microbiol. Biotechnol. 2*, 191-194.

27. Ando, N., Brignole, E. J., Zimanyi, C. M., Funk, M. A., Yokoyama, K., Asturias, F. J., Stubbe, J., and Drennan, C. L. (2011) Structural Interconversions Modulate Activity of *Escherichia coli* Ribonucleotide Reductase, *Proc. Natl. Acad. Sci. U. S. A. 108*, 21046-21051.

28. Ando, N., Li, H. R., Brignole, E. J., Thompson, S., McLaughlin, M. I., Page, J. E., Asturias, F. J., Stubbe, J., and Drennan, C. L. (2016) Allosteric Inhibition of Human Ribonucleotide Reductase by dATP Entails the Stabilization of a Hexamer, *Biochemistry 55*, 373-381.

29. Brown, N. C., and Reichard, P. (1969) Ribonucleoside Diphosphate Reductase - Formation of Active and Inactive Complexes of Proteins B1 and B2, *J. Mol. Biol. 46*, 25-38.

30. Rofougaran, R., Crona, M., Vodnala, M., Sjöberg, B. M., and Hofer, A. (2008) Oligomerization Status Directs Overall Activity Regulation of the *Escherichia coli* Class Ia Ribonucleotide Reductase, *J. Biol. Chem. 283*, 35310-35318.

31. Rhodes, D. V., Crump, K. E., Makhlynets, O., Snyder, M., Ge, X. C., Xu, P., Stubbe, J., and Kitten, T. (2014) Genetic Characterization and Role in Virulence of the Ribonucleotide Reductases of *Streptococcus sanguinis*, *J. Biol. Chem. 289*, 6273-6287.

32. Fairman, J. W., Wijerathna, S. R., Ahmad, M. F., Xu, H., Nakano, R., Jha, S., Prendergast, J., Welin, R. M., Flodin, S., Roos, A., Nordlund, P., Li, Z., Walz, T., and Dealwis, C. G. (2011) Structural Basis for Allosteric Regulation of Human Ribonucleotide Reductase by Nucleotide-Induced Oligomerization, *Nat. Struct. Mol. Biol. 18*, 316-322.

33. Jonna, V. R., Crona, M., Rofougaran, R., Lundin, D., Johansson, S., Brännström, K., Sjöberg, B. M., and Hofer, A. (2015) Diversity in Overall Activity Regulation of Ribonucleotide Reductase, *J. Biol. Chem. 290*, 17339-17348.

34. Kashlan, O. B., Scott, C. P., Lear, J. D., and Cooperman, B. S. (2002) A Comprehensive Model for the Allosteric Regulation of Mammalian Ribonucleotide Reductase. Functional Consequences of ATP- and dATP-Induced Oligomerization of the Large Subunit, *Biochemistry 41*, 462-474.

35. Scott, C. P., Kashlan, O. B., Lear, J. D., and Cooperman, B. S. (2001) A Quantitative Model for Allosteric Control of Purine Reduction by Murine Ribonucleotide Reductase, *Biochemistry 40*, 1651-1661.

36. Korbie, D. J., and Mattick, J. S. (2008) Touchdown PCR for Increased Specificity and Sensitivity in PCR Amplification, *Nat. Protoc. 3*, 1452-1456.

37. Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D., and Bairoch, A. (2005) Protein Identification and Analysis Tools on the ExPASy Server., In *The Proteomics Protocols Handbook* (Walker, J. M., Ed.), pp 571-607, Humana Press, Totowa, NJ.

38. Smith, P. K., Krohn, R. I., Hermanson, G. T., Mallia, A. K., Gartner, F. H., Provenzano, M. D., Fujimoto, E. K., Goeke, N. M., Olson, B. J., and Klenk, D. C. (1985) Measurement of Protein Using Bicinchoninic Acid, *Anal. Biochem. 150*, 76-85.

39. Mayhew, S. G., and Massey, V. (1969) Purification and Characterization of Flavodoxin from *Peptostreptococcus elsdenii*, *J. Biol. Chem. 244*, 794-802.

40. Edelhoch, H. (1967) Spectroscopic Determination of Tryptophan and Tyrosine in Proteins, *Biochemistry 6*, 1948-1954.

41. Gill, S. C., and von Hippel, P. H. (1989) Calculation of Protein Extinction Coefficients from Amino Acid Sequence Data, *Anal. Biochem. 182*, 319-326.

42. Steeper, J. R., and Steuart, C. D. (1970) A Rapid Assay for CDP Reductase Activity in Mammalian Cell Extracts, *Anal. Biochem. 34*, 123-130.

43. Segel, I. H. (1993) Multisite and Allosteric Enzymes, In *Enzyme Kinetics: Behavior and Analysis of Rapid Equilibrium and Steady-State Enzyme Systems* Wiley Classics Library ed., pp 346-464, John Wiley & Sons, Inc., New York.

44. Motulsky, H., and Christopoulos, A. (2004) Comparing Models Using the Extra Sum-of-squares F Test, In *Fitting Models to Biological Data Using Linear and Nonlinear Regression*, pp 138-142, Oxford University Press, New York.

45. (1996) Characteristics of Nucleic Acids, In *Current Protocols in Molecular Biology*, pp A.1D.1-A.1D.11, John Wiley & Sons, Inc.

46. Zhao, H. Y., Ghirlando, R., Piszczek, G., Curth, U., Brautigam, C. A., and Schuck, P. (2013) Recorded Scan Times can Limit the Accuracy of Sedimentation Coefficients in Analytical Ultracentrifugation, *Anal. Biochem. 437*, 104-108.

47. Zhao, H. Y., Brown, P. H., Balbo, A., Fernandez-Alonso, M. D., Polishchuck, N., Chaudhry, C., Mayer, M. L., Ghirlando, R., and Schuck, P. (2010) Accounting for Solvent Signal Offsets in the Analysis of Interferometric Sedimentation Velocity Data, *Macromol. Biosci. 10*, 736-745.

48. Laue, T. M., Shah, B. D., Ridgeway, T. M., and Pelletier, S. L. (1992) Computer-aided Interpretation of Analytical Sedimentation Data for Proteins, In *Analytical Ultracentrifugation in Biochemistry and Polymer Science* (Harding, S. E., Rowe, A. J., and Horton, J. C., Eds.), pp 90-125, Royal Society of Chemistry, Cambridge.

49. Butt, T. R., Edavettal, S. C., Hall, J. P., and Mattern, M. R. (2005) SUMO Fusion Technology for Difficult-to-express Proteins, *Protein Expr. Purif. 43*, 1-9.

50. Marblestone, J. G., Edavettal, S. C., Lim, Y., Lim, P., Zuo, X., and Butt, T. R. (2006) Comparison of SUMO Fusion Technology with Traditional Gene Fusion Systems: Enhanced Expression and Solubility with SUMO, *Protein Sci. 15*, 182-189.

51. Bendezú, F. O., Hale, C. A., Bernhardt, T. G., and de Boer, P. A. J. (2009) RodZ (YfgA) is Required for Proper Assembly of the MreB Actin Cytoskeleton and Cell Shape in *E. coli*, *Embo J. 28*, 193-204.

52. Murray, T. A., and Swenson, R. P. (2003) Mechanism of Flavin Mononucleotide Cofactor Binding to the *Desulfovibrio vulgaris* Flavodoxin. 1. Kinetic Evidence for Cooperative Effects Associated with the Binding of Inorganic Phosphate and The 5'-Phosphate Moiety of the Cofactor, *Biochemistry 42*, 2307-2316.

53. Larsson, A., and Reichard, P. (1966) Enzymatic Synthesis of Deoxyribonucleotides. X. Reduction of Purine Ribonucleotides; Allosteric Behavior and Substrate Specificity of the Enzyme System from *Escherichia coli* B, *J. Biol. Chem. 241*, 2540-2549.

54. Ge, J., Yu, G. X., Ator, M. A., and Stubbe, J. (2003) Pre-Steady-State and Steady-State Kinetic Analysis of *E. coli* Class I Ribonucleotide Reductase, *Biochemistry 42*, 10071-10083.

55. Hofer, A., Ekanem, J. T., and Thelander, L. (1998) Allosteric Regulation of *Trypanosoma brucei* Ribonucleotide Reductase Studied *in vitro* and *in vivo*, *J. Biol. Chem. 273*, 34098-34104.

56. Holmgren, A., Reichard, P., and Thelander, L. (1965) Enzymatic Synthesis of Deoxyribonucleotides, VIII. Effects of ATP and dATP in the CDP Reductase System from *E. coli*, *Proc. Natl. Acad. Sci. U. S. A. 54*, 830-836.

57. Larsson, A., and Reichard, P. (1966) Enzymatic Synthesis of Deoxyribonucleotides. IX. Allosteric Effects in the Reduction of Pyrimidine Ribonucleotides by the Ribonucleoside Diphosphate Reductase System of *Escherichia coli*, *J. Biol. Chem. 241*, 2533-2539.

58. Eriksson, S., Thelander, L., and Åkerman, M. (1979) Allosteric Regulation of Calf Thymus Ribonucleoside Diphosphate Reductase, *Biochemistry 18*, 2948-2952.

59. Berglund, O. (1972) Ribonucleoside Diphosphate Reductase Induced by Bacteriophage T4. II. Allosteric Regulation of Substrate Specificity and Catalytic Activity, *J. Biol. Chem. 247*, 7276-7281.

60. The physiological significance of dCTP-stimulated NDP reduction by the T4 phage Ia RNR is unlikely given that the virus uses hydroxymethyl-dCTP and, therefore, expresses a dCTP pyrophosphohydrolase to prevent dCTP incorporation during DNA replication.[97] On the other hand, the physiological significane of dCTP-stimulated NDP reduction by the *T. acidophilum* class II RNR is unknown.

61. Eliasson, R., Pontis, E., Jordan, A., and Reichard, P. (1999) Allosteric Control of Three B_{12}-dependent (Class II) Ribonucleotide Reductases - Implications for the Evolution of Ribonucleotide Reduction, *J. Biol. Chem. 274*, 7182-7189.

62. Zimanyi, C. M., (2013) *Structural Studies of Allosteric Regulation in the Class Ia Ribonucleotide Reductase from Escherichia coli*. Ph.D. Thesis, Massachusetts Institute of Technology.

63. Zimanyi, C. M., Chen, P. Y. T., Kang, G., Funk, M. A., and Drennan, C. L. (2016) Molecular Basis for Allosteric Specificity Regulation in Class Ia Ribonucleotide Reductase from *Escherichia coli*, *eLife 5*, 23.

64. Ortega, A., Amorós, D., and de la Torre, J. G. (2011) Prediction of Hydrodynamic and Other Solution Properties of Rigid Proteins from Atomic- and Residue-Level Models, *Biophys. J. 101*, 892-898.

65. Zimanyi, C. M., Ando, N., Brignole, E. J., Asturias, F. J., Stubbe, J., and Drennan, C. L. (2012) Tangled Up in Knots: Structures of Inactivated Forms of *E. coli* Class Ia Ribonucleotide Reductase, *Structure 20*, 1374-1383.

66. Brown, N. C., Larsson, A., and Reichard, P. (1967) On Subunit Structure of Ribonucleoside Diphosphate Reductase, *J. Biol. Chem. 242*, 4272-4273.

67. Ormö, M., and Sjöberg, B. M. (1990) An Ultrafiltration Assay for Nucleotide Binding to Ribonucleotide Reductase, *Anal. Biochem. 189*, 138-141.

68. von Döbeln, U., and Reichard, P. (1976) Binding of Substrates to *Escherichia coli* Ribonucleotide Reductase, *J. Biol. Chem. 251*, 3616-3622.

69. Schuck, P. (2010) Sedimentation Patterns of Rapidly Reversible Protein Interactions, *Biophys. J. 98*, 2005-2013.

70. Thelander, L. (1973) Physicochemical Characterization of Ribonucleoside Diphosphate Reductase from *Escherichia coli*, *J. Biol. Chem. 248*, 4591-4601.

71. Schuck, P. (2003) On the Analysis of Protein Self-Association by Sedimentation Velocity Analytical Ultracentrifugation, *Anal. Biochem. 320*, 104-124.

72. Schuck, P., and Rossmanith, P. (2000) Determination of the Sedimentation Coefficient Distribution by Least-Squares Boundary Modeling, *Biopolymers 54*, 328-341.

73. Minnihan, E. C., Ando, N., Brignole, E. J., Olshansky, L., Chittuluru, J., Asturias, F. J., Drennan, C. L., Nocera, D. G., and Stubbe, J. (2013) Generation of a Stable, Aminotyrosyl Radical-Induced $\alpha_2\beta_2$ Complex of *Escherichia coli* Class Ia Ribonucleotide Reductase, *Proc. Natl. Acad. Sci. U. S. A. 110*, 3835-3840.

74. Uhlin, U., and Eklund, H. (1994) Structure of Ribonucleotide Reductase Protein R1, *Nature 370*, 533-539.

75. Crona, M., Furrer, E., Torrents, E., Edgell, D. R., and Sjöberg, B. M. (2010) Subunit and Small-Molecule Interaction of Ribonucleotide Reductases Via Surface Plasmon Resonance Biosensor Analyses, *Protein Eng. Des. Sel. 23*, 633-641.

76. Meyer, F. M., and Stülke, J. (2013) Malate Metabolism in *Bacillus subtilis*: Distinct Roles for Three Classes of Malate-Oxidizing Enzymes, *FEMS Microbiol. Lett. 339*, 17-22.

77. Meyer, H., Weidmann, H., Mäder, U., Hecker, M., Völker, U., and Lalk, M. (2014) A Time Resolved Metabolomics Study: The Influence of Different Carbon Sources During Growth and Starvation of *Bacillus subtilis*, *Mol. Biosyst. 10*, 1812-1823.

78. Møllgaard, H., and Neuhard, J. (1978) Deoxycytidylate Deaminase from *Bacillus subtilis* - Purification, Characterization, and Physiological Function, *J. Biol. Chem. 253*, 3536-3542.

79. Møllgaard, H. (1980) Deoxyadenosine/Deoxycytidine Kinase from *Bacillus subtilis*: Purification, Characterization, and Physiological Function, *J. Biol. Chem. 255*, 8216-8220.

80. Neuhard, J., Price, A. R., Schack, L., and Thomassen, E. (1978) Two Thymidylate Synthetases in *Bacillus subtilis*, *Proc. Natl. Acad. Sci. U. S. A. 75*, 1194-1198.

81. Saxild, H. H., and Nygaard, P. (1987) Genetic and Physiological Characterization of *Bacillus subtilis* Mutants Resistant to Purine Analogs, *J. Bacteriol. 169*, 2977-2983.

82. Saxild, H. H., and Nygaard, P. (1991) Regulation of Levels of Purine Biosynthetic Enzymes in *Bacillus subtilis* - Effects of Changing Purine Nucleotide Pools, *J. Gen. Microbiol. 137*, 2387-2394.

83. Soga, T., Ohashi, Y., Ueno, Y., Naraoka, H., Tomita, M., and Nishioka, T. (2003) Quantitative Metabolome Analysis Using Capillary Electrophoresis Mass Spectrometry, *J. Proteome Res. 2*, 488-494.

84. Hofer, A., Schmidt, P. P., Gräslund, A., and Thelander, L. (1997) Cloning and Characterization of the R1 and R2 Subunits of Ribonucleotide Reductase from *Trypanosoma brucei*, *Proc. Natl. Acad. Sci. U. S. A. 94*, 6959-6964.

85. Benner, S. A. (1989) Enzyme Kinetics and Molecular Evolution, *Chem. Rev. 89*, 789-806.

86. Cornish-Bowden, A. (1976) The Effect of Natural Selection on Enzymic Catalysis, *J. Mol. Biol. 101*, 1-9.

87. Chimploy, K., and Mathews, C. K. (2001) Mouse Ribonucleotide Reductase Control - Influence of Substrate Binding Upon Interactions with Allosteric Effectors, *J. Biol. Chem. 276*, 7093-7100.

88. Hendricks, S. P., and Mathews, C. K. (1997) Regulation of T4 Phage Aerobic Ribonucleotide Reductase - Simultaneous Assay of the Four Activities, *J. Biol. Chem. 272*, 2861-2865.

89. Hendricks, S. P., and Mathews, C. K. (1998) Allosteric Regulation of Vaccinia Virus Ribonucleotide Reductase, Analyzed by Simultaneous Monitoring of its Four Activities, *J. Biol. Chem. 273*, 29512-29518.

90. Weiss, B., and Wang, L. H. (1994) De Novo Synthesis of Thymidylate via Deoxycytidine in *dcd* (dCTP Deaminase) Mutants of *Escherichia coli*, *J. Bacteriol. 176*, 2194-2199.

91. Maaβ, S., Sievers, S., Zühlke, D., Kuzinski, J., Sappa, P. K., Muntel, J., Hessling, B., Bernhardt, J., Sietmann, R., Völker, U., Hecker, M., and Becher, D. (2011) Efficient, Global-Scale Quantification of Absolute Protein Amounts by Integration of Targeted Mass Spectrometry and Two-Dimensional Gel-Based Proteomics, *Anal. Chem. 83*, 2677-2684.

92. Maaβ, S., Wachlin, G., Bernhardt, J., Eymann, C., Fromion, V., Riedel, K., Becher, D., and Hecker, M. (2014) Highly Precise Quantification of Protein Molecules per Cell During Stress and Starvation Responses in *Bacillus subtilis*, *Mol. Cell. Proteomics 13*, 2260-2276.

93. Muntel, J., Fromion, V., Goelzer, A., Maaβ, S., Mäder, U., Büttner, K., Hecker, M., and Becher, D. (2014) Comprehensive Absolute Quantification of the Cytosolic Proteome of *Bacillus subtilis* by Data Independent, Parallel Fragmentation in Liquid Chromatography/Mass Spectrometry (LC/MS^E), *Mol. Cell. Proteomics 13*, 1008-1019.

94. O'Connell, J. D., Zhao, A., Ellington, A. D., and Marcotte, E. M. (2012) Dynamic Reorganization of Metabolic Enzymes into Intracellular Bodies, In *Annual Review of Cell and Developmental Biology, Vol 28* (Schekman, R., Ed.), pp 89-111, Annual Reviews, Palo Alto.

95. Aye, Y., Brignole, E. J., Long, M. J. C., Chittuluru, J., Drennan, C. L., Asturias, F. J., and Stubbe, J. (2012) Clofarabine Targets the Large Subunit (α) of Human Ribonucleotide Reductase in Live Cells by Assembly into Persistent Hexamers, *Chem. Biol. 19*, 799-805.

96. Aye, Y., and Stubbe, J. (2011) Clofarabine 5'-Di and -Triphosphates Inhibit Human Ribonucleotide Reductase by Altering the Quaternary Structure of its Large Subunit, *Proc. Natl. Acad. Sci. U. S. A. 108*, 9815-9820.

97. Warren, R. A. J. (1980) Modified Bases in Bacteriophage DNAs, *Annu. Rev. Microbiol. 34*, 137-158.

Chapter 4

Discovery of a Tightly Bound Equivalent of dAMP Associated with *Bacillus subtilis* NrdE that Modulates Inhibition of Ribonucleotide Reductase by dATP

4.1. INTRODUCTION

Presented in this chapter are the results of studies examining the structural and functional consequences of a tightly bound equivalent of 2′-deoxyadenosine 5′-monophosphate (dAMP) that was found associated with *Bacillus subtilis* NrdE on RNR activity and regulation. The discovery of this nucleotide resulted from an alternative interpretation of the SV-AUC data reported in Chapter 3 for the sedimentation of NrdE in the presence of nucleotides. As described there, these experiments were monitored using interference optics, which is a higher resolution detection method relative to A_{280}. Consequently, a comparison of $s_{20,w}$ distributions generated for ~2 μM samples of NrdE revealed two peaks when monitoring by interference (**Figure 3.12A**, top distribution) versus only one peak when monitoring by A_{280} (**Figure 3.11A**, black trace). For the single species detected by A_{280}, the hydrodynamic properties calculated from the fits to the c(s) model are consistent with NrdE monomer. The same is true for the lower molecular weight species detected by interference (**Figure 3.12A**, peak **1**), which constitutes the majority (~70 – 75%) of the total protein. This observation raised questions about the nature of the structure adopted by the remaining 25 – 30% of the NrdE that sedimented as a larger species (**Figure 3.12A**, peak **2**).

The results of the SV-AUC experiments examining NrdE with increasing concentrations of dATP clearly demonstrated that the larger species corresponded to NrdE dimer (**Figure 3.12A**). Prior to obtaining these results, however, the species was initially proposed to be an alternative conformer of NrdE monomer. This interpretation was largely based on the SV-AUC experiments examining NrdE with increasing GDP concentrations (**Figure 3.13C**), the results of which showed that saturating amounts of GDP cause the larger NrdE species to fully convert into the monomeric state. Such a result is very unusual for class I RNRs, and it was difficult to conceive that substrate could cause NrdE to monomerize, let alone alter its tertiary structure significantly. An altered

conformer of NrdE monomer was also appealing because it might provide an explanation for the unexpected dATP inhibition of the *B. subtilis* RNR discussed in the previous chapters.

In considering the possibility of an altered monomer conformation, one major question that needed to be addressed was what caused such a large change in the tertiary structure of NrdE. An increase in the $s_{20,w}$ (~6.4 S in the experiments reported in Chapter 3) requires the alternate state to be significantly more compact relative to structures of NrdE obtained by X-ray crystallography (PDB 1PEM[1]). One hypothesis put forth that was consistent with compaction of the tertiary structure was that a tightly associated ligand, presumably dATP, was bound to NrdE. A simple test of this hypothesis was to denature NrdE and determine if any nucleotide could be identified in the supernatant. Unexpectedly, significant absorption at 260 nm, a λ_{max} consistent with adenine-containing nucleotides, was observed with such a supernatant, thus launching the studies reported in this chapter.

The hypothesis that a ligand tightly bound to NrdE might be the cause of an alternate conformation of the protein arose from previous work by Aaron Hoskins, a former Stubbe laboratory member, and the lab of Steven Ealick (Cornell University) on the enzyme formylglycinamide ribonucleotide amidotransferase (FGAR-AT, also known as PurL). PurL catalyzes the conversion of FGAR, ATP, and Gln to formylglycinamidine ribonucleotide (FGAM), ADP, P_i, and Glu in the fourth step of purine biosynthesis (**Figure 4.1**). Two forms of PurLs, large and small, have been described in the literature. Large PurLs (lgPurL), exemplified by the *Salmonella enterica* Serovar Typhimurium enzyme, are a single polypeptide of 140 kDa with three domains: a glutaminase domain, a FGAM synthetase domain, and a N-terminal domain.[2] Small PurLs (smPurL), exemplified by the *B. subtilis* and *Thermotoga maritima* enzymes, are 60 – 80 kDa proteins that are structurally homologous to the FGAM synthetase domain of lgPurL.[3] The

active FGAR-AT requires formation of a protein complex consisting of smPurL, PurQ (homologous to the glutaminase domain of lgPurL), and two copies of PurS, a protein for which the function is currently unknown.[4, 5]

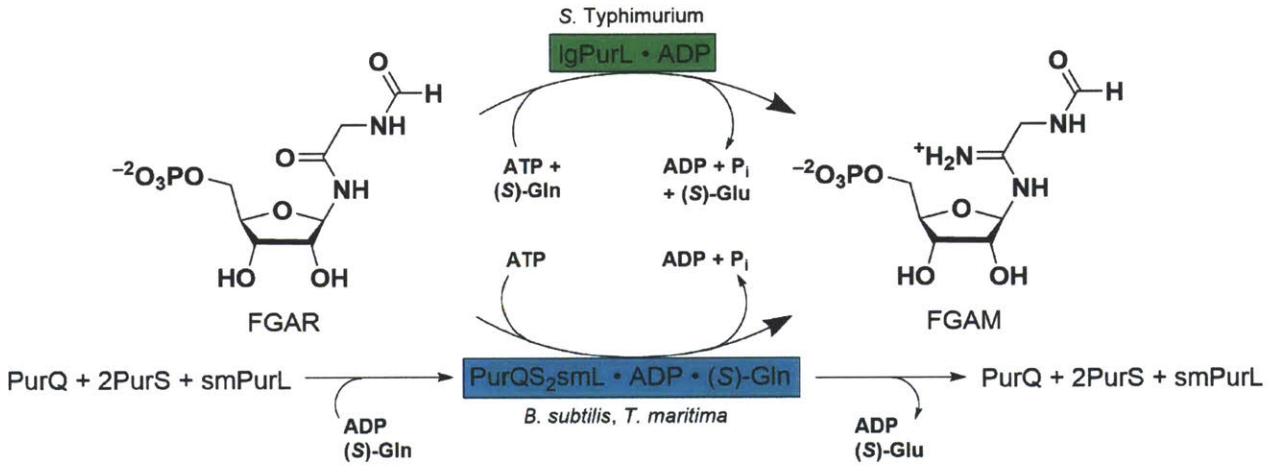


Figure 4.1. The reaction catalyzed by formylglycinamide amidotransferase (FGAR-AT), an enzyme involved in *de novo* purine biosynthesis. Different organisms utilize one of two related enzyme systems, both of which bind one equivalent of ADP. In the *S.* Typhimurium large PurL (lgPurL), the ADP is buried within the protein tertiary structure and, thus, is thought to be inserted during protein synthesis and/or folding and plays a purely structural role in the enzyme. For the enzymes from *B. subtilis* and *T. maritima*, ADP, in conjunction with glutamine, mediates the formation of the active FGAR-AT complex from smPurL, PurQ, and two copies of PurS.

Besides structural homology, both forms of PurL also share a common attribute in that each binds an equivalent of Mg^{2+}-ADP (or ATP, in the case of the *T. maritima* smPurL[3]) in an auxiliary site on the enzyme (**Figure 4.1**).[2-4, 6] Comparison of the *T. maritima* smPurL structure in the absence versus presence of the other subunits (PDB 2HS0[3] and PDB 3D54[5], respectively) reveals that the nucleotide occupying the auxiliary site stabilizes the positions of several key residues contacting the glutaminase subunit, PurQ, in the FGAR-AT complex.[5] This observation, in conjunction with the demonstrated absolute requirement of ADP and Gln in the assembly of a tight *B. subtilis* FGAR-AT complex,[4] indicates this nucleotide (*B. subtilis*) or ATP (*T. maritima*) is critical for the formation of multi-subunit FGAR-AT complexes. In contrast, the ADP molecule

in lgPurL from *S.* Typhimurium (PDB 1T3T[2]) is deeply buried within the protein tertiary structure and can only be released by denaturation,[7] indicating that it is likely playing a more structural role in this enzyme.[5]

Given the peculiarities of the *B. subtilis* RNR reported in the previous chapters, it seemed likely that the compound associated with NrdE played a role in the overall function of the enzyme. Therefore, studies were carried out to determine the identity of the compound and its effects on RNR activity, regulation, and structure. The compound was isolated from NrdE by heat denaturation of the protein, and subsequent characterization by polyethyleneimine (PEI)-cellulose thin-layer chromatography (TLC), UV-visible spectrophotometry, and ^{1}H-NMR spectroscopy led to the surprising result that the compound was a mixture largely composed of 2′-deoxyadenosine 5′-monophosphate (dAMP), but also containing smaller amounts of (deoxy)nucleoside di- and triphosphates whose identity has not been firmly established. Quantitation of the relative amounts of dAMP and protein in recombinant NrdE preparations revealed that the former is present in sub-stoichiometric amounts (0.5 – 0.7 dAMP/α) and cannot be installed or removed by non-denaturing methods. However, NrdE loaded with a full complement of one equivalent of dAMP per monomer can be separated from apo-NrdE using anion exchange chromatography. Characterization of the effects of apo- and holo-NrdE on the function of *B. subtilis* RNR revealed subtle differences between the two forms: the presence of dAMP renders *B. subtilis* RNR more susceptible to dATP inhibition and perhaps increases the propensity of NrdE to oligomerize. Finally, preliminary crystallographic evidence (A. Maggiolo, A. Boal, Pennsylvania State University) suggests the dAMP binding site is solvent exposed and can induce formation of alternative NrdE quaternary structures that may be inhibitory in nature.

4.2. EXPERIMENTAL

4.2.1. Materials and methods. Chemicals, equipment, and routine methods were the same as described in Chapters 2 and 3. Primers were from Invitrogen. Plastic-backed PEI – cellulose TLC plates impregnated with short- and long-wave UV fluorescent compounds were from Sigma, whereas non-fluorescent PEI-cellulose plates were from Macherey-Nagel. DpnI restriction endonuclease was from Promega. Slide-A-Lyzer dialysis cassettes with 3.5 or 7 kDa MWCO regenerated cellulose membranes were from Thermo Fisher Scientific and were used as received. His_6-tagged NrdE (720 nmol min^{-1} mg^{-1} α_2), NrdF, TrxA (200 nmol min^{-1} mg^{-1}), and TrxB (18 μmol min^{-1} mg^{-1}) were prepared as described in Chapter 2. Tagless *B. subtilis* NrdE (710 nmol min^{-1} mg^{-1} α_2), NrdF, and NrdI were prepared as described in Chapter 3. Reconstitution of the Fe- and $Mn(III)_2$-Y• cofactors and purification of holo-protein (His_6-Feβ_2 (1.2 Y•/β_2) = 150 nmol min^{-1} mg^{-1} β_2, Mn-β_2 (0.85 Y•/β_2) = 1300 nmol min^{-1} mg^{-1} β_2) was performed as described in Chapter 2. Unless noted otherwise, the concentrations of NrdE are relative to monomer.

4.2.2. Identification of a tightly bound equivalent of dAMP associated with NrdE.

4.2.2A. Quantitation of dAMP bound to NrdE. Prior to quantitation, DTT was removed by exchanging aliquots of NrdE into 50 mM sodium phosphate, pH 7.6, 150 mM NaCl immediately before analysis using three to four rounds of dilution-concentration with Amicon Ultra-0.5 YM30 microcentrifugal filters. From this buffer-exchanged stock, replicate samples of NrdE (~90 μM, ~100 μL) were prepared. Each sample was immersed in a boiling water bath for 10 min and then cooled on ice. The precipitated protein was pelleted by centrifugation (21130 x g, 10 min, 4 °C) and the supernatant was carefully transferred to a new container by pipetting. The pellet was centrifuged again and any remaining liquid was carefully removed by pipetting and recombined with the supernatant collected from the first spin. The total volume of the supernatant

was estimated by pipetting before transferring it to a sub-micro quartz cuvette (Starna) and measuring the A_{260} (ε_{260} = 15400 M^{-1} cm^{-1}).[8] The pellet was dissolved in 200 µL Edelhoch buffer[9, 10] and the amount of protein quantitated by A_{280} (ε_{280} = 70830 M^{-1} cm^{-1}).[11] The average ratio of dAMP:NrdE was as follows: *as-isolated NrdE* = 0.66 ± 0.11 (n = 18), *apo-NrdE* = 0.02 ± 0.02 (n = 5), *holo-NrdE* = 1.05 ± 0.18 (n = 8).

4.2.2B. Large-scale isolation of dAMP. Two preparations of as-isolated His_6-NrdE (*preparation 1*: 145 µM, 3.3 mL; *preparation 2*: 260 µM, 7.4 mL) were loaded into Slide-A-Lyzer cassettes and dialyzed against three 4 L volumes of water over 8 h at 4 °C with dialysate exchanges every 2 h during the first 4 h of dialysis. The small molecule(s) were separated from NrdE and quantitated (**Table 4.1**) as described in the previous section. The recovered supernatants were pooled, filtered through Amicon Ultra-0.5 YM10 microcentrifugal filters, and lyophilized.

Table 4.1. Total dAMP recovery after dialysis of as-isolated NrdE into water.

Preparation	$V_{initial}$ (mL)	nmol NrdE	nmol dAMP	dAMP:NrdE[a]	V_{final} (mL)	nmol NrdE	nmol dAMP	dAMP:NrdE[a]
1	3.3	480	255	0.53	3.5	480	255	0.53
2	7.4	1900	1280	0.68	9.2	1900	1100	0.57

[a] Ratio expressed as mol dAMP bound per mol NrdE monomer.

4.2.2C. Determination of the nucleoside phosphorylation state by TLC. PEI-cellulose plates measuring 4 x 5 or 4 x 10 cm were pre-developed in ddH_2O, dried at room temperature in the dark for 24 – 36 h, and then stored in the dark at 4 °C until use. In order to get straight solvent fronts during development, the recommendations by Randerath and Randerath[12] were used: dividing lines parallel to the direction of solvent flow were scratched into the layer 4 – 5 mm from either side of the plate, and at the bottom a series of lines ~2 mm in length and parallel to the direction of solvent migration were also scratched into the layer. The origin was placed ~8 – 9 mm from the bottom of the plate.

Solutions of (d)AMP, (d)ADP, and (d)ATP for nucleotide standards were prepared in 1.8 mM Tris, pH 7.5. For fluorescent (non-fluorescent) plates, the origin was spotted with 9 – 20 nmol (50 – 300 pmol) of nucleotide standards or unknown, applied in 0.5 μL additions with subsequent drying using a gentle stream of compressed air after each application. For each analysis two plates were prepared: one plate was developed in 0.5 M formic acid/formate, pH 3.4 to resolve mono- and diphosphates, and the other developed in a 3:1 mixture of 4 M formic acid/formate, pH 3.4 : 4% (w/v) boric acid to resolve di- and triphosphates. The 4 M formic acid/formate stock was prepared by combining the required amounts of > 95% reagent grade formic acid and ammonium formate to yield a pH of 3.4. If necessary, the pH was adjusted with 1 M NH_4OH and 0.5 M acetic acid. The 0.5 M formic acid/formate solvent was prepared by diluting 4 M solution with distilled water; pH adjustment was found to be unnecessary.

Fluorescent plates were dried with a heat gun and spots subsequently visualized under short-wave UV light. Non-fluorescent plates were dried with a gentle stream of compressed air, taped to a glass plate, and subjected to staining with Hanes-Isherwood reagent following the procedure described by Rowley and Kenyon.[13]

4.2.2D. Characterization of the NrdE-associated nucleotide by ^{1}H-NMR. The nucleotide-containing supernatants recovered from the two batches of His_6-NrdE were pooled (10 mL, ~1.4 μmol, pH 7.8) and loaded onto a small bed of DEAE-A25 resin (3.5 x 0.8 cm, ~1.76 mL) equilibrated in water. The column was washed with 7 mL water and then eluted using 7 mL 500 mM NH_4HCO_3, pH ~7.8. The eluate (~1.1 μmol, 77% recovery) was exchanged into D_2O and the NH_4HCO_3 removed by repeated rounds of lyophilization and dissolution in deuterated solvent. The final sample (700 μL) had a concentration of ~1.5 mM based on $\varepsilon_{260} = 15400\ M^{-1}\ cm^{-1}$. Proton NMR spectra were acquired on a Varian Inova-500 MHz NMR spectrometer in the Department of

Chemistry Instrumentation Facility. Spectra of standard samples of dAMP, dADP, and dATP prepared in water and exchanged into D_2O by lyophilization were acquired as references. Other than drawing from the same D_2O stock, no other efforts were made to equalize the pH of the sample and standards.

4.2.2E. Separation of (deoxy)nucleoside mono-, di-, and triphosphates by anion exchange chromatography. The separation of the nucleotide mixture isolated from His_6-NrdE was carried out by Stubbe group member Albert Kim using the NMR sample prepared in the previous section. Five hundred microliters of the NMR sample (0.75 μmol) were loaded onto a DEAE Sephadex A25 column (9 x 1 cm, ~7 mL) equilibrated in 50 mM triethylammonium bicarbonate (TEABC), pH ~7.5. The column was washed with 7 CVs of equilibration buffer before being developed with a 75 mL x 75 mL linear gradient from 50 mM to 400 mM TEABC, followed by a 30 mL x 30 mL linear gradient from 400 mM to 600 mM TEABC. Fractions were screened for A_{259} and revealed three peaks eluting at ~230 mM, ~390 mM, and ~550 mM TEABC. Fractions encompassing each peak were pooled separately and the TEABC was removed by repeated rounds of lyophilization followed by dissolution in D_2O. The samples were analyzed by ^{1}H-NMR as described above and on fluorescent PEI-cellulose TLC plates using 1 M LiCl as the chromatographing solvent.

4.2.2F. Nucleotide isolation from NrdE using perchloric acid precipitation and assessment for non-specific phosphate hydrolysis. The following sets of experiments were carried out by Albert Kim. A sample of as-isolated NrdE (10 mL, 132 μM) in Tris buffer (50 mM Tris, pH 7.6, 100 mM NaCl, 5% (v/v) glycerol, 10 mM DTT) was precipitated by addition of $HClO_4$ to a final concentration of 1% (v/v). The sample was incubated on ice for 5 min before pelleting the precipitated protein by centrifugation (21130 x g, 10 min, 4 °C). The pH of the

supernatant was neutralized by addition of 800 μL 2 M KOH prior to freezing the solution at -20 °C for 1 h. The sample was then thawed on ice, divided into ten 1 mL aliquots, and centrifuged to pellet the $KClO_4$ precipitate (21130 x g, 10 min, 4 °C). The compounds recovered from as-isolated NrdE were separated by anion exchange chromatography, quantitated by UV spectrophotometry, and analyzed by ^{1}H-NMR as described above. A control sample (10 mL) containing 80 μM dATP and 160 μM BSA was subjected to the same procedure to assess the extent of non-specific hydrolysis of phosphate from the nucleotide that occurred under the conditions of the experiment.

4.2.3. Efforts to load or remove dAMP from as-isolated NrdE. As-isolated NrdE (48 or 63 nmol, 200 μL, 0.67 dAMP/α) in 50 mM HEPES, pH 7.6, 150 mM NaCl, 15 mM $MgCl_2$, 5% (w/v) glycerol, 5 mM DTT was incubated with 10 equiv of dAMP or dATP either at 4 °C for 2 h on a nutating platform or at 25 °C for 20 – 30 min. To see if GDP could remove the bound dAMP, another sample (63 nmol, 200 μL) was incubated with 1 mM GDP for 30 min at 25 °C. Small amounts of precipitate were observed at the end of all incubations; thus, samples were centrifuged (21130 x g, 1 min, 4 °C) before being loaded onto and eluted from a Sephadex G25 column (28 x 1.5 cm, ~49 mL bed volume) equilibrated in the same buffer. Protein-containing fractions were pooled and concentrated prior to quantitating the dAMP load as described in **Section 4.2.2A**.

4.2.4. Separation of holo- from apo-NrdE by MonoQ anion exchange chromatography. As-isolated NrdE (70 nmol, 1 mL, 0.5 – 0.7 dAMP/α) in storage buffer (50 mM sodium phosphate, pH 7.6, 150 mM NaCl, 5% (w/v) glycerol, 5 mM DTT) was centrifuged (21130 x g, 4 min, 4 °C) and loaded onto a MonoQ 10/100 GL column (8 mL) equilibrated in 50 mM Tris, pH 7.6, 100 mM NaCl, 5% (w/v) glycerol, 1 mM tris(carboxyethyl)phosphine (TCEP). The column was washed with 10 mL of equilibration buffer prior to development with a 160 mL linear gradient from 100 – 500 mM NaCl in Tris buffer at a flow rate of 1 mL min^{-1}. Apo-NrdE eluted between 260 – 270

mM NaCl and holo-NrdE between 330 – 350 mM NaCl. Protein containing fractions from each group were pooled, concentrated, and stored at -80 °C. At a later date, samples of apo- or holo-NrdE recovered from multiple fractionations as described above were pooled, concentrated, and rechromatographed on the MonoQ column to obtain a homogenous preparation of the desired protein. These samples were exchanged into storage buffer by gel filtration on a Sephadex G25 column. The dAMP load of the protein was quantitated as described in **Section 4.2.2A**.

4.2.5. Assessment of the effects of dAMP loading on NrdE quaternary structure.

4.2.5A. SEC. Samples of apo- or holo-NrdE (0.5 nmol, 200 μL, < 0.01 and 1.06 dAMP/α, respectively) were loaded onto an analytical Superdex 200 10/100 GL column (~24 mL bed volume) connected to an ÄKTA Purifier 100 FPLC system (GE Healthcare). Protein was eluted at 4 °C at a flowrate of 0.25 mL min^{-1} using 50 mM HEPES, pH 7.6, 150 mM NaCl, 15 mM magnesium acetate, 5% (w/v) glycerol. The column was calibrated with a High Molecular Weight Gel Filtration Calibration kit (GE Healthcare) containing the following standards: Blue Dextran 2000 (void volume determination), thyroglobulin (669 kDa), ferritin (440 kDa), aldolase (158 kDa), conalbumin (75 kDa), and ovalbumin (44 kDa).

4.2.5B. SV-AUC. A sample of freshly rechromatographed holo-NrdE (110 μM, 650 μL, 1.01 dAMP/α) was exchanged into AUC buffer (50 mM HEPES, pH 7.6, 150 mM NaCl, 15 mM $MgCl_2$, 1 mM TCEP) using gel filtration on a Sephadex G25 column (28 x 1.5 cm, ~49 mL). Protein-containing fractions were pooled and concentrated. A 500-μL sample of 2 μM holo-NrdE was prepared by mixing the appropriate volumes of concentrated protein and AUC buffer. Cells assembled with Epon charcoal double sector centerpieces and sapphire windows were loaded with approximately equal amounts of sample and reference buffer (~440 μL). Laser delay adjustments and radial calibration were completed with the rotor spinning at 3000 rpm; nothing sedimented

during this procedure. The system was thermally equilibrated to 20 °C for 1.5 h before sedimenting holo-NrdE for 17 h at 42000 rpm. Sedimentation was monitored by interference with scans collected every 1.5 min. The first 220 scans were loaded into Sedfit, time stamp corrected,[14] and fit to the c(s) model with a resolution of 100 and a regularization factor of 0.68 using alternating rounds of Marquardt-Levenberg and Simplex optimization algorithms until the fits converged. The buffer mismatch module[15] was implemented in data fitting as well. Sedimentation coefficient distributions were corrected to standard state using ρ = 1.009 g mL^{-1}, η = 1.054 cP, and $\bar{v}$ = 0.7324 cm^3 g^{-1}, all of which were calculated and temperature corrected in the program Sednterp.[16]

4.2.6. Assessing the effects of dAMP loading on RNR activity. NDP reduction activity catalyzed by apo- and holo-NrdE (0.05 and 1.20 dAMP/α, respectively) was assessed using the spectrophotometric assay. Assays under routine conditions (500 μL) were conducted at 37 °C and contained 50 mM HEPES, pH 7.6, 15 mM $MgCl_2$, 1 mM EDTA, 40 μM TrxA, 0.4 μM TrxB, 0.2 mM NADPH, 0.5 μM Mn-β_2 (0.9 Y•/β_2), 1 μM NrdE, 3 mM ATP, and 1 mM CDP. In the examination of the effects of dAMP on dATP inhibition, ATP was substituted with dATP at concentrations ranging from 250 nM to 1 mM. For reassessment of the steady-state kinetics with effector nucleotide as the variable component, the appropriate substrate/effector pair was used in place of ATP/CDP. Substrate was used at a final concentration of 1 mM while the concentration of effector varied as follows: ATP = 25 μM – 4 mM, dGTP = 0.25 – 30 μM, and TTP = 0.1 – 30 μM.

4.2.7. Assessing the effects of dAMP loading on dATP binding to NrdE. The binding of [2,8-^{3}H]-dATP (Moravek Biochemicals) to apo- and holo-NrdE (0.02 and 0.97 dAMP/α, respectively) in assay buffer (50 mM HEPES, pH 7.6, 150 mM NaCl, 15 mM magnesium acetate) was assessed

using the ultrafiltration technique described by Ormö and Sjöberg.[17] Nanosep® 30 kDa MWCO Omega filter cartridges (Pall Corporation) with polyethersulfone membranes were used in lieu of the discontinued Millipore Ultrafree-MC PTTK devices. [2,8-^{3}H]-dATP was exchanged into assay buffer and diluted with dATP to a specific activity of 90 – 83700 cpm nmol^{-1}. The absence of non-specific binding of [2,8-^{3}H]-dATP to the membrane and/or filter device was determined in control assays in the presence or absence of 3 μM BSA.

A typical assay was conducted at 25 °C in a total volume of 200 μL. Assay mixtures consisted of assay buffer, 5 mM DTT, apo- or holo-NrdE (1, 5, or 12 μM), and [2,8-^{3}H]-dATP (0.1 – 170 μM). The NrdE concentration was chosen such that no more than 80% and no less than 10% of the total [2,8-^{3}H]-dATP was bound. Buffer, DTT, and nucleotide were combined and thermally equilibrated for 2 min. NrdE was subsequently added and the assays continued to incubate at 25 °C for 5 min. A 30 μL aliquot was withdrawn and counted to determine the total concentration of [2,8-^{3}H]-dATP in the assay mixture. One hundred fifty microliters of the remaining solution were transferred to a thermally equilibrated filter device and centrifuged (4500 – 7600 x g, 1 min, 25 °C). Thirty microliters of the filtrate were taken and counted to determine the amount of free [2,8-^{3}H]-dATP. The amount of bound nucleotide was calculated as the difference between the total and free [2,8-^{3}H]-dATP concentrations. The data were plotted as a saturation binding curve and fit to equation 4.1 using non-linear regression in Igor Pro.

$$\frac{[dATP]_{bound}}{[NrdE]_{monomer}} = \frac{n \times [dATP]_{free}}{K_d + [dATP]_{free}} \tag{4.1}$$

In the case of the data set collected with holo-NrdE, the data were also fit to equation 4.2 to account for the apparent non-specific binding (Q_{nsb}).[18]

$$\frac{[dATP]_{bound}}{[NrdE]_{monomer}} = \frac{n \times [dATP]_{free}}{K_d + [dATP]_{free}} + (Q_{nsb} \times [dATP]_{free}) \tag{4.2}$$

4.2.8. Site-directed mutagenesis, expression, and purification of tagless holo-NrdE-$C_{382}S$. The vector pE-SUMO-*nrdE* (Chapter 3) was amplified by PCR using Phusion® High Fidelity DNA polymerase (New England Biolabs) and the primers 5′-CCTT**AGC**TCTGAAGTGCTGCAG GCATCTCAAGTTTCTTCATAC-3′ and 5′-CAGA**GCT**AAGGTTAGAAAATTTGACCTTGG AAATATGGTTATTCGC-3′ to introduce the desired mutation (mutated codon indicated in boldface font). The product was incubated with 10 U DpnI at 37 °C for 75 min and subsequently transformed into XL10 Gold ultra-competent *E. coli* cells (Stratagene) and selected for on LB agar plates containing 100 μg mL^{-1} Kan. Single colonies were grown to saturation overnight in LB at 37 °C and the plasmid was subsequently isolated with a mini prep kit (Qiagen) before being submitted for sequencing (Quintara Biosciences), which confirmed the successful mutation of the C_{382} codon in *nrdE*. The mutant vector was transformed into BL21 (DE3) cells and tagless NrdE-$C_{382}S$ was expressed and purified as described in Chapter 3. Holo-NrdE-$C_{382}S$ (0.95 dAMP/α) was obtained by anion exchange chromatography as described in **Section 4.2.4**.

4.2.9. Determination of the effects of dAMP loading on $\alpha_2\beta_2$ complex formation.

4.2.9A. Assessing the strength of the interaction between tagless apo- or holo-NrdE and Fe(III)$_2$-Y• His_6-NrdF using pull-down assays. The pull-down assay procedure was adapted from those described by Kanchana Ravichandran[19] with minimal modification. Samples (300 μL) contained 50 mM HEPES, pH 7.6, 150 mM NaCl, 15 mM magnesium acetate, 5 mM DTT, 2.5 μM Fe(III)$_2$-Y• His_6-Fβ_2, and 5 μM of either apo- or holo-tagless NrdE (< 0.01 and 1.06 dAMP/α, respectively). The reaction mixtures were incubated in a 25 °C water bath for 15 min prior to adding 100 μL of homogeneous HisPur Ni-NTA resin slurry (Thermo Fisher) equilibrated in the same buffer, but with DTT omitted. Samples were affixed to a rotating platform and incubated for 10 min at RT. The resin was gently pelleted (100 x g, 30 s, 25 °C) and the supernatant collected

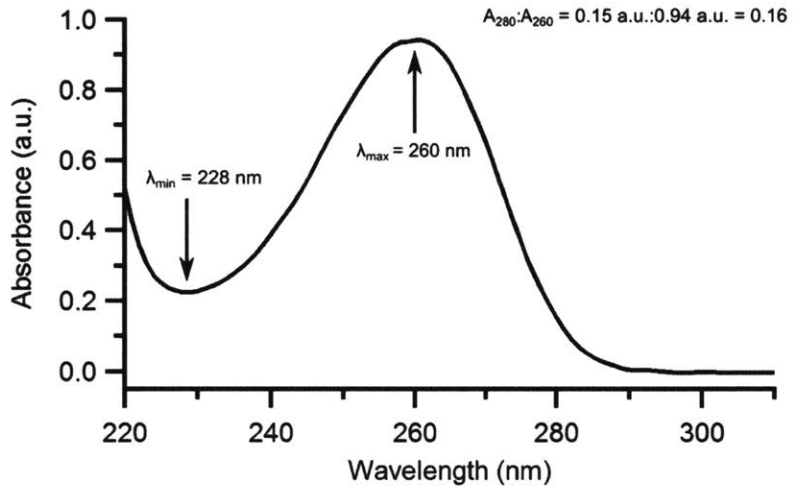
as the "flow through" fraction. The resin was resuspended in 600 μL of buffer supplemented with 15 mM imidazole, mixed for 10 s by inversion, and pelleted again to give "wash 1." The wash step was repeated a second time before eluting the protein with 600 μL of buffer supplemented with 250 mM imidazole following the same procedure. Twenty microliters of the flow through, wash 1, wash 2, and elution fractions were then analyzed on a 12.5% (w/v) SDS-PAGE gel.

4.2.9B. Assessing the quaternary structure of 1:1 mixtures of holo-NrdE-$C_{382}S$ and $Mn(III)_2$-Y• β_2 in the absence and presence of nucleotides by SV-AUC. Aliquots of $Mn(III)_2$-Y• NrdF and NrdE-$C_{382}S$ (0.95 dAMP/α) were exchanged into 50 mM HEPES, pH 7.6, 150 mM NaCl, 15 mM $MgCl_2$, 1% (w/v) glycerol, 1 mM TCEP using an Econo-Pac® 10DG desalting column (Bio-Rad) equilibrated with > 20 CVs of buffer. Samples of 2 μM α_2-$C_{382}S$ and Mn-β_2 with or without 250 μM GDP and 12 μM TTP were prepared in at total volume of 500 μL and were analyzed by SV-AUC as described in **Section 4.2.5B**.

4.3. RESULTS

4.3.1. Identification of dAMP as the major compound associated with heterologously-expressed as-isolated NrdE. The unusual dATP inhibition exhibited by the *B. subtilis* RNR suggested that a tightly bound compound associated with NrdE might be a 2′-deoxyadenosine analog. This hypothesis was supported by the results of TLC and UV analysis of protein-free samples of the compound. For UV analysis, as-isolated NrdE was first exchanged into water or phosphate buffer and then heat denatured to release the bound nucleotide. The precipitated protein was removed by centrifugation and a UV spectrum of the supernatant recorded between 210 – 320 nm as shown in **Figure 4.2**. The spectrum revealed a λ_{max} of 259 – 260 nm, a λ_{min} of 227 – 229 nm, and an A_{280}:A_{260} ratio of 0.16 – 0.19. These parameters are consistent with values reported in the literature for either adenosine or 2′-deoxyadenosine analogs,[8] therefore indicating that the

compound associated with NrdE was likely an adenine-containing mononucleoside or nucleotide rather than residual nucleic acids carried over from protein purification.



Literature values for pure compounds

	λ_{max} (nm)	λ_{min} (nm)	A_{280}:A_{260}
ATP	259	227	0.15
ADP	259	227	0.16
AMP	259	227	0.16
A	260	227	0.14
dATP	259	226	0.15
dAMP	259	226	0.15
dA	260	225	0.15

Figure 4.2. Representative UV spectrum of compound(s) released from as-isolated His_6-NrdE by heat denaturation into 50 mM sodium phosphate, pH 7.6, 150 mM NaCl. Tabulated to the right are values for λ_{max}, λ_{min}, and the A_{280}:A_{260} reported in the literature[8] for pure compound dissolved in water at pH 7.0.

To qualitatively determine if the putative (2′-deoxy)adenosine analog mixture was phosphorylated and, if so, the number of phosphates present, it was examined by PEI-cellulose TLC using the procedures of Randeranth and Randeranth.[12] An image of a representative chromatogram run on a fluorescent plate using phosphorylated adenosine analogs (AMP, ADP, and ATP) as standards is shown in **Figure 4.3A**, and the corresponding R_f values of the spots are tabulated in **Table 4.2**. The lane in which the NrdE-associated compounds were run revealed three spots of increasing intensity relative to the origin that had similar R_f values to AMP, ADP and ATP (**Table 4.2**), indicating that mono-, di-, and triphosphates were present. To confirm this result, the (deoxy)- nucleoside mixture was separated on a DEAE A25 anion exchange column using a linear gradient of 50 – 600 mM TEABC. The chromatogram revealed three peaks eluting at ~230 mM, ~390 mM, and ~550 mM TEABC, and reanalysis of each compound by TLC revealed

that the material from each migrated respectively as (deoxy)nucleoside mono-, di-, and triphosphates (**Figure 4.3B**). Quantitation of the amount of each compound recovered using UV spectrophotometry (ε_{260} = 15400 M^{-1} cm^{-1}) revealed a 3:2:1 ratio of (d)NMP:(d)NDP:(d)NTP, which is consistent with the spot intensities observed by TLC for the pre-separated mixture (**Figure 4.3A**). Therefore, the TLC results, coupled with the UV spectrum (**Figure 4.2**), suggested that the compounds associated with NrdE were a heterogeneous mixture of phosphorylated (2′-deoxy)adenosine analogs, with the monophosphate being present in the largest amount.

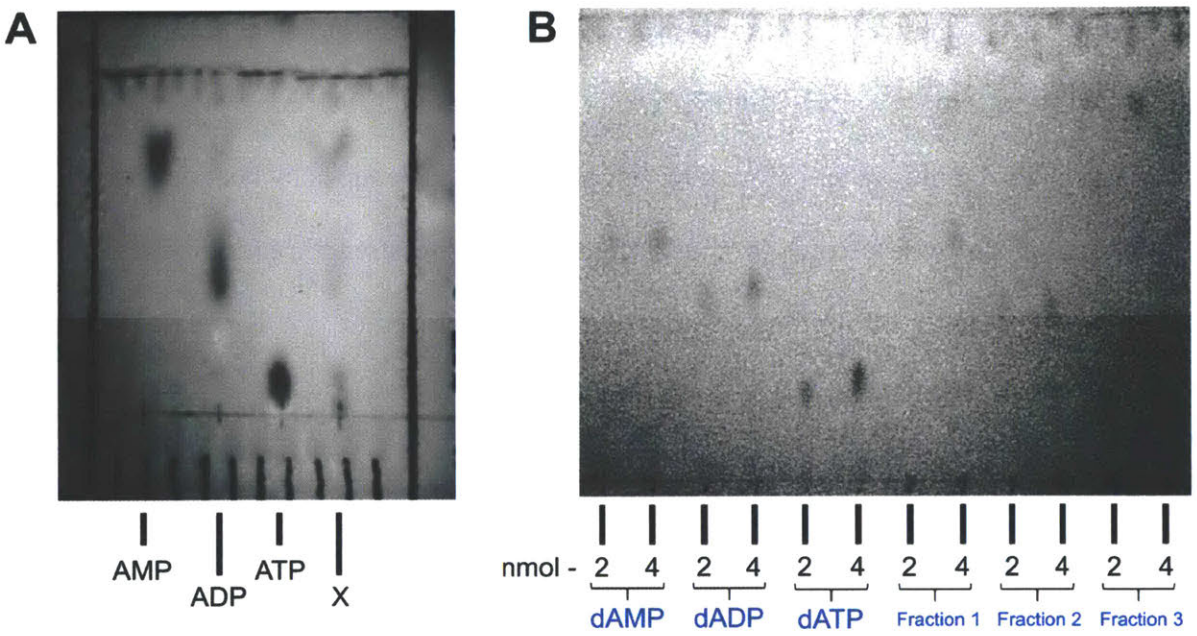


Figure 4.3. Analysis of compound(s) released from as-isolated His_6-NrdE by PEI-cellulose TLC. (**A**) Analysis immediately after isolation from NrdE (in the lane denoted X). The amounts of standards and sample loaded were as follows: AMP = 12 nmol, ADP = 10 nmol, ATP = 11.5 nmol, compound X = 3.3 nmol. The plate was developed in a 3:1 mixture of 4 M formic acid/formate, pH 3.4 : 4% (w/v) boric acid. (**B**) Analysis of fractions from anion exchange chromatography. *Fraction 1* – compound eluting at ~240 mM TEABC, *Fraction 2* – compound eluting at ~390 mM TEABC, *Fraction 3* – compound eluting at ~550 mM TEABC. The plate was developed with 1 M LiCl. After the plates had dried, they were illuminated with short wave UV and photographed with a smart phone. The images were rendered black and white, and the brightness and contrast have been adjusted so that all spots are clearly visible.

Table 4.2. R_f values for the TLC analyses shown in **Figure 4.3**.

Figure 4.3A		Figure 4.3B	
Compound	**R_f**	**Compound**	**R_f**
AMP	0.75	dAMP	0.51
ADP	0.42	dADP	0.41
ATP	0.08	dATP	0.20
	0.75	Fraction 1	0.52
X	0.42	Fraction 2	0.37
	0.08	Fraction 3	0.17

The identity of the (deoxy)adenosine analogs was established by ^{1}H-NMR spectroscopy using a combined sample prepared from two large scale isolates from different as-isolated NrdE preparations. To remove salts and buffer molecules that would interfere with the analysis, the protein was dialyzed into water before releasing the bound nucleotides by heat denaturation. The protein was surprisingly tolerant to non-buffered, low ionic strength conditions as very little precipitation was observed during and after the dialysis procedure. After heat denaturing the protein and removing the visible precipitate by centrifugation, the supernatants were pooled and filtered using Amicon microcentrifugal filters to remove any last traces of protein in the sample. Analysis of this material by NMR revealed significant amounts of contaminating glycerol that was likely picked up during the Amicon filtration step. The mixture of compounds was loaded onto a small bed of DEAE A25 and washed with water to remove the glycerol before being eluted with 500 mM NH_4HCO_3. After removal of NH_4HCO_3 by repeated rounds of lyophilization, NMR analysis revealed that the glycerol was successfully removed, but the presence of multiple species resulted in a complicated spectrum. Therefore, the nucleoside mono-, di-, and triphosphates

were separated using anion exchange chromatography as described above and were analyzed individually by ^{1}H-NMR after removal of TEABC by lyophilization.

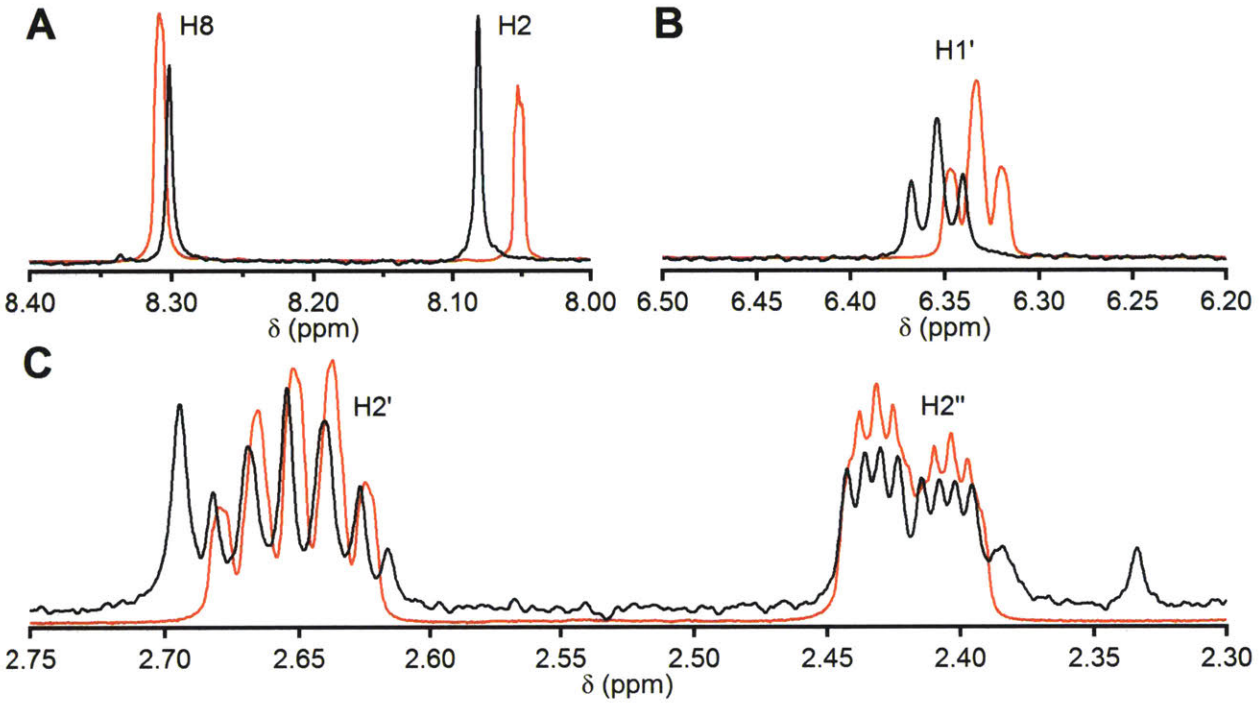


Figure 4.4. Comparison of ^{1}H-NMR spectra of the nucleoside monophosphate purified from as-isolated His_6-NrdE (black trace) and a dAMP standard (red trace). Panels shows protons bonded to (**A**) the adenine base, (**B**) the anomeric carbon (C1′), and (**C**) the 2′ carbon. Differences in the chemical shifts of the two spectra are a result of no pH control in the samples.

Table 4.3. Chemical shifts for dAMP and the nucleoside monophosphate recovered from as-isolated His_6-NrdE.

	Chemical shift (ppm), splitting pattern, and integration results[a,b]								
Compound	**H2**	**H8**	**H1′**	**H2′**	**H2″**	**H3′**	**H4′**	**H5′**[c]	**Ref.**
dAMP	8.12 (s)	8.50 (s)	6.43 (t)	2.81 (m)	2.59 (m)	4.73 (m)	4.26 (m)	3.95 (m)	*20*
	8.05 (s, 1H)	8.31 (s, 1H)	6.33 (t, 1H)	2.65 (m, 1H)	2.42 (m, 1H)	4.55 (m, 1H)	4.10 (m, 1H)	3.84 (m, 2H)	This work
Nucleotide from NrdE	8.07 (s, 1H)	8.30 (s, 1H)	6.35 (t, 1H)	2.64 (m, 1H)	2.41 (m, 1H)	4.54 (m, 1H)	4.10 (m, 1H)	3.86 (m, 2H)	This work

[a] Peak splitting patterns: s = singlet, t = triplet, and m = multiplet.
[b] In the study by Davies and Danyluk,[20] peak integrations were not reported.
[c] Signals from the two 5′ protons are coalesced.

Expansions of the regions in the ^{1}H-NMR spectrum of the NrdE-associated nucleoside monophosphate that are important for characterization of adenosine and deoxyadenosine analogs are shown in **Figure 4.4**, and the experimentally measured chemical shifts are reported in **Table 4.3**. For comparison, a spectrum of authentic dAMP was acquired and is superimposed on that of the NrdE-associated compound in **Figure 4.4**. The small differences in chemical shifts (**Table 4.3**) of the dAMP standard and the NrdE-associated (d)AMP are attributable to a lack of pH control and using the HOD peak as an internal reference. Regardless of these differences, the ^{1}H-NMR analysis demonstrated that dAMP is the identity of the nucleoside monophosphate associated with NrdE. The most diagnostic features of adenine-containing nucleotides are associated with the H2 and H8 protons of the base (**Figure 4.4A**). Diagnostic features of 2′-deoxyribonucleotides include splitting of the anomeric proton (H1′) signal into a doublet of doublets (an apparent triplet) indicative of coupling to two protons (**Figure 4.4B**) and, in the case of deoxyadenosine analogs, the appearance of two well-resolved low field multiplets centered at 2.53 ppm that are indicative of the H2′ and H2″ protons (**Figure 4.4C**).[20] The J^2 and J^3 coupling constants for H1′, H2′, H2″, and H3′ measured for the NrdE-associated deoxynucleoside monophosphate are consistent with previously reported values for dAMP (**Table 4.4**),[20, 21] further substantiating the identity of the sugar moiety as 2′-deoxyribose. Finally, evidence that the major compound is the nucleoside monophosphate is provided by the nearly coincident chemical shifts (**Table 4.3**) of the H3′ signals from the NrdE-associated deoxynucleotide monophosphate and the dAMP standard, and the similar splitting patterns and chemical shifts of the C5′ protons for both nucleotides (**Figure 4.5**). The full spectrum for dAMP isolated from NrdE is presented in Appendix 4. Altogether, the UV, TLC, and ^{1}H-NMR analyses strongly agree with each other and support the conclusion that the majority (0.5 equiv) of the mixture of the compounds associated with NrdE is dAMP.

Table 4.4. J^2 and J^3 proton coupling constants for dAMP and the nucleoside monophosphate recovered from as-isolated His_6-NrdE.

Nucleotide	Coupling constants (Hz)				
	$J_{1',2'(S)}$	$J_{1',2'(R)}$	$J_{2'(R),2'(S)}$	$J_{2'(S),3'}$	$J_{2'(R),3'}$
dAMP	6.9[a]	6.6[a]	-	5.8[a]	3.8[a]
	7.3[b]	6.2[b]	14.0[b]	6.1[b]	3.4[b]
	7.0	6.5	14.0	-	-
Nucleotide from NrdE	7.0	6.5	14.0	6.2	3.0

[a] From Fang *et al.*, 1971.[21]
[b] From Davies and Danyluk, 1974.[20]

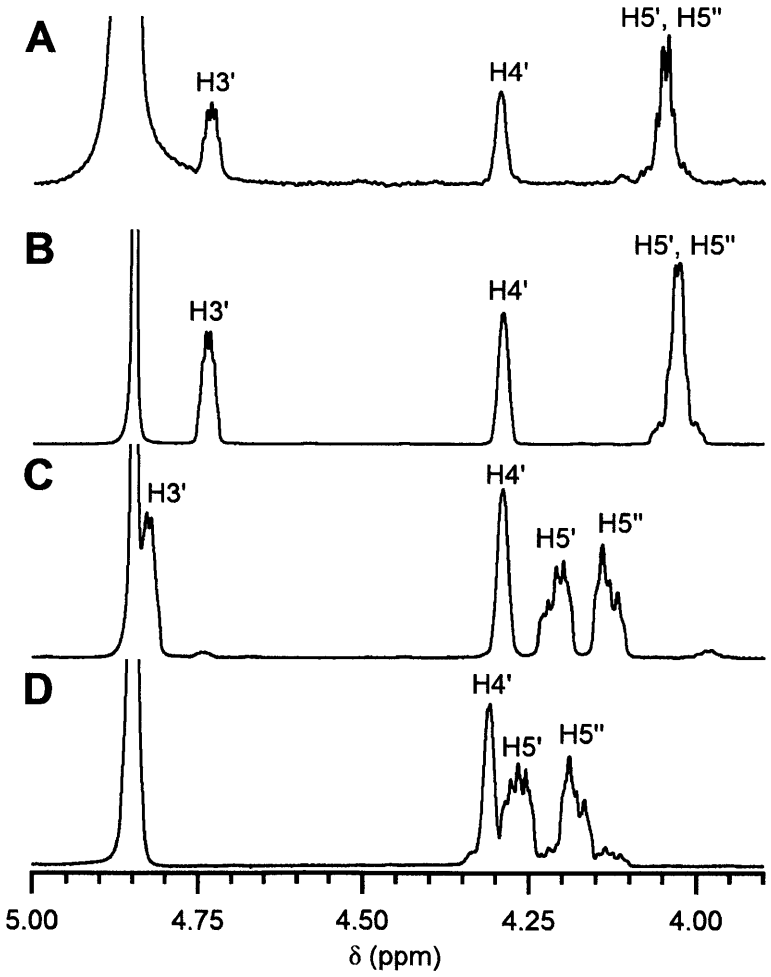


Figure 4.5. Comparison of the 3.8 – 5.0 ppm region of the ^{1}H-NMR spectra of (**A**) the sample of nucleoside monophosphate purified from as-isolated His_6-NrdE, (**B**) dAMP standard, (**C**) dADP standard, and (**D**) dATP standard (H3′ of dATP is masked by the H_2O/HOD solvent peak). Note the differences in the splitting patterns of the C5′ and C5″ protons for dADP and dATP compared to dAMP.

The ^{1}H-NMR spectra of the NrdE-associated nucleoside di- and triphosphates recovered from the DEAE A25 column are shown in Appendix 4. The UV spectrum (not shown) of these compounds had a single peak with a λ_{max} at 260 nm, indicating that both were likely (deoxy)adenosine analogs. The amount of nucleoside triphosphate recovered was too small to get good signal-to-noise using standard NMR probes (Appendix 4, **Figure A4.11**) and, therefore, was not examined any further. The nucleoside diphosphate, on the other hand, appeared to be a mixture of a ribonucleoside and deoxyribonucleoside analog, as indicated by the appearance of two H1′ signals (**Figure A4.8**), one being a triplet (6.35 ppm, indicative of 2′-deoxyribose) and the other a doublet (5.97 ppm, indicative of ribose). The aromatic region of the spectrum exhibited two peaks (**Figure A4.7**) that appeared to consist of partially-resolved signals from multiple species, and the signal intensity of the lower field protons was much more prominent than that from the higher field ones. While the chemical shifts of the peaks are most consistent with adenine, the unusual ratio of the intensity for the H2 and H8 peaks prevented strong conclusions from being made about the identity of the base. The low field regions (**Figure A4.9** and **A4.10**) showed peaks consistent with the presence of a sugar moiety, but the lack of pH control in the sample made it difficult to conclude whether or not the material was mostly ribose- or deoxyribose-based analogs. Taken together, the UV and NMR analyses of the nucleoside di- and triphosphates isolated from NrdE suggest that they correspond to ADP, dADP, and (d)ATP, but further verification is required.

4.3.2. The presence of dAMP is not a result of non-specific hydrolysis of dATP during protein denaturation. The identification of dAMP as the major nucleotide associated with as-isolated NrdE was a highly unexpected result. Given that nucleoside di- and triphosphates were also isolated from NrdE and that they were most likely dADP and dATP, it was possible that the presence of dAMP might have resulted from non-specific hydrolysis of dADP and/or dATP during

protein purification and/or isolation of the nucleotides by heat denaturation. To assess the likelihood that the latter scenario had occurred, two control experiments were carried out using perchloric acid precipitation as the protein denaturation method. In the first, a mixture of 160 μM BSA and 80 μM dATP was treated with 1% (v/v) $HClO_4$ on ice for 5 min before pH neutralization with KOH and removal of the protein and $KClO_4$ precipitates by centrifugation. Almost complete recovery of the starting amount of dATP (99% yield) was achieved after chromatographing the supernatant on a DEAE A25 anion exchange column, thus indicating that the nucleotide was inert to non-specific hydrolysis under the conditions of the experiment.

The $HClO_4$ precipitation procedure was subsequently used in the second control to isolate the nucleotide mixture from a new batch of as-isolated NrdE (132 μM in 10 mL). The chromatogram of this material again revealed three peaks, each eluting at a similar concentration of TEABC as was reported above for the nucleotide mixture isolated from NrdE by heat denaturation. The UV spectra of these compounds exhibited a λ_{max} at 260 nm, and quantitation of the amount of each nucleotide revealed an approximately 2:3:1 ratio of mono-, di-, and triphosphate, respectively (580 nmol recovered in total). The identity of the nucleoside monophosphate was shown to be dAMP by ^{1}H-NMR analysis, which revealed similar chemical shifts and coupling constants to those listed in **Table 4.3** and **Table 4.4**, respectively. The fact that the (d)NTP was again the least abundant component of the mixture demonstrates that non-specific hydrolysis of (d)NDPs and (d)NTPs as a result of the protein denaturation method utilized is not responsible for the presence of (d)NMP associated with NrdE. However, whether or not hydrolysis is occurring during protein purification remains to be determined. The fact that nucleoside diphosphate was the major component of the mixture isolated from the new batch of

as-isolated NrdE could be indicative of non-specific hydrolysis, but also could be attributable to variation between protein preparations; further studies are required to resolve this issue.

4.3.3. dAMP is tightly bound and cannot be installed or removed from as-isolated NrdE *in vitro* by non-denaturing methods. The most extraordinary outcome of the ability to dialyze as-isolated NrdE into water was the nearly quantitative recovery of the starting amount of dAMP, despite dialyzing each protein preparation against a total volume of 12 L. This phenomenon was determined by measuring the amount of dAMP associated with the protein before and after dialysis, and the results are tabulated in **Table 4.1**. This result indicated that the dAMP bound to NrdE with either an extremely tight binding constant or that it was buried within the protein tertiary structure. Quantitation of the amount of dAMP initially bound to as-isolated NrdE revealed a loading factor of 0.5 – 0.7 mol dAMP per mol NrdE (range observed over seven preparations), suggesting that the protein could associate with up to one equivalent. It was hypothesized that if the dAMP binding site was surface-exposed, then incubation of as-isolated NrdE with dAMP should increase the loading to one dAMP per NrdE peptide.

To test this hypothesis, loading experiments in which samples of as-isolated NrdE were incubated with 10 equivalents of dAMP at 4 °C or 25 °C were carried out. As a control, as-isolated NrdE was also incubated with 10 equivalents of dATP under the same conditions given that a nucleoside or deoxynucleoside triphosphate was observed by TLC (**Figure 4.3**). The treated protein samples were separated from free nucleotide by Sephadex G25 gel filtration and the amount of deoxyadenosine analog re-quantitated. The results (**Table 4.5**) revealed that the load of nucleotide bound to NrdE increased in both cases, regardless of temperature. However, the target of one equivalent nucleotide bound per peptide was not achieved. It is noteworthy that a small amount of precipitated protein was observed after the incubation period had expired,

suggesting that apo-NrdE was unstable, which is consistent with additional data presented in **Section 4.3.5**.

Table 4.5. Ratio of dAMP:NrdE after treatment of as-isolated protein with dAMP, dATP, GDP, or freeze-thaw cycling.

Treatment	Incubation conditions		nmol protein	nmol dAMP	dAMP:NrdE
	Temp. (°C)	Time (min)			
Pre-treatment	-	-	9.1	5.9	0.65
dAMP	4	120	7.9	6.2	0.78
dAMP	25	30	8.4	6.5	0.78
dATP	25	20	8.3	6.5	0.79
GDP	25	30	5.5	4.3	0.75
Freeze-thaw	-	-	7.4	5.1	0.69

Saturating concentrations of GDP, as shown in Chapter 3, caused all protein in NrdE SV-AUC samples to convert to a state with a tertiary structure that is consistent with a traditional class I RNR α monomer (**Figure 3.13C**). Working under the assumption that dAMP binding to NrdE causes the tertiary structure of the protein to undergo significant compaction, as discussed in this introduction to this chapter, a possible explanation for the observation described above is that GDP binding to the catalytic site of NrdE induces a protein conformational change that relaxes the compacted state. Consequently, the affinity of NrdE for dAMP is weakened and, perhaps, results in the dissociation of dAMP from the protein. If correct, this hypothesis suggests that incubation of as-isolated NrdE with GDP should lower the dAMP content of the protein. Therefore, as-isolated NrdE was incubated with 1 mM GDP and the amount of dAMP quantitated after removal of free nucleotides by gel filtration. Unexpectedly, the results indicated that the amount of dAMP bound per NrdE increased (**Table 4.5**), despite the absence of any exogenously added deoxyadenosine analog. Also similar to the previous experiments was the fact that a small amount of protein precipitate was observed after incubating the protein with GDP.

Given that in all the previously described experiments an increase in the dAMP load was measured and protein precipitation was noted after each incubation period, an interpretation of the data is that apo-NrdE is not as stable under the experimental conditions and, therefore, precipitated and was removed by centrifugation. The hypothesis of apo-protein instability was supported by the measured increases in dAMP load of stocks of NrdE that had been subjected to several freeze-thaw cycles (**Table 4.5**) and was confirmed, as discussed subsequently, with purified apo-NrdE. Therefore, the apparent increase in dAMP bound per NrdE for all these samples is the result of the loss of ~15 – 20% of the total starting amount of NrdE to precipitation due to apo-NrdE instability. Consequently, the results of these experiments suggest that the dAMP binding site is buried within the tertiary structure of NrdE and cannot acquire its ligand from free solvent pools.

4.3.4. Apo- and holo-NrdE can be separated by anion exchange chromatography. Since the dAMP loading of as-isolated NrdE could not be increased by addition of exogenous nucleotide, a method was sought to separate holo-NrdE from apo-protein. Previous analysis of the His_6-tagged protein by anion exchange chromatography (Appendix 2, **Figure A2.4**) demonstrated the ability to resolve at least two distinct forms of NrdE. This observation prompted a reexamination of this method to see if the different forms of NrdE corresponded to different loading states. Shown in **Figure 4.6A** is a representative chromatogram of a sample of as-isolated NrdE run on the MonoQ column. The complexity of the chromatogram is likely due to the heterogeneity of the protein itself since the sample is a mixture of apo-NrdE and protein bound to dAMP and other (deoxy)-nucleoside di- and triphosphates (**Figure 4.3**). Furthermore, with a loading concentration of 70 μM, a variety of NrdE homo- and heterodimers can form and interconvert on the timescale of the separation.

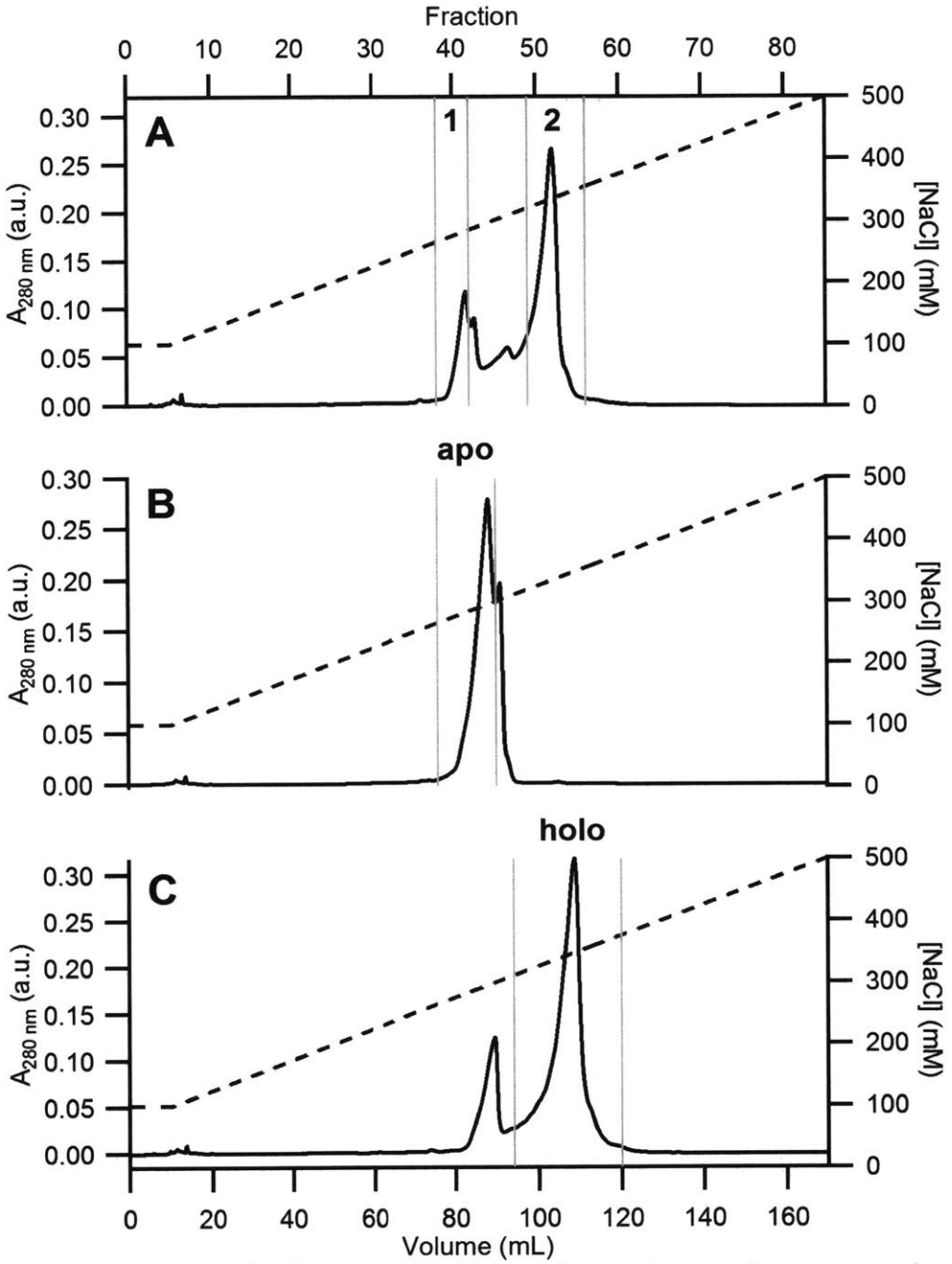


Figure 4.6. Separation of holo- from apo-NrdE by anion exchange chromatography. The solid line is the A_{280} and the dashed line the NaCl gradient. Note that all experiments were performed with the same gradient. Vertical grey bars indicate the limits of fraction pooling for each type of separation. (**A**) Separation of as-isolated NrdE. **1** = apo-NrdE fraction pool, **2** = holo-NrdE fraction pool. (**B**) Second round of chromatography for an apo-NrdE sample composed of **1** pools from multiple fractionations of as-isolated protein. (**C**) Second round of chromatography for a holo-NrdE sample composed of **2** pools from multiple fractionations of as-isolated protein.

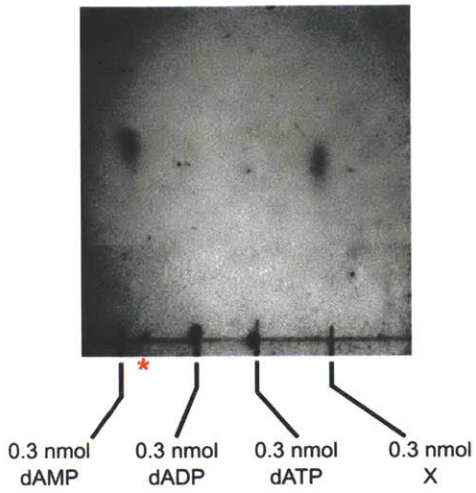


Figure 4.7. PEI-cellulose TLC analysis of nucleotide isolated from a sample of "homogeneous" holo-NrdE. The indicated amount of deoxynucleotide standards or the sample isolated from NrdE (lane X) were spotted on a 5 x 4 cm plate (non-fluorescent) that was developed in 0.5 M formic acid/formate, pH 3.4 and stained with Hanes-Isherwood reagent. R_f values: dAMP = 0.53, X = 0.51. The red * indicates an intensely staining artifact on the plate.

As an initial simplification, the chromatogram was treated as having three regions: peak **1**, peak **2**, and the zone in between the two peaks (**Figure 4.6A**). Quantitation of the amount of dAMP bound to the protein eluting in peak **1** (~260 – 270 mM NaCl) and peak **2** (330 – 350 mM NaCl) revealed that the former pool had less than 0.1 mol dAMP associated with the protein, whereas the latter pool had close to 1 mol dAMP bound per mol NrdE. Protein from the zone between the peaks had close to 0.5 dAMP bound per NrdE. This result demonstrates that the anion exchange chromatography is adequate for separating apo- from holo-NrdE. For their subsequent characterization, samples of apo- or holo-protein recovered from multiple fractionations of as-isolated NrdE were combined, concentrated, and rechromatographed to generate preparations of the desired protein in a more homogenous state (**Figure 4.6B** and **C**). Isolation and analysis of the nucleotide(s) bound to holo-NrdE by PEI-cellulose TLC revealed only the presence of dAMP (**Figure 4.7**), suggesting that peak **2** (**Figure 4.6A**) was mostly dAMP-loaded protein. Nucleotide

quantitation of the homogeneous stocks yielded 0.02 ± 0.02 mol dAMP bound per mol apo-NrdE and 1.0 ± 0.2 mol dAMP bound per mol holo-NrdE.

4.3.5. Bound dAMP appears to stabilize NrdE tertiary structure. As mentioned above, it was proposed that apo-NrdE is unstable, giving rise to the apparent increases in the amount of bound dAMP observed during the *in vitro* nucleotide loading experiments. This hypothesis is supported by the following observations made with apo-NrdE during the course of the subsequent studies described in this section. Apo-NrdE on ice is relatively stable in storage buffer (50 mM sodium phosphate, pH 7.6, 150 mM NaCl, 5% (w/v) glycerol, 10 mM DTT) for at least 3 h at concentrations up to ~50 μM monomer as judged by activity assays, which showed minimal activity loss over this timeframe. However, it precipitated readily during concentration procedures unless buffers are supplemented with 5% (w/v) glycerol and 150 mM NaCl. For the nucleotide binding experiments described later, in which glycerol was omitted from the buffer, the osmolyte was removed by gel filtration from stored samples of apo-NrdE immediately before use. In samples of apo-protein in glycerol-free buffer that were kept on ice, small amounts of protein precipitation were observed within 30 mins after the removal of glycerol. In contrast, holo-NrdE was observed to be stable in all of the conditions used in the subsequent experiments. Taken together, the results suggest dAMP binding can stabilize the tertiary structure of NrdE in solution, but the physiological relevance of this is unknown.

4.3.6. NrdE exhibits a higher propensity to oligomerize with dAMP bound. The effects of dAMP loading on NrdE quaternary structure were examined by SEC. As shown previously in Chapter 2, as-isolated NrdE behaves as a monomer at physiologically relevant concentrations (a representative example is shown in **Figure 4.8A**). Examination of apo- and holo-NrdE under similar conditions yielded nearly identical results (**Figure 4.8B**). The retention volumes (14.8 mL

for apo-NrdE and 14.1 mL for holo-protein) of the main peak in each experiment corresponded to molecular weights of 65 kDa (apo-NrdE) and 76 kDa (holo-NrdE), thus both forms of the enzyme appeared to be mainly monomeric. However, a small but significant amount of holo-NrdE eluted at a retention volume corresponding to a molecular weight of 133 kDa, consistent with the presence of a dimeric species. This observation plus the slightly earlier retention volume of holo-NrdE monomer suggested dAMP binding to NrdE imparts on the protein a higher propensity to dimerize relative to apo-protein.

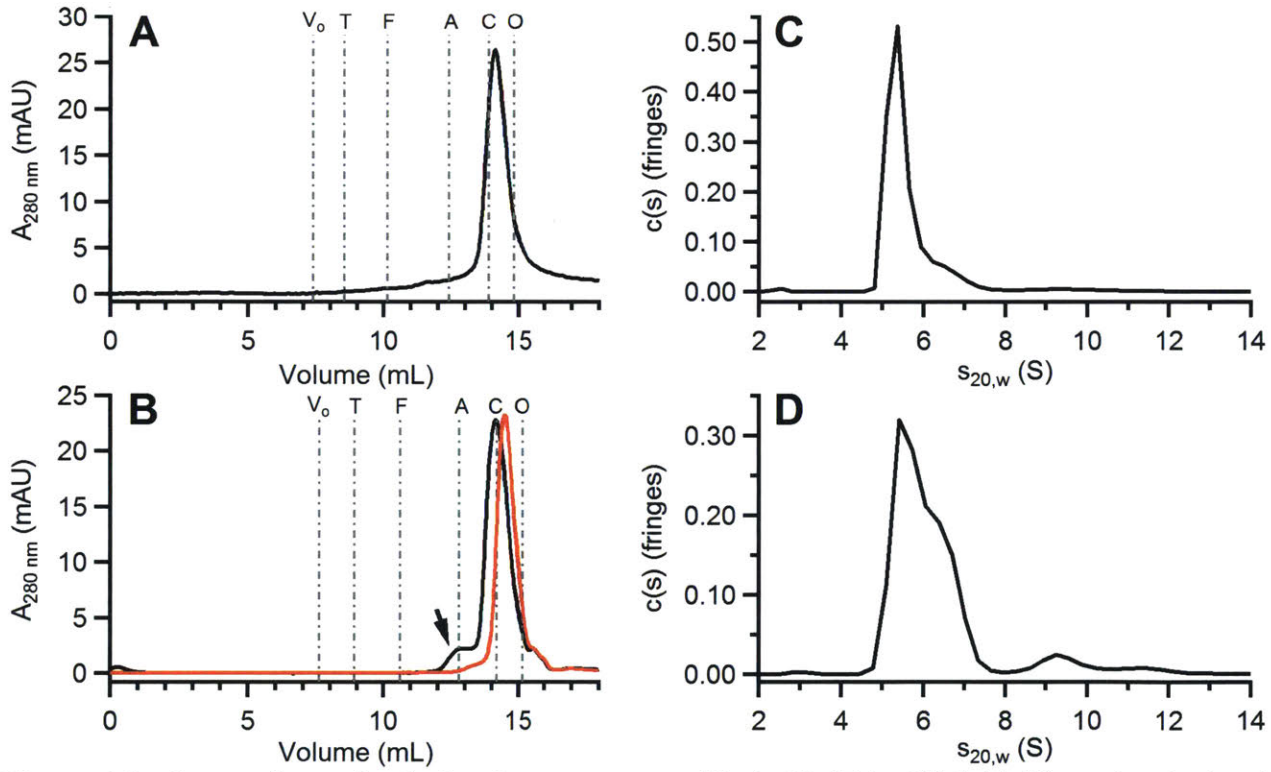


Figure 4.8. Comparison of as-isolated versus apo- and holo-NrdE by SEC (**A**, **B**), and as-isolated and holo-NrdE by SV-AUC (**C**, **D**). (**A**) SEC chromatogram for 2.7 μM as-isolated NrdE. (**B**) SEC chromatogram for 2.5 μM holo-NrdE (black trace) and apo-NrdE (red trace). The arrow points to the small amount of holo-NrdE eluting as a dimer. Vertical lines in the chromatograms indicate the retention volumes of molecular weight calibrants: V_o = void volume, T = thyroglobulin, F = ferritin, A = aldolase, C = conalbumin, O = ovalbumin. (**C**) Sedimentation coefficient distribution for 2.1 μM as-isolated NrdE. (**D**) Sedimentation coefficient distribution for 2.0 μM holo-NrdE.

The effect of dAMP on the quaternary structure of NrdE was further investigated using SV-AUC. The stability issues exhibited by apo-NrdE precluded its analysis by this technique, therefore comparisons are made only between holo- and as-isolated NrdE. As shown in Chapters 2 and 3, as-isolated NrdE sedimented as a mixture of monomers and dimers at physiological concentrations, with the majority (~80%) of the protein behaving as monomer (a representative example is shown in **Figure 4.8C**). The sedimentation coefficient distributions for samples of holo-NrdE run under identical conditions are distinctly different (**Figure 4.8D**). The amount of dimeric NrdE has increased substantially (~41% dimer and ~49% monomer), resulting in a rather broad peak encompassing $s_{20,w}$ values from 5.0 – 7.5 S. In addition, roughly 10% of the protein exists in the form of higher order oligomers appearing at ~9.5 S and ~11.6 S. The results of the SEC and SV-AUC analyses thus indicate that dAMP binding to NrdE increases the ability of the protein to oligomerize, but does not stabilize a particular quaternary structure of the protein.

4.3.7. Under non-inhibitory conditions, RNR activity is not affected by the presence of dAMP. Given that 2′-deoxyribonucleotides affect the specificity and overall activity of the *B. subtilis* RNR, as described in Chapter 3, it seemed feasible that the tightly bound dAMP could also alter one or both attributes of the enzyme. Thus, the effects of apo- versus holo-NrdE on the activity of the *B. subtilis* RNR were assessed using the spectrophotometric assay. Under routine assay conditions (1:1 ratio of 0.5 μM subunits, 1 mM CDP, 3 mM ATP), the specific activities of apo- and holo-NrdE (340 ± 4 and 350 ± 2 nmol min^{-1} mg^{-1} $\alpha_2\beta_2$, respectively) were indistinguishable from as-isolated protein (340 ± 2 nmol min^{-1} mg^{-1} $\alpha_2\beta_2$), thus the binding of dAMP appeared to have no effect on the overall activity of the enzyme.

Table 4.6. Steady-state parameters for *B. subtilis* class Ib RNR using apo- and holo-NrdE.[a]

NrdE form	V_{max} (nmol min^{-1} mg^{-1})[b]	K_m (μM)	k_{cat} (s^{-1})	k_{cat} K_m^{-1} (M^{-1} s^{-1})
1 mM UDP and variable ATP				
As-isolated[c]	520 ± 10	350 ± 40	2.1 ± 0.1	[6.0 ± 0.6] x 10^3
Apo[d]	320 ± 9	150 ± 20	1.3 ± 0.1	[8.7 ± 1.2] x 10^3
Holo[e]	460 ± 10	260 ± 30	1.8 ± 0.1	[7.0 ± 0.7] x 10^3
1 mM CDP with variable ATP				
As-isolated[c]	400 ± 20	200 ± 30	1.6 ± 0.1	[8.3 ± 1.3] x 10^3
Holo[e]	260 ± 9	60 ± 10	1.0 ± 0.1	[1.6 ± 0.3] x 10^4
1 mM ADP with variable dGTP				
As-isolated[c]	430 ± 6	1.7 ± 0.2	1.8 ± 0.1	[1.0 ± 0.1] x 10^6
Holo[e]	310 ± 10	1.1 ± 0.2	1.2 ± 0.1	[1.1 ± 0.2] x 10^6
1 mM GDP with variable TTP				
As-isolated[c]	220 ± 3	1.2 ± 0.1	0.9 ± 0.1	[7.8 ± 0.6] x 10^5
Holo[e]	240 ± 8	1.1 ± 0.1	1.0 ± 0.1	[9.0 ± 0.8] x 10^5

[a] Measurements made using tagless proteins unless otherwise indicated.
[b] Specific activity reported in terms of $\alpha_2\beta_2$.
[c] Data listed are for His_6-tagged proteins (from **Table 3.3**).
[d] Tagless apo-protein contained 0.02 mol dAMP per mol NrdE monomer.
[e] Tagless holo-protein contained 1.1 mol dAMP per mol NrdE monomer.

The steady-state kinetic measurements reported in Chapter 3 were made using as-isolated NrdE; therefore, any effect that dAMP has on k_{cat} and K_m were partially masked due to the presence of the apo-form of NrdE. With homogeneous samples of apo- and holo-NrdE now available, some of the steady-state parameters of *B. subtilis* RNR were reassessed to examine the role dAMP might have on activity. Effects on the potency of the effector nucleotides was examined by varying the concentrations of the effectors in the presence of 1 mM of the appropriate substrate. As apo-NrdE was obtained in small quantities, only the assessment of variable ATP concentrations in the presence of 1 mM UDP was completed. The results, tabulated in **Table 4.6**, are fairly similar to those observed previously for as-isolated NrdE (**Table 3.3**). In general, the values of both k_{cat} and K_m are lower for both apo- and holo-NrdE as compared to as-isolated protein, with the effects more pronounced with apo-NrdE versus holo-protein. However, the resulting catalytic efficiencies

are still identical, within error, to that observed with as-isolated NrdE; thus, the presence of dAMP appears to have no effect on the potency of the effectors in stimulating NDP reduction by *B. subtilis* RNR.

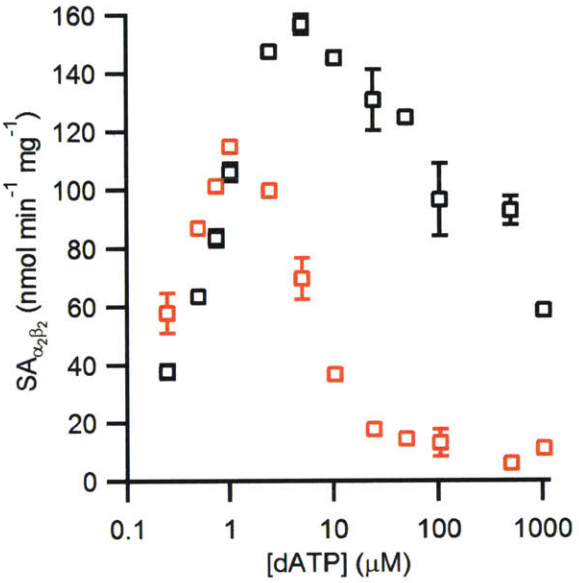


Figure 4.9. dATP inhibition of CDP reduction activity for apo-NrdE (black squares) and holo-NrdE (red squares). Assays were conducted in a total volume of 500 μL at 37 °C using a 1:1 ratio of apo- or holo-NrdE : $Mn(III)_2$-Y• NrdF, 1 mM CDP, and the indicated dATP concentration. Two replicates at each dATP concentration were run and the results averaged. In most cases, the error bars (± 1 standard deviation) are smaller than the size of the markers.

4.3.8. Apo- and holo-NrdE exhibit differences in dATP inhibition. As the *B. subtilis* Ib RNR is the first member of this subclass reported to exhibit inhibition by dATP, the effects of bound dAMP on this process were investigated. The results of activity assays examining the inhibitory activity of dATP on CDP reduction are shown in **Figure 4.9** and reveal a clear difference between apo- and holo-NrdE. Both forms of the protein exhibit nearly identical stimulation of CDP reduction in the range of 250 nm – 1 μM dATP. Between 1 – 5 μM dATP, CDP reduction by apo-NrdE continues to increase whereas inhibition of the holo-protein sets in. Both proteins are inhibited above 5 μM dATP, but apo-NrdE retains an overall higher activity (approximately 7.5-fold on average) at a given dATP concentration in comparison to holo-protein. The inhibition of

holo- NrdE is also much more potent, as only ~15% of its maximal activity is observed at a final dATP concentration of 25 μM. In contrast, apo-NrdE still exhibits 37% of maximal activity at 1 mM dATP. Therefore, the presence of dAMP associated with NrdE appears to amplify the inhibitory effects of dATP on the CDP reduction activity of the *B. subtilis* RNR.

4.3.9. Binding experiments suggest apo-NrdE can bind two equivalents of dATP/α whereas holo-NrdE can only bind one. The binding of [2,8-^{3}H]-dATP to apo- and holo-NrdE was investigated in an effort to better understand the results of the activity analysis shown in **Figure 4.9**. The results of these experiments show saturable dATP binding to apo-NrdE (**Figure 4.10A**), whereas with holo-NrdE, the binding curve is more complex (**Figure 4.10B**). To avoid overfitting, a rule of 20 data points per fittable parameter was adopted in considering which models to use to fit the binding data collected with apo- and holo-NrdE.[22] Therefore, given the number of data points collected with each protein (apo-NrdE = 60, holo-NrdE = 65), only models with two or three fittable parameters were considered.

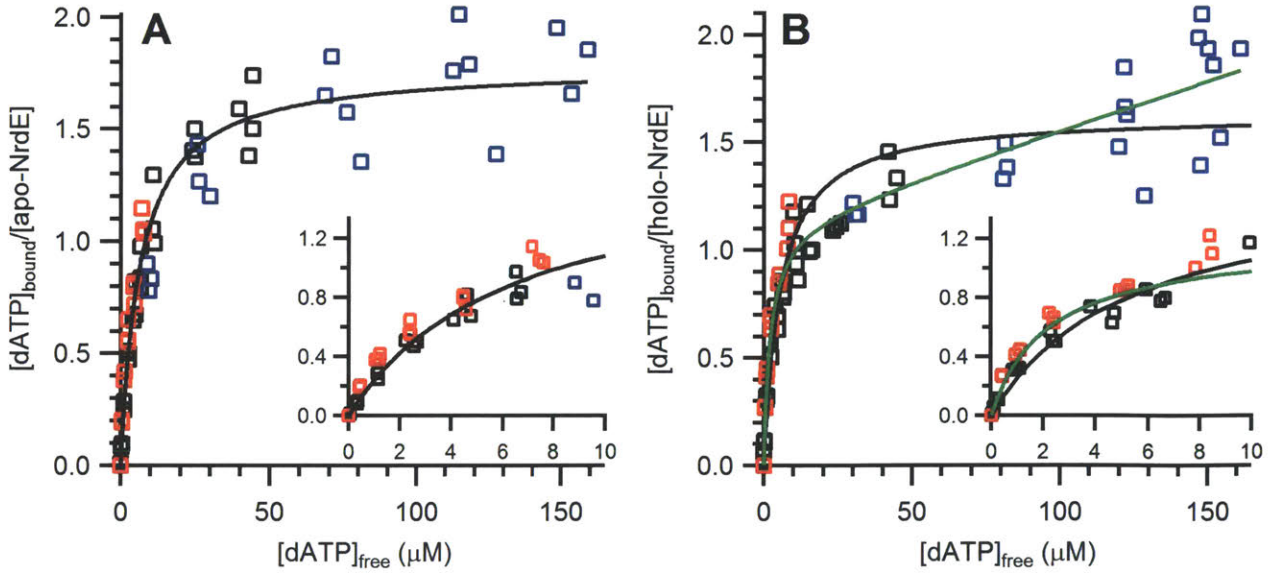


Figure 4.10. Saturation curves for dATP binding to (**A**) apo-NrdE, and (**B**) holo-NrdE. Insets show blow ups of the 0 – 10 μM free dATP concentration region for each plot. Each square represents one measurement. The concentration of NrdE used was 1 μM (black squares), 5 μM (red squares), and 12 μM (blue squares). Solid black lines indicate best global fit to eq 4.1. The solid green line in (**B**) represents the best global fit to eq 4.2.

Table 4.7. Best fit parameters for dATP binding to apo- and holo-NrdE.

Fit equation	Apo-NrdE		Holo-NrdE	
	n	***K*d (μM)**	***n***	***K*d (μM)**
4.1	1.8	6.4 ± 0.5	1.6	5.4 ± 0.6
4.2	-	-	1.1[a]	2.1 ± 0.4[a]

[a] Q_{nsb} = [4.0 ± 0.5] x 10^{-3} μM.

Both data sets were initially fit to the simplest binding model (eq 4.1), which assumes NrdE has *n* equivalent dATP binding sites with identical K_d values. The best fits (**Table 4.7**) suggest both apo- and holo-NrdE can bind up to two equivalents of [2,8-^{3}H]-dATP per monomer with an apparent K_d of roughly 6 μM. The fit to the holo-NrdE data set, however, is problematic given that a disproportionate number of points fall below the line in the range of 10 – 100 μM free dATP and above the line at free dATP concentrations greater than 100 μM; therefore, an alternative model was sought for the fitting of these data. In contrast, the model adequately fits the apo-NrdE data set and was accepted. Given the absence of an ATP-cone domain on the protein, this result is extremely interesting as it indicates that, in addition to the specificity site, *B. subtilis* NrdE possesses an additional site to which dATP can bind. This additional site may, perhaps, be the source of the inhibitory effects of dATP on RNR activity.

The non-saturating behavior of [2,8-^{3}H]-dATP binding to holo-NrdE is indicative of non-specific binding. This effect seems to be attributable to only holo-NrdE as it was not observed with apo-protein or in control assays using [2,8-^{3}H]-dATP alone or in the presence of BSA, which has no known nucleotide binding capacity. As noted in Chapter 3, aggregates had formed in SV-AUC samples of as-isolated NrdE in the presence of 100 μM and 500 μM dATP. Therefore, the appearance of non-specific binding with holo-protein might be caused by dATP-induced aggregation of holo-NrdE in solution. It has been noted previously that attempts to experimentally measure the non-specific binding of ligands often leads to overestimation of the number of binding

sites and underestimation of the K_d of the binding interaction.[18] These issues have led to the advocacy of incorporating non-specific binding as a fittable term in the total binding of a system.[18] In general, the model is described by equation 4.3, where n_{nsb} and $K_{d(nsb)}$ represent the number of non-specific sites with an overall apparent K_d.

$$\frac{[dATP]_{bound}}{[NrdE]_{monomer}} = \frac{n \times [dATP]_{free}}{K_d + [dATP]_{free}} + \frac{n_{nsb} \times [dATP]_{free}}{K_{d(nsb)} + [dATP]_{free}} \tag{4.3}$$

However, it is often the case in binding studies that the non-specific binding can be treated as unsaturable ($K_{nsb} >> [dATP]_{free}$) within the concentration range of ligand used. Therefore, the non-specific binding term can be simplified to one fittable parameter, Q_{nsb} (= n_{nsb}/K_{nsb}), as shown in eq 4.2.[18]

The holo-NrdE data set fits significantly better to eq 4.2 (**Figure 4.10B**), and the resulting parameters (**Table 4.7**) suggest that holo-NrdE can only bind one equivalent of dATP specifically with an apparent K_d of ~2 μM. When coupled with the results obtained with apo-NrdE, this interpretation of the data suggests that one site can bind both dATP and dAMP and is, therefore, surface exposed. If this interpretation of the data is correct, then it starkly contrasts with the results of the *in vitro* loading experiments reported in **Section 4.3.3**, and, furthermore, raises the question of how the site can bind dAMP (and possibly dATP) with apparently high affinity, but cannot acquire the nucleotide from free solvent pools. Further studies are required in order to establish whether or not dAMP and dATP can bind to the same site and the basis for the high affinity of the site for dAMP. As discussed later, the dAMP binding site is tentatively identifiable in the preliminary crystal structure of NrdE. Thus, repeating these binding experiments with NrdE mutants in which residues appearing to be important for nucleotide binding are removed may provide further insight into these perplexing observations. Regardless, from the activity and

binding data presented here, it can be concluded that the association of dAMP with NrdE can modulate the ability of dATP to inhibit RNR.

4.3.10. dAMP loading does not influence the ability of NrdE and NrdF to form a qualitatively "tight" complex. As shown by Zhang *et al.*, NrdE and NrdF co-purified in an approximately 1:1 ratio through three columns when isolated from *B. subtilis* cells.[23] This unprecedented result starkly contrasts with studies of other class Ib RNRs, where the subunits, although forming a qualitatively "tight" interaction *in vitro*,[24] still separate during purification from the endogenous host and are individually isolated.[25-28] The discovery of a tightly bound dAMP associated with NrdE, although not yet confirmed to be physiologically relevant in *B. subtilis*, raises the question of whether or not the nucleotide is the factor responsible for the unusually tight interaction between the subunits of RNR isolated from the endogenous host.

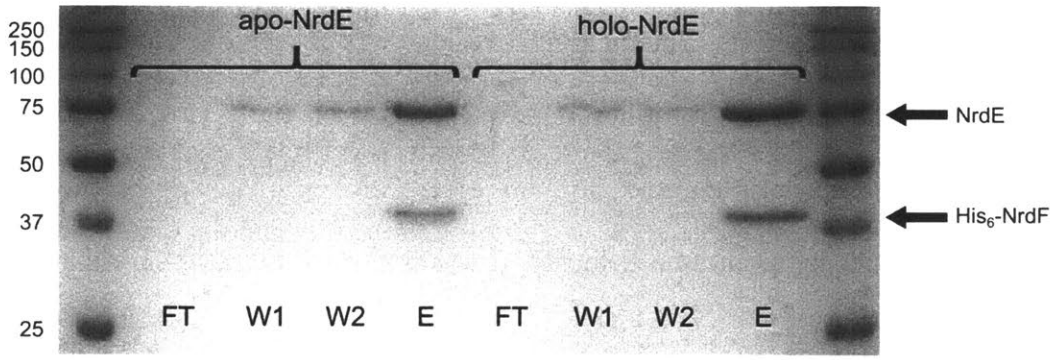


Figure 4.11. Qualitative pull-down of apo- and holo-NrdE by $Fe(III)_2$-Y• His_6-NrdF. The individual steps are indicated: FT = flow through, W1 = first wash, W2 = second wash, and E = elution. If one form of NrdE interacted preferentially with NrdF, it would be expected to show up in the elution fraction. If the other form of NrdE did not, it would have been expected to end up entirely in the flow through fraction.

To test this hypothesis, a qualitative pull-down assay using Ni-NTA affinity resin was carried out using a 1:1 ratio of recombinant tagless apo- or holo-NrdE as the prey and $Fe(III)_2$-Y• His_6-NrdF as the bait. If His_6-$Fe\beta_2$ is able to pull one form of NrdE down, but not the other, the

results expected from these experiments were that apo- and holo-NrdE would elute in different fractions, as judged by SDS-PAGE. The form of NrdE that could interact with His_6-$Fe\beta_2$ should be retained with the resin until elution with buffer containing 250 mM imidazole, whereas the non-interacting form would wash out using buffers with low imidazole concentrations. The results are shown in **Figure 4.11** and revealed that His_6-tagged NrdF was equally capable of pulling down both forms of NrdE, indicating that the "tight" interaction between the two subunits occurs independently of dAMP bound to NrdE.

4.3.11. Purification and characterization of NrdE-$C_{382}S$. In order to determine if the presence of substrate and effector nucleotides together are capable of stabilizing a NrdE/NrdF complex, the *B. subtilis* RNR needed to be inactivated without introducing major perturbations to the overall quaternary structure of the enzyme. For this purpose, site-directed mutagenesis was used to substitute C_{382} in NrdE, which is the essential Cys that is transiently oxidized to a thiyl radical each turnover (Table 1.2), with serine. This particular residue was chosen over other radical transfer pathway residues because C_{382} is the furthest removed residue from the α/β interface and, therefore, should have minimal impact on complex formation. However, as pathway blocked mutants of class Ib RNRs have not be reported previously, it was important to make sure that NrdE-$C_{382}S$ behaved like wildtype enzyme with the exception that it was inactive. Both $C_{382}S$ and wildtype NrdE behaved similarly during purification, and the mutant was found to still have 0.6 equiv of tightly bound dAMP associated with it. Furthermore, the apo- and holo-forms of NrdE-$C_{382}S$ could be separated from one another by anion exchange chromatography, which produced chromatograms similar to those shown in **Figure 4.6** for as-isolated and holo-NrdE. Quantitation of the amount of protein and dAMP loading revealed similar yields relative to wildtype (holo-NrdE-C382S loading = 0.95 dAMP/α, only measured on one preparation). The mutant was able

to form a tight complex with $Fe(III)_2$-Y• NrdF as indicated by using pull-down assays, which produced data analogous to that shown in **Figure 4.11**. Finally, NrdE-$C_{382}S$ did not measurably reduce CDP as judged by the radioactive assay, indicating the mutant was inactive as expect.

4.3.12. Biophysical analysis of holo-NrdE:NrdF complexes indicate dAMP does not influence overall RNR quaternary structure. Even though the interaction between as-isolated wildtype NrdE and NrdF appears to be tight, the complexes formed from the two proteins rapidly interconvert on the timescales of SEC and SV-AUC as shown in Chapter 2. It is possible, however, that dAMP might have an effect on complex stability which was masked by peak broadening resulting from structural interconversions involving apo-NrdE. To test this hypothesis, a 1:1 mixture of 2 μM tagless holo-α_2-$C_{382}S$ and $Mn(III)_2$-Y• β_2 were analyzed by SV-AUC in the absence of nucleotides. The results of this experiment are shown in **Figure 4.12** along with a $s_{20,w}$ distribution generated from a sample run under similar conditions using as-isolated wildtype NrdE. Although the detection system used in each case to monitor sedimentation was different (A_{280} for samples using as-isolated NrdE and interference optics for samples using holo-NrdE), the results from both experiments are similar. The $\alpha_m(\beta_2)_n$ complexes formed with holo-NrdE are still unstable and interconvert on the timescale of sedimentation, as evidenced by the reaction boundary between the $s_{20,w}$ values expected for the individual subunits and that expected for the $\alpha_2\beta_2$ complex. The peak at $s_{<Wt>}[20,w] = 5.9$ S (best fit $M_W = 78$ kDa) is consistent with NrdE monomer whereas the peak at $s_{<Wt>}[20,w] = 10.1$ S (best fit $M_W = 175$ kDa) is consistent with an asymmetric $\alpha_2\beta_2$ complex similar to the holo-RNR structure (PDB 2BQ1) of the *S.* Typhimurium class Ib RNR.[29] The higher order complex sedimenting at ~15 S (0.4% of the total signal) observed in the sample prepared with holo-NrdE (**Figure 4.12B**) may indicate that dAMP binding facilitates formation of larger $\alpha_m(\beta_2)_n$ complexes. However, because interference optical detection is much

more sensitive than A_{280} detection, it is likely that a similar complex existed in the sample prepared with as-isolated wildtype NrdE, but that its concentration was too low to be detected by A_{280}.

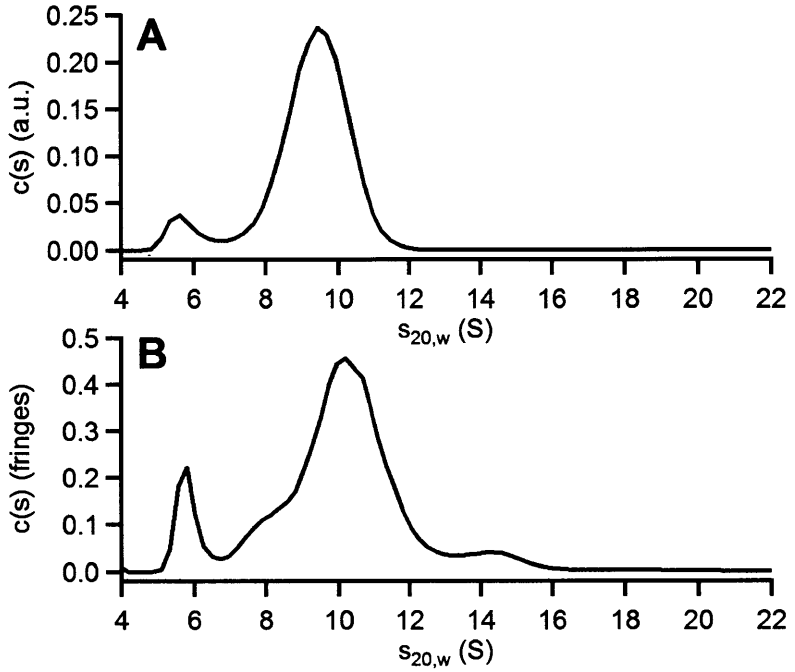


Figure 4.12. Sedimentation coefficient distributions for 1:1 mixtures of ~2μM $Mn(III)_2$-Y• β_2 and (**A**) as-isolated α_2 (monitored by A_{280}), or (**B**) holo-α_2-$C_{382}S$ (monitored by interference).

4.3.13. SV-AUC analysis of a 1:1 mixture of holo-NrdE-$C_{382}S$:$Mn(III)_2$-Y• NrdF in the presence of GDP and TTP reveals $\alpha_2\beta_2$ and higher order complexes. As dAMP alone is unable to stabilize a complex of NrdE and NrdF, a SV-AUC experiment with substrate (GDP), effector (TTP), and a 1:1 mixture of 2 μM holo-α_2-$C_{382}S$:$Mn(III)_2$-Y• β_2 was run to determine if a stable RNR complex could be observed under these conditions. The results are shown in **Figure 4.13** and reveal the formation of very large complexes of NrdE and/or NrdF ($s_{<Wt>}$[20,w] spanning 13.6 – 37.3 S, M_{WS} spanning 310 kDa – 1 MDa) that interconvert on the timescale of sedimentation. The distribution is reminiscent of those observed in similar experiments run in the presence of dATP (**Figure 3.15**). However, unlike these previous results, two prominent peaks with a

$s_{<Wt>}$[20,w] = 6.3 S and 10.0 S containing a substantial fraction (42%) of the total protein were observed (**Figure 4.13**). The 10 S species is the largest contributor to the total signal (34%) of any species in the distribution. The peak at 6.3 S (best fit M_W = 69 kDa) is most consistent with NrdE monomer, while the peak at 10.0 S (best fit M_W = 139 kDa) is consistent with the asymmetric holo-RNR complex.[29] The results are encouraging, but still show that *B. subtilis* NrdE:NrdF complexes are temporally unstable, even in the presence of substrate and effector nucleotides. Furthermore, the tightly bound dAMP associated with NrdE does not play an observable role in the formation or stabilization of active NrdE:NrdF complexes.

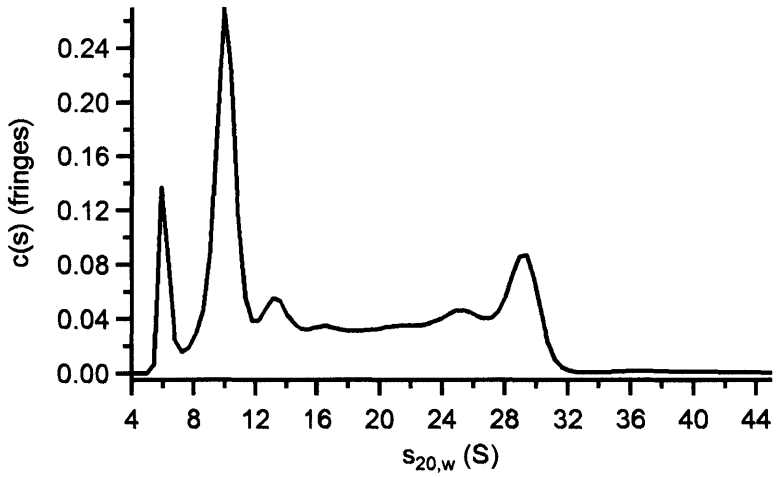


Figure 4.13. Sedimentation coefficient distribution for a 1:1 mixture of 2 μM holo-α_2-$C_{382}S$: $Mn(III)_2$-Y• β_2 supplemented with 12 μM TTP and 250 μM GDP.

4.3.14. A tightly bound dAMP associated with NrdE appears to be unique to the *B. subtilis* system. The discovery of a tightly bound equivalent of dAMP associated with *B. subtilis* NrdE prompted an examination of additional class Ib NrdE proteins available in the laboratory to see if this phenomenon was observed with other class Ib enzymes. Samples of *E. coli*[30] and *S.* Typhimurium (Marica Merkel Johannson, Parker, and Stubbe, unpublished) NrdE (each was 9 nmol in 100 μL) were isolated, denatured with heat, and the supernatants examined subsequent to

centrifugation. No detectable UV absorption features from 250 – 300 nm were observed in either case. If the nucleotide was present, it would correspond to a loading of less than 0.07 dAMP per NrdE. As noted in previous chapters, the *E. coli* and *S.* Typhimurium Ib RNRs are unique in that they come from organisms in which a class Ia RNR is the main system for dNTP synthesis and are not representative of the Ib enzymes in general. Examination of NrdEs from other Ib systems which are used solely as the aerobic RNR that is required for DNA replication and repair will establish whether a tightly associated dAMP is unique to the *B. subtilis* RNR, or is a more widespread phenomenon.

4.4. DISCUSSION

RNRs, because of their role as the only *de novo* source of dNTPs for replication and repair, are the subjects of several mechanisms of regulation to ensure that balanced pools of the building blocks of DNA are maintained.[31] One such mechanism is to tune overall activity of RNR in response to the dATP/ATP ratio inside the cell. The basis for this regulation has recently been shown to be interconversion between an active $\alpha_2\beta_2$ complex (low dATP/ATP ratio) and inhibitory quaternary structures, such as the *E. coli* $\alpha_4\beta_4$ ring[32] and human α hexamer,[33] when the dATP/ATP ratio is high. These interconversions are mediated by ATP-cone domains[34] fused to the N-terminus of the α subunit (and in a few cases β[35]) of RNR. Until now, overall activity regulation of RNR activity has been attributed to the presence of an ATP-cone[31] since systems lacking these structures, such as the class Ib RNRs from *S.* Typhimurium, *Lactococcus lactis*, *Mycobacterium tuberculosis*, and *Streptococcus sanguinis*, are not inhibited by dATP.[24, 25, 36-38]

The potent inhibition of the *B. subtilis* class Ib RNR by dATP reported in the previous chapters of this thesis showed, for the first time, that overall activity regulation of RNR is not solely mediated by ATP-cone domains. Similar to the class Ia RNRs, the mechanism of inhibition

again appears to be formation of an alternative quaternary structure (**Figure 3.15**), the structural basis for which is currently unknown because the *B. subtilis* RNR does not possess an ATP-cone domain. The results presented in this chapter add a new level of complexity to the novel mechanism of overall activity control of the *B. subtilis* RNR as a hitherto unrecognized tightly bound equivalent of dAMP associated with NrdE has been identified and found to augment the susceptibility of the enzyme to dATP inhibition. This discovery came about due to the observation of a previously undetected quaternary structure of NrdE which was interpreted to be an altered conformer of the monomeric protein that was, perhaps, related in some way to the ability of dATP to inhibit the enzyme. Influenced by previous observations from the Stubbe group on another enzyme involved in nucleotide metabolism, FGAR-AT,[2, 4] it was initially thought that the alternative structure resulted from a tightly bound equivalent of dATP. While the hypothesis turned out to be correct in some regards, the isolation of nucleotide mixtures from recombinant NrdE that were predominantly composed of dAMP and the inability to increase the dAMP load of NrdE by addition of exogenous nucleotide, despite the apparently solvent-exposed nature of the binding site, revealed a more complicated system than originally expected.

4.4.1. A role for tightly bound nucleotides in RNRs is unclear. The tight association of nucleotides with RNRs is rather unprecedented given the transient nature of substrate and effector binding that is required in order for the timely adjustment of the enzyme's specificity and activity in response to changes in the relative ratios and amounts of the dNTP pools. Although rare, the existence of tightly-bound nucleotides has been inferred previously for the T4 phage class Ia and the *E. coli* and *L. lactis* class III RNRs.[39, 40] In these cases, the ligand is suspected to be a tightly bound equivalent of dATP. However, rather than being an intrinsic attribute of the enzyme, the presence of the nucleotide has been, at least in the case of the class III enzymes, disregarded as an

artifact resulting from the use of a dATP-Sepharose affinity column during protein purification.[40] In this affinity resin, dATP is linked to the Sepharose backbone via a phosphoanhydride bond (**Figure 4.14**) which is susceptible to hydrolysis and, therefore, could explain the isolation of protein associated with nucleotide.

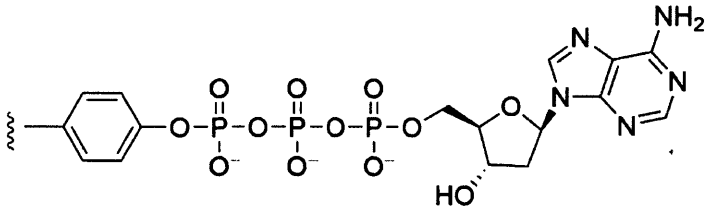


Figure 4.14. The linkage between dATP and the resin in dATP-Sepharose.

In the cases of the T4 phage Ia and the *E. coli* and *L. lactis* class III RNRs, the presence of the putative tightly bound dATP initially went unrecognized for quite some time and, consequently, the results of early studies of these enzymes may have been misinterpreted. In the case of the T4 phage class Ia RNR, the subunits co-purified with one another, which was distinctly different from the *E. coli* Ia RNR and led to the conclusion that the subunit interaction in the T4 phage enzyme was "tight."[41] Later studies by Christopher Mathews' group on this RNR revealed that the subunits readily dissociated when the enzyme was purified without the use of a dATP-Sepharose affinity column, and that a "tight" complex could be reconstituted with the purified subunits using any of the dNTP effectors.[39] Thought not explicitly stated, the results suggest that the T4 phage RNR might have acquired dATP on the affinity column, leading to the purification of a tight complex. The results obtained with the *E. coli* and *L. lactis* class III RNRs were more complicated.[42, 43] Both enzymes possessed two classes of effector binding sites, but unlike the *E. coli* Ia RNR, ATP was able to bind to only one of these sites and stimulated the reduction of pyrimidine substrates. Addition of TTP and dGTP, which bound to the second allosteric site, appeared to switch the substrate specificity to purines, whereas dATP could bind to both classes

of sites and inhibited RNR activity. From these results, it was concluded that these class III enzymes had mutually exclusive specificity sites controlling the reduction of purine and pyrimidine substrates, and that dATP binding to either site could inhibit the enzyme. It was stated, but not shown, that subsequent studies revealed that these class III RNRs have a single type of specificity site and a single type of overall activity similar to the *E. coli* Ia RNR.[40]

The discussion of the T4 phage and bacterial class III RNRs above illustrates some of the difficulties that are encountered when working with proteins that can bind multiple nucleotides with distinctly different affinities. Several factors, including the *in vivo* concentrations of the protein, the concentrations of nucleotide ligands, the temporal expression of the protein, environmental cues, and, in eukaryotes, sub-cellular localization, all contribute to regulating the ligand binding capabilities of a protein. By removal of the protein from the context of the natural cellular environment, these regulatory mechanisms are lost and, thus, the protein will bind those nucleotides for which it has the highest affinity. This is the case for RNRs that have ATP-cone domains and bind dATP with a much higher affinity than ATP. Inside the cell, the relative concentrations of these two nucleotides indicate that ATP-cone domains on RNRs will most often be occupied by ATP rather than dATP

The work reported in this chapter is the first systematic study examining the effects of a tightly bound nucleotide on the structure and function of any RNR. These studies were pursued due to the unusual dATP inhibition exhibited by the *B. subtilis* enzyme and to the fact that the most abundant of the nucleotides bound to *B. subtilis* NrdE was dAMP. dAMP is a product of DNA degradation and currently there is no data supporting a possible role for it in regulating dNTP pools in *B. subtilis*. The fact that the observation of dAMP bound to NrdE has been reproduced by another member of the Stubbe lab suggests the nucleotide may be physiologically relevant,

although this remains to be determined. A hypothesis in support of a role for dAMP in *B. subtilis* RNR function is discussed in Chapter 5. YmaB, the fourth protein of the Ib RNR operon in *B. subtilis*, is shown to have the ability to cleave dATP into dAMP and pyrophosphate. Given that genes organized into operons are often functionally associated,[44] YmaB could act as the source of the NrdE-associated dAMP. Ultimately, NrdE will need to be isolated from *B. subtilis* and analyzed for tightly bound nucleotides in order to determine whether or not dAMP is associated with the protein. It may also be possible that (d)ADP and/or dATP is the physiologically relevant nucleotide given that they were also observed to be associated with NrdE. These results are particularly intriguing given recent studies that have established that (deoxy)nucleoside mono-, di-, and triphosphates are bound to NrdR, a transcription factor that universally regulates expression of RNR operons in response to *in vivo* dNTP concentrations.[45]

4.4.2. A preliminary crystal structure of NrdE suggests the putative dAMP binding site is solvent exposed. As mentioned in the Results section, one of the most surprising observations was the fact that the dAMP was quite inert to removal by dialysis (**Table 4.1**). This observation, the observed sub-stoichiometric loading of dAMP relative to NrdE, and the fact that exogenously added dAMP (or dATP) was unable to fully complement NrdE, strongly suggested that the nucleotide binding site is buried in the protein tertiary structure. In contrast, the binding assay results for apo-NrdE and holo-NrdE, as mentioned above, indicate dATP and dAMP can bind to the same site on NrdE, suggesting that it is solvent exposed.

In an attempt to resolve this puzzling issue, crystallographic studies in collaboration with the Boal group at Pennsylvania State University have been recently initiated to determine the location of the dAMP binding site in NrdE. This work was made possible by the ability to separate holo- from apo-NrdE by anion exchange chromatography (**Figure 4.6**), thus allowing

homogeneous samples to be prepared in order to maximize the chances of finding dAMP. The strategy worked as a preliminary structure of *B. subtilis* NrdE co-crystallized with the dAMP has been solved by Ailiena Maggiolo (Boal group) to 3.1 Å. The protein was crystallized from a solution of 0.1 M sodium citrate, pH 4.0, 1.8 M NH_4SO_4 in an anaerobic glovebox. Though the pH at first glance is concerning, it appears to have not effected the overall protein structure as the *B. subtilis* NrdE is nearly superimposable with the previously reported *S.* Typhimurium NrdE structure (PDB 1PEM[1]) (**Figure 4.15**). The model is missing the last 13 residues of the C-terminus, which is common with class I α structures given that the tail is flexible so that it can enter the catalytic site and reduce the disulfide bond after each turnover.[46] There are also three other disordered regions (**Figure 4.15B**), two of which are localized at the specificity site and are likely disordered due to the absence of effector nucleotide. The disorder in the remaining region may reflect the preliminary nature of the model.

The dAMP binding site is located in the N-terminal domain of NrdE in a small pocket located between the second and fifth α-helices (*S.* Typhimurium secondary structure numbering[1]) (**Figure 4.15**). The most surprising aspect of the site is that it is clearly solvent exposed with the only real anchor holding the nucleotide in place being the adenine base (**Figure 4.16A**). The base is lodged in a hydrophobic pocket composed of V_{33}, F_{37}, the β – δ methylene groups of R_{90}, and F_{91}. Hydrogen bonding contacts between the protein and the nucleobase are made by the backbone amide and carbonyl group of F_{91} with N1 and the exocyclic N6 amine, respectively (**Figure 4.16A**). These backbone interactions fit the definition of a "reverse" adenine-binding motif found in many proteins that can bind adenine-containing nucleotides and cofactors.[47] Further stabilization of the base comes from hydrophobic stacking interactions with the sidechain of F_{37} and the γ and δ methylene groups of R_{90} on opposite sides of the adenine ring. Three other contacts

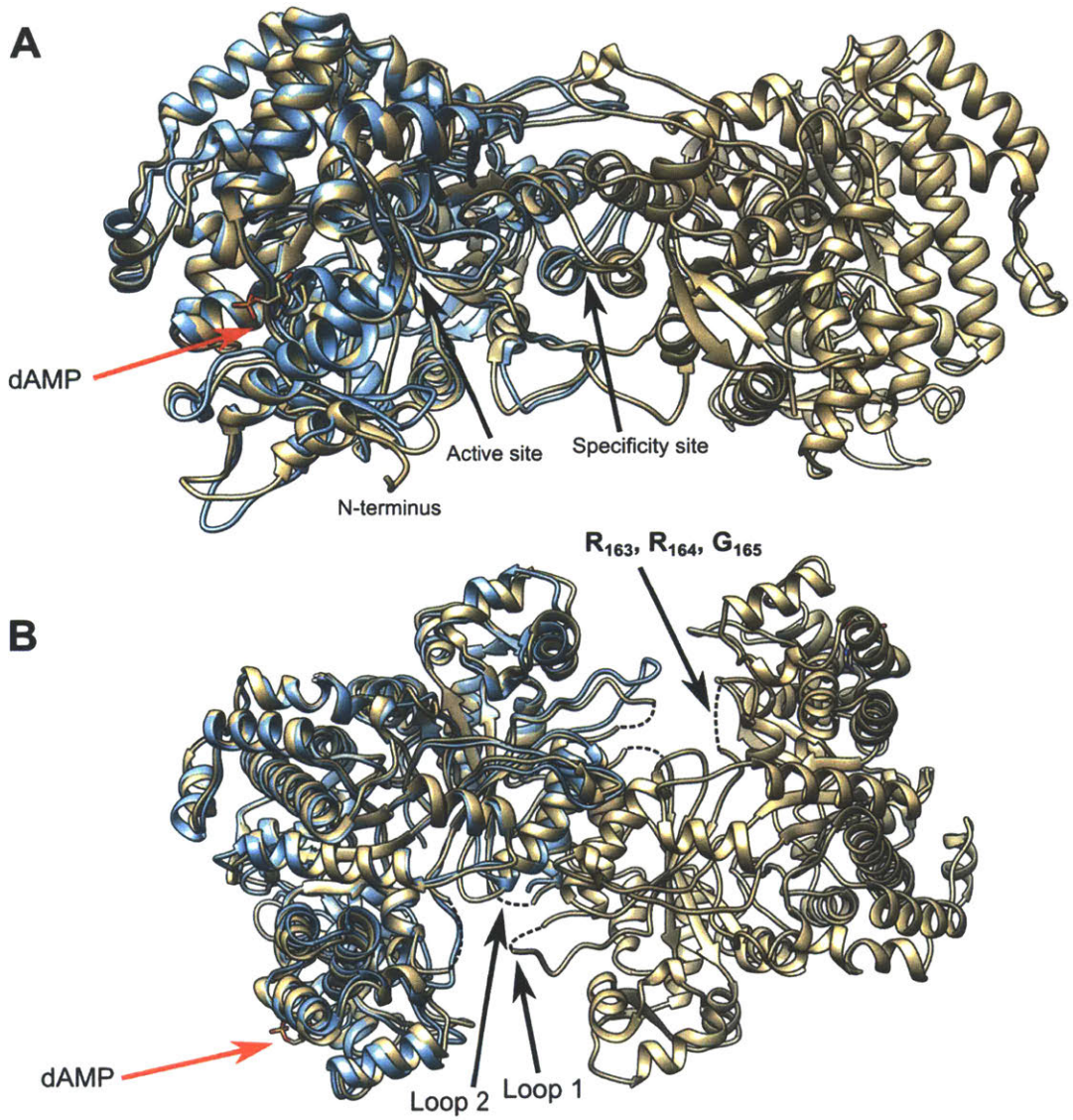


Figure 4.15. Three-dimensional structure of *B. subtilis* NrdE (tan) solved to 3.1 Å by Ailiena Maggiolo (Amie Boal group, Pennsylvania State University). Two views are presented: (**A**) front, and (**B**) top. The structure of *S.* Typhimurium NrdE (blue, PDB 1PEM) has been overlaid on one monomer of the *B. subtilis* protein to illustrate the structural homology of the two proteins (RMSD = 1 Å over 576 atom pairs). The dashed lines shown in (**B**) indicate disorder regions in the structure. Two regions are located in loops 1 and 2 of the specificity site and are likely disordered due to the absence of an effector nucleotide. The third region is indicated and includes three residues, R_{163}, R_{164}, and G_{165}.

between the protein and the rest of the dAMP molecule can be discerned in the structure. The 3′-OH of the ribofuranose moiety forms a hydrogen bonding interaction with sidechain amide

nitrogen of N_{42}, and the phosphate forms hydrogen bonds with the Nε of R_{90} and the imidazole sidechain of H_{34} (**Figure 4.16A**).

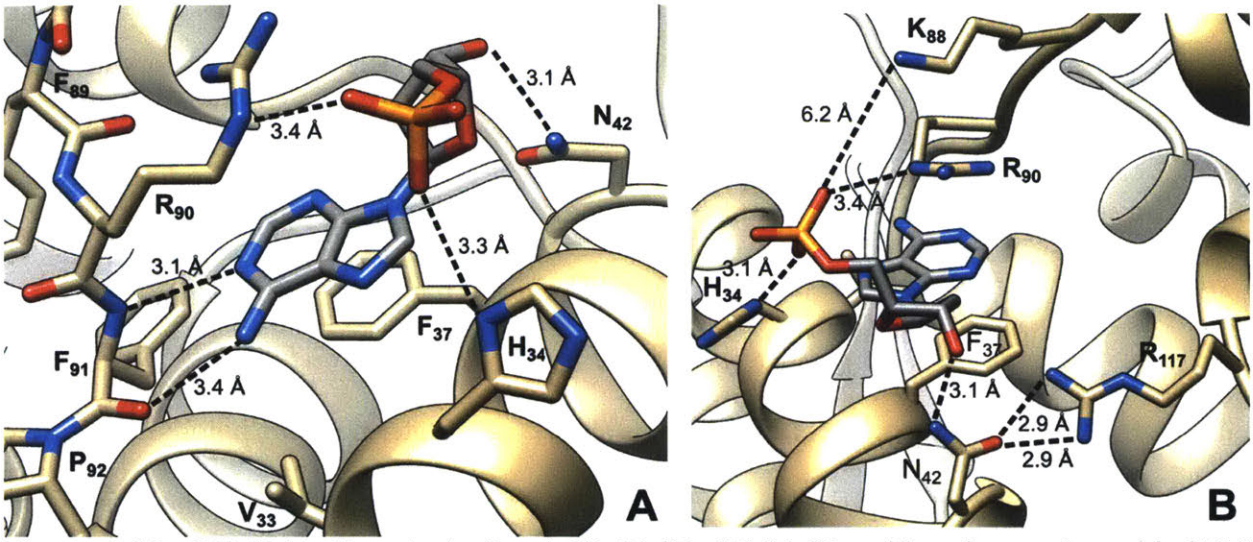


Figure 4.16. dAMP binding site in *B. subtilis* NrdE. (**A**) NrdE residues interacting with dAMP. Dashed lines indicate hydrogen bonding interactions with the indicated distances. (**B**) Potential ligands from NrdE for dAMP that could be affected by the acidic pH of crystallization are shown in boldfaced font. Dashed lines indicate hydrogen bonding interactions except for K_{88}, where the distance reported is the shortest distance of approach between the Nζ and the phosphate O atom.

The large solvent exposure of the binding site raises several questions, the foremost of which is how the nucleotide is held so tightly that it cannot be removed with dialysis and why is it not possible to increase the basal dAMP load of the protein with exogenously added nucleotide. No clear reasons for this discrepancy are offered by the structure. One possible explanation is that the non-physiological pH of the crystallization conditions might have altered the protonation state of other residues responsible for binding dAMP. For example, although H_{34} is within hydrogen bonding distance of the phosphate group of dAMP, the sidechain does not seem positioned appropriately to form a strong hydrogen bond with the ligand (**Figure 4.16A**). Protonation of the phosphate group of dAMP, too, might affect the ability of residues like K_{88} and R_{90} (**Figure 4.16B**) to form salt bridges due to the neutralization of the negative charges. There is also a possibility

that residues deeper in the cleft could also participate in binding dAMP, such as R_{117} (**Figure 4.16B**) and E_{119} (not shown). If either of these residues does participate in binding dAMP to NrdE at neutral pH, the nucleotide would need to be reoriented so that potential hydrogen bonds between the side chain and dAMP could form. Such a reorientation of the nucleotide would bury it deeper in the protein structure, thereby making it more difficult to get dAMP in or out of the binding site. Efforts are currently underway to find more physiologically relevant crystallization conditions for *B. subtilis* NrdE to investigate this possibility.

4.4.3. dAMP increases the susceptibility of B. subtilis RNR to dATP inhibition. Despite being unable to load NrdE with dAMP *in vitro*, the ability to separate apo- from holo-NrdE allowed investigations to commence to determine what affect the dAMP molecule has on the activity and structure of NrdE and/or the *B. subtilis* RNR. The main conclusion that can be drawn from the biochemical and biophysical data presented in this chapter is that the presence versus absence of the nucleotide affects the ability of dATP to inhibit the reduction of CDP and, by inference, the other NDPs. With holo-NrdE, dATP was found to be a more potent inhibitor (**Figure 4.9**), suggesting that the presence of dAMP resulted in pre-formation of the dATP binding site. While this site may be separate from the dAMP site observed in the crystal structure, the dATP binding experiments, as described above, suggest that the dAMP and dATP binding sites are one and the same. If this is true, then one possible reason for the increased potency of dATP in the presence of holo-NrdE could be that a cooperative binding effect is occurring where initial occupation of the allosteric site by dAMP preforms it and makes the exchange of dAMP for dATP easier to come about. This effect is similar to that reported for phosphate ion and FMN binding to *Desulfovibrio vulgaris* flavodoxin[48] which inspired the FMN reconstitution procedure for NrdI described in Chapter 3. Such a scenario might also explain the non-specific dATP binding observed with holo-

NrdE (**Figure 4.10B**) since increasing dATP concentrations would make the exchange of dAMP for dATP in the binding site occur more readily.

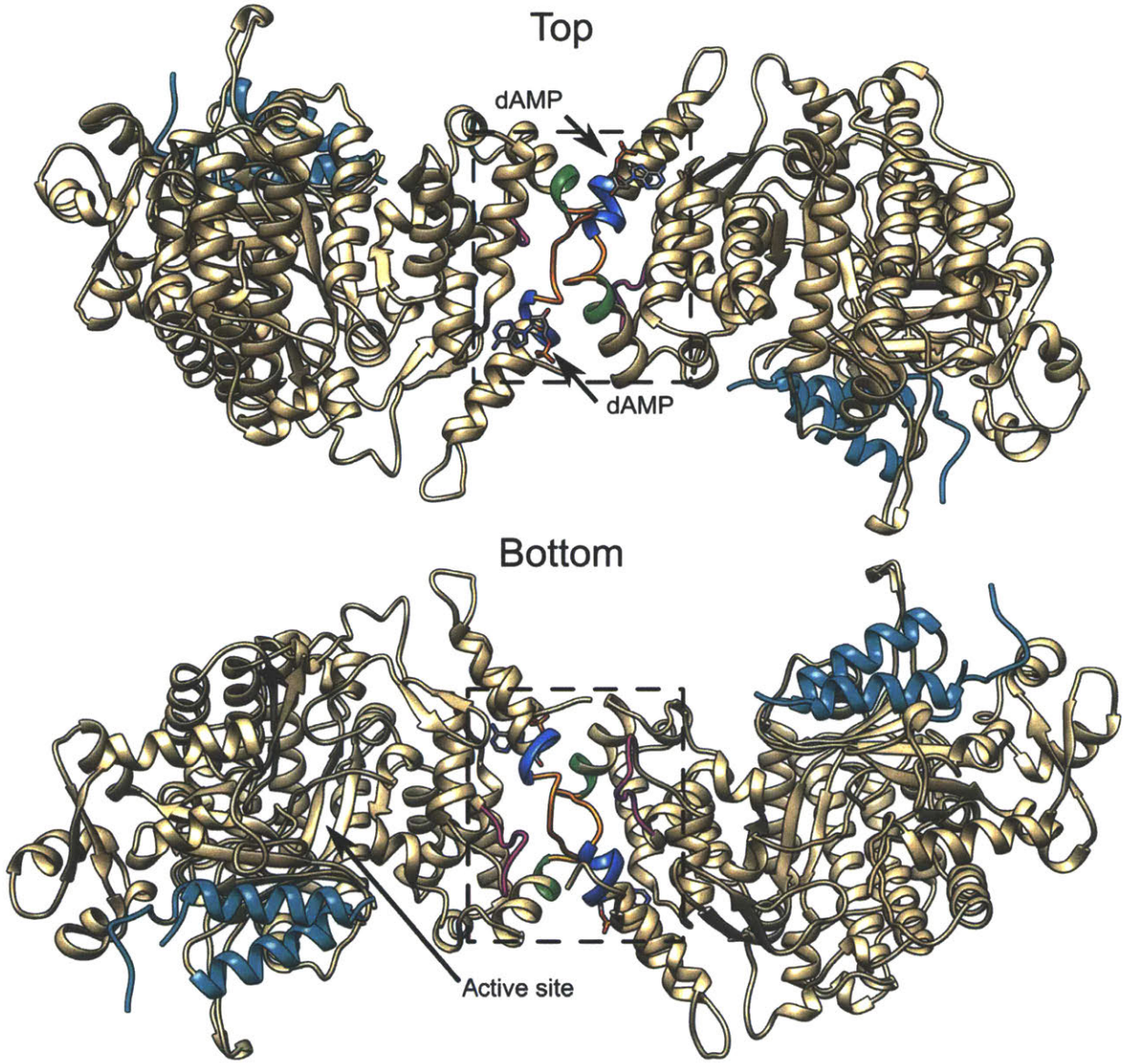


Figure 4.17. *B. subtilis* NrdE dimer induced by dAMP binding. Residues participating in the canonical dimer interface (**Figure 4.15**) are colored cyan. The dashed boxes indicate the regions blown up in **Figure 4.18** to show the contacts made at this atypical dimer interface. Colored secondary structures are as follows: blue = V_{38} – I_{41} or Q_{43}, orange = N_{42} or N_{44} – N_{50}, green = L_{51} – E_{53}, magenta = E_{110} – R_{117}.

4.4.4. dAMP may cause NrdE to form an atypical dimer with its N-terminal domain. Little else about the activity of the *B. subtilis* RNR was affected by the presence of dAMP associated with NrdE, and the nucleotide was also found to have little influence on the ability of stable NrdE:NrdF complexes to form. In contrast, holo-NrdE alone showed a higher propensity to oligomerize as compared to apo-protein (**Figure 4.8**), and the hydrodynamic properties of the quaternary structure(s) formed are consistent with a dimer. However, it is difficult imagine how the binding of dAMP at the N-terminal end of NrdE would facilitate the formation of a dimer similar to the one shown in **Figure 4.15**. Instead, the crystal structure of NrdE suggests that an overall different and unexpected quaternary structure may form if dAMP is bound to NrdE.

Crystallographic symmetry mates of NrdE protomers were found to have crystallized with their N-terminal domains juxtaposed with one another (**Figure 4.17**), resulting in ~650 $Å^2$ of both protomers being buried at the interface. The interaction surface encompasses the C-terminal end of α-helix #2 (V_{38} – I_{41} or Q_{43}), the loop connecting helix #2 and #3 (N_{42} or N_{44} – N_{50}), the N-terminal end of helix #3 (L_{51} – E_{53}), and a loop composed of residues D_{110} – R_{117} (**Figure 4.18**). The two dAMP molecules from each monomer sit at the edges of the N-terminal dimer interface, somewhat akin to effector binding in the specificity site at a canonical α_2 interface (**Figure 4.19**), and it is through these molecules that most of the inter-protomer contacts are made. At least one hydrogen bonding interaction is observed between E_{53} of one monomer and the 3′-OH group of the dAMP molecule bound to the opposite monomer (**Figure 4.19**). The position of H_{49} suggests that this residue could also possibly form a hydrogen bonding interaction with the 3′-OH of dAMP. Alternatively, the residue might form a hydrogen bonding interaction with guanidinium group of R_{90} in the opposite protomer (~4 Å away), thereby closing the hydrophobic pocket around the adenine moiety of dAMP. It is likely that protonation of H_{49} at the pH of the crystallization

solution prevents it from forming a hydrogen bond with either dAMP or R_{90}. Finally, the hydrogen bonds formed between N_{42} and R_{117} (**Figure 4.20**) orient the sidechain of the latter such that it can potentially form a π-cation or hydrophobic interaction with F_{47} on the opposite protomer that could contribute to the stabilization of the NrdE N-terminal dimer. Other hydrogen bonding interactions that are potentially relevant, but are not discernable in the current structure, could occur between Q_{43} and N_{44} of one protomer and E_{110}, K_{112}, and K_{113} on the other; again, the conditions of the crystallization solution may have prevented these interactions.

The possibility that this intriguing NrdE dimer can form has potential ramifications in both the interpretation of the biophysical data collected with *B. subtilis* NrdE and in the mechanism of RNR inhibition by the combination of dAMP and/or dATP. HYDROPRO[49] calculations predict the $s_{20,w}$ value of the N-terminal NrdE dimer to be 7.7 S, which is very similar to the value (8.0 S) for the canonical NrdE dimer shown in **Figure 4.15**. Under the conditions of the SV-AUC experiments run in Chapter 3 (low protein concentrations, non-optimal angular velocity, background signals from buffer and menisci mismatches), the resolution would not be high enough

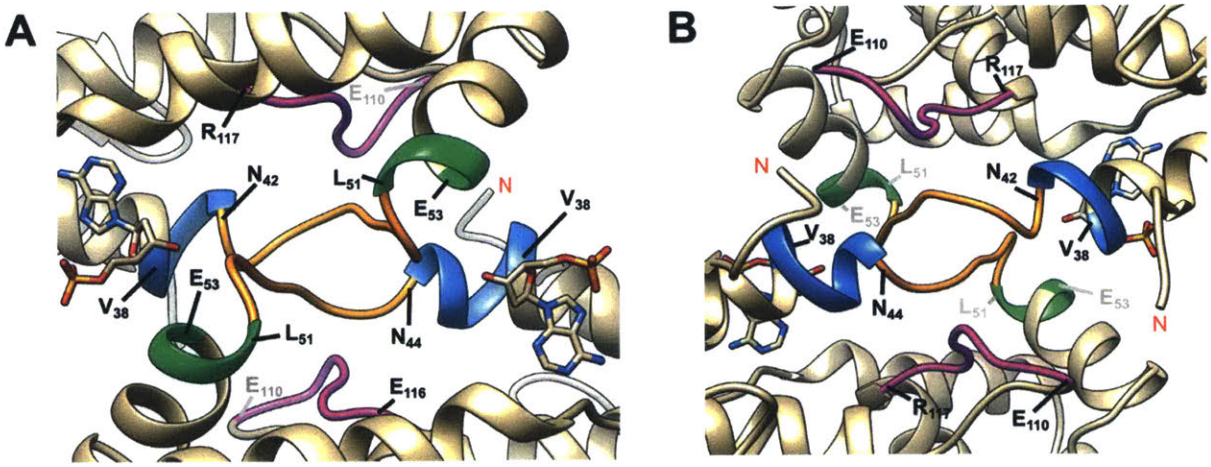


Figure 4.18. Blow up images of the interface between the N-terminal ends of the two NrdE protomers: (**A**) top side, (**B**) bottom side. The red N indicates the first N-terminal residue (S_2 or Q_3) visible in the electron density. Colored secondary structures are as follows: blue = V_{38} – I_{41} or Q_{43}, orange = N_{42} or N_{44} – N_{50}, green = L_{51} – E_{53}, magenta = E_{110} – R_{117}.

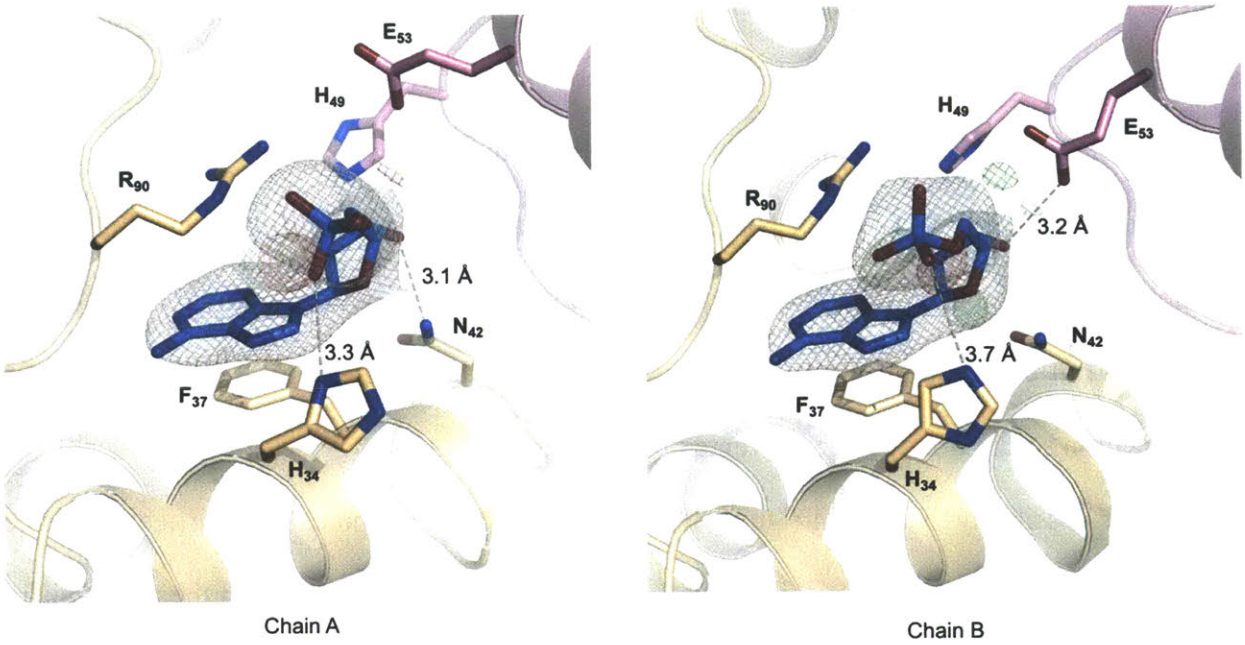


Figure 4.19. The two binding sites for dAMP at the *B. subtilis* NrdE N-terminal dimer interface. One protomer is tan and the other purple. The electron density for dAMP at each site is shown. Dashed lines indicate hydrogen bonding interactions with the indicated distances.

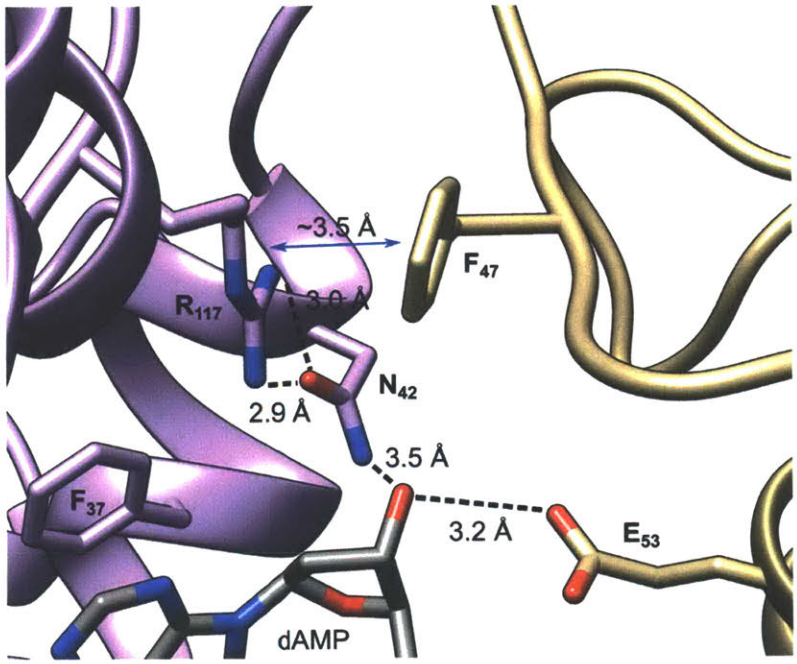


Figure 4.20. Hydrogen bonding network setting up a potential inter-protomer π-cation interaction between R_{117} and F_{47} at the N-terminal dimer interface of *B. subtilis* NrdE. Dashed lines indicate hydrogen bonding interactions with the given distances. The blue double-sided arrow indicates the rough distance separating R_{117} in one protomer from F_{47} in the opposite protomer.

to clearly distinguish the N-terminal and canonical NrdE dimers from each other. Therefore, it may be possible that peak **2** in the $s_{20,w}$ distributions (**Figure 3.12A**, for example) reported in Chapter 3 corresponds to the N-terminal rather than the canonical NrdE dimer. Furthermore, dATP concentrations that were inhibitory by activity assays appeared to cause NrdE to dimerize by SV-AUC (**Figure 3.12A**). The results of the nucleotide binding assays demonstrated that dATP and dAMP can occupy the same site on NrdE; thus, the dimeric species observed in the presence of dATP by SV-AUC could correspond to the N-terminal NrdE dimer instead of the canonical one. These hypotheses can be tested using SV-AUC and other biophysical techniques to analyze NrdE mutants with impaired dAMP binding to see if the dimeric species (peak **2**) is still observed and, if so, whether the amount of it formed is increased by the presence of dATP.

If these hypotheses are correct, then the formation of an N-terminal NrdE dimer provides a plausible explanation for the ability of dAMP and/or dATP to inhibit *B. subtilis* RNR. This explanation is illustrated in **Figure 4.21**, beginning with a recreation of the N-terminal NrdE dimer in a docking model of the *B. subtilis* RNR $\alpha_2\beta_2$ complex. Using the appended NrdE protomer as a template, a second $\alpha_2\beta_2$ complex is built into the structure. Examination of this $\alpha_4\beta_4$ model reveals the β_2 subunits sterically clash with one another. It is this clashing that could be responsible for inhibition of RNR because it prevents both $\alpha_2\beta_2$ pairs from forming a complex that is competent in NDP reduction. The mechanism is consistent with an interconversion of the active enzyme into an inhibited quaternary structure. Furthermore, as mentioned in Chapter 3, the *B. subtilis* RNR was observed to form a fibril-like structure in the presence of dATP by cryo-electron microscopy (Brignole, Zhang, Drennan, Stubbe, unpublished). One hypothesis regarding the structure of this fibril is that it could consist of repeating α_2 or $\alpha_2\beta_2$ units with alternating N-terminal and canonical

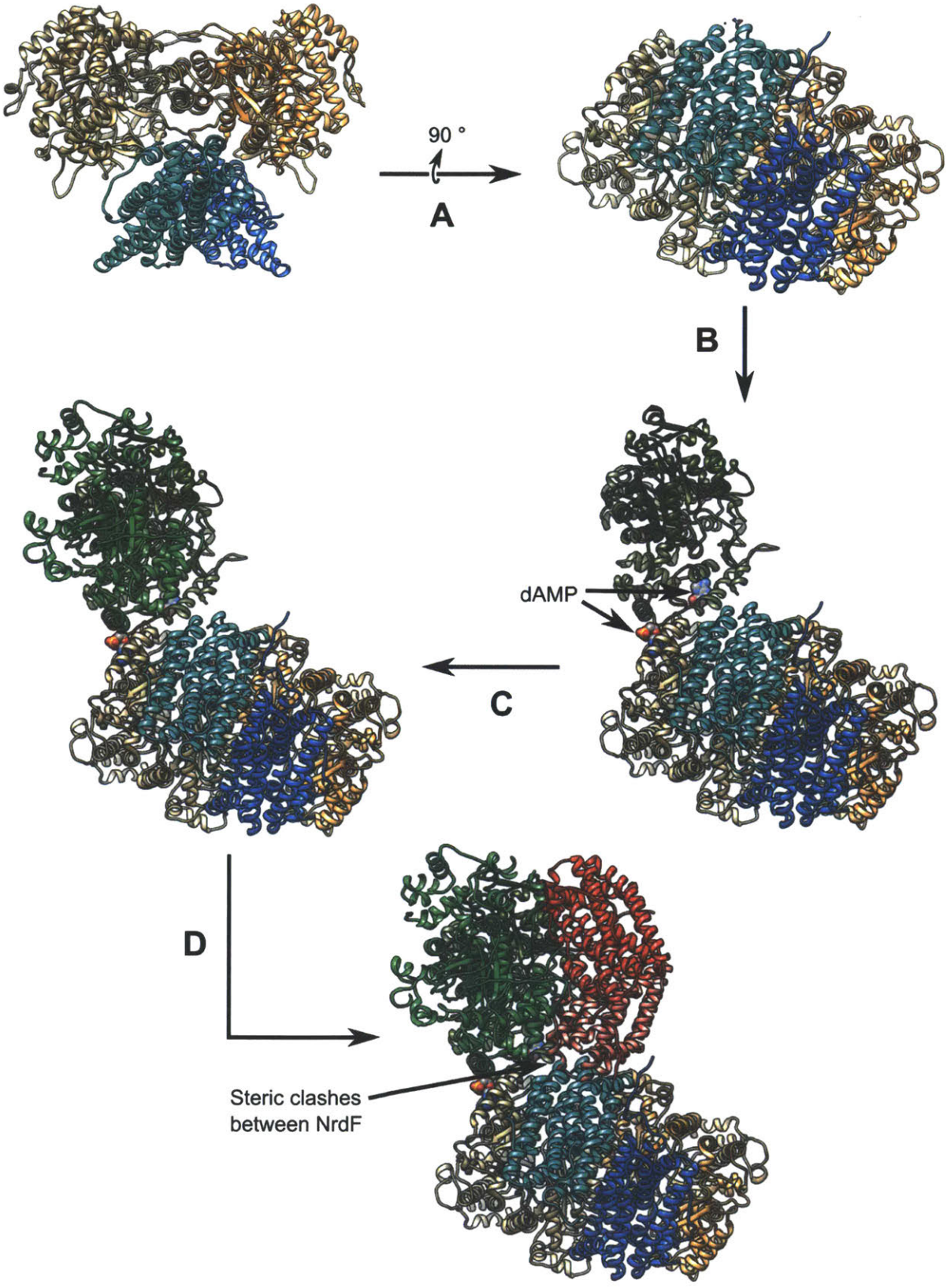

90 °
A
B
dAMP
C
D
Steric clashes
between NrdF

Figure 4.21 (opposite page). The N-terminal dimer of *B. subtilis* NrdE induced by dAMP binding might inhibit RNR activity by preventing $\alpha_2\beta_2$ complex formation. A model of the inhibited complex is built up beginning with a docking model of the *B. subtilis* RNR created from the previously reported $Mn(II)_2$-NrdF dimer (blue and cyan, PDB 4DRO[50]) and two copies of the preliminary NrdE structure (tan and light orange). The dAMP molecules are initially hidden for clarity. (**A**) The docking model is rotated by 90 ° so that the bottom of the β_2 subunit is facing the reader. (**B**) Recreation of the N-terminal NrdE dimer. The dAMP molecules are now shown as spheres and colored according to element. The third NrdE protomer in the model that is participating in the N-terminal dimer is colored olive green. Using this protomer, another copy of the $\alpha_2\beta_2$ complex is going to be built into the model. (**C**) Addition of a fourth NrdE protomer (green) to complete the second canonical α_2. (**D**) Addition of another copy of β_2 to complete the second $\alpha_2\beta_2$ complex. It can be seen from this model that formation of a N-terminal NrdE dimer would introduce steric clashes between the β_2 subunits of the adjacent $\alpha_2\beta_2$ complexes. These steric clashes could prevent competent $\alpha_2\beta_2$ complexes from forming and, thus, inhibit NDP reduction.

dimer interfaces if NrdE, with either dAMP or dATP bound, is capable of forming both types of dimers, as suggested by the model shown in **Figure 4.21**. Future studies of the *B. subtilis* Ib RNR using a combination of cyro-electron microscopy and X-ray crystallography with apo- and holo-NrdE should help unravel the possible importance of fibril formation and the tightly-bound dAMP in the inhibition of this RNR by dATP.

4.5. REFERENCES

1. Uppsten, M., Färnegårdh, M., Jordan, A., Eliasson, R., Eklund, H., and Uhlin, U. (2003) Structure of the Large Subunit of Class Ib Ribonucleotide Reductase from *Salmonella typhimurium* and Its Complexes with Allosteric Effectors, *J. Mol. Biol. 330*, 87-97.
2. Anand, R., Hoskins, A. A., Stubbe, J., and Ealick, S. E. (2004) Domain Organization of *Salmonella typhimurium* Formylglycinamide Ribonucleotide Amidotransferase Revealed by X-ray Crystallography, *Biochemistry 43*, 10328-10342.
3. Morar, M., Anand, R., Hoskins, A. A., Stubbe, J., and Ealick, S. E. (2006) Complexed Structures of Formylglycinamide Ribonucleotide Amidotransferase from *Thermotoga maritima* Describe a Novel ATP Binding Protein Superfamily, *Biochemistry 45*, 14880-14895.
4. Hoskins, A. A., Anand, R., Ealick, S. E., and Stubbe, J. (2004) The Formylglycinamide Ribonucleotide Amidotransferase Complex from *Bacillus subtilis*: Metabolite-Mediated Complex Formation, *Biochemistry 43*, 10314-10327.

5. Morar, M., Hoskins, A. A., Stubbe, J., and Ealick, S. E. (2008) Formylglycinamide Ribonucleotide Amidotransferase from *Thermotoga maritima*: Structural Insights into Complex Formation, *Biochemistry 47*, 7816-7830.

6. Anand, R., Hoskins, A. A., Bennett, E. M., Sintchak, M. D., Stubbe, J., and Ealick, S. E. (2004) A Model for the *Bacillus subtilis* Formylglycinamide Ribonucleotide Amidotransferase Multiprotein Complex, *Biochemistry 43*, 10343-10352.

7. Hoskins, A. A., (2006) *Channeling in Purine Biosynthesis: Efforts to Detect Interactions between PurF and PurD and Characterization of the FGAR-AT Complex*. PhD Thesis, Massachusetts Institute of Technology.

8. (1996) Characteristics of Nucleic Acids, In *Current Protocols in Molecular Biology*, pp A.1D.1-A.1D.11, John Wiley & Sons, Inc.

9. Edelhoch, H. (1967) Spectroscopic Determination of Tryptophan and Tyrosine in Proteins, *Biochemistry 6*, 1948-1954.

10. Gill, S. C., and von Hippel, P. H. (1989) Calculation of Protein Extinction Coefficients from Amino Acid Sequence Data, *Anal. Biochem. 182*, 319-326.

11. The ε_{280} was calculated using eq 2.1. NrdE has 3 W and 42 Y, and all Cys residues weere assumed to be reduced.

12. Randerath, K., and Randerath, E. (1964) Ion-Exchange Chromatography of Nucleotides on Poly-(Ethyleneimine)-Cellulose Thin Layers, *J. Chromatogr. A 16*, 111-125.

13. Rowley, G. L., and Kenyon, G. L. (1974) PEI-Cellulose Thin-Layer Chromatography Product Studies of Creatine Kinase and Pyruvate Kinase Reactions, *Anal. Biochem. 58*, 525-533.

14. Zhao, H. Y., Ghirlando, R., Piszczek, G., Curth, U., Brautigam, C. A., and Schuck, P. (2013) Recorded Scan Times can Limit the Accuracy of Sedimentation Coefficients in Analytical Ultracentrifugation, *Anal. Biochem. 437*, 104-108.

15. Zhao, H. Y., Brown, P. H., Balbo, A., Fernandez-Alonso, M. D., Polishchuck, N., Chaudhry, C., Mayer, M. L., Ghirlando, R., and Schuck, P. (2010) Accounting for Solvent Signal Offsets in the Analysis of Interferometric Sedimentation Velocity Data, *Macromol. Biosci. 10*, 736-745.

16. Laue, T. M., Shah, B. D., Ridgeway, T. M., and Pelletier, S. L. (1992) Computer-aided Interpretation of Analytical Sedimentation Data for Proteins, In *Analytical Ultracentrifugation in Biochemistry and Polymer Science* (Harding, S. E., Rowe, A. J., and Horton, J. C., Eds.), pp 90-125, Royal Society of Chemistry, Cambridge.

17. Ormö, M., and Sjöberg, B. M. (1990) An Ultrafiltration Assay for Nucleotide Binding to Ribonucleotide Reductase, *Anal. Biochem. 189*, 138-141.

18. Mendel, C. M., and Mendel, D. B. (1985) Non-Specific Binding - The Problem, and a Solution, *Biochem. J. 228*, 269-272.

19. Ravichandran, K., (2016) *Mechanistic Investigations of the Radical Transport Pathway in Fluorotyrosine-substituted Class Ia Ribonucleotide Reductase*. Ph.D. Thesis, Massachusets Institute of Technology.

20. Davies, D. B., and Danyluk, S. S. (1974) Nuclear Magnetic Resonance Studies of 5'-Ribo- and Deoxyribonucleotide Structures in Solution, *Biochemistry 13*, 4417-4434.

21. Fang, K. N., Kondo, N. S., Miller, P. S., and Ts'o, P. O. P. (1971) Conformation and Interaction of Dinucleoside Monophosphates and Diphosphates. 4. Proton Magnetic Resonance Study on Adenine Dideoxynucleoside Monophosphate with Emphasis on Furanose Conformation, *J. Am. Chem. Soc. 93*, 6647-6656.

22. Harrell, F. E., Jr. (2015) Multivariable Modeling Strategies, In *Regression Modeling Strategies: With Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis* Second Edition ed., pp 63-102, Springer International Publishing, New York.

23. Zhang, Y., and Stubbe, J. (2011) *Bacillus subtilis* Class Ib Ribonucleotide Reductase is a Dimanganese(III)-Tyrosyl Radical Enzyme, *Biochemistry 50*, 5615-5623.

24. Eliasson, R., Pontis, E., Jordan, A., and Reichard, P. (1996) Allosteric Regulation of the Third Ribonucleotide Reductase (NrdEF Enzyme) from Enterobacteriaceae, *J. Biol. Chem. 271*, 26582-26587.

25. Jordan, A., Pontis, E., Atta, M., Krook, M., Gibert, I., Barbé, J., and Reichard, P. (1994) A Second Class I Ribonucleotide Reductase in *Enterobacteriaceae* - Characterization of the *Salmonella typhimurium* Enzyme, *Proc. Natl. Acad. Sci. U. S. A. 91*, 12892-12896.

26. Stolle, P., Barckhausen, O., Oehlmann, W., Knobbe, N., Vogt, C., Pierik, A. J., Cox, N., Schmidt, P. P., Reijerse, E. J., Lubitz, W., and Auling, G. (2010) Homologous Expression of the *nrdF* Gene of *Corynebacterium ammoniagenes* Strain ATCC 6872 Generates a Manganese-metallocofactor (R2F) and a Stable Tyrosyl Radical (Y•) Involved in Ribonucleotide Reduction, *FEBS J. 277*, 4849-4862.

27. Willing, A., Follmann, H., and Auling, G. (1988) Ribonucleotide Reductase of *Brevibacterium ammoniagenes* is a Manganese Enzyme, *Eur. J. Biochem. 170*, 603-611.

28. Yang, F. D., Lu, G. Z., and Rubin, H. (1994) Isolation of Ribonucleotide Reductase from *Mycobacterium tuberculosis* and Cloning, Expression, and Purification of the Large Subunit, *J. Bacteriol. 176*, 6738-6743.

29. Uppsten, M., Färnegårdh, M., Domkin, V., and Uhlin, U. (2006) The First Holocomplex Structure of Ribonucleotide Reductase Gives New Insight into Its Mechanism of Action, *J. Mol. Biol. 359*, 365-377.

30. Cotruvo, J. A., and Stubbe, J. (2008) NrdI, A Flavodoxin Involved in Maintenance of the Diferric-Tyrosyl Radical Cofactor in *Escherichia coli* Class Ib Ribonucleotide Reductase, *Proc. Natl. Acad. Sci. U. S. A. 105*, 14383-14388.

31. Hofer, A., Crona, M., Logan, D. T., and Sjöberg, B. M. (2012) DNA Building Blocks: Keeping Control of Manufacture, *Crit. Rev. Biochem. Mol. Biol. 47*, 50-63.

32. Ando, N., Brignole, E. J., Zimanyi, C. M., Funk, M. A., Yokoyama, K., Asturias, F. J., Stubbe, J., and Drennan, C. L. (2011) Structural Interconversions Modulate Activity of *Escherichia coli* Ribonucleotide Reductase, *Proc. Natl. Acad. Sci. U. S. A. 108*, 21046-21051.

33. Ando, N., Li, H. R., Brignole, E. J., Thompson, S., McLaughlin, M. I., Page, J. E., Asturias, F. J., Stubbe, J., and Drennan, C. L. (2016) Allosteric Inhibition of Human Ribonucleotide Reductase by dATP Entails the Stabilization of a Hexamer, *Biochemistry 55*, 373-381.

34. Aravind, L., Wolf, Y. I., and Koonin, E. V. (2000) The ATP-Cone: An Evolutionarily Mobile, ATP-Binding Regulatory Domain, *J. Mol. Microbiol. Biotechnol. 2*, 191-194.

35. Lundin, D., Torrents, E., Poole, A. M., and Sjöberg, B. M. (2009) RNRdb, a Curated Database of the Universal Enzyme Family Ribonucleotide Reductase, Reveals a High Level of Misannotation in Sequences Deposited to Genbank, *BMC Genomics 10*, 8.

36. Jordan, A., Pontis, E., Åslund, F., Hellman, U., Gibert, I., and Reichard, P. (1996) The Ribonucleotide Reductase System of *Lactococcus lactis* - Characterization of an NrdEF Enzyme and a New Electron Transport Protein, *J. Biol. Chem. 271*, 8779-8785.

37. Makhlynets, O., Boal, A. K., Rhodes, D. V., Kitten, T., Rosenzweig, A. C., and Stubbe, J. (2014) *Streptococcus sanguinis* Class Ib Ribonucleotide Reductase: High Activity with Both Iron and Manganese Cofactors and Structural Insights, *J. Biol. Chem. 289*, 6259-6272.

38. Yang, F. D., Curran, S. C., Li, L. S., Avarbock, D., Graf, J. D., Chua, M. M., Lui, G. Z., Salem, J., and Rubin, H. (1997) Characterization of Two Genes Encoding the *Mycobacterium tuberculosis* Ribonucleotide Reductase Small Subunit, *J. Bacteriol. 179*, 6408-6415.

39. Hanson, E., and Mathews, C. K. (1994) Allosteric Effectors are Required for Subunit Association in T4 Phage Ribonucleotide Reductase, *J. Biol. Chem. 269*, 30999-31005.

40. Nordlund, N., and Reichard, P. (2006) Ribonucleotide Reductases, In *Annual Review of Biochemistry*, pp 681-706, Annual Reviews, Palo Alto.

41. Berglund, O. (1972) Ribonucleoside Diphosphate Reductase Induced by Bacteriophage T4. I. Purification and Characterization, *J. Biol. Chem. 247*, 7270-7275.

42. Eliasson, R., Pontis, E., Sun, X. Y., and Reichard, P. (1994) Allosteric Control of the Substrate-Specificity of the Anaerobic Ribonucleotide Reductase from *Escherichia coli*, *J. Biol. Chem. 269*, 26052-26057.

43. Torrents, E., Buist, G., Liu, A., Eliasson, R., Kok, J., Gibert, I., Gräslund, A., and Reichard, P. (2000) The Anaerobic (Class III) Ribonucleotide Reductase from *Lactococcus lactis* - Catalytic Properties and Allosteric Regulation of the Pure Enzyme System, *J. Biol. Chem. 275*, 2463-2471.

44. Osbourn, A. E., and Field, B. (2009) Operons, *Cell. Mol. Life Sci. 66*, 3755-3775.

45. McKethan, B. L., and Spiro, S. (2013) Cooperative and Allosterically Controlled Nucleotide Binding Regulates the DNA Binding Activity of NrdR, *Mol. Microbiol. 90*, 278-289.

46. Mao, S. S., Holler, T. P., Yu, G. X., Bollinger, J. M., Booker, S., Johnston, M. I., and Stubbe, J. (1992) A Model for the Role of Multiple Cysteine Residues Involved in Ribonucleotide Reduction - Amazing and Still Confusing, *Biochemistry 31*, 9733-9743.

47. Denessiouk, K. A., Rantanen, V. V., and Johnson, M. S. (2001) Adenine Recognition: A Motif Present in ATP-, CoA-, NAD-, NADP-, and FAD-Dependent Proteins, *Proteins 44*, 282-291.

48. Murray, T. A., and Swenson, R. P. (2003) Mechanism of Flavin Mononucleotide Cofactor Binding to the *Desulfovibrio vulgaris* Flavodoxin. 1. Kinetic Evidence for Cooperative Effects Associated with the Binding of Inorganic Phosphate and The 5'-Phosphate Moiety of the Cofactor, *Biochemistry 42*, 2307-2316.

49. Ortega, A., Amorós, D., and de la Torre, J. G. (2011) Prediction of Hydrodynamic and Other Solution Properties of Rigid Proteins from Atomic- and Residue-Level Models, *Biophys. J. 101*, 892-898.

50. Boal, A. K., Cotruvo, J. A., Stubbe, J., and Rosenzweig, A. C. (2012) The Dimanganese(II) Site of *Bacillus subtilis* Class Ib Ribonucleotide Reductase, *Biochemistry 51*, 3861-3871.

Chapter 5

Investigation of the Role of YmaB in the Regulation of the *Bacillus subtilis* Class Ib RNR

5.1. INTRODUCTION

Reported in this chapter are the results of experiments probing the function of YmaB, the fourth protein of the *B. subtilis* class Ib RNR operon, and efforts to obtain evidence supporting or refuting the hypothesis that the function of YmaB is linked to RNR regulation. YmaB is a small, 206 residue protein of unknown function encoded just downstream of the *nrdF* gene in *B. subtilis* (**Figure 5.1**). Its location adjacent to the class Ib RNR genes is conserved in immediate phylogenetic relatives (e.g. *Bacillus amyloliquifaciens*, *Bacillus thuringiensis*, etc.) and independently encoded homologs can be found in a few other Gram positive and extremophilic bacteria such as *Staphylococcus aureus* (42% identity, 62% with positive substitutions) and *Thermotoga maritima* (35% identity, 48% with positive substitution), respectively. Evolutionarily, the protein does not associate with a specific RNR class since interrogation of the genomes of *S. aureus*, *T. maritima*, and other organisms encoding a YmaB homolog reveals up to two encoded RNRs, one being class Ib or class II and the other, when present, a class III system. The essentiality of YmaB for the viability of *B. subtilis* is debatable as attempts to grow knockout mutants under standard laboratory conditions (LB medium, 37 °C) have produced conflicting results.[1, 2] In contrast, the *S. aureus* YmaB has been found to be essential for cell survival under standard laboratory conditions (BHI medium, 37 °C).[3]

Functionally related genes in prokaryotic genomes are often clustered closely together in groups called operons.[4] Genes belonging to an operon are transcribed from a single promoter and are coordinately regulated by the same transcription factors in response to different environmental stimuli. The *ymaB* gene, for the most part, meets the criteria commonly used to classify a gene as part of an operon.[4] *nrdF* and *ymaB* are closely spaced (50 base pairs), and the four genes of the putative operon, *nrdI-nrdE-nrdF-ymaB*, are bracketed by a well-defined promoter upstream of

nrdI and a strong transcriptional terminator immediately downstream of *ymaB* (**Figure 5.1**).[2, 5] The genes are similarly regulated in response to the oxygen tension of the environment[5] and the results of a bioinformatics study[6] predicts that they are likewise co-regulated in response to *in vivo* deoxyribonucleotide levels via the transcriptional regulator NrdR.[7]

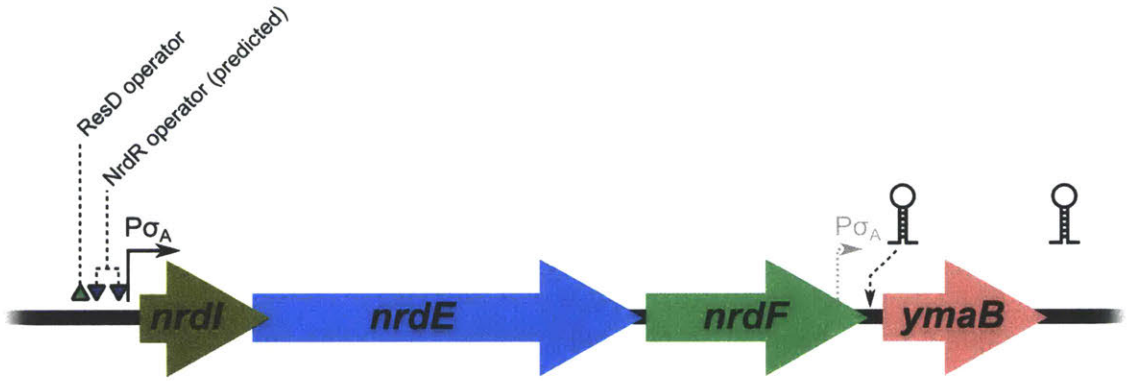


Figure 5.1. Organization of the class Ib RNR operon in the genome of *B. subtilis* and its immediate phylogenetic relatives. Note that *nrdI* overlaps the *nrdE* coding region by 41 bp, hence the reason for the overlap of the arrows in this representation. The relative positions of the *ResD* (green triangle)[5] and predicted *NrdR* (blue triangles)[6] operator sequences upstream of the promoter (Pσ_A) are indicated. The black stem-loops indicate experimentally verified terminators for *nrdI-nrdE-nrdF* and *nrdI-nrdE-nrdF-ymaB*.[5, 8, 9] The gray-colored Pσ_A indicates a partially conserved promoter that could drive separate *ymaB* transcription.[2]

At the transcriptional level, another characteristic of prokaryotic operons is that all members are usually transcribed on a single polycistronic mRNA.[4] The best evidence for mRNAs encoding both *ymaB* and the RNR genes comes from Northern blotting experiments probing RNR transcription in response to oxygen tension.[5] A *nrdE*-specific probe detected two transcripts ~3500 and ~4100 nucleotides (nt) long in *B. subtilis* cells grown either aerobically or anaerobically in minimal medium. The shorter one is consistent with a transcript encoding *nrdI*, *nrdE*, and *nrdF* alone (3461 nt, including intergenic regions), while the longer one correlates with the size expected if *ymaB* was included in the message (4131 nt). This interpretation of the data was strengthened by mapping the transcriptional start sites of both mRNAs to the intergenic region upstream of

nrdI.[5] These data also confirm the prediction[2] of a termination site in the intergenic region between *nrdF* and *ymaB* that leads to transcription of *nrdI-nrdE-nrdF* separately (**Figure 5.1**). In a more recent study,[8] polycistronic transcripts encoding *ymaB* and the RNR genes were reported to be observed in *B. subtilis* cells grown under 104 different conditions, indicating their co-expression is a universal phenomenon.

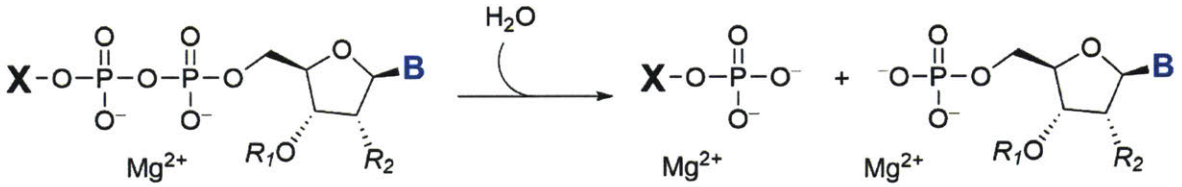


Figure 5.2. General reaction scheme for hydrolysis of (d)NDP-X substrates by Nudix hydrolases. **B** = base (A, C, G, T, U and derivatives thereof). R_1 = H, PO_3^{2-}. R_2 = H, OH, OPO_3^{2-}. **X** can be any of a number of chemical functional groups, including phosphate, nucleoside 5′-polyphosphates, sugars, and the functional ends of cofactors (for example, nicotinamide riboside or pantethine).

BLAST searches using *B. subtilis* YmaB as a query returns MutT and other members of the Nudix hydrolase superfamily as the top homologous hits. As discussed in **Section 1.6**, Nudix hydrolases are small enzymes (16 – 21 kDa) that hydrolyze substrates typically consisting of a (deoxy)nucleotide 5′-diphosphate linked to a functional group X (**Figure 5.2**), including canonical (d)NTPs and their oxidized or chemically modified derivatives, NDP-sugars, cofactors, dinucleoside polyphosphates, and mRNA caps.[10] These enzymes can be identified by a conserved amino acid sequence, called the Nudix box, consisting of Gx_5Ex_5[UA]xREx_2EExGU, where x can be any amino acid and U is a bulky hyrdophobic residue.[10] The REx_2EE unit is essential for catalysis and ligates two to three Mg^{2+} or Mn^{2+} ions that activate water for nucleophilic attack typically at the α or β phosphate group of the substrate.[11, 12] Nudix hydrolases furthermore adopt a characteristic α/β/α tertiary fold known as the Nudix fold.[10] As Nudix hydrolases tend to be highly promiscuous towards substrates, the *in vivo* function of these enzymes is often difficult to

establish. The two best studied cases are MutT, which cleaves mutagenic 8-oxo-(d)GTP to sanitize dNTP pools,[13, 14] and RppH, which hydrolyzes the 5′-triphosphate caps of mRNAs in prokaryotes to initiate their degradation.[15, 16] The broad substrate range for the remaining studied Nudix hydrolases has led to the belief that these enzymes act as "house keepers" by removing cytotoxic/mutagenic nucleotides, degrading signaling molecules, and maintaining balanced pools of nucleotide-based building blocks for anabolic processes.[17]

The unusual allosteric and quaternary structural properties of the *B. subtilis* Ib RNR, including the discovery of a tightly bound equivalent of dAMP associated with NrdE as reported in Chapter 4, prompted an examination of a possible functional relationship between YmaB activity and RNR regulation. Specifically, studies reported in this chapter were conducted in order to obtain fundamental information regarding the protein's function and structure, and to determine if YmaB could play a role in the installation of dAMP into NrdE. Screening the protein with potential substrates demonstrated YmaB is indeed a Nudix hydrolase that possesses a specificity for the canonical dNTPs with a higher preference for pyrimidines. Concentration-dependent SV-AUC experiments revealed that the protein is a monomer in the absence of substrates. Finally, coexpression of YmaB and NrdE together in *E. coli* under standard growth conditions did not significantly increase the amount of dAMP associated with the latter. Thus, a functional link between YmaB and the *B. subtilis* Ib RNR remains enigmatic.

5.2. EXPERIMENTAL

5.2.1. General methodology. Chemicals, equipment, and routine methods were essentially the same as described in the previous chapters. Yeast inorganic pyrophosphatase (~200 U mg^{-1}) was from Roche. **Nudix assay buffer** consisted of 50 mM Tris, pH 8.0, 5 mM $MgCl_2$.

5.2.2. Cloning, expression, and purification of tagless YmaB. *ymaB* was amplified from *B. subtilis* JH624 genomic DNA with the touchdown PCR technique[18] using the primers 5′-AGGAGGGGTCTCTAGGT**ATG**GGAAAAATGGACG-3′ and 5′-CCGGCTCGAG**TTA**TTCA AGAATATCAACAACAAACTGTG-3′ to introduce a 5′ *BsaI* and 3′ *XhoI* restriction site (underlined, start and stop codons indicated in boldface font) into the amplicon. The amplicon was digested with BsaI and XhoI and ligated into similarly digested pE-SUMO (LifeSensors, Kan^r) using T4 DNA ligase to generate the vector pE-SUMO-*ymaB*. Sequencing (MIT Biopolymers Laboratory) confirmed the successful cloning of *ymaB*.

The vector was transformed into BL21 (DE3) cells and the His_6-Smt3-tagged protein was expressed following the standard growth procedure (Chapter 2) to give 2.45 g cell paste L^{-1}. The purification of His_6-Smt3-YmaB was the same as the standard procedure (Chapter 2) up through the IMAC step, except that the lysis and Ni-NTA column buffers were supplemented with 100 mM NaCl. Protein-containing fractions from the Ni-NTA column were pooled, diluted 5-fold in Tris buffer (50 mM Tris, pH 7.6, 5% (w/v) glycerol) supplemented with 100 mM NaCl, and loaded onto a Q-Sepharose column (5.7 x 2.5 cm, 28 mL). The column was washed with 2 CVs of diluent before being developed with a 300 mL linear gradient from 100 mM to 500 mM NaCl in Tris buffer. His_6-Smt3-YmaB eluted at ~250 mM NaCl. Protein-containing fractions were pooled, concentrated to ~1.5 mg mL^{-1}, aliquoted (~1.7 mL), and stored at -80 °C. A total of ~65 mg His_6-Smt3-YmaB (6.5 mg g^{-1} cell paste) was recovered (> 90% pure as judged by SDS-PAGE). Protein concentrations were determined by UV-vis using $\varepsilon_{280} = 19900\ M^{-1}\ cm^{-1}$.[19]

For preparation of the tagless YmaB used in all experiments reported in this chapter, His_6-Smt3-YmaB (15 mg) was digested with SUMO protease (prepared as described in Chapter 3) using a 25:1 mole ratio of substrate:protease. Prior to protease addition, DTT was added to a final

concentration of 5 mM. The samples were incubated at 4 °C on a nutating platform for 8 h and then rapidly frozen and stored at -80 °C. The samples at a later time were thawed, pooled, and loaded onto a Ni-NTA column (3.5 x 1.5 cm, 2.75 mL) equilibrated in Nudix assay buffer supplemented with 20 mM imidazole. The column was washed with 60 mL equilibration buffer and twenty 2 mL fractions were collected. Protein-containing fractions were pooled, concentrated to ~ 1 mL, and desalted on a Sephadex G25 column (28 x 1.5 cm, 49 mL) equilibrated in Nudix assay buffer. A total of 11 mg (75% yield) tagless YmaB was recovered (≥ 97% pure as judged by SDS-PAGE). Protein concentrations were determined using $\varepsilon_{280} = 18500\ M^{-1}\ cm^{-1}$.[19]

5.2.3. Examination of YmaB for Nudix hydrolase activity by phosphate detection. Substrates for YmaB were identified by screening for released inorganic phosphate using the phosphomolybdate-malachite green assay.[20] Assay solutions of 200 μL contained 1 mM substrate, 0.5 U yeast inorganic pyrophosphatase (when assaying (d)NTPs (**Figure 5.3A**)) or 4 U calf alkaline phosphatase (when assaying all other substrates (**Figure 5.3B**)), 50 mM Tris, pH 8.0, 5 mM $MgCl_2$, and 1.5 μM YmaB. Solutions were pre-incubated at 37 °C for 2 min prior to the addition of YmaB to initiate turnover. Twenty-five microliter aliquots of assay solution were withdrawn every 5 min in a 20 min incubation period and mixed with 45 μL 1.6% (w/v) $HClO_4$ to quench the reaction. After sitting for 30 s at room temperature (RT), the pH of quenched aliquots was neutralized with 33 μL 0.4 M KOH before freezing the samples in liquid N_2. Control reactions in which Nudix assay buffer was added in place of YmaB, substrate, or phosphatase were run in parallel. All control and experimental assays were run in duplicate. When assaying NADPH, assay solutions were prepared as described above except that alkaline phosphatase was omitted. Every 5 min, 25 μL of assay solution were withdrawn, mixed with 10 μL Nudix assay buffer containing 20 U alkaline phosphatase, and incubated at 37 °C for 4 min before quenching with 35

μL 2% (w/v) $HClO_4$. Subsequent pH neutralization and freezing were completed as described above.

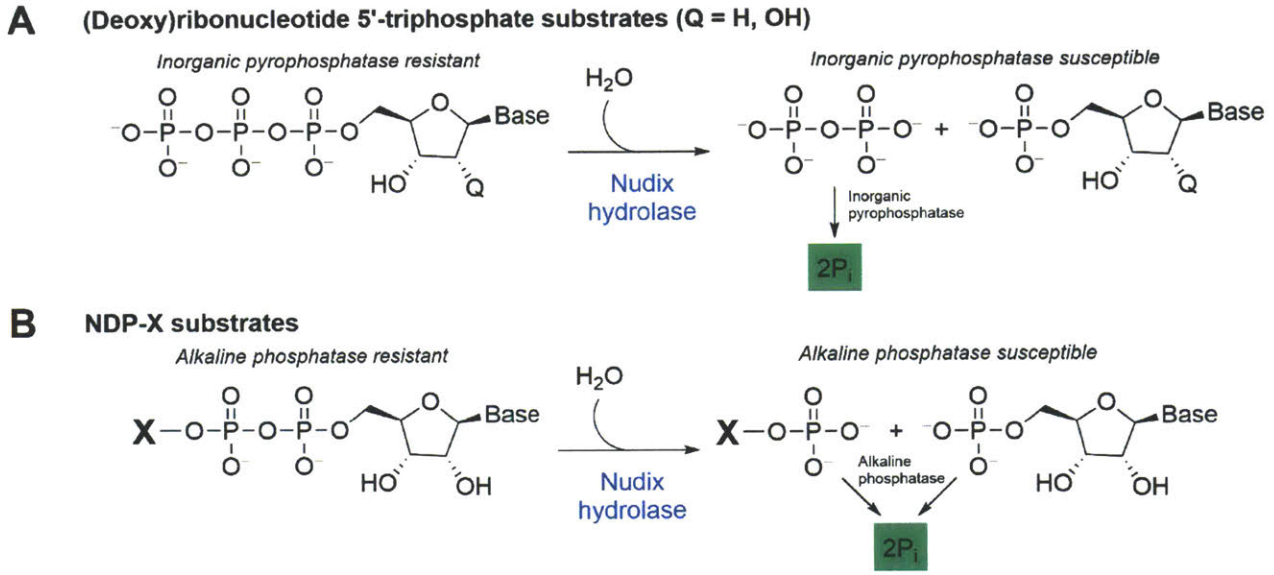


Figure 5.3. The logic behind the use of phosphate detection to measure the activity of Nudix hydrolases. Although not explicitly shown, nucleotides are associated with Mg^{2+} ions. (**A**) Cleavage of (d)NTPs by a Nudix hydrolase results in generation of inorganic pyrophosphate, which is subsequently cleaved by inorganic pyrophosphatase to yield two equivalents of Pi detectable by the phosphomolybdate assay. (**B**) Cleavage of all other potential Nudix hydrolase substrates (NDP-sugars, cofactors, dinucleoside polyphosphates, etc.) results in alkaline phosphatase susceptible phosphate groups that, when hydrolyzed from product, are detected by the phosphomolybdate assay.

Assay work-up. Frozen aliquots (103 μL) were thawed on ice and ~0.4 – 0.6 mg of activated charcoal (Aldrich) was added. The samples were vortexed for ~10 s before pelleting the charcoal and $KClO_4$ precipitate by centrifugation (20817 x g, 10 min, 4 °C). Supernatant (80 μL) was mixed with 20 μL working color reagent (described below) and the samples were incubated in the dark for 30 min at RT prior to measuring A_{630}. If required, dilutions of the supernatant were made with deionized water. Phosphate concentrations were determined using a standard curve of 2 – 20 μM sodium phosphate that was prepared from a 1000 ± 4 mg L^{-1} IC stock standard solution (Sigma). Working color reagent was prepared immediately before use by mixing 1 mL Reagent

A (6 M H_2SO_4, 0.12% (w/v) malachite green), 250 μL Reagent B (7.5% (w/v) sodium molybdate), and 20 μL Reagent C (11% (v/v) Tween 20).[20]

5.2.4. Qualitative examination of YmaB for Nudix hydrolase activity by PEI-cellulose TLC. Assay solutions were prepared as described for the phosphomolybdate assays except that inorganic pyrophosphatase and alkaline phosphatase were omitted. The assays were incubated at 37 °C for 20 min before transferring the solutions to Amicon Ultra-0.5 YM10 microcentrifugal filters and centrifuging (14000 x g, 5 min, 4 °C). Control reactions in which YmaB was omitted were run in parallel. A total of 3 μL of filtrate from the control and experimental reactions were spotted in 0.5 μL increments on two 5 x 4 cm PEI-cellulose TLC plates along with 0.5 μL of 10 – 30 mM solutions of authentic substrate and the expected deoxynucleoside 5′-monophosphate product. Each plate was developed at RT with a different solvent: Solvent 1 was a 3:1 mixture of 4 M formic acid/formate, pH 3.4:4% (w/v) boric acid and Solvent 2 contained 0.5 M formic acid/formate, pH 3.4. After development and drying, chromatograms were visualized by irradiation under a short wave (254 nm) UV lamp and marked using a soft-leaded pencil before photographing the plates with a Gel Doc™ 2000 Imager (BioRad).

5.2.5. Determination of YmaB quaternary structure by SV-AUC. Six 500 μL samples of YmaB in Nudix assay buffer at concentrations spanning 10 – 58 μM were prepared. Cells assembled with Epon charcoal double sector centerpieces (12 mm) and quartz windows were loaded with ~440 μL Nudix assay buffer in the reference sector and ~430 μL protein solution in the sample sector. The samples were radially calibrated and then thermally equilibrated to 20 °C for 1.75 h prior to beginning the experiment. Complete sedimentation was measured by A_{280} over 18 h at 20 °C and an angular velocity of 42000 rpm using a Beckman XL-I analytical

ultracentrifuge (MIT Biophysical Instrumentation Facility) and scans were collected every 1.2 min.

The first 100 scans collected for each cell, covering the entire sedimentation of YmaB, were loaded into Sedfit, time-stamp corrected,[21] and fit to the c(s) model with a resolution of 100 and a regularization factor of 0.68 using alternating rounds of Simplex and Marquardt-Levenberg optimization algorithms until best fit values converged. Distributions were corrected to the standard state ($s_{20,w}$) using a $\rho = 1.006$ g cm^{-3}, $\eta = 1.033$ cP, and $\bar{v} = 0.7371$ mL g^{-1} for YmaB. The solvent parameters and partial specific volume were calculated and temperature corrected using the program Sednterp.[22] Weight-averaged $s_{20,w}$ values were determined by integrating the distributions over the $s_{20,w}$ range containing all species, and error in this value was assessed using 5000 rounds of Monte-Carlo simulations. These values were plotted as a function of YmaB concentration to generate a weight-averaged $s_{20,w}$ isotherm, which indicated that YmaB was not self-associating into larger oligomers. Therefore, the individual data sets were exported into the program Sedphat and globally fit for molecular weight and $s_{20,w}$ using the hybrid local continuous distribution/global discrete species model with a regularization factor of 0.68 and alternating rounds of Simplex and Marquardt-Levenberg optimization until convergence. The initial guesses for the molecular weight (23400 Da) and $s_{20,w}$ (2.3 S) of the global discrete species were respectively based on the YmaB amino acid sequence and HYDROPRO predictions[23] using a threading model of YmaB generated with the program PHYRE[24] using the *T. maritima* YmaB structure (PDB 3E57) as a template. Two segments of local continuous distribution were set between 0 – 1.3 S and 3.3 – 20 S. The 68% confidence interval (i.e. one standard deviation) for the best fit molecular weight and $s_{20,w}$ was determined using automated error surface projection

searches in which the parameter for which the error was being determined was fixed and the other fit parameters floated.

5.2.6. Determination of the ability of YmaB to affect the dAMP load of NrdE *in vivo*. The *ymaB* gene was amplified from genomic DNA using the forward primer 5′-GGAGGACGCAGTA CCCATGGGAAAAATGGACG-3′ and the reverse primer described in **Section 5.2.2** to introduce a 5′ *NcoI* (underlined, start codon indicated in boldface font) and 3′ *XhoI* restriction site into the amplicon. The amplicon was digested with NcoI and XhoI and subsequently ligated into similarly digested pBAD/Myc-HisA (Invitrogen, Amp^r). The successful cloning of *ymaB* was determined by sequencing (Quintara Biosciences). This vector places the *ymaB* gene under the control of an arabinose-inducible promoter which, when induced, drives the production of tagless YmaB.

The vectors pE-SUMO-*nrdE* (**Section 3.2.3**) and pBAD-*ymaB* were simultaneously transformed into BL21 (DE3) cells and selected for on LB agar plates containing 100 μg mL^{-1} Kan and 100 μg mL^{-1} Amp. Cultures of single colonies grown to saturation overnight (~12 – 14 h) were diluted 200-fold into 1 L volumes of antibiotic-supplemented LB medium in 6 L Erlenmeyer flasks and grown to an OD_{600} of 0.3 – 0.4 at 37 °C with 150 rpm shaking. Sterile-filtered arabinose solution (2% (w/v)) was subsequently added to a final concentration of 0.02% (w/v) and the cultures incubated for an additional 1 h prior to inducing His_6-Smt3-NrdE expression with 0.2 mM IPTG. The induced cultures were incubated for an additional 2 h and then were harvested by centrifugation (3500 x g, 15 min, 4 °C). Cell pellets were frozen (liq. N_2) and stored at -80 °C until purification. A control strain of BL21 (DE3) cells transformed with pE-SUMO-*nrdE* and pBAD/Myc-HisA empty vector was propagated separately under identical conditions. A total of 4.2 g cell paste L^{-1} was recovered for each strain.

The purification of His_6-Smt3-NrdE and removal of the tag were completed as described in **Section 3.2.3**. Quantitation of the basal amount of dAMP bound to NrdE isolated from the control or experimental strain was completed in triplicate or quadruplicate, respectively, as described in **Section 4.2.2A**. The statistical significance of the difference in dAMP loading between the control and experimental strains was assessed using a two-tailed, two independent sample Student's t-test for data sets with equal variance (equation 5.1).

$$t = \frac{< x_1 > - < x_2 >}{\sqrt{\frac{(n_1 - 1) * s_{x_1}^2 + (n_2 - 1) * s_{X_2}^2}{n_1 + n_2 - 2}} * \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \qquad (5.1)$$

In eq 5.1, $<x_1>$ is the average dAMP load, $s_{x_1}^2$ is the standard deviation, and n_1 the number of replicate measurements made for the experimental strain and $<x_2>$, $s_{x_2}^2$, and n_2 have the same meanings for the control strain. The value t was converted into a probability (p) by referring to a t-table. A $p < 0.01$ was considered statistically significant difference.

5.3. RESULTS

5.3.1. Expression and purification of tagless YmaB. His_6-Smt3-YmaB was overexpressed and purified to near homogeneity with good yield following standard protocols (**Figure 5.4A**). Incubation of the protein with a 25:1 mole ratio of His_6-Smt3-YmaB:SUMO protease for 8 h at 4 °C resulted in > 95% cleavage of the tag from YmaB. High recoveries of nearly homogenous tagless YmaB (≥ 97% purity) were obtained after subtractive IMAC and desalting procedures (**Figure 5.4B**). Both His_6-Smt3-tagged and tagless YmaB appear to run ~2000 – 3000 Da larger on SDS-PAGE gels than expected based on sequence. However, the AUC data reported in **Section 5.3.3** clearly indicates that unmodified tagless protein was obtained by the procedures described in the experimental section. Therefore, the basis for the aberrant migration of the protein by SDS-PAGE is currently unknown.

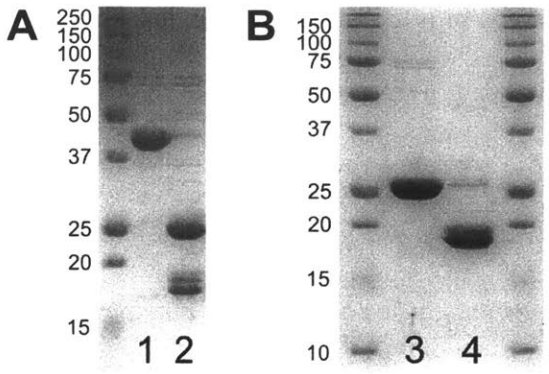


Figure 5.4. Preparation of tagless YmaB (23.4 kDa). The images shown are representative SDS-PAGE gels (12.5% (w/v)) illustrating the His_6-Smt3 tag removal procedure for YmaB. (**A**) *Lane 1*: Purified His_6-Smt3-YmaB (35.8 kDa), *lane 2*: digest product after an 8 h incubation with a 25:1 mol ratio of substrate:SUMO protease at 4 °C. (**B**) *Lane 3*: tagless YmaB after removal of His_6-tagged components by subtractive IMAC, l*ane 4*: His_6-tagged components removed by subtractive IMAC. The ~18 kDa species in *lanes 2* and *4* is His_6-Smt3.

5.3.2. YmaB is a Nudix hydrolase with specificity towards dNTPs. YmaB was first examined for its ability to hydrolyze (d)NDP-X substrates to establish whether or not the protein was a Nudix hydrolase as suggested by the BLAST results. Canonical (d)NTPs, NAD(P)H, and FAD were screened as potential substrates by monitoring the release of phosphate from functional groups that are initially alkaline phosphatase or inorganic pyrophosphatase resistant, but become susceptible upon cleavage of the phosphoanhydride bond between the α and β phosphorus atoms (**Figure 5.3**). The phosphomolybdate assay with enhanced sensitivity (ε_{630} = 90000 M^{-1} cm^{-1}) by the inclusion of malachite green and Tween 20 was used for subsequent detection of the released inorganic phosphate.[20]

Time-dependent phosphate release was detected when YmaB was incubated with dATP and TTP (**Figure 5.5**), but not with NDPs (ATP, CTP, GTP), NAD(P)H, and FAD (data not shown). The activities are stable as indicated by the linearity of the data over 20 min, and specific activities of 560 ± 50 nmol min^{-1} mg^{-1} and 3500 ± 400 nmol min^{-1} mg^{-1} were calculated for dATP and TTP, respectively (corrected for background, see subsequent discussion). Controls in which

YmaB or pyrophosphatase were omitted from the reaction also exhibited low rates of phosphate release (0.03 – 1.3 nmol min^{-1} mg^{-1}) over time, but not in assays in which substrate was omitted (**Figure 5.5**). Given the acidic nature of the phosphomolybdate reagent, the temperature at which the workup was completed (incubation at RT for > 30 min), and the presence of small amounts of Mg^{2+} required for YmaB activity, these low rates of phosphate release are likely attributable to background decomposition of inorganic pyrophosphate and/or residual substrate/product not removed by the activated charcoal (**Figure 5.5**). The data, therefore, show that phosphate release depends on the presence of both YmaB and inorganic pyrophosphatase, consistent with YmaB-catalyzed hydrolysis of dATP/TTP to dAMP/TMP and pyrophosphate, followed by hydrolysis of the latter by inorganic pyrophosphatase to yield inorganic phosphate (**Figure 5.3A**). The results thus support the functional prediction, based on the BLAST search results, that YmaB is a Nudix hydrolase. The specific activities for dATP and TTP hydrolysis by YmaB are lower than those reported for other Nudix hydrolases catalyzing the same reactions (1100 – 36000 nmol min^{-1} mg^{-1} for dATP and 1000 – 31000 nmol min^{-1} mg^{-1} for TTP), but are likely more physiologically relevant given that the results from previous studies were obtained at pH values of 8.5 – 9.[25-27]

The specificity of YmaB for nucleotides containing 2′-deoxyribose was further examined qualitatively using PEI-cellulose TLC to monitor dNMP formation. Photographs of the developed plates are shown in **Figure 5.6** and the R_f values are tabulated in **Table 5.1**. Only spots with R_f values similar to that of the dNMP standards were observed in assays in which the pyrimidine dNTPs were incubated with YmaB for 20 min at 37 °C. The results indicate that both dCTP and TTP are efficiently hydrolyzed by YmaB and that the rate of dCTP hydrolysis was similar to or faster than that observed with TTP. dATP and dGTP were also cleaved, but are poorer substrates for the enzyme given that significant amounts of starting material, as judged by the comparable R_f

values of the dNTP standards with the slower migrating material observed in the products from the complete assay, remained after incubation for the same amount of time. Of the purines, dGTP appeared to be the poorest substrate for YmaB because the spot corresponding to dGMP produced from substrate cleavage was the faintest under UV irradiation as compared with the other nucleotides. Altogether, the results of the TLC assays establish YmaB is a promiscuous dNTP pyrophosphohydrolase with qualitative specificity in the order of TTP = dCTP > dATP > dGTP.

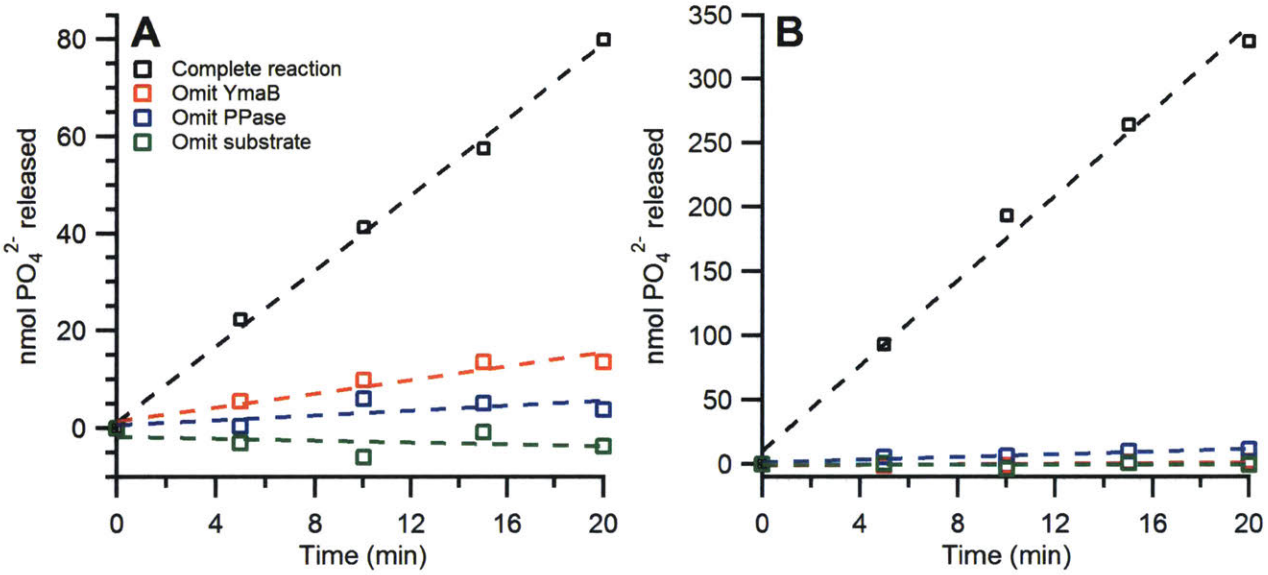


Figure 5.5. Analysis of YmaB for Nudix hydrolase activity by the detection of phosphate with the phosphomolybdate assay. Shown are representative examples of the data generated. Assays solutions (200 µL) consisting of 50 mM Tris, pH 8.0, 5 mM $MgCl_2$, 1.5 µM YmaB, 0.5 U yeast inorganic pyrophosphatase (PPase) and either (**A**) 1 mM dATP, or (**B**) 1 mM TTP were incubated at 37 °C for 20 min. Every 5 min, aliquots were withdrawn and quenched before being worked up as described in the Experimental section. Control assays omitting the indicated component of the complete reaction were run in parallel. Note the y-axis scale differences between (**A**) and (**B**).

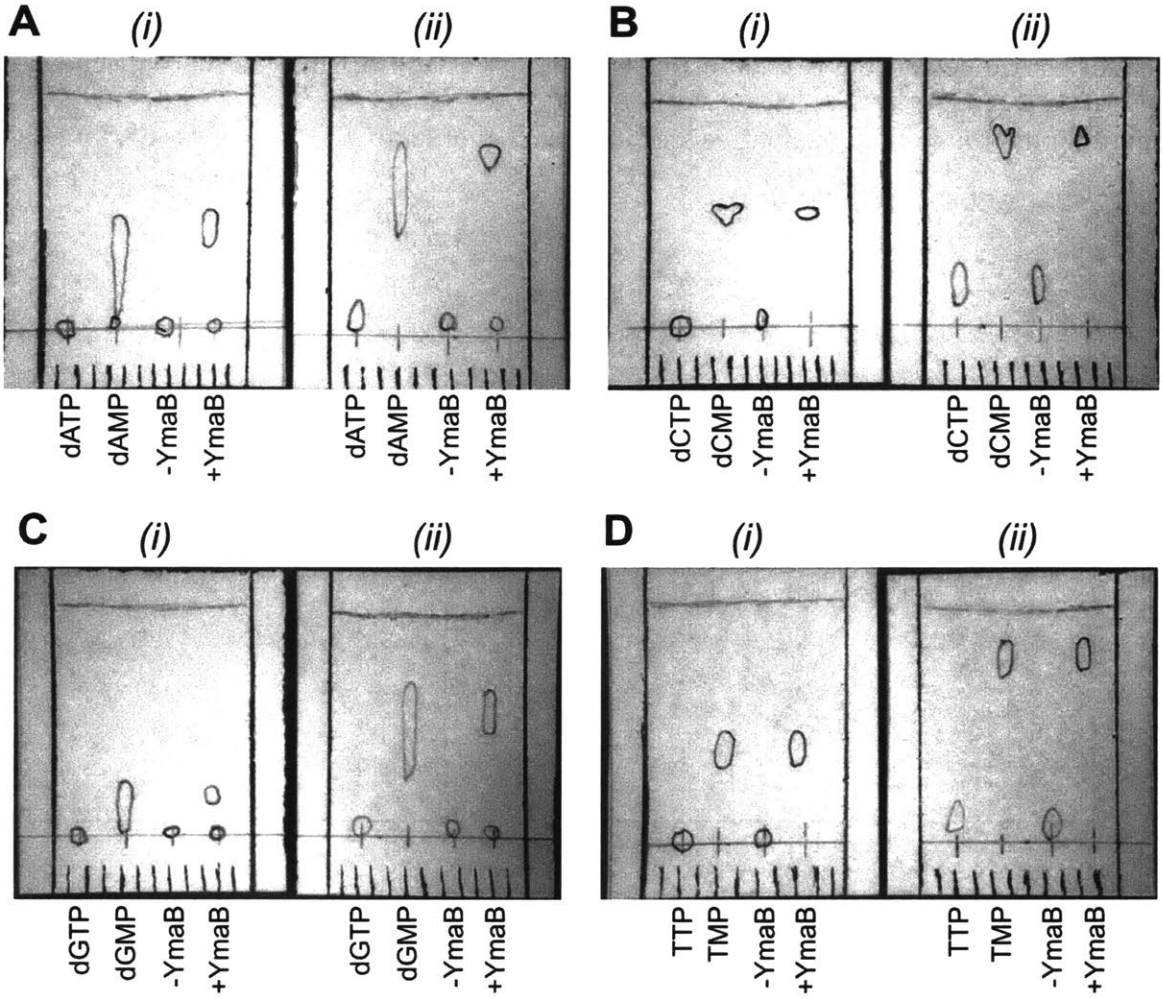


Figure 5.6. Qualitative demonstration by PEI-cellulose TLC of the ability of YmaB to hydrolyze (**A**) dATP, (**B**) dCTP, (**C**) dGTP, and (**D**) TTP. Two plates (5 x 4 cm) for each substrate were prepared and run in either *(i)* 0.5 M formic acid, pH 3.4 to resolve monophosphates, or *(ii)* a 3:1 mixture of 4 M formic acid, pH 3.4 : 4% (w/v) boric acid to resolve di- and triphosphates. –YmaB = control assay in which YmaB was omitted, and +YmaB =complete assay.

Table 5.1. R_f values for the TLC analyses shown in **Figure 5.6**.

Lane	Sample	R_f	
		Solvent (i)	Solvent (ii)
Figure 5.6A: dATP analysis			
1	dATP standard	NM[a]	0.07
2	dAMP standard	0.41	0.73
3	–YmaB control	NM[a]	0.07
4	Complete assay	0.43	0.76 (upper), 0.07 (lower)
Figure 5.6B: dCTP analysis			
1	dCTP standard	NM[a]	0.18
2	dCMP standard	0.50	0.80
3	–YmaB control	NM[a]	0.18
4	Complete assay	0.50	0.85
Figure 5.6C: dGTP analysis			
1	dGTP standard	NM[a]	0.05
2	dGMP standard	0.15	0.45
3	–YmaB control	NM[a]	0.05
4	Complete assay	0.17	0.55 (upper), 0.05 (lower)
Figure 5.6D: TTP analysis			
1	TTP standard	NM[a]	0.11
2	TMP standard	0.40	0.80
3	–YmaB control	NM[a]	0.10
4	Complete assay	0.50	0.81

[a] No migration.

5.3.3. YmaB exists as a monomer in the absence of substrate. The quaternary structure of YmaB was examined using concentration-dependent SV-AUC experiments. A single major species with $s_{<WT>}[20,w] \approx 2.4$ S was observed in the concentration range of 10 – 58 μM (**Figure 5.7A**). This value is consistent with the $s_{20,w}$ (2.3 S) predicted for a YmaB monomer using HYDROPRO[23] and a threading model of the protein. Small amounts of higher molecular weight species were also observed in the range of 4 – 18 S (not shown). The contribution of these species to the total detected signal (~9%) remained constant with increasing protein concentration, indicating that these species were discrete and independent from YmaB. It is likely that the high molecular weight species correspond to co-purifying contaminants, some of which are visible in **Figure 5.4B**.

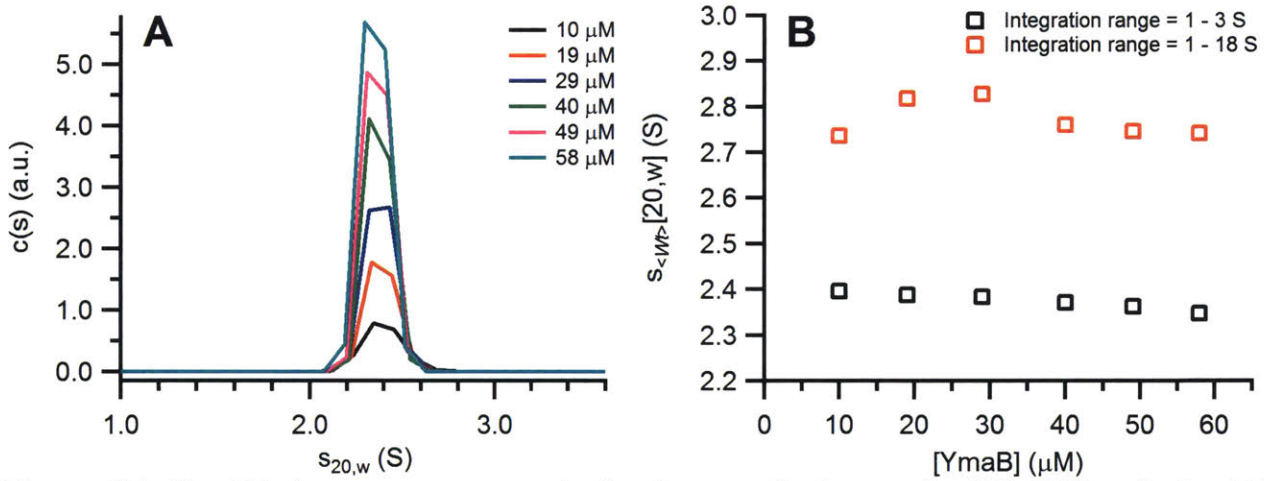


Figure 5.7. YmaB behaves as a monomer in the absence of substrates by SV-AUC analysis. (**A**) Overlay of c(s) sedimentation coefficient distributions in the range of 1 – 3.5 S for samples of YmaB at the indicated concentrations. In all cases, the peak had a $s_{<WT>}[20,w]$ of ~2.4 S. (**B**) Weight-averaged $s_{20,w}$ between 1 – 3 S (black boxes) and 0 – 18 S (red boxes) for the distributions shown in (**A**) plotted as a function of YmaB concentration. The isotherms show that $s_{<WT>}[20,w]$ remains relatively constant with increasing protein concentrations, indicating YmaB does not undergo self-association reactions.

Three pieces of evidence provided by the SV-AUC data indicate that YmaB is a discrete, non-self-associating monomer in the absence of any potential substrates. (1) The contribution of

the major species sedimenting at $s_{<w_f>}[20,w] \approx 2.4$ S to the total detected signal (~91%) remained constant as the protein concentration increased. (2) The position of the major peak visually did not appear to shift to higher sedimentation coefficients with increasing protein concentrations (**Figure 5.7A**). (3) The weight averaged $s_{20,w}$ isotherms for the peak (1 – 3 S) or the entire distribution (0 – 18 S), although exhibiting slightly declining values, are overall relatively constant as the protein concentration increased (**Figure 5.7B**). The declining $s_{<w_f>}[20,w]$ values with increasing concentration suggested YmaB is behaving non-ideally at high concentrations. Given these observations, the six data sets were imported into Sedphat and fit to the "hybrid local continuous distribution/global discrete species" model to rigorously determine the molecular weight and $s_{20,w}$ for YmaB. The best global fit yielded 23200 Da, 68.3% confidence interval (CI) [22300, 23900], and 2.35 S, 68.3% CI [2.33, 2.37], respectively, both of which are within 2% error of the expected (23400 Da) or predicted (2.3 S) values for YmaB monomer.

5.3.4. Testing of the ability of YmaB to increase *in vivo* loading of NrdE with dAMP. The discovery of a tightly bound equivalent of dAMP associated with NrdE (Chapter 4) prompted an examination of the possibility that YmaB, with its ability to hydrolyze dATP, was responsible for generating and loading NrdE with dAMP in *B. subtilis*. As described in Chapter 4, the apparent strength of binding and the inability to remove or install dAMP *in vitro* suggested the nucleotide was buried within the tertiary structure of NrdE and, thus, was likely incorporated during ribosomal synthesis and/or folding. To test this hypothesis, YmaB and NrdE were co-expressed in *E. coli* to see if the basal dAMP load of latter was increased relative to that of NrdE expressed in *E. coli* in the absence of the hydrolase. The amount of dAMP bound to NrdE in the absence of YmaB is fairly high to begin with (0.5 – 0.7 equiv), therefore full complementation of NrdE with one equivalent of nucleotide was the desired goal.

The preliminary experiment reported here was completed by co-expressing YmaB and His_6-Smt3-NrdE in *E. coli* BL21 (DE3) cells following a previously described procedure that was used for inducing the expression of protein folding chaperones prior to the induction of the protein of interest.[28] The procedure relies on expressing the helper proteins and the protein of interest from mutually exclusive inducible promoters, allowing for tight control of the expression of each protein independently from the other. Therefore, the *ymaB* gene was cloned without an affinity tag into pBAD/Myc-HisA so that its expression could be controlled by an arabinose-inducible promoter and so that the enzyme could be separated from NrdE during purification. *E. coli* cells co-transformed with pBAD-*ymaB* and pE-SUMO-*nrdE* were propagated following the same growth scheme as originally reported:[28] cultures were grown at 37 °C with shaking (150 rpm) to an OD_{600} of ~0.35 before adding arabinose (0.02% (w/v) final concentration) to induce YmaB expression. An hour was allotted between arabinose and IPTG induction to allow the *in vivo* concentration of YmaB to accumulate to steady levels prior to inducing His_6-Smt3-NrdE expression, for which the same concentration of IPTG (0.2 mM) as reported in the previous chapters was used instead of 1 mM as described by Zhang *et al.*[28] To ensure that any effects on the dAMP loading of NrdE was due to YmaB, a control strain co-transformed with pBAD/Myc-HisA empty vector was propagated under similar conditions.

SDS-PAGE analysis of the induction samples from the control and experimental strains is shown in **Figure 5.8**. A protein migrating just above the 25 kDa marker, consistent with YmaB (**Figure 5.4B**), is clearly induced by arabinose in the cells transformed with pE-SUMO-*nrdE* and pBAD-*ymaB*, but not in the control strain. One potential issue with BL21 (DE3) cells is that this strain can utilize arabinose as a carbon source; therefore, the expression of proteins from an arabinose-inducible promoter can be expected to decrease over time as the monosaccharide is

consumed by the growing culture. This appeared not to be a problem in the experiment reported here as the expression level of YmaB appeared to persist, as indicated by the similar staining intensity of the bands in the pre-IPTG and post-induction samples in **Figure 5.8**, within the 3 h time frame between addition of arabinose and cell harvest. The intensely stained band in the range of 90 – 100 kDa appearing after IPTG addition is consistent with the size observed for purified His_6-Smt3-NrdE (**Figure 3.1A**). The analysis, therefore, indicates that His_6-Smt3-NrdE and YmaB were successfully co-expressed together in *E. coli*.

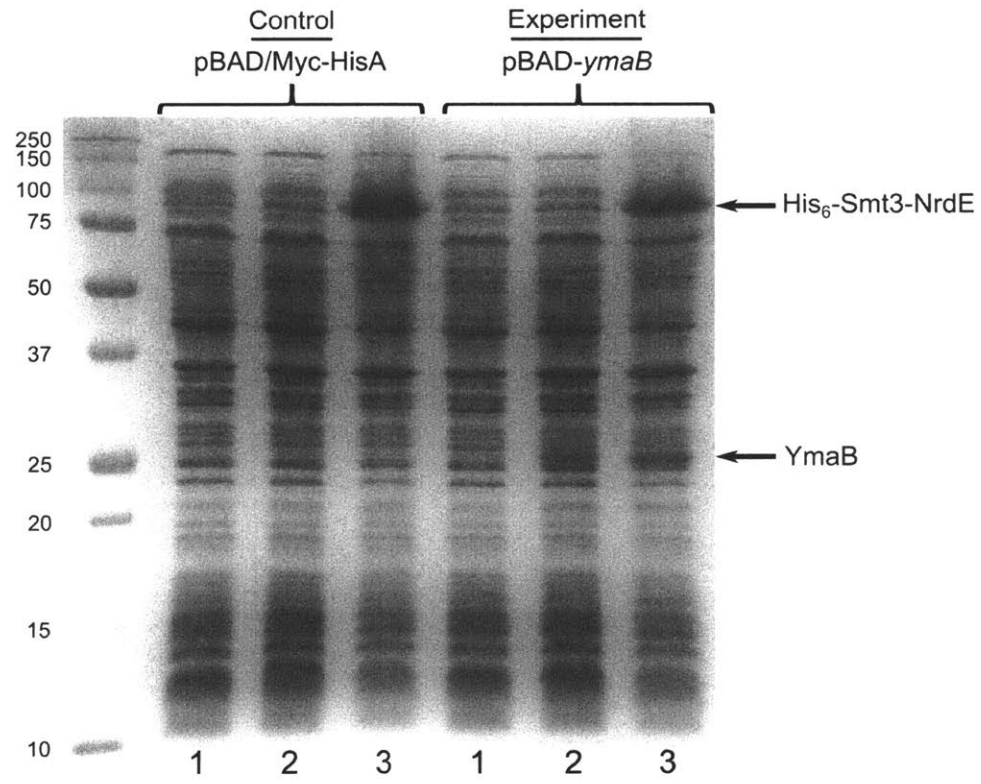


Figure 5.8. Confirmation of the co-expression of YmaB and His_6-Smt3-NrdE in BL21 (DE3) *E. coli* cells. A SDS-PAGE gel (12.5% (w/v)) showing the protein content of cells co-transformed with pE-SUMO-*nrdE* and either pBAD/Myc-HisA (left three lanes) or pBAD-*ymaB* (right three lanes). *Lane 1* – pre-arabinose induction sample, *lane 2*, pre-IPTG induction sample, *lane 3* = pre-harvest sample.

NrdE was purified from both strains (Chapter 3) and the dAMP load of the as-isolated proteins was quantified following the procedure described in Chapter 4. The results of the quantitative analysis are tabulated in **Table 5.2**. The data are encouraging, as co-expression of YmaB and NrdE appeared to have slightly increased the dAMP load of the latter relative to the control strain. However, the results are not statistically different from each other ($p = 0.02$). Given the sample size, more data and additional experiments examining different expression levels of NrdE and YmaB are needed in order to determine if co-expression of the two proteins in *E. coli* can truly increase the amount of dAMP bound to NrdE.

Table 5.2. Quantitation of dAMP bound to NrdE with or without co-expression of YmaB.[a]

Sample	Replicate	$[NrdE]_{initial}$ (μM)	Volume (μL)	nmol dAMP detected	nmol NrdE detected	dAMP:NrdE ratio
NrdE from control strain[b]	1	70	84	3.3	5.4	0.62
	2	86	86	4.3	6.1	0.71
	3	85	84	4.1	5.8	0.71
					Average	0.68 ± 0.05
NrdE from experimental strain[c]	1	76	93	4.2	5.5	0.77
	2	78	93	4.7	5.4	0.86
	3	84	87	4.7	6.0	0.78
	4	75	96	4.5	5.1	0.87
					Average	0.82 ± 0.05

[a] The amounts of NrdE are reported as monomer.
[b] BL21 (DE3) transformed with pE-SUMO-*nrdE* and pBAD/Myc-HisA.
[c] BL21 (DE3) transformed with pE-SUMO-*nrdE* and pBAD-*ymaB*.

A flaw in the experimental design that needs rectification prior to repeating these experiments is that the two plasmids utilized are incompatible with each other due to their utilization of the same replication of origin. Both pBAD/Myc-HisA and pE-SUMO use a pBR322 origin of replication and, thus, will compete for the same replication machinery inside the cell. Such competition can result in partitioning of the plasmids into different daughter cells during cell division unless both plasmids have mutually exclusive selection markers that can be utilized to

ensure cells retain them. The fact that pE-SUMO and pBAD/Myc-HisA have different antibiotic resistance markers that were used during the successful cultivation of the co-transformed *E. coli* strain suggests that the plasmid incompatibility might not have been an issue. However, the lifetime of kanamycin and ampicillin in these cultures is not known; thus, the possibility exists that incompatibility issues arose later during cultivation. Re-cloning one of the genes into a vector that is compatible with pE-SUMO or pBAD/Myc-HisA and repeating the experiment should establish whether or not plasmid incompatibility was a problem not accounted for in the results presented in this chapter.

5.4. DISCUSSION

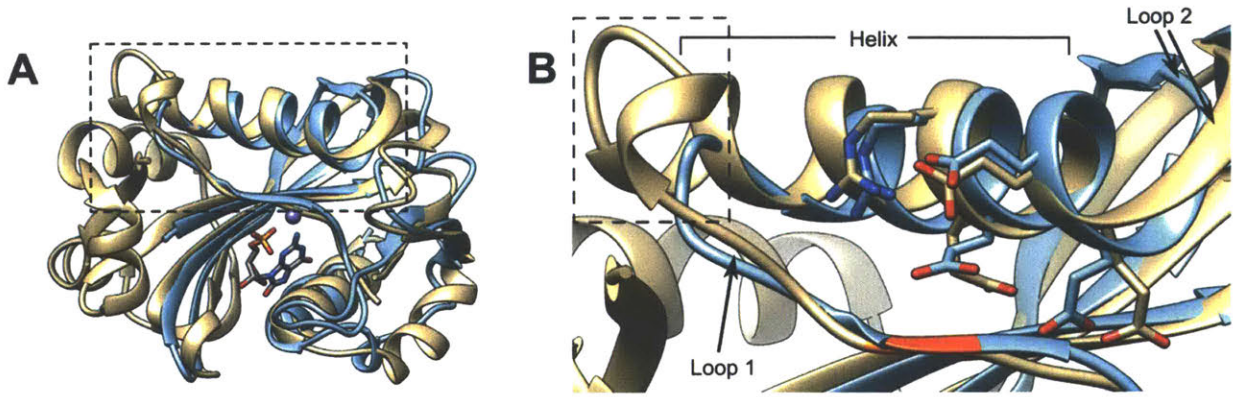
5.4.1. dNTPs link YmaB activity with RNR regulation. The concurrence in *B. subtilis* of a unique protein in the operon of a class Ib RNR that exhibits unusual allosteric and quaternary structural properties suggests that YmaB, provisionally identified as a Nudix hydrolase by BLAST searches, could have a role in the function of RNR. Given the fact that RNR activity is modulated by dNTPs, it seemed logical that a link between YmaB and RNR function could be of a regulatory nature, perhaps by dNTP hydrolysis, since most Nudix hydrolases have specificity for nucleotides of the form (d)NDP-X (**Figure 5.2**). The results of the activity analyses presented in this chapter (**Figure 5.5** and **5.6**) confirmed YmaB's functional prediction, and the identification of dNTPs as its preferred substrates provides the first, albeit indirect, biochemical link between YmaB and the regulation of the *B. subtilis* RNR. If the *in vivo* functions of Nudix hydrolases are "house-keepers" as proposed,[17] then YmaB may simply help control the levels and ratios of the dNTP pools, thereby indirectly affecting RNR activity by changing effector concentrations. However, the sensitivity of RNR to dATP inhibition, despite lacking an ATP-cone domain,[29] and the tightly bound dAMP associated with NrdE suggests that YmaB might play a more direct role in regulating the enzyme.

The ability of Nudix hydrolases to regulate the activity of another enzyme is a relatively unexplored topic to date as most studies have focused on the ability of these enzymes to hydrolyze free substrate molecules in solution. Altering the free concentrations of allosteric effectors for an enzyme, as mentioned above for RNR, could be one route for Nudix hydrolase-based regulation, but so far no studies have attempted to investigate such links. It was recently reported that poly-ADP-ribose covalently bound to proteins can serve as the substrate for the human enzyme hNUDT16 *in vitro*, therefore demonstrating for the first time that a Nudix hydrolase could act on protein-bound ligands.[30] ADP-ribosylation is a protein post-translational modification found in all domains of life that largely plays a role in cellular signaling, but also has been shown to inhibit the activity of certain enzymes including glutamate dehydrogenase and glutamine synthetase.[31-33] The poly-ADP-ribosylated proteins tested in the study by Palazzo *et al.*,[30] however, had no known enzymatic activities, and the physiological relevance of the observed ADP-ribose hydrolytic activity catalyzed by hNUDT16 was not examined. It has also been shown recently that the binding of other proteins can regulate the activity of a Nudix hydrolase. In *E. coli*, diaminopimelate epimerase (DapF) was found to form a tight complex with the Nudix hydrolase RppH (K_d = 5 nM) and increased the latter's mRNA decapping activity both *in vitro* and *in vivo*.[34] This result suggests that the reverse scenario, a Nudix hydrolase directly regulating the activity of another enzyme, is possible. The establishment of a direct link between the activity of a Nudix hydrolase and the regulation of another protein *in vivo* would be a first for this superfamily and, thus, be a valuable contribution to the understanding of physiological roles of Nudix hydrolases inside the cell.[10]

5.4.2. YmaB homologs forms a structurally distinct subgroup of Nudix hydrolases. A structure of YmaB from *T. maritima* (PDB 3E57) has been determined by a structural consortium

at 1.89 Å and offers a comparison with the 2.56 Å structure of the prototypical Nudix hydrolase MutT (PDB 3A6U[35]) from *E. coli* (**Figure 5.9A**). A superimposition of the structures (RMSD = 1.65 Å) reveals striking divergences in the secondary structure of the Nudix box of the two proteins (**Figure 5.9B**). These observations are corroborated by a structure-based sequence alignment (**Figure 5.9C**) which shows the YmaB Nudix box is lengthened by six residues, resulting in a $GHx_5GAx_{3-4}Ex_5UxRExxEExxU$ sequence motif that is manifested as a longer first loop and an extension of the α-helix N-terminus in the loop-helix-loop structural fold that Nudix boxes adopt. However, the RExxEE segment critical for the catalytic activity of other Nudix hydrolases (**Figure 5.9B**) appears to be structurally unaffected.[11, 12] Another noteworthy divergence, but not uncommon among the Nudix hydrolases, is that YmaB has an N-terminal extension of approximately 55 amino acids relative to MutT. In other enzymes, such extensions are responsible for dimerization,[36, 37] contributions to substrate specificity,[38] and, in one case, imparts the enzyme with a secondary activity.[36]

As YmaB is the first identified Nudix hydrolase to have a lengthened Nudix box, the function of this extension is currently unknown. The Gly-Ala dyad in the extension that is conserved in the three YmaB sequences (G_{100} and A_{101} in *T. maritima*, **Figure 5.9C**) packs in closely next to the N-terminal end of the α-helix bearing the RExxEE motif in the *T. maritima* structure, which may reduce the conformational flexibility of the loop. Such a packing interaction might aid in positioning a conserved His residued (H_{94} in *T. maritima*) in the substrate binding site and, therefore, help govern the specificity of the enzyme. Further studies with mutants of the RExxEE motif and other conserved residues in these proteins should help determine the relevance of each to specificity and catalysis in YmaB.



C

```
TmYmaB   92 GGHVREGDGATPREAFL-KGLEREVNEEVDV 121
SaYmaB   98 GGHMNDVPGAESINEVLRVNAQRELEEEVGL 128
BsYmaB   97 GGHMNAIEGAASFAEVLKLNTDRELEEELQI 127
SaMutT   30 GGKIEE-------GESQVHALLREVKEELNL  53
BsMutT   32 GGRVDP-------GESAEEAAVREILEETGY  55
EcMutT   37 GGKIEM-------GETPEQAVVRELQEEVGI  60
            **                   **  **  :
```

Figure 5.9. Structural comparison of *E. coli* MutT (PDB 3A6U, blue) and *T. maritima* YmaB (PDB 3E57, tan). (**A**) Overlay of MutT and YmaB (RMSD = 1.65 Å for the overlaid portions of the structures). Mn^{2+} and 8-oxo-dGMP from the MutT structure indicate the location of the active and substrate recognition sites in Nudix hydrolases, respectively. The ion and compound are shown as sticks and colored according to element: gray = C, blue = N, red = O, orange = P, purple = Mn. The dashed box indicates the loop-helix-loop motif adopted by the Nudix box. (**B**) Zoom in on the Nudix box loop-helix-loop structural motifs to show the secondary structure of the seven residues (dashed box) present in YmaB proteins. The glycine residues initiating the Nudix boxes are colored orange. The side chains of the RExxEE segment are shown as sticks and colored according to heteroatom. (**C**) Structure-based sequence alignment of the Nudix box of representative MutT and YmaB proteins. Note the large gap in the MutT sequences as compared to the YmaB sequences. Residues conserved in all sequences and conserved with positive substitutions are colored red and blue, respectively. Conserved residues in either YmaBs or MutTs are highlighted in yellow or cyan, respectively.

A particularly intriguing result from a recent study of *B. cereus* CDP-choline pyrophosphatase (CDP-Chase)[36] prompted an examination of the quaternary structure of YmaB in order to assess whether or not its long N-terminus caused the protein to dimerize and/or gave YmaB a secondary function. Like YmaB, CDP-Chase has an N-terminal extension fused to the core Nudix fold. In this case, the extension not only causes CDP-Chase to dimerize, but also

imparts on the enzyme single stranded RNA 3′ → 5′ exonuclease activity in addition to its ability to hydrolyze CDP-choline. Furthermore, the exonuclease activity was reported to be accelerated in the presence of CDP-choline. The expression of CDP-Chase in sporulating cells suggested that the enzyme played an important role in the commitment of *B. cereus* to sporulation by degrading both building blocks for membrane expansion and mRNAs no longer needed by the cell.[36, 39] Critical to these observations was the dependence of the exonuclease activity of *B. cereus* CDP-Chase on the enzyme's ability to form an asymmetric dimer.[36]

The results of the study by Duong-Ly *et al.*[36] raised the questions of whether or not other as yet uncharacterized Nudix hydrolases might form similar asymmetric dimers via N-terminal extensions, and, if so, what other types of activities could be imparted on these enzymes. In the case of YmaB, it seemed feasible that if it formed an asymmetric dimer, then YmaB could have a secondary function, in addition to its ability to degrade dNTPs, that might contribute to the regulation of RNR under different environmental conditions. To investigate this possibility, SV-AUC experiments were carried out (**Figure 5.7**), which revealed that YmaB is a monomer, indicating perhaps that it does not possess a secondary activity if dimerization is key for such a phenomenon to occur. However, these experiments were run in the absence of nucleotides, and it might be possible that YmaB forms a dimer only in the presence of the appropriate substrate(s). Further biophysical studies with catalytically inactive mutants or using metal ions that cannot substitute for the requirement of Mg^{2+} or Mn^{2+} could address this possibility.

5.4.3. Effort to test the proposed role for YmaB as a dAMP insertase in E. coli. It seems more than a mere coincidence that YmaB is associated with a class Ib RNR that exhibits dATP inhibition and has dAMP tightly bound to its α subunit, both of which are highly unusual in comparison to other Ib systems. Therefore, it seems likely that YmaB's function is going to be

linked somehow with one or both of these attributes, even though the apparent preference of the enzyme for pyrimidine dNTPs seems at odds with this proposal. The substrate promiscuity often exhibited by Nudix hydrolases makes establishing the main function of the enzyme *in vivo* difficult. For example, *B. subtilis* RppH exhibits a high specific activity for dNTP hydrolysis,[27, 40] despite having 5′-mRNA decapping activity which is demonstrably more important for the cell.[16] As YmaB and RNR are also removed from the context of their native environment, there could be as yet unknown proteins and/or small molecule effectors in *B. subtilis* that might stimulate dATP hydrolysis while suppressing other hydrolytic activities under certain conditions.

YmaB's ability to hydrolyze dATP can be envisioned to accomplish one of two mutually exclusive goals: (1) reduction of local dATP concentrations to alleviate RNR inhibition, or (2) generate and insert dAMP into NrdE, thus making RNR more susceptible to dATP inhibition. The latter proposal is most appealing, given that NrdE was unable to be loaded with exogenously supplied dAMP. As mentioned in Chapter 4, this result suggested that dAMP was buried within NrdE and, thus, was required to be installed during ribosomal synthesis and/or folding. No examples of Nudix hydrolases performing a similar function currently exist, but the recent report of *E. coli* RppH forming a tight complex with DapF[34] is encouraging and suggests that an interaction of YmaB with NrdE, whether folded or unfolded, falls in the realm of possibility.

The result of co-expressing YmaB and NrdE together in *E. coli*, while promising, was inconclusive as to the ability of the former to install dAMP on the latter. Regardless of plasmid incompatibility noted earlier, the experiment presents other issues that need to be addressed. If dATP is utilized as the precursor for NrdE-associated dAMP, then the *in vivo* rate of expression and the total amounts of NrdE and YmaB need to be finely tuned to the *in vivo* concentrations of dATP in order to maintain balanced dNTP pools and prevent the inhibition of the *B. subtilis* RNR.

The relative concentrations of NrdE and YmaB do not differ by more than 2-fold during exponential and stationary phase (Appendix 2, **Table A2.4** and **A2.5**), therefore the use of dual expression vectors to produce roughly equivalent levels of both proteins may be warranted. In addition to the vector(s) used, the types and amounts of inducer used and the length of induction will need to be optimized as well in order for the co-expression experiment to successfully executed. One final important issue requiring consideration is that factors critical for the functional association of the two proteins may be missing in the heterologous host. Since *B. subtilis* is one of the best studied Gram-positive bacteria and is genetically tractable, examination of an association of YmaB and NrdE within their native environment may provide the best chance of demonstrating a role for the latter in the installation of dAMP into NrdE.

5.4.4. Can YmaB help insert dAMP into folded NrdE? The recently obtained structure of *B. subtilis* NrdE with dAMP bound to its N-terminal domain suggests that installation of the nucleotide might not necessarily depend on NrdE being in an unfolded state. Further support for this statement is provided by the results of the nucleotide binding studies, which showed apo-NrdE could bind two equivalents of dATP whereas holo-NrdE bound only one. This result indicates that dAMP blocks one of the dATP binding sites in holo-protein, and perhaps suggests that a dATP molecule occupying the site could be the immediate precursor for the more tightly bound dAMP via hydrolysis of its γ- and β-phosphates. If this hypothesis is correct, then YmaB could be the agent responsible for the hydrolysis reaction.

A possible model that accommodates the data presented in Chapter 4 and the hypothesized function of YmaB is shown in **Figure 5.10**. Under conditions of dNTP demand, the *in vivo* dATP concentration remains low enough that occupancy of the dAMP/dATP binding site is minimized. Once the demands of the cell are met and the dNTP pools in general begin to increase in size, the

occupancy of the site by dATP increases in frequency (**Figure 5.10**, step A). Under these conditions, YmaB recognizes and binds to NrdE at the site where the dATP is bound (**Figure 5.10**, step B) and hydrolyzes the nucleotide to dAMP (step C). This reaction triggers a conformational change in NrdE, possibly powered by hydrolysis of pyrophosphate, that renders the site high affinity for dAMP. Dissociation of YmaB (**Figure 5.10**, step D) releases dAMP-loaded NrdE which, being more susceptible to dATP inhibition, binds to NrdF and begins a general shutdown of dNTP production.

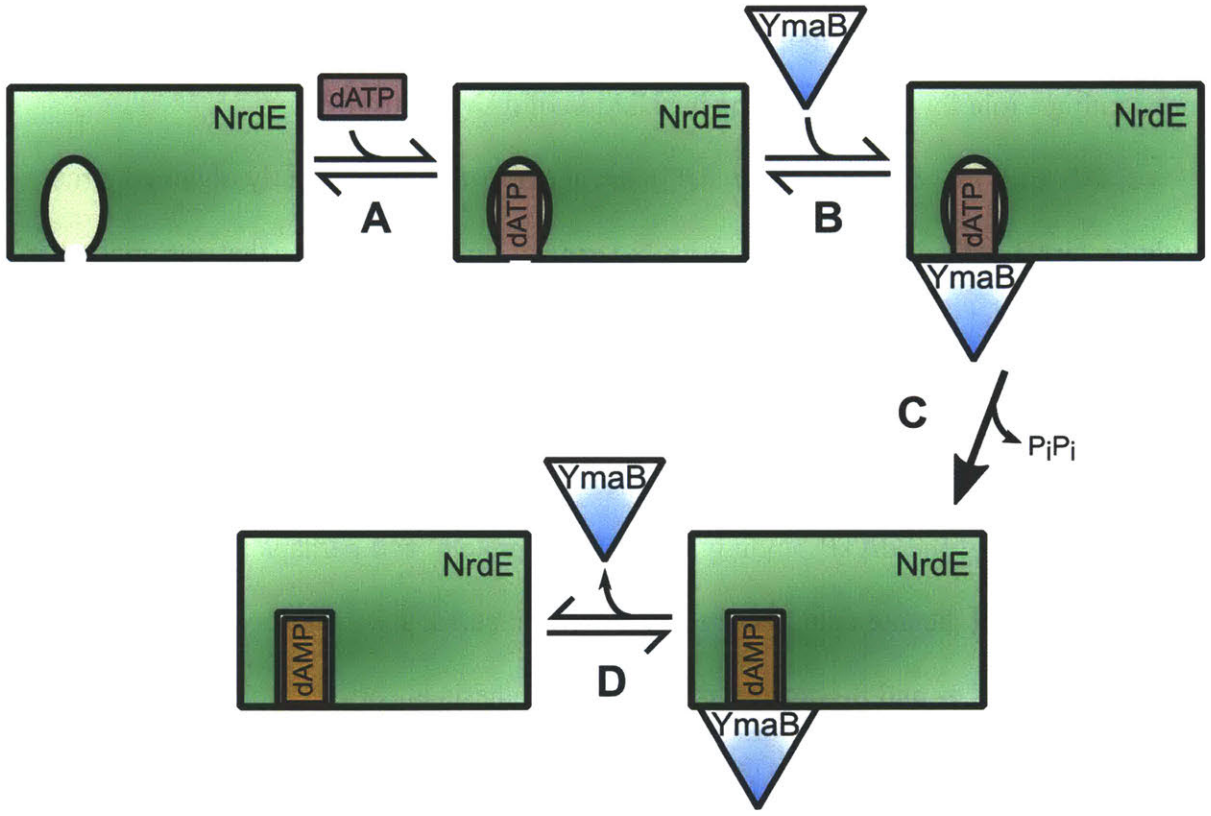


Figure 5.10. A possible model for YmaB-facilitated insertion of dAMP into folded NrdE. *Step A*: the open binding site on NrdE can bind dATP with relatively low affinity. Increased occupancy of the site can result from a buildup of dATP, indicating that the cell is replete with dNTPs. *Step B*: YmaB recognizes the dATP-occupied site and binds to NrdE. *Step C*: YmaB catalyzes the hydrolysis of dATP to dAMP and pyrophosphate. This reaction causes a protein conformational change in NrdE, possibly driven by free energy derived from pyrophosphate hydrolysis, that renders the site high affinity for dAMP. *Step D*: YmaB dissociates from dAMP-loaded NrdE, which is now more susceptible to dATP inhibition.

This model presents several predictions that further studies can address. The most salient feature is that YmaB and NrdE form a complex, but only when the latter has dATP bound in the overall activity allosteric site. The use of non-hydrolysable dATP analogs, catalytically inactive YmaB mutants, and/or metal ions that cannot support YmaB catalysis in conjunction with biophysical techniques should be able to establish if such a complex can form. Regardless of the result, such studies should go a long way in understanding any functional link between YmaB activity and RNR function.

5.4.5. Redundancy of functions in B. subtilis will make establishing an in vivo role for YmaB challenging. The final challenge, regardless of whether or not YmaB supports RNR function, will be relating any discoveries made with YmaB *in vitro* to the enzyme's physiological function in *B. subtilis*. As with all Nudix hydrolases, the establishment of an *in vivo* function for YmaB will be a difficult task due to substrate promiscuity that results in a redundancy of function between several enzymes. With the caveat that only a limited number of compounds have been tested, YmaB seems to have a narrower substrate specificity relative to other reported Nudix hydrolases since it was found to be only able to hydrolyze the canonical dNTPs. However, *B. subtilis* has at least five other Nudix hydrolases that have been characterized or annotated (**Table 5.3**), one of which, RppH, has been reported to be fairly efficient at catalyzing the pyrophosphohydrolysis of dNTPs[27, 40] as mentioned above, and another, MutT, can be inferred to have such activity as well based on studies of *E. coli* MutT.[13] Furthermore, several other phosphatases have been either characterized or annotated in the *B. subtilis* genome, and one, YpgQ, has recently been shown to have Mn^{2+}-dependent (d)NTP pyrophosphohydrolase activity.[41] If YmaB's function is tied to RNR, however, then this link provides a potential handle that could distinguish

YmaB-related activity from that of other Nudix hydrolases and phosphatases and, therefore, provides a better chance at identifying the function of YmaB *in vivo*.

Table 5.3. Additional predicted and confirmed Nudix hydrolases in *B. subtilis*.

Enzyme	Activities detected	Proposed function	Reference
MutT	-	Anti-mutator	*42*
RppH (aka YtkD, MutTA)	8-oxo-dGTPase (d)NTPase Ap_4Aase mRNA 5′-exonuclease	Anti-mutator mRNA turnover Regulate metabolite levels	*16, 27, 40, 42*
NudF	ADP-ribose	Regulate metabolite levels Isopentenol biosynthesis	*43, 44*
YjhB[a]	CDP-choline RNA nuclease (predicted)	Regulate sporulation and cell division	*36*
YvcI	-	-	-

[a] Activities and functions are predictions given that YjhB is 44% identical (65% with positive substitutions) to the *B. cereus* CDP-Chase.

5.5. REFERENCES

1. Kobayashi, K., Ehrlich, S. D., Albertini, A., Amati, G., Andersen, K. K., Arnaud, M., Asai, K., Ashikaga, S., Aymerich, S., Bessieres, P., Boland, F., Brignell, S. C., Bron, S., Bunai, K., Chapuis, J., Christiansen, L. C., Danchin, A., Débarbouillé, M., Dervyn, E., Deuerling, E., Devine, K., Devine, S. K., Dreesen, O., Errington, J., Fillinger, S., Foster, S. J., Fujita, Y., Galizzi, A., Gardan, R., Eschevins, C., Fukushima, T., Haga, K., Harwood, C. R., Hecker, M., Hosoya, D., Hullo, M. F., Kakeshita, H., Karamata, D., Kasahara, Y., Kawamura, F., Koga, K., Koski, P., Kuwana, R., Imamura, D., Ishimaru, M., Ishikawa, S., Ishio, I., Le Coq, D., Masson, A., Mauël, C., Meima, R., Mellado, R. P., Moir, A., Moriya, S., Nagakawa, E., Nanamiya, H., Nakai, S., Nygaard, P., Ogura, M., Ohanan, T., O'Reilly, M., O'Rourke, M., Pragai, Z., Pooley, H. M., Rapoport, G., Rawlins, J. P., Rivas, L. A., Rivolta, C., Sadaie, A., Sadaie, Y., Sarvas, M., Sato, T., Saxild, H. H., Scanlan, E., Schumann, W., Seegers, J., Sekiguchi, J., Sekowska, A., Séror, S. J., Simon, M., Stragier, P., Studer, R., Takamatsu, H., Tanaka, T., Takeuchi, M., Thomaides, H. B., Vagner, V., van Dijl, J. M., Watabe, K., Wipat, A., Yamamoto, H., Yamamoto, M., Yamamoto, Y., Yamane, K., Yata, K., Yoshida, K., Yoshikawa, H., Zuber, U., and Ogasawara, N. (2003) Essential *Bacillus subtilis* Genes, *Proc. Natl. Acad. Sci. U. S. A. 100*, 4678-4683.

2. Scotti, C., Valbuzzi, A., Perego, M., Galizzi, A., and Albertini, A. M. (1996) The *Bacillus subtilis* Genes for Ribonucleotide Reductase are Similar to the Genes for the Second Class I NrdE/NrdF Enzymes of *Enterobacteriaceae*, *Microbiology 142*, 2995-3004.

3. Chaudhuri, R. R., Allen, A. G., Owen, P. J., Shalom, G., Stone, K., Harrison, M., Burgis, T. A., Lockyer, M., Garcia-Lara, J., Foster, S. J., Pleasance, S. J., Peters, S. E., Maskell, D. J., and Charles, I. G. (2009) Comprehensive Identification of Essential *Staphylococcus aureus* Genes Using Transposon-Mediated Differential Hybridisation (TMDH), *BMC Genomics 10*, 291.

4. Osbourn, A. E., and Field, B. (2009) Operons, *Cell. Mol. Life Sci. 66*, 3755-3775.

5. Härtig, E., Hartmann, A., Schätzle, M., Albertini, A. M., and Jahn, D. (2006) The *Bacillus subtilis nrdEF* Genes, Encoding a Class Ib Ribonucleotide Reductase, are Essential for Aerobic and Anaerobic Growth, *Appl. Environ. Microbiol. 72*, 5260-5265.

6. Rodionov, D. A., and Gelfand, M. S. (2005) Identification of a Bacterial Regulatory System for Ribonucleotide Reductases by Phylogenetic Profiling, *Trends Genet. 21*, 385-389.

7. Torrents, E., Grinberg, I., Gorovitz-Harris, B., Lundström, H., Borovok, I., Aharonowitz, Y., Sjöberg, B. M., and Cohen, G. (2007) NrdR Controls Differential Expression of the *Escherichia coli* Ribonucleotide Reductase Genes, *J. Bacteriol. 189*, 5012-5021.

8. Nicolas, P., Mäder, U., Dervyn, E., Rochat, T., Leduc, A., Pigeonneau, N., Bidnenko, E., Marchadier, E., Hoebeke, M., Aymerich, S., Becher, D., Bisicchia, P., Botella, E., Delumeau, O., Doherty, G., Denham, E. L., Fogg, M. J., Fromion, V., Goelzer, A., Hansen, A., Härtig, E., Harwood, C. R., Homuth, G., Jarmer, H., Jules, M., Klipp, E., Le Chat, L., Lecointe, F., Lewis, P., Liebermeister, W., March, A., Mars, R. A. T., Nannapaneni, P., Noone, D., Pohl, S., Rinn, B., Rügheimer, F., Sappa, P. K., Samson, F., Schaffer, M., Schwikowski, B., Steil, L., Stülke, J., Wiegert, T., Devine, K. M., Wilkinson, A. J., van Dijl, J. M., Hecker, M., Völker, U., Bessières, P., and Noirot, P. (2012) Condition-Dependent Transcriptome Reveals High-Level Regulatory Architecture in *Bacillus subtilis*, *Science 335*, 1103-1106.

9. Rasmussen, S., Nielsen, H. B., and Jarmer, H. (2009) The Transcriptionally Active Regions in the Genome of *Bacillus subtilis*, *Mol. Microbiol. 73*, 1043-1057.

10. McLennan, A. G. (2006) The Nudix Hydrolase Superfamily, *Cell. Mol. Life Sci. 63*, 123-143.

11. Mildvan, A. S., Xia, Z., Azurmendi, H. F., Saraswat, V., Legler, P. M., Massiah, M. A., Gabelli, S. B., Bianchet, M. A., Kang, L. W., and Amzel, L. M. (2005) Structures and Mechanisms of Nudix Hydrolases, *Arch. Biochem. Biophys. 433*, 129-143.

12. Ooga, T., Yoshiba, S., Nakagawa, N., Kuramitsu, S., and Masui, R. (2005) Molecular Mechanism of the *Thermus thermophilus* ADP-Ribose Pyrophosphatase from Mutational and Kinetic Studies, *Biochemistry 44*, 9320-9329.

13. Bhatnagar, S. K., and Bessman, M. J. (1988) Studies on the Mutator Gene, *mutT* of *Escherichia coli* - Molecular Cloning of the Gene, Purification of the Gene Product, and Identification of a Novel Nucleoside Triphosphatase, *J. Biol. Chem. 263*, 8953-8957.

14. Maki, H., and Sekiguchi, M. (1992) MutT Protein Specifically Hydrolyzes a Potent Mutagenic Substrate for DNA Synthesis, *Nature 355*, 273-275.

15. Deana, A., Celesnik, H., and Belasco, J. G. (2008) The Bacterial Enzyme RppH Triggers Messenger RNA Degradation by 5' Pyrophosphate Removal, *Nature 451*, 355-358.

16. Richards, J., Liu, Q. S., Pellegrini, O., Celesnik, H., Yao, S. Y., Bechhofer, D. H., Condon, C., and Belasco, J. G. (2011) An RNA Pyrophosphohydrolase Triggers 5 '-Exonucleolytic Degradation of mRNA in *Bacillus subtilis*, *Mol. Cell 43*, 940-949.

17. Bessman, M. J., Frick, D. N., and O'Handley, S. F. (1996) The MutT Proteins or "Nudix" Hydrolases, a Family of Versatile, Widely Distributed, "Housecleaning" Enzymes, *J. Biol. Chem. 271*, 25059-25062.

18. Korbie, D. J., and Mattick, J. S. (2008) Touchdown PCR for Increased Specificity and Sensitivity in PCR Amplification, *Nat. Protoc. 3*, 1452-1456.

19. Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D., and Bairoch, A. (2005) Protein Identification and Analysis Tools on the ExPASy Server., In *The Proteomics Protocols Handbook* (Walker, J. M., Ed.), pp 571-607, Humana Press, Totowa, NJ.

20. Baykov, A. A., Evtushenko, O. A., and Avaeva, S. M. (1988) A Malachite Green Procedure for Ortho-Phosphate Determination and its Use in Alkaline Phosphatase-Based Enzyme Immunoassay, *Anal. Biochem. 171*, 266-270.

21. Zhao, H. Y., Ghirlando, R., Piszczek, G., Curth, U., Brautigam, C. A., and Schuck, P. (2013) Recorded Scan Times can Limit the Accuracy of Sedimentation Coefficients in Analytical Ultracentrifugation, *Anal. Biochem. 437*, 104-108.

22. Laue, T. M., Shah, B. D., Ridgeway, T. M., and Pelletier, S. L. (1992) Computer-aided Interpretation of Analytical Sedimentation Data for Proteins, In *Analytical Ultracentrifugation in Biochemistry and Polymer Science* (Harding, S. E., Rowe, A. J., and Horton, J. C., Eds.), pp 90-125, Royal Society of Chemistry, Cambridge.

23. Ortega, A., Amorós, D., and de la Torre, J. G. (2011) Prediction of Hydrodynamic and Other Solution Properties of Rigid Proteins from Atomic- and Residue-Level Models, *Biophys. J. 101*, 892-898.

24. Kelley, L. A., and Sternberg, M. J. E. (2009) Protein Structure Prediction on the Web: A Case Study Using the Phyre Server, *Nat. Protoc. 4*, 363-371.

25. O'Handley, S. F., Frick, D. N., Bullions, L. C., Mildvan, A. S., and Bessman, M. J. (1996) *Escherichia coli orf17* Codes for a Nucleoside Triphosphate Pyrophosphohydrolase Member of the MutT Family of Proteins - Cloning, Purification, and Characterization of the Enzyme, *J. Biol. Chem. 271*, 24649-24654.

26. Xu, W. L., Dunn, C. A., Jones, C. R., D'Souza, G., and Bessman, M. J. (2004) The 26 Nudix Hydrolases of *Bacillus cereus*, a Close Relative of *Bacillus anthracis*, *J. Biol. Chem. 279*, 24861-24865.

27. Xu, W. L., Jones, C. R., Dunn, C. A., and Bessman, M. J. (2004) Gene *ytkD* of *Bacillus subtilis* Encodes an Atypical Nucleoside Triphosphatase Member of the Nudix Hydrolase Superfamily, *J. Bacteriol. 186*, 8380-8384.

28. Zhang, X., Liu, Y., Genereux, J. C., Nolan, C., Singh, M., and Kelly, J. W. (2014) Heat-Shock Response Transcriptional Program Enables High-Yield and High-Quality Recombinant Protein Production in *Escherichia coli*, *ACS Chem. Biol. 9*, 1945-1949.

29. Aravind, L., Wolf, Y. I., and Koonin, E. V. (2000) The ATP-Cone: An Evolutionarily Mobile, ATP-Binding Regulatory Domain, *J. Mol. Microbiol. Biotechnol. 2*, 191-194.

30. Palazzo, L., Thomas, B., Jemth, A. S., Colby, T., Leidecker, O., Feijs, K. L. H., Zaja, R., Loseva, O., Puigvert, J. C., Matic, I., Helleday, T., and Ahel, I. (2015) Processing of Protein ADP-Ribosylation by Nudix Hydrolases, *Biochem. J. 468*, 293-301.

31. Choi, M. M., Huh, J. W., Yang, S. J., Cho, E. H., Choi, S. Y., and Cho, S. W. (2005) Identification of ADP-Ribosylation Site in Human Glutamate Dehydrogenase Isozymes, *FEBS Lett. 579*, 4125-4130.

32. Hassa, P. O., Haenni, S. S., Elser, M., and Hottiger, M. O. (2006) Nuclear ADP-Ribosylation Reactions in Mammalian Cells: Where are We Today and Where are We Going?, *Microbiol. Mol. Biol. Rev. 70*, 789-829.

33. Moss, J., Watkins, P. A., Stanley, S. J., Purnell, M. R., and Kidwell, W. R. (1984) Inactivation of Glutamine Synthetases by an NAD-Arginine ADP-Ribosyltransferase, *J. Biol. Chem. 259*, 5100-5104.

34. Lee, C. R., Kim, M., Park, Y. H., Kim, Y. R., and Seok, Y. J. (2014) RppH-Dependent Pyrophosphohydrolysis of mRNAs is Regulated by Direct Interaction with DapF in *Escherichia coli*, *Nucleic Acids Res. 42*, 12746-12757.

35. Nakamura, T., Meshitsuka, S., Kitagawa, S., Abe, N., Yamada, J., Ishino, T., Nakano, H., Tsuzuki, T., Doi, T., Kobayashi, Y., Fujii, S., Sekiguchi, M., and Yamagata, Y. (2010) Structural and Dynamic Features of the MutT Protein in the Recognition of Nucleotides with the Mutagenic 8-Oxoguanine Base, *J. Biol. Chem. 285*, 444-452.

36. Duong-Ly, K. C., Gabelli, S. B., Xu, W. L., Dunn, C. A., Schoeffield, A. J., Bessman, M. J., and Amzel, L. M. (2011) The Nudix Hydrolase CDP-Chase, a CDP-Choline Pyrophosphatase, is an Asymmetric Dimer with Two Distinct Enzymatic Activities, *J. Bacteriol. 193*, 3175-3185.

37. Gabelli, S. B., Bianchet, M. A., Bessman, M. J., and Amzel, L. M. (2001) The Structure of ADP-Ribose Pyrophosphatase Reveals the Structural Basis for the Versatility of the Nudix Family, *Nat. Struct. Biol. 8*, 467-472.

38. Kang, L. W., Gabelli, S. B., Bianchet, M. A., Xu, W. L., Bessman, M. J., and Amzel, L. M. (2003) Structure of a Coenzyme A Pyrophosphatase from *Deinococcus radiodurans*: a Member of the Nudix Family, *J. Bacteriol. 185*, 4110-4118.

39. Schaeffer, P. (1969) Sporulation and Production of Antibiotics, Exoenzymes, and Exotoxins, *Bacteriological Reviews 33*, 48-71.

40. Ramírez, M. I., Castellanos-Juárez, F. X., Yasbin, R. E., and Pedraza-Reyes, M. (2004) The *ytkD* (*mutTA*) Gene of *Bacillus subtilis* Encodes a Functional Antimutator 8-oxo-(dGTP/GTP)ase and is Under Dual Control of Sigma A and Sigma F RNA Polymerases, *J. Bacteriol. 186*, 1050-1059.

41. Jeon, Y. J., Park, S. C., Song, W. S., Kim, O. H., Oh, B. C., and Yoon, S. I. (2016) Structural and Biochemical Characterization of Bacterial YpgQ Protein Reveals a Metal-Dependent Nucleotide Pyrophosphohydrolase, *J. Struct. Biol. 195*, 113-122.

42. Castellanos-Juárez, F. X., Álvarez-Álvarez, C., Yasbin, R. E., Setlow, B., Setlow, P., and Pedraza-Reyes, M. (2006) YtkD and MutT Protect Vegetative Cells but Not Spores of *Bacillus subtilis* from Oxidative Stress, *J. Bacteriol. 188*, 2285-2289.

43. Dunn, C. A., O'Handley, S. F., Frick, D. N., and Bessman, M. J. (1999) Studies on the ADP-ribose Pyrophosphatase Subfamily of the Nudix Hydrolases and Tentative Identification of *trgB*, a Gene Associated with Tellurite Resistance, *J. Biol. Chem. 274*, 32318-32324.

44. Withers, S. T., Gottlieb, S. S., Lieu, B., Newman, J. D., and Keasling, J. D. (2007) Identification of Isopentenol Biosynthetic Genes from *Bacillus subtilis* by a Screening Method Based on Isoprenoid Precursor Toxicity, *Appl. Environ. Microbiol. 73*, 6277-6283.

Appendix 1, Section 1

In silico structural studies of the *E. coli* class Ia RNR

A1.1. Introduction. Reported in this appendix are intellectual contributions, based off of structural observations, that were made for other members of the Stubbe group who are studying the radical transfer pathway and the α/β interface of the *E. coli* class Ia RNR. This enzyme, like the *B. subtilis* Ib RNR, is composed of two subunits, α_2 and β_2, that form an $\alpha_2\beta_2$ complex during turnover. The α_2 subunit binds substrates and effectors and houses the essential cysteine, C_{439}, that is transiently oxidized to a thiyl radical during each turnover. The catalytic radical is housed in the β_2 subunit on a tyrosine residue, Y_{122}, located deep in the protein tertiary structure adjacent to a diferric cluster. The oxidation of C_{439} in α by Y_{122}• in β occurs over a distance of ~35 Å by multiple proton-coupled electron transfer (PCET) steps involving three redox active tyrosine residues (Y_{356} – β, Y_{730} and Y_{731} – α) and possibly one tryptophan (W_{48} – β). The radical transfer pathway of the *E. coli* Ia RNR is the most extensively studied of all RNRs.[1] The working hypothesis of its mechanism is shown in **Figure A1.1**.

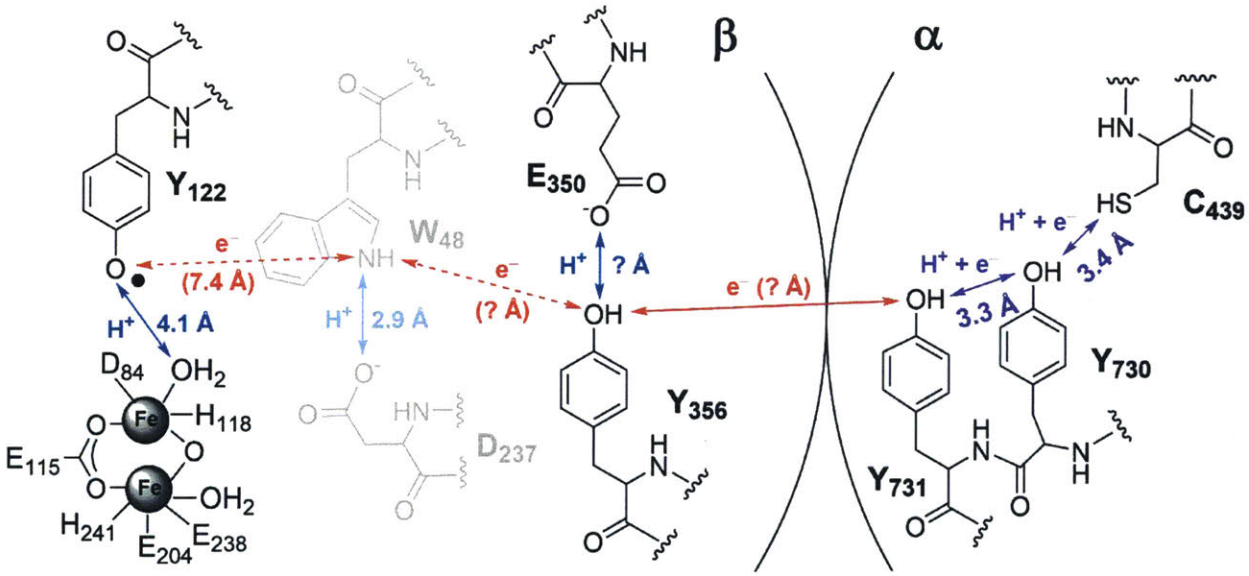


Figure A1.1. Proposed radical transfer mechanism in the *E. coli* class Ia RNR. Electron movement is shown as red arrows, proton movement as blue arrows, and co-linear movement of both proton and electron as purple arrows. Y_{356} and its putative proton acceptor E_{350} are located on the flexible C-terminal tail of β and are disordered in all published structures. W_{48} and its putative proton acceptor D_{237} are shown in gray as there currently is no evidence supporting their involvement in the radical transfer mechanism.

Though many attempts have been made, a structure of the active complex of the *E. coli* Ia RNR has still not been obtained by X-ray crystallography or NMR spectroscopy. Low resolution structures have been obtained by cryo-electron microscopy[2] and, in general, support an $\alpha_2\beta_2$ complex as the active form of the enzyme as originally proposed in a docking model of the individual subunits based on shape and charge complementarity.[3] Unfortunately, several critical components of the radical transfer pathway are located at the $\alpha_2\beta_2$ interface, a region of the complex which remains poorly characterized in all class I RNRs. In particular, the C-terminal tail of β, which is responsible for the majority of the interactions between the subunits,[4] is located at the interface and contains the two conserved residues, Y_{356} and E_{350}, that facilitate radical transfer between α and β.[1, 5] The former is transiently oxidized during each turnover to Y_{356}• while the latter was initially proposed to be the proton acceptor for this reaction, but is now thought to be important in initiation of radical transfer during turnover. In all structures of class I RNR β_2 subunits reported to date, the last 30+ residues of the C-terminal tail are disordered.

Although the physical locations of Y_{356} and E_{350} in the $\alpha_2\beta_2$ interface are currently unknown, the ability to site-specifically incorporate unnatural amino acids that can trap pathway radicals has enabled spectroscopic distance measurements to be made that put some constraints on the possible location of Y_{356}.[1] These measurements are physically possible due to the half-of-sites reactivity of RNR: the radical of one α/β pair completes forward PCET, nucleotide reduction, and reverse PCET back to Y_{122} before the second pair is permitted to react. The coupling of the electronic spin of Y_{122}• in the unreacted α/β pair with radicals trapped on pathway residues in the reacted pair has allowed diagonal distance measurements to be made using pulsed electron-electron double resonance (PELDOR) spectroscopy.[1] These diagonal distances are shown in **Figure A1.2**.

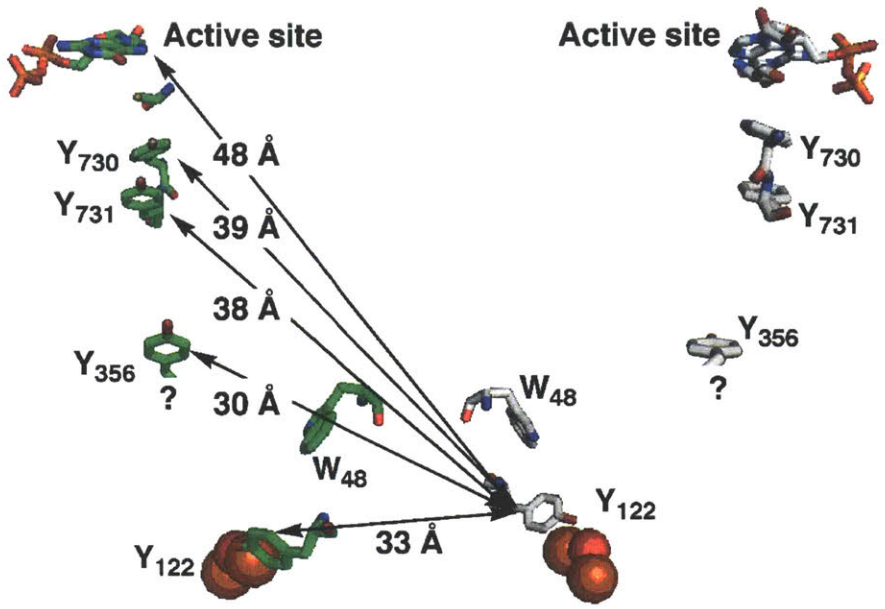


Figure A1.2. Diagonal distances measured by PELDOR spectroscopy in wildtype RNRs and mutants in which radical traps were installed on pathway. The residues of one α/β pair are colored green and those in the other pair gray. Reproduced with permission from Yokoyama, K.; Smith, A.A.; Corzilius, B.; Griffin, R.G.; Stubbe, J. (**2011**) Equilibration of Tyrosyl Radicals (Y_{356}•, Y_{731}•, Y_{730}•) in the Radical Propagation Pathway of the *Escherichia coli* Class Ia Ribonucleotide Reductase. *J. Am. Chem. Soc.* **133**: 18420-18432. Copyright 2011 American Chemical Society.

In addition to the distance measurements, the enormous body of literature on class I RNRs, replete with spectroscopic, kinetic, and structural data, is full of clues that can help establish new structural links or support existing ones if an effort is made to put together a working model of the $\alpha_2\beta_2$ interface in lieu of an actual structure. The structural observations reported here are the preliminary results of such an effort. Specifically, an attempt is made to predict the location of Y_{356} in the $\alpha_2\beta_2$ complex in order to identify residues that may alter its electrostatic environment and possibly serve as the proton acceptor for its transient oxidation during turnover. A second objective was to establish the probable position and function of E_{350} at the interface since studies from former Stubbe group members Ellen Minnihan[6] and Kanchana Ravichandran[7] indicate that this residue is not acting as the proton acceptor for Y_{356} oxidation. Finally, unpublished work by Jimin Shao's group at Zhejiang University School of Medicine implicates a hitherto

uncharacterized residue, E_{52}, as important in RNR catalytic activity, and an attempt is made to establish a structural basis for this essentiality. The structural observations and manipulations described in the following discussions were made using the program Chimera, available from the University of California – San Francisco (cgl.ucsf.edu).[8]

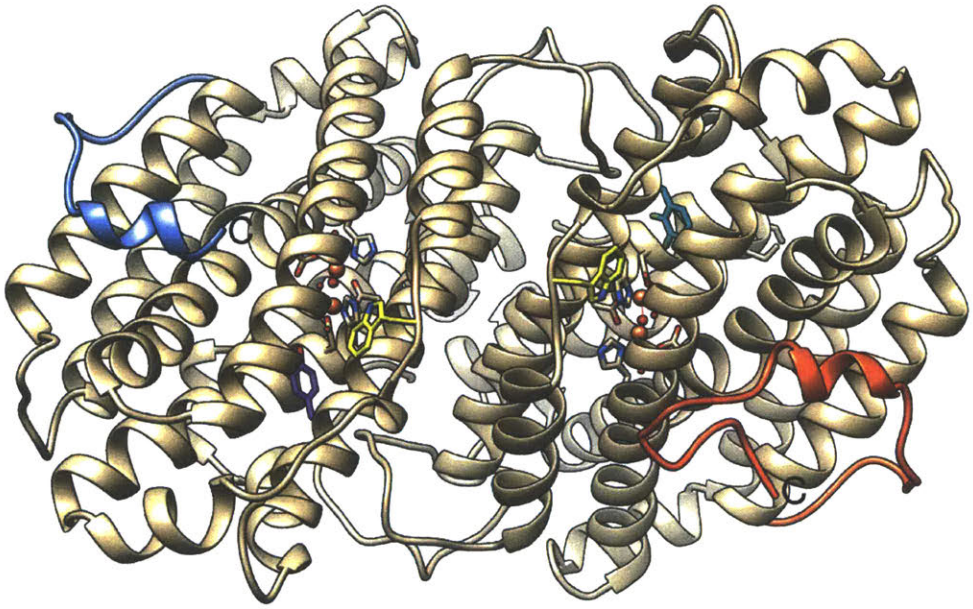


Figure A1.3. Comparison of the C-terminal tails of wildtype (left protomer) and F_2Y_{122} (right protomer) *E. coli* β. This view shows the surface of β_2 that interfaces with α_2 in the $\alpha_2\beta_2$ complex. The C-terminal tails of each monomer are colored blue (wildtype) and red (F_2Y_{122}), the ends of which are marked with the letter "C." Landmark residues are shown as sticks and colored as follows: W_{48} = yellow, Y_{122} = purple, F_2Y_{122} = cyan.

A1.2. General considerations. Much of this work was prompted by the recent release of crystal structures of the *E. coli* β_2 subunit with unnatural fluorotyrosine analogs (F_nY, n = 2 – 3) incorporated at position 122, which have revealed more residues of the C-terminal tail of this subunit than previously observed.[9] **Figure A1.3** presents a β_2 model composed of one wildtype protomer (PDB 1MXR[10]) and one F_2Y_{122} protomer (PDB 5CI1[9]) to illustrate this comparison; other than the lengths of the C-terminal tails, the structures are nearly identical (0.59 Å rmsd over 334

atom pairs). The ordered region of the wildtype structure ends at L_{339}, whereas in the F_2Y_{122} monomer, ordered residues extend out to Q_{349}. Assuming the position of residues 340 – 349 in the F_2Y_{122} structure are reflective of that in the $\alpha_2\beta_2$ complex and combining this new data with the structural information known previously about residues 360 – 375, which were used as a peptide in the crystallization of α_2,[3, 11] only ten residues in the C-terminal tail of β remain structurally unaccounted for. It has been noted, however, that these newly visible residues have high B factors (100 – 150 Å) relative to the rest of the structure (30 – 70 Å) and, thus, their position is likely going to change when in the $\alpha_2\beta_2$ complex.[9]

One of these new structures was used to make a docking model of the *E. coli* $\alpha_2\beta_2$ complex (**Figure A1.4**) from which the subsequent structural observations and discussions are made. As shown, each protomer in the docking model has been assigned a number that will be used in the following discussions. The model consists of the α_2 structure with GDP bound at the active site and TTP at the effector site (PDB 4R1R[11]) and the β_2 structure with 2,3-difluorotyrosine incorporated at position 122 (F_2-Y_{122} β_2, PDB 5CI1[9]). Comparisons will also be made using a recently released structure of an $\alpha_4\beta_4$ complex in which GDP, TTP, and dATP are bound to α_2 at the catalytic, specificity, and activity sites, respectively (PDB 5CNV[12]). This structure is also unique in that large domain movements in α have occurred that result in the active site clamping down on GDP and, thus, reflects α_2 in a primed state that is likely similar to that expected in the active $\alpha_2\beta_2$ complex.[12] In the recent study by Zimanyi *et al.*,[12] the clamped-down α structure was referred to as the "substrate-bound" state and the relaxed structures the "substrate-free" state. This terminology will be adopted in the subsequent discussions.

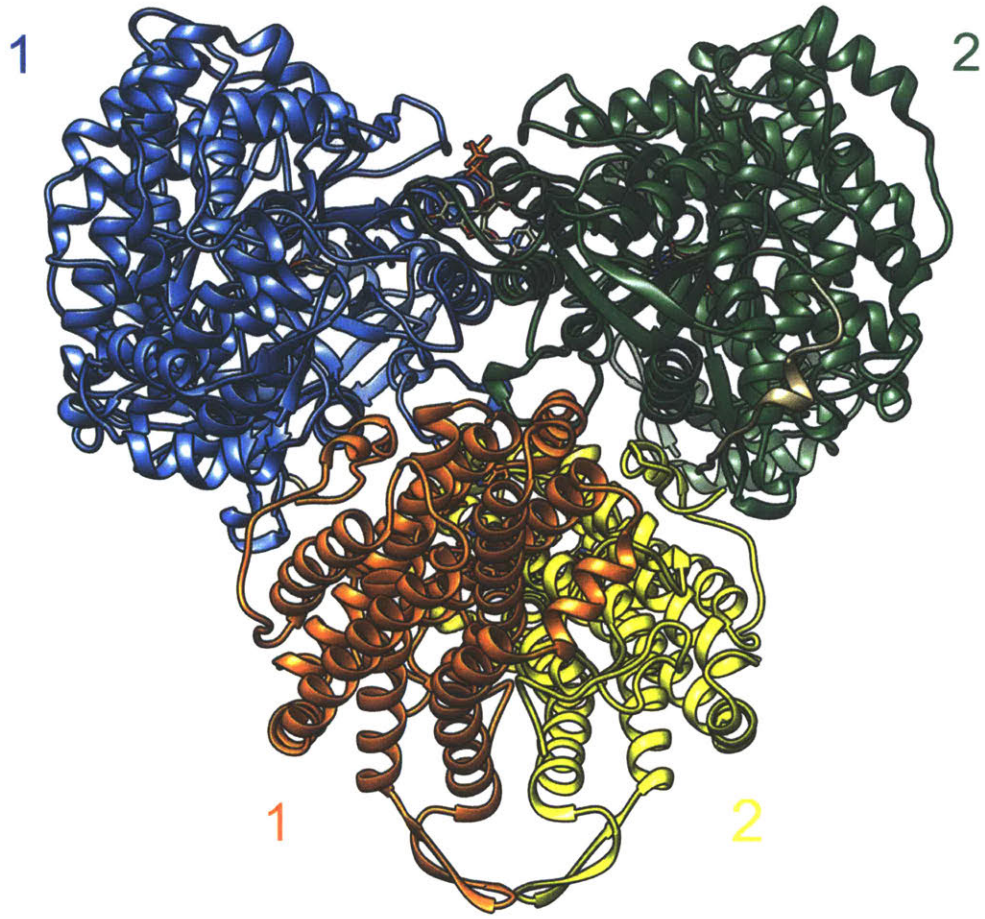


Figure A1.4. A docking model of the *E. coli* class Ia RNR prepared from PDB 4R1R[11] (α_2) and two copies of PDB 5CI1[9] (F_2Y_{122}-β). α is colored green and blue, and β orange and yellow. Each protomer has been assigned a number as indicated to facilitate the discussions in the main text. The peptides corresponding to the C-terminal tail of β used in the crystallization of α_2 are colored tan. The specificity and catalytic sites are occupied with TTP and GDP, respectively, which are shown as stick models and colored as follows: tan = C, red = O, blue = N, orange = P.

Lastly, many of the residues discussed in this section would be classified as non-conserved using the original sequencing profiles of the class Ia RNR α and β subunits available in the RNR database.[13] Recently, the database has undergone refinement and revision and, as a result, the Ia RNRs have been sub-categorized into groups using Hidden Markov Model profiling to find and elucidate different amino acid consensus sequences among the class Ia enzymes (http://rnrdb.pfitmap.org/). This sub-categorization establishes more sequence conservation between inherently similar RNRs and, thus, provides a starting point for investigating the subtleties

of activity, regulation, and structure of enzymes from the different Ia RNR groups. Based on these profiles, the subunits of *E. coli* class Ia RNR fall into group G (NrdAg and NrdBg), which is largely composed of γ-proteobacteria, but also includes some β-proteobacteria, cyanobacteria, and viruses. The sequence alignments for the subunits of the group G α and β subunits are presented at the end of this appendix.

A1.3. E_{350} may play a structural role in the conformational gating of radical transfer. Despite not directly participating in the radical transfer process, E_{350} is still essential for RNR because mutation of this residue abolishes RNR activity.[14] If it is assumed that the position of residues 340 – 349 observed in the recently released F_nY_{122}-β structures[9] reflects their actual position in the $\alpha_2\beta_2$ complex, E_{350} will be far removed (~10 Å) from the radical transfer pathway and, therefore, have a distinctly different role in the function of the enzyme. Previous studies by Climent, Sjöberg, and Haung indicated that E_{350} is likely not playing a major structural role in the stabilization of the $\alpha_2\beta_2$ complex as the K_d for the subunit interaction between α_2 and E_{350}A-β_2 exhibited only a 2.5-fold decrease in strength relative to wildtype (0.5 μM versus 0.2 μM, respectively).[14] One of the last remaining parts of the radical transfer mechanism (**Figure A1.1**) that has yet to be understood is the protein conformational change that gates radical injection onto the pathway. Therefore, given that E_{350} appears to be critical for catalysis by RNR, but does not function as a proton acceptor or as a structural component in the formation of the $\alpha_2\beta_2$ complex, a role for the residue in the conformational gating of radical transfer is plausible.

To begin exploring the idea that E_{350} participates in the conformational gating of radical transfer, it was initially assumed that the positions of $L_{340} - Q_{349}$ observed in the F_nY_{122}-β structures was reflective of that in the $\alpha_2\beta_2$ complex. E_{350} was built into the β #1 protomer of the docking model (**Figure A1.4**) and then the region surrounding it was examined for potential hydrogen

bonding and/or salt bridge partners. Only residues with high conservation within the group G Ia RNR sequences were considered. The results of this analysis are shown in **Figure A1.5**. R_{315}, Q_{345}, and Q_{349} in β and R_{639} and R_{735} in α were all reasonably close to E_{350} spatially. Both Gln residues and R_{639} were conserved in all sequences in the group G alignments whereas R_{315} and R_{735} were conserved in all but a few of the viral sequences.

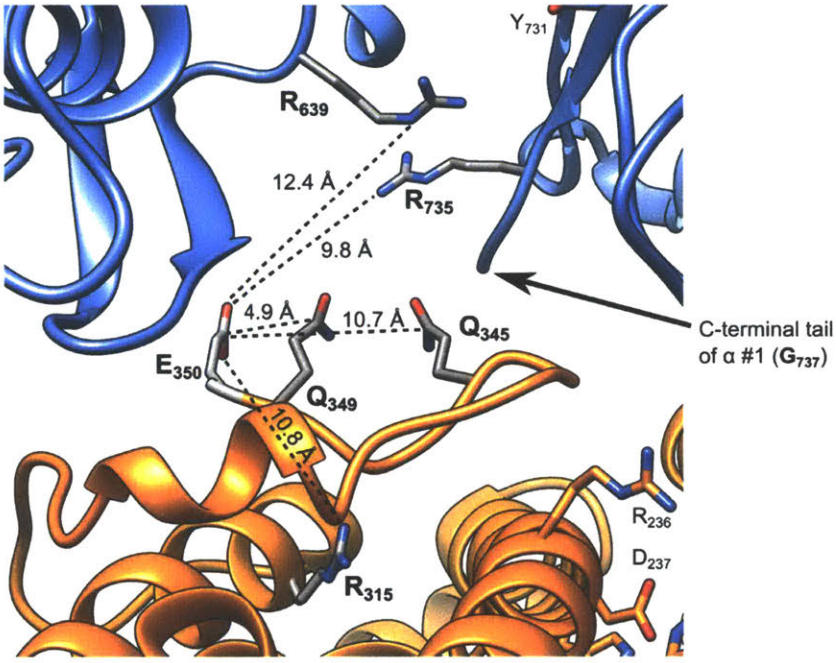


Figure A1.5. Potential conserved hydrogen bonding and salt bridge partners for E_{350} in the β subunit of the *E. coli* class Ia RNR. Potential partners are labeled in boldfaced text and the distances between them and E_{350} are indicated. The sidechains of E_{350}, Q_{345}, and Q_{349} are not visible in the source PDB file for F_2Y_{122}-β_2 (5CI1[9]) and are built in manually for visualization purposes only.

Out of the five, only for R_{735} did a role in the conformational gating of radical transfer seem obvious due to its location on the same strand as Y_{730} and Y_{731}. A salt bridge interaction between E_{350} in β and R_{735} in α could orient and stabilize the position of the strand relative to Y_{356} such that Y_{731} and Y_{730} are optimally organized for propagation of the radical between α and β. It also could be possible that E_{350} interacts with R_{735} to position the latter in the vicinity of a pathway Tyr residue

and, thus, alter the electrostatic environment. As such, it was recommended that studies of α with mutations made at position 735 be undertaken to investigate a possible E_{350}/R_{735} interaction. The $R_{735}Q$ mutant exhibited a specific activity of 1100 nmol min^{-1} mg^{-1}, ~60% of that measured for wildtype protein (Albert Kim, unpublished results). Although this is a fairly significant loss in activity, it was expected that if R_{735} did interact with E_{350}, then removal of the former would completely abolish RNR activity given the results of analogous studies with E_{350} mutants. Therefore, R_{735} is not essential for *E. coli* RNR catalytic activity.

A salt bridge interaction between R_{639} in α and E_{350} in β was considered next, although it was unclear how such an interaction could lead to injection of the radical onto the pathway. Interestingly, the $R_{639}Q$ mutant of α was catalytically inactive and had a weaker interaction with β_2 (K_d = ~4 μM versus 0.2 μM for wildtype, Kim, Qinghui Lin, unpublished results), therefore indicating that R_{639} is essential for RNR function. To determine if these observations could possibly be the result of a disrupted E_{350}/R_{639} interaction, a comparison of the substrate-free (PDB 4R1R[11]) and substrate-bound (PDB 5CNV[12]) α structures was made to examine the conformational changes that R_{639} undergoes between the two state. As mentioned previously, the α tertiary structure has clamped down on the substrate nucleotide in the substrate-bound state and likely reflects the protein in the primed $\alpha_2\beta_2$ complex, thus the conformation of R_{639} in this structure should provide clues as to its function during active turnover. The comparison revealed that the sidechain of R_{639} rotates about 5° counterclockwise about the Cβ – Cα bond and 47° counter-clockwise about the Cδ – Cγ bond in going from the substrate-free to -bound state (**Figure A1.6A**). The result of this conformational change is that the guanidinium group of R_{639} forms hydrogen bonding interactions (3.2 – 3.3 Å) with the backbone carbonyl of I_{644} (**Figure A1.6A**). I_{644} is located in a flexible loop composed of residues 641 – 656 that forms a β-hairpin secondary

structure adjacent to the N-terminal ATP-cone domain (**Figure A1.6B**). Compared to the substrate-free state of α_2,[11] this loop undergoes a 7.7 Å movement (Cα – Cα distance for K_{648}) in towards the center of the TIM barrel upon substrate binding.[12] The hydrogen bonds between R_{639} and the backbone carbonyl of I_{644} appears to be one of the main interactions holding the β-hairpin in this closed position, perhaps explaining why the R_{639} mutants were inactive and exhibited weaker affinity for β_2. If E_{350} in β plays any sort of stabilizing role by hydrogen bonding to R_{639} in α in the closed state, this interaction is not critical, therefore an interaction between these two residues is unlikely.

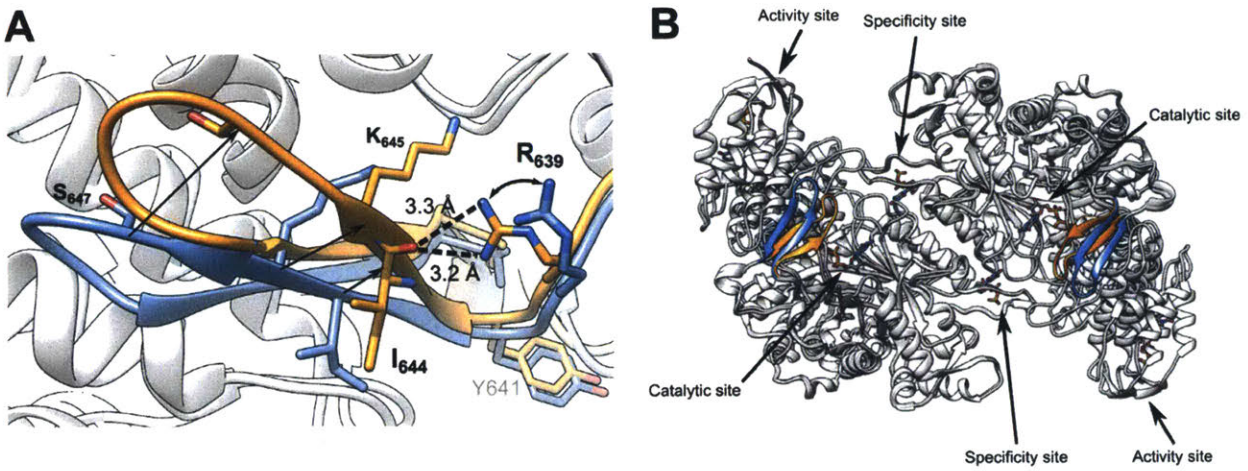


Figure A1.6. (**A**) R_{639} rotates from its position in the substrate-free state (blue) and stabilizes the β-hairpin loop conformation in the substrate-bound state (orange) by forming hydrogen bonds with the backbone carbonyl of I_{644}. Also shown are K_{645} and S_{647}, two residues conserved in the Group G Ia RNRs that undergo large movements between the substrate-free and substrate-bound states (indicated by arrows). The Val residue in the background behind R_{639} is V_{656}. (**B**) Location of residues 641 – 655 in the substrate-free (blue, PDB 4R1R[11]) and substrate-bound (orange, PDB 5CNV[12]) states of *E. coli* α_2. The catalytic, specificity, and activity sites are occupied by GDP, TTP, and dATP, respectively.

The remaining three residues in the vicinity of E_{350} in β (R_{315}, Q_{345}, and Q_{349}) have yet to be explored. Both Gln residues were only recently observed in wildtype and F_nY β structures, but their side chains are disordered and the high B factors of the remaining atoms indicates that they

are very flexible in the absence of α.[9] The mutation of either residue to Cys followed by labeling with a benzophenone (BP) crosslinking agent resulted in nearly complete elimination of RNR activity (Q_{345}C-BP = 0.1% as active as wildtype, Q_{349}C-BP = 0.07% as active as wildtype).[15] The corresponding control reaction with the unlabeled Cys mutants was not run; thus, the essentiality of these residues to RNR activity or structure are currently unknown. In the case of R_{315}, mutation of the equivalent residue in the human RNR β subunit to alanine resulted in only a 10% drop in activity,[16] thus indicating the residue is likely not critical for RNR activity.

All of these results support the conclusion that the newly observed residues in the C-terminal tail of the F_nY β structures are not located in positions reflective of the situation in the $\alpha_2\beta_2$ complex. The location and function of E_{350} remains obscure, but the results from Ravichandran[7] and Climent[14] point towards a structural role for this residue in the conformational gating of radical injection onto the pathway. One possibility is presented at the end of the next section. However, the structural predictions reported here and the subsequent experimentation testing these predictions have helped in beginning to piece together a picture of what may be going on in the interface during turnover of the *E. coli* Ia RNR. The essentiality of R_{639} for activity by likely acting as a buttress for the closed conformation of a β-hairpin loop (**Figure A1.6**) suggests that components of this loop may be essential for complex formation and/or catalytic activity. Indeed, there are two conserved residues, K_{645} and S_{647} (**Figure A1.6A**), positioned at the end of the loop that undergoes the largest movement. Further studies of these residues and the other observations reported here will help further advance our understanding of this complex enzyme.

A1.4. Mapping plausible locations of Y_{356} in the $\alpha_2\beta_2$ complex structural model. Tyrosine 356 in the β_2 subunit of the *E. coli* class Ia RNR (**Figure A1.1**) is the last major redox-active residue of the radical transfer pathway that has not been observed structurally. Currently there are

two constraints for the location of Y_{356} in the $\alpha_2\beta_2$ complex. *Constraint 1:* the residue is located in the α/β interface region based on (*i*) theoretical considerations (Y_{356}• would be quickly quenched if exposed to solvent), and (*ii*) the positions, in the docking model,[3] of the last ordered C-terminal residue of β_2 (V_{340}[17] or Q_{349} in recently released structures[9]), the N-terminal residue (Q_{360}) of the peptides bound to α_2 that mimic the C-terminal tail of β,[3, 11] and the other radical transfer pathway participants (C_{439}, Y_{730}, and Y_{731} in α). *Constraint 2:* Radicals trapped at Y_{356} in one β protomer are 30 Å away from Y_{122}• in the opposite, unreacted protomer as measured by PELDOR spectroscopy (**Figure A1.2**).[18, 19] Logically it seemed reasonable that a rough idea of the possible location of Y_{356} could be obtained by first centering a sphere with a radius of 30 Å on Y_{122} in one β protomer of the $\alpha_2\beta_2$ docking model and then, using information from other sources, rationally reduce the total search space (i.e. the surface area of the sphere) to specific elements at the interface in the opposite α/β pair. An assumption made here is that the β protomers do not rotate relative to one another along the β_2 interface in the $\alpha_2\beta_2$ complex. This assumption seems valid given that the average distance between both Y_{122} residues in every *E. coli* β_2 crystal structure, including those in the $\alpha_4\beta_4$ inhibited complexes,[12, 20, 21] is 33.1 ± 0.8 Å, which is in excellent agreement with the PELDOR measured distances for β_2 alone and in complex with α_2.[19, 22]

One subtlety of the PELDOR measurements is that the distances are not atom-to-atom distances *per se*. Since the unpaired electron of a tyrosyl radical is delocalized over the phenol ring, with the majority of the spin density residing on the O and carbons 1, 3, and 5 (see **Figure A1.7A** for atom numbering scheme),[23] the measured distances between trapped pathway radicals and Y_{122}• in the opposite α/β pair is the distance between the average position of the spin density on both residues. Therefore, it was important to make distance measurements in the docking model from a location on tyrosine residues that best represented the probable "center" of the spin density.

For the modeling described here, the center of the spin density was taken as the intersection of a line drawn between the OH group and C1 and a line drawn between C3 and C5 (**Figure A1.7B**), similar to that described previously.[22] Because structure viewing programs like Chimera can only measure distances between two atoms with explicit x,y,z coordinates in the PDB file, dummy atoms were built *in silico* and were placed into the docking model at the desired location within the phenol ring for each tyrosine of interest (**Figure A1.7C**). This process for one tyrosine is illustrated in **Figure A1.7**. Using this definition of the average spin density, distances of 38.2 Å and 39.8 Å were measured between Y_{122} in β #2 and Y_{731} and Y_{730}, respectively, in the opposite α protomer (α #1) in the docking model. These values are in excellent agreement with the distances measured between these pairs of residues previously by PELDOR.[18, 19]

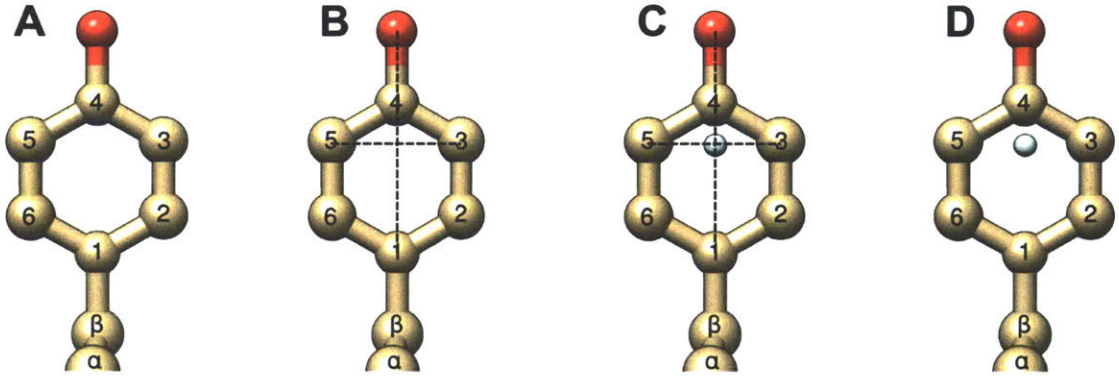


Figure A1.7. Illustration of defining in the "center" of spin density on a tyrosyl radical. (**A**) Number scheme for tyrosine. (**B**) A guide line is drawn between C3 and C5 and between C1 and the OH group. (**C**) A dummy atom is built into the atomic structure and placed at the intersection of the lines. (**D**) The final structure after removing the guide lines. This position of the center of spin density on Y• is similar to that described previously.[22]

The result of placing a sphere 30 Å in radius at the center of spin density on Y_{122} in β #2 in the docking model of *E. coli* Ia RNR is shown in **Figure A1.8**. By analogy to the PELDOR measurements, the α/β pair (α #2 and β #2) housing the Y_{122} residue from which the measurement is being made is considered to be unreacted, thus any of the sphere's surface area that falls within

these protomers can be eliminated from consideration immediately. In the opposite α/β pair (α #1 and β #1), because Y_{356} will be located within the interfacial region, any surface area falling within the tertiary structure of β #1 and outside of the active site in α #1 can also be disregarded. What remains is the spherical element shown in **Figure A1.9**. The corners of the element can very roughly be defined by Cα of E_{51} and Nε of H_{247} in β #1, Cα of R_{406} in α #1, and Nζ of K_{283} in α #2. The furthest reach of the element into the α/β interface is roughly defined by Cα and Cβ of R_{236} in all structures of β, and additionally by Cα of D_{342} and the amide nitrogen atoms of Q_{343} and V_{344} in the recent F_nY_{122} β structures.[9] Based on the sequence alignments of the group G class Ia RNRs, noteworthy highly conserved residues that fall on or close to the element include the stretch of six residues from $F_{47} - E_{52}$ (FWRPEE), R_{236}, and D_{237} in β #1; R_{411} in α #1; and R_{329} in α #2 (discussed later in this appendix). Many of these residues have previously been established or are predicted to be important in cluster assembly[24] (W_{48}) and/or in radical transfer between α and β (W_{48},[3] R_{236},[25] D_{237},[26] and R_{411}[27]).

Data collected by former Stubbe group members Kenichi Yokoyama (unpublished) and Kanchana Ravichandran[7] strongly indicate that W_{48} in β does not serve as a transiently oxidized intermediate on the radical transfer pathway as originally proposed.[3, 5] Therefore, radical propagation in the *E. coli* Ia RNR requires Y_{122}• to directly oxidize Y_{356}, which imposes additional constraints for the possible position of the latter in the $\alpha_2\beta_2$ complex based on the parameters that define the rate of electron transfer. The electron transfer rate (k_{ET}, s^{-1}) depends on the distance (R, in Å) between the donor and acceptor, the driving force (ΔG°, i.e. the relative reduction potential between donor and acceptor), and the reorganization energy (λ, i.e. the energy required to reorganize nuclei from the equilibrium positions of the reactants to the products) and can be defined by equation 1 for an endergonic reaction.[28]

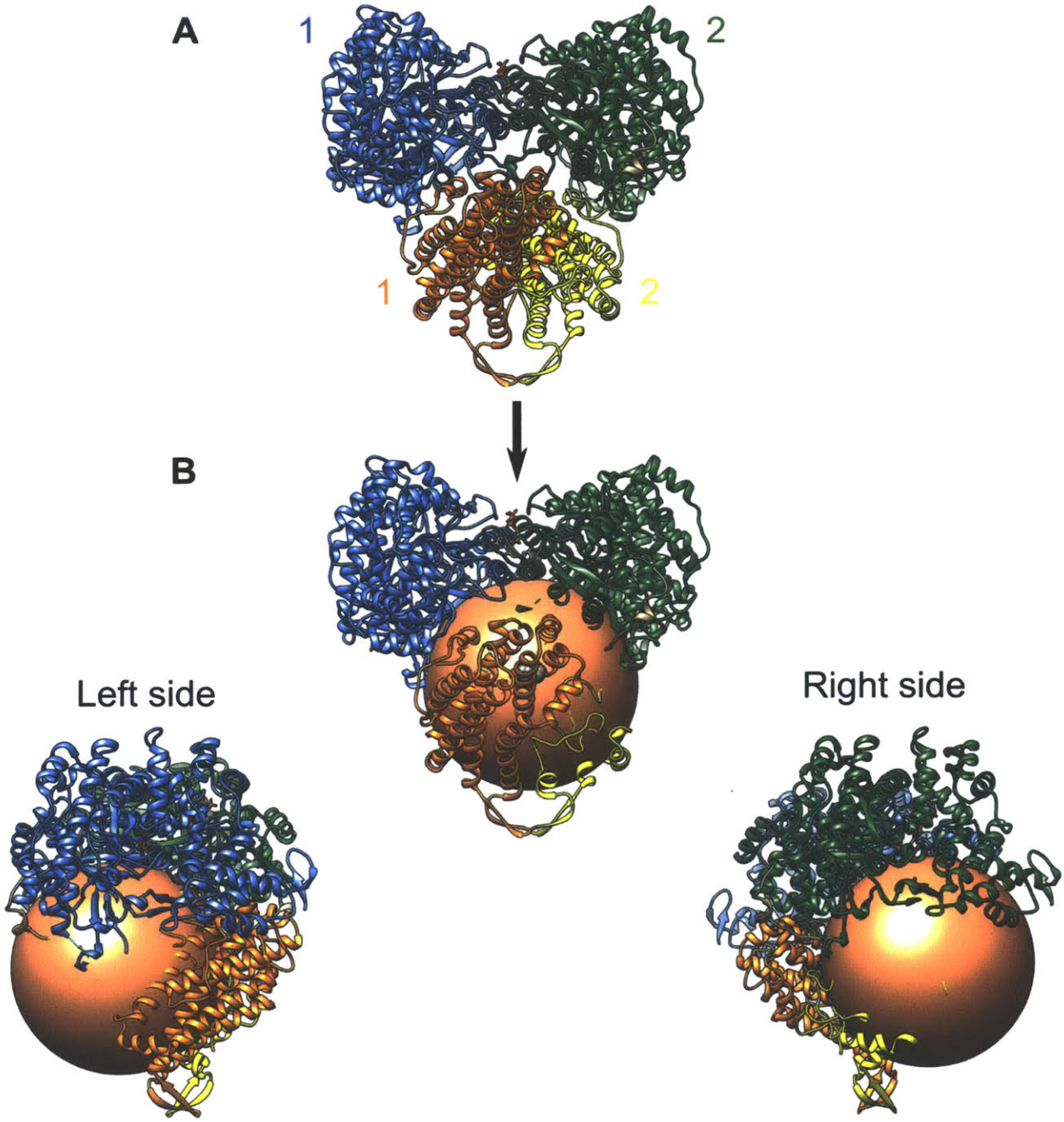


Figure A1.8. (**A**) Docking model of the *E. coli* class Ia RNR. Each α and β protomer has been colored differently and assigned a number as indicated. (**B**) The result of placing a sphere with a radius of 30 Å at the center of spin density on Y_{122} in β protomer #2. The right and left sides of the $\alpha_2\beta_2$ complex are shown to illustrate the extent of coverage of the sphere.

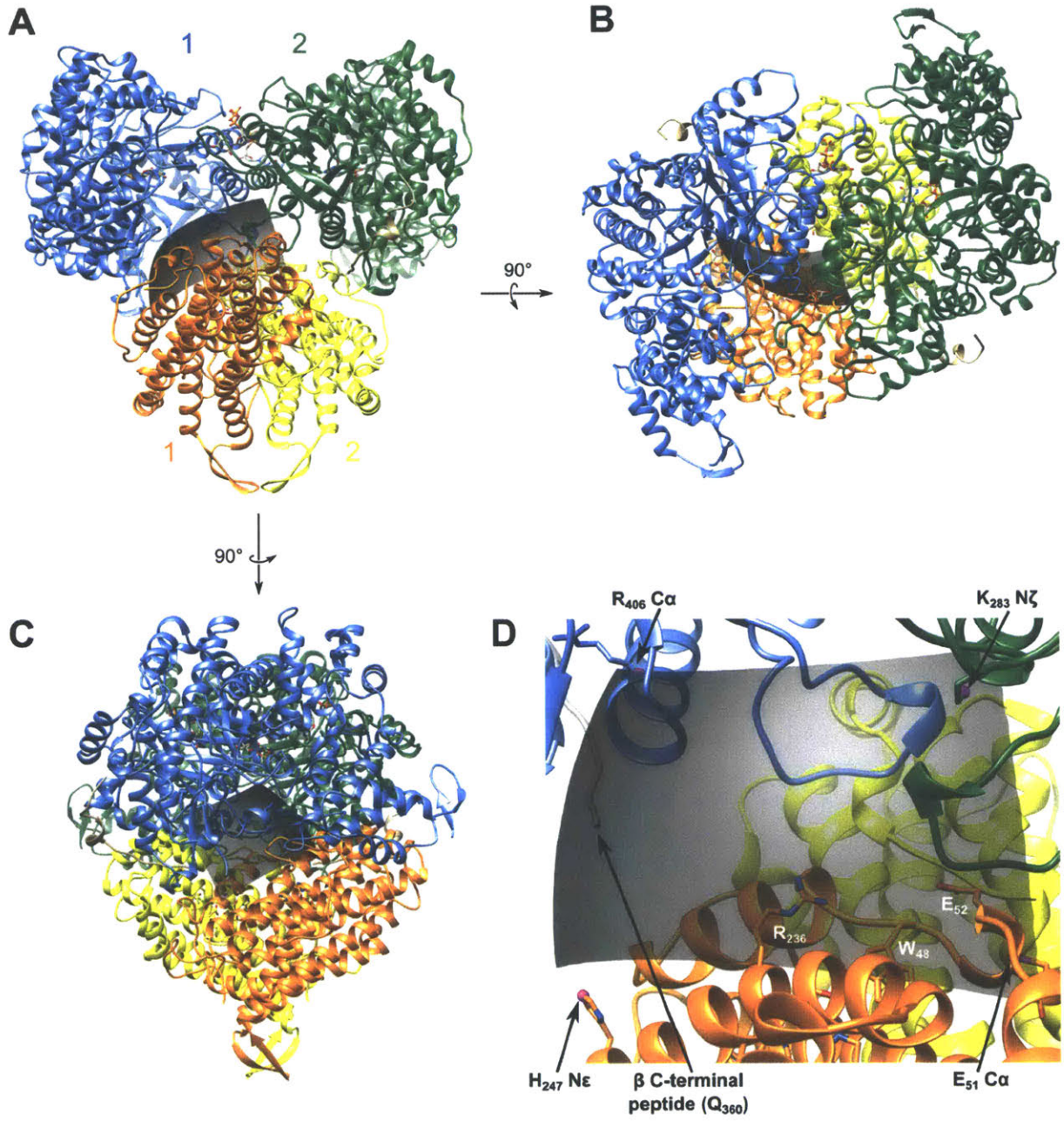


Figure A1.9. Reduction of the search space where Y_{356} could feasibly be located at the interface of the *E. coli* class Ia RNR $\alpha_2\beta_2$ complex. By constraining the plausible space where Y_{356} could be located at the interface between α #1 and β #1, the surface area of the 30 Å sphere displayed in **Figure A1.8B** was reduced to the spherical element (translucent gray) shown in the panels above. Three different views of the complex are presented to illustrate the size of the spherical element: (**A**) front, (**B**) top, and (**C**) left side. (**D**) Zoom in on the spherical element at the interfacial region between α #1 and β #1. Atoms roughly demarking the corners of the element are represented as balls colored magenta. All other atoms are colored according to element. The landmark residues W_{48}, E_{52}, and R_{236} are indicated. Residues 332 – 349 on the C-terminal tail of β have been omitted for clarity in (**D**).

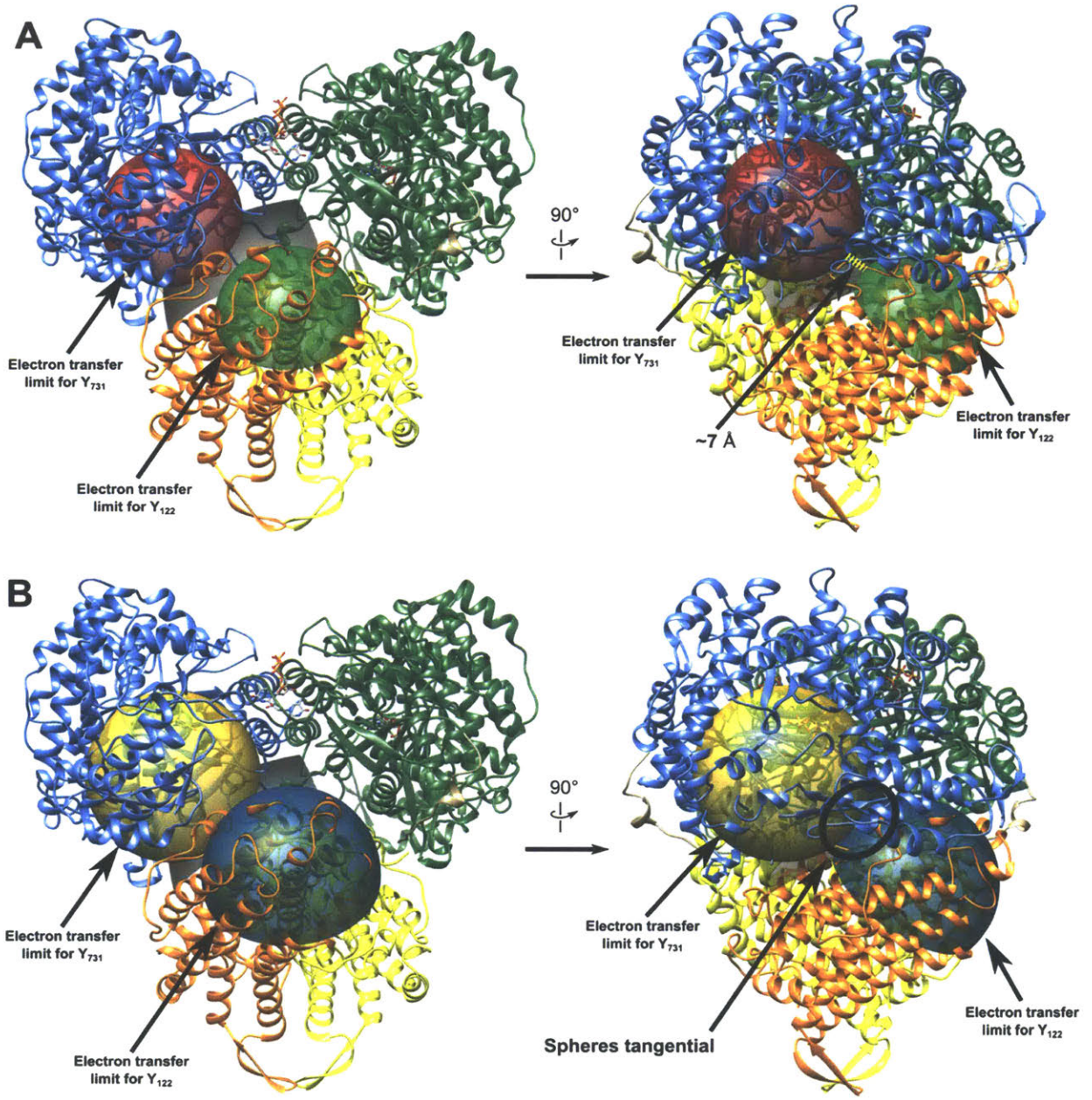


Figure A1.10. Predicting the position of Y_{356} in the *E. coli* Ia RNR $\alpha_2\beta_2$ complex from the distance dependence expected for radical transfer between $Y_{122} - Y_{356}$ and $Y_{356} - Y_{731}$ in a highly polar (**A**) or non-polar (**B**) protein environment. As described in the main text, the distance between each pair of Tyr residues is expected to be ~14 Å in a polar environment and 17.5 Å in a non-polar environment. These distances are indicated in the pictures shown above by spheres of the proper radius placed at the center of spin density for Y_{122} in β #1 ((**A**) = green, (**B**) = blue) and Y_{731} in α #1 ((**A**) = red, (**B**) = yellow) in the model presented in **Figure A1.9**. The point where the spheres on Y_{122} and Y_{371} overlap with one another and the spherical element described in **Figure A1.9** will indicate a plausible position for Y_{356} in the $\alpha_2\beta_2$ complex. In (**B**), such a point exists without requiring movement of the subunits relative to one another. In (**A**), 7 Å separates the spheres on Y_{122} and Y_{731}, requiring some subunit reorganization to bring the spheres into contact.

$$\log_{10}(k_{ET}) = 15 - 0.6R - \frac{3.1(-\Delta G^{\circ} + \lambda)^2}{\lambda} - \frac{\Delta G^{\circ}}{0.06} \quad (1)$$

Since the radical in RNR is propagated by PCET, proton transfer occurs concomitant with electron transfer and, thus, the proton acceptor should be close enough to the OH group of the transiently oxidized Try residues that movement of H^+ is not rate-limiting. If this is assumed to be true, eq 1 can yield an estimate of the distances between the pathway tyrosine residues provided reasonable estimates of k_{ET}, $\Delta G°$, and λ are provided. Studies by Yokoyama and Ravichandran have estimated that the oxidation of Y_{356} by Y_{122}• and Y_{730} (α) by Y_{356}• (β) are each approximately 0.1 eV uphill and, thus, provide an estimate of $\Delta G°$.[7, 19] The value for λ in RNR is unknown, thus the calculation will utilize estimates of 0.7 – 1.4 eV which reflect typical protein environments ranging from non-polar to polar, respectively.[28] Finally, previous work by Yokoyama suggests that k_{ET} in RNR occurs on the order of 300 s^{-1}.[19] Plugging these value into eq 1 and solving for R yields distances of 17.5 Å (λ = 0.7 eV) and 14.3 Å (λ = 1.4 eV). Therefore, Y_{356} should be located at the interface within this range of distances from both Y_{122} in β and Y_{731} in α.

The results of placing spheres with radii of 14 Å or 17.5 Å on the center of spin density for Y_{122} in β #1 and Y_{731} in α #1 are shown in **Figure A1.10**. In both cases, the spheres have significant overlap with the element defining the plausible space for the location of Y_{356} previously discussed in **Figure A1.9**. A further refinement for the position of Y_{356} can be defined by points of overlap between the two spheres on Y_{122} and Y_{731} and the spherical element. If it is assumed that the radical transfer pathway is largely non-polar in nature, then a point exists in the docking model where all three overlap and, thus requires no reorganization of the subunits from their respective positions. On the other hand, if the pathway is largely polar, then a 7 Å gap exists between the closest approach of the 14 Å spheres on Y_{122} and Y_{731} (**Figure A1.10A**). An overlap of all three elements in this case can be achieved if the subunits approach each other more closely and/or by a

slight rotation of β_2 such that the lobe of β #1 tips further into the active site of α #1 (**Figure A1.11**).

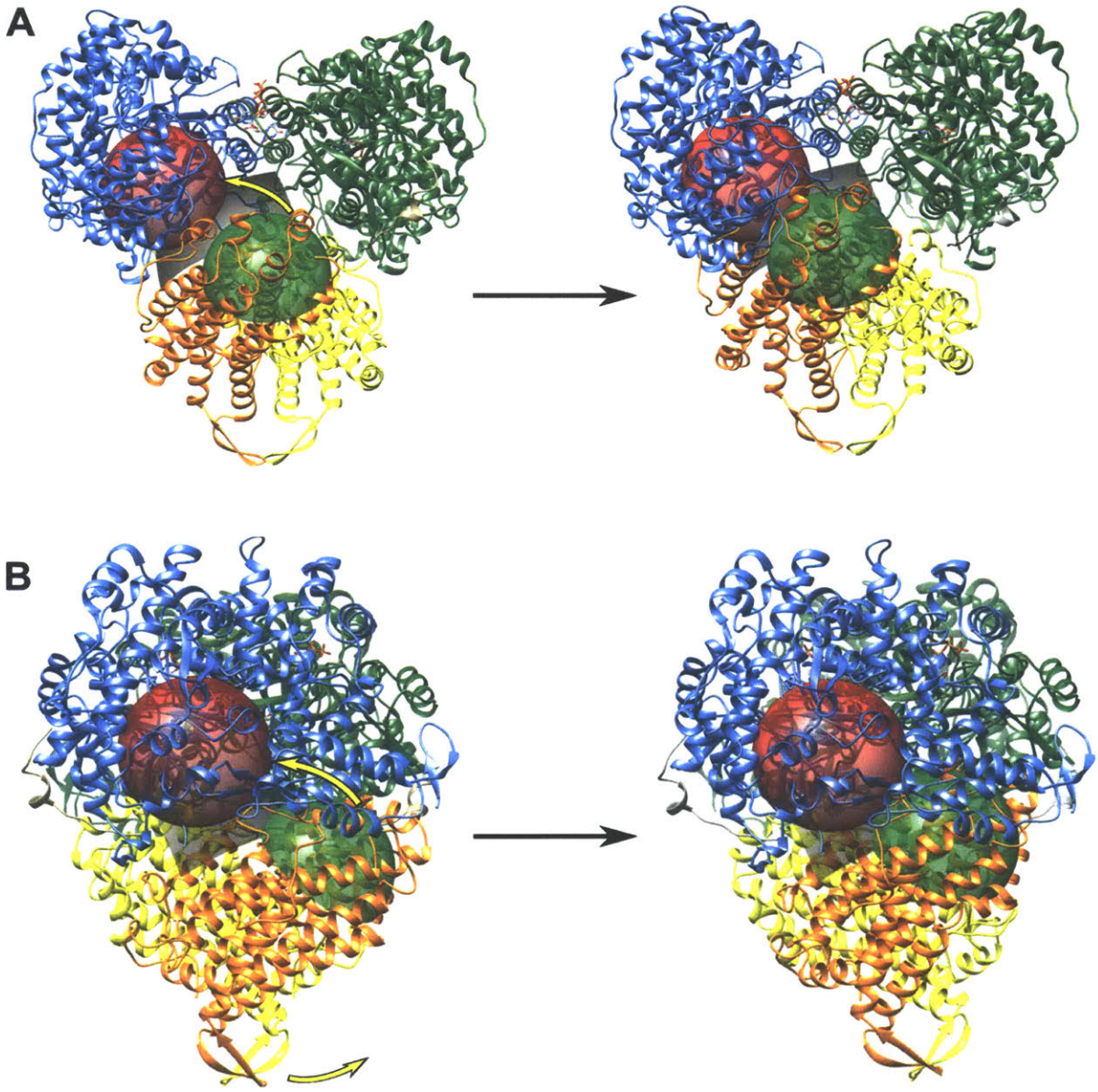


Figure A1.11. Reorganization of β_2 required to make the 14 Å spheres on Y_{122} in β #1 and Y_{731} in α #1 tangential to one another in the *E. coli* Ia RNR $\alpha_2\beta_2$ complex. This model assumes that the radical transfer pathway is polar in character. Yellow arrows indicate the direction of rotation of β required to produce the desired result. Two views of the complex are presented: (**A**) front and (**B**) left side.

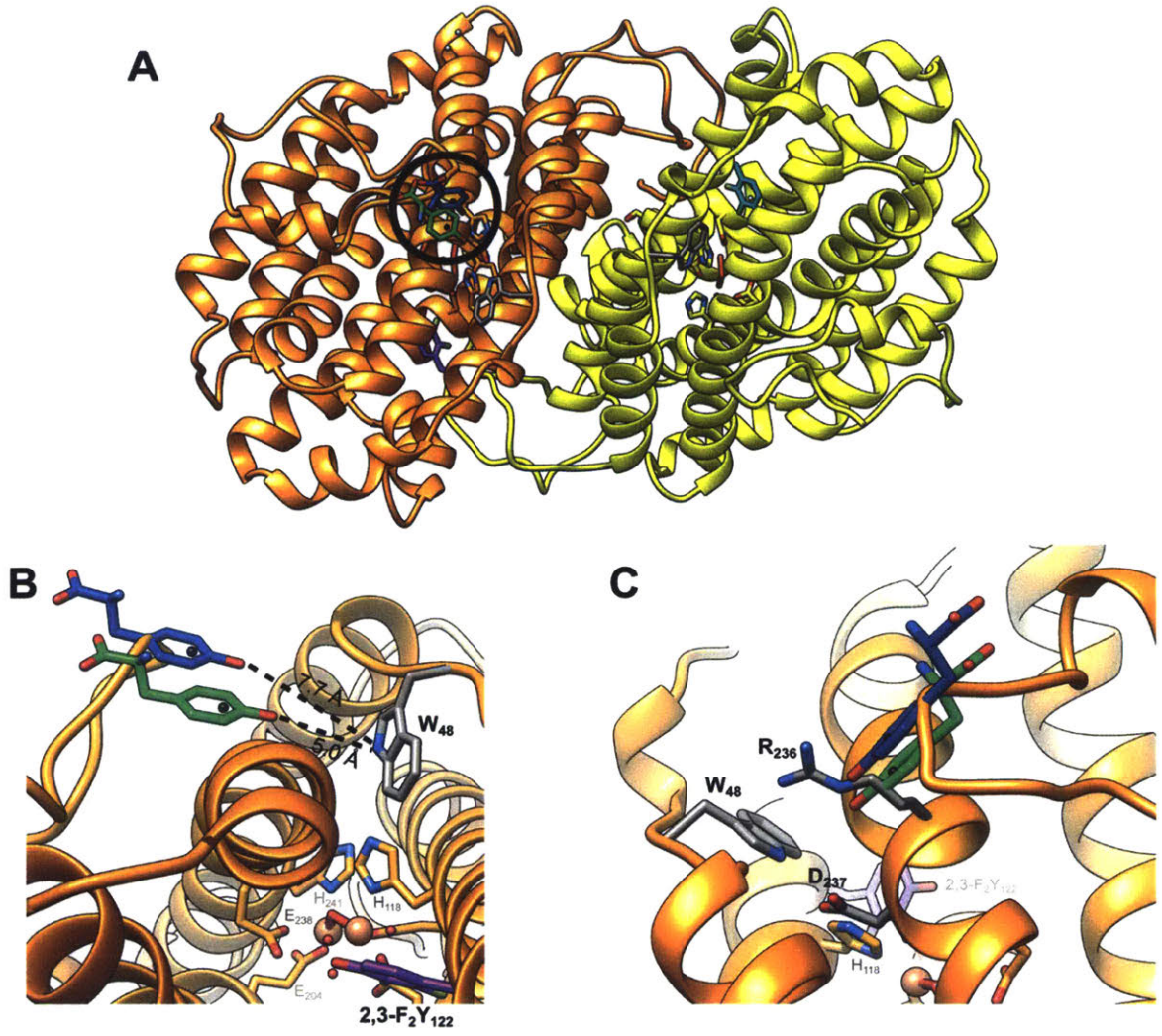


Figure A1.12. Plausible location for Y_{356} in the *E. coli* Ia RNR $\alpha_2\beta_2$ complex. For clarity, α_2 has been omitted from the figures. Two Tyr residues are built into the model to represent where Y_{356} would be located assuming a non-polar radical pathway (blue) or a polar pathway (green). Note that the conformations of each Tyr residue shown is not assumed to be the actual conformation in the $\alpha_2\beta_2$ complex. The black dots in the center of the modeled Tyr indicates the center of spin density that is 30 Å away from Y_{122} in β #2 and the proper distance from Y_{122} in β #1 and Y_{731} in α #1 (non-polar = 17.5 Å, polar = 14 Å). (**A**) Overview of β_2 looking down on the surface that interacts with α_2. The position of Y_{356} is indicated with a circle. Landmark residues are shown and colored as follows: W_{48} = gray, 2,3-F_2Y_{122} = purple (β #1) and cyan (β #2). (**B**) Position of Y_{356} relative to Y_{122} and the $Fe(III)_2$ cluster. (**C**) Position of Y_{356} relative to R_{236} and D_{237}, which are previously implicated to be important in the radical transfer mechanism out of β.[25, 26] Both residues overlap/sterically clash with R_{236} and the C-terminal tail of β (S_{341} – V_{344}), whereas they are ~3 – 5 Å above the position of D_{237}.

Regardless of whether the radical transfer pathway is polar or non-polar, the predicted position for Y_{356} based on the constraints applied is quite similar (**Figure A1.12**). The residue sits right on top of R_{236}, which has a similar sidechain conformation in all β_2 structures currently available, and sterically clashes with the newly visible portion (specifically between $S_{341} - V_{344}$) of the C-terminal tail observed in the F_nY_{122}-β structures.[9] At first glance, it would seem like there would be too many steric clashes for Y_{356} to occupy this space. However, as mentioned earlier, these newly visible residues ($S_{341} - Q_{349}$) are likely not in a position that reflects their actual location in the $\alpha_2\beta_2$ complex and, therefore, clashes with these residues will likely not be an issue in the active enzyme.[9] On the other hand, R_{236} still present a problem, thus it is likely that the true location of Y_{356} will be somewhere in the vicinity of R_{236} rather than right on top of it.

Interestingly, the predicted position of Y_{356} coincides with several pieces of evidence that implicate the W_{48}-R_{236}-D_{237} triad as a critical component of the radical transfer pathway out of β.[25, 26] R_{236} and D_{237} equivalents are absolutely conserved in the group G Ia RNRs and highly conserved in many other class I enzymes. Mutation of D_{237} to Glu preserves RNR activity whereas the Asn mutant is inactive, but in both cases cluster assembly and the interactions with α_2 were minimally perturbed.[26] For R_{236}, mutants of the analogous residue in the mouse Ia RNR were inactive except for the Glu substitution, which showed 40% of wildtype activity.[25] Further examination of the mouse β_2 mutants revealed that these proteins also retained the ability to assemble the $Fe(III)_2$-Y• cluster and had an unaltered binding affinity to α_2.[25] These results indicated that both residues likely played a catalytic rather than structural role. As mentioned above, current data indicates W_{48} is not transiently oxidized during radical transfer out of β (Yokoyama, unpublished),[7] but the fact that it forms a hydrogen bond network with R_{236} and D_{237} suggests that it may still play a role in the radical transfer mechanism. Most importantly, the triad

is connected to the diferric cluster via a hydrogen bond between D_{237} and H_{118}, one of the ligands for the Fe atom closest to Y_{122}.

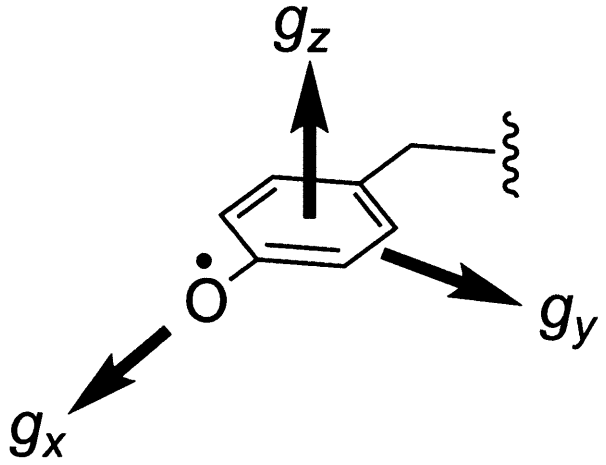


Figure A1.13. Principal components of the EPR g-tensor for tyrosyl radicals. The g_x and g_y vectors are parallel to the aromatic plane of the ring while g_z is orthogonal to it.

Evidence also exists for a possible interaction of Y_{356} with a positively charged residue or metal ion in the $\alpha_2\beta_2$ complex. The data comes from a pulsed high frequency EPR spectroscopic study of mutants of α and β in which the radical trap 3-aminotyrosine (NH_2Y) replaced Y_{356} in β and Y_{731} in α on the radical transfer pathway.[29] The use of high frequency microwaves allows for the resolution of the g-tensor of the radical into its principal components, g_x, g_y, and g_z (**Figure A1.13**). Pertinent to this discussion is the g_x component, which runs along the C – O bond of the phenol and can report on hydrogen bonding to and/or the electrostatic environment around this group. For Y_{356}, the value of g_x was found to be significantly perturbed from that expected for a free NH_2Y•, and the magnitude of the perturbation suggested the presence of multiple hydrogen bonding partners and/or positive charges in the vicinity of the residue. The use of deuterium electron-nuclear double resonance (ENDOR) spectroscopy, however, revealed no exchangeable deuterons that were strongly coupled with the radical, indicating that NH_2Y_{356}• lacked any

moderately strong hydrogen bonding interactions with other residues at the α/β interface. The study concluded that the electrostatic environment around Y_{356} needed to be altered by positive charges and/or water molecules/networks in order to account for the perturbed g_x value.[29] Given the predicted position for Y_{356} reported here, a contributor to this altered electrostatic environment could be R_{236}.

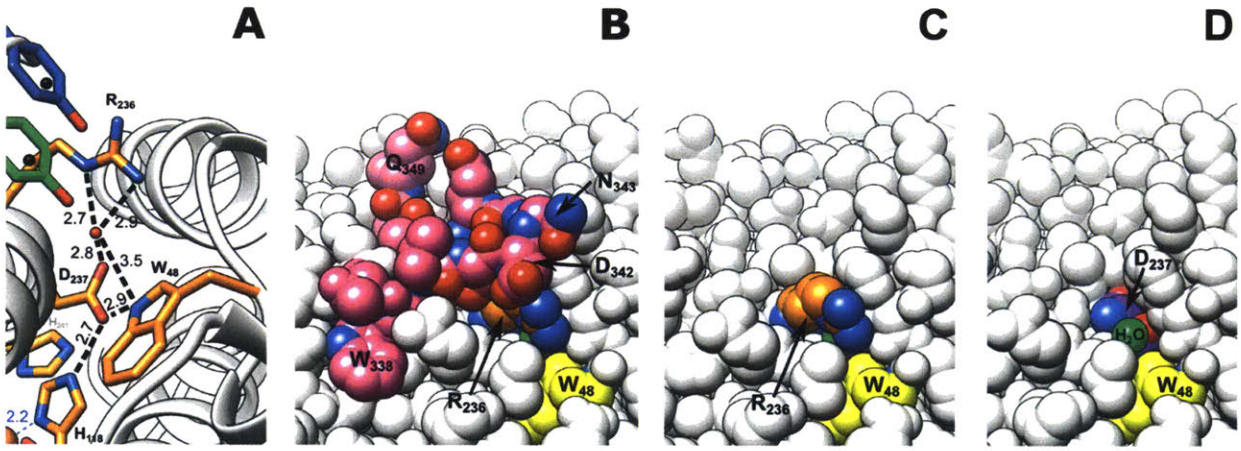


Figure A1.14. A possible proton acceptor for Y_{356} during radical propagation in the *E. coli* Ia RNR. (**A**) A water molecule hydrogen bonded to R_{236}, D_{237}, and possibly W_{48} is observed in nearly all *E. coli* β_2 structures as well as those of other class I RNR β subunits. This network is in contact with the diferric cluster via the interaction between D_{327} and H_{118}. Dashed lines indicate hydrogen bonds and numbers correspond to the interatomic distance in Å. The positions of Y_{356} expected in polar (green) or non-polar environments (blue) are also indicated, and their OH groups are, respectively, 1.7 Å and 4.3 Å away from the water molecule. (**B-D**) A possible mechanism for the conformational gate of radical transfer between Y_{122} and Y_{356}. Shown are space filling models with the following color schemes for carbon: β C-terminal tail (W_{338} – Q_{349}) – pink, R_{236} – orange, W_{48} – yellow, D_{237} – purple, water molecule from (**A**) – green. The Y_{356} models have been removed for clarity. (**B**) The C-terminal tail in the F_nY_{122} β structures is noted to have high B-factors and, therefore, is likely going to move in the $\alpha_2\beta_2$ complex.[9] (**C**) Removal of the C-terminal tail exposes R_{236}, which appears to be capping the hydrogen bond network to prevent solvent exposure and, therefore, is acting as a gate to the network. (**D**) In the $\alpha_2\beta_2$, the sidechain of R_{236} is moved by an as yet unidentified mechanism, thus exposing the water molecule to act as the proton acceptor for Y_{356}.

Finally, as mentioned earlier, extensive studies by Minnihan and Ravichandran have demonstrated that E_{350}, another highly conserved residue in the structurally uncharacterized part

of the C-terminal tail of β, does not appear to be acting as the proton acceptor for the transient oxidation of Y_{356} during RNR turnover as originally proposed.[5] In relation to probable alternate proton acceptors, further work by Ravichandran indicated that in $\alpha_2\beta_2$ complexes composed of F_3-Y_{122}•-β and $Y_{730/731}$F-α mutants, the radical could equilibrate between Y_{122} and Y_{356}.[7] This equilibration was found to be pH-dependent, suggesting that the proton lost from Y_{356} upon oxidation ended up in a solvent exchangeable position or was taken up immediately by a water molecule that then rapidly exchanges with bulk solvent.[7] If the latter were true, then water molecules, clusters, or networks that crystallize in nearly the same spot in many or all available β_2 structures could potentially indicate feasible proton acceptors for Y_{356} oxidation.

A particularly intriguing possibility for the proton acceptor of Y_{356} could be a water molecule located the middle of the W_{48}-R_{236}-D_{237} triad (**Figure A1.14A**). It has been noted previously[25] that this space in nearly every crystal structure of a class I β subunit is occupied with a water molecule that is hydrogen bonded to R_{236}, D_{237} and sometimes W_{48}. Comparison of β_2 structures also reveals that the sidechain conformations of the triad are nearly identical. This observation, in conjunction with the link between the triad and the diferric cluster via hydrogen bonding between D_{237} and H_{118}, indicates that this network of hydrogen bonds is important for some aspect of RNR function. In the F_nY_{122}-β structures, the water is buried in β beneath both R_{236} and the newly visible portion of the C-terminal tail (**Figure A1.14B**), indicating significant conformational changes need to occur if the water molecule does indeed act as the proton acceptor for Y_{356}. As noted earlier, the C-terminal tail of β is expected to be in a different conformation in the $\alpha_2\beta_2$ complex and, thus, will expose R_{236} to the interfacial space between α and β (**Figure A1.14C**). Uncovering the water molecule requires a conformational change in R_{236} (**Figure A1.14D**), which could possibly be induced by steric repulsion if Y_{356} is forced into the location

predicted form the modeling reported here. Alternatively, other residues of the C-terminal tail of β currently invisible in crystal structures could cause the conformational change by forming hydrogen bonding contacts with R_{236} that lift the sidechain off of the water molecule, thus exposing the latter to Y_{356}. One candidate is E_{350}, which as mentioned in the previous section, is suspected to play a role in lifting the conformational gate for radical injection onto the pathway. Therefore, future experiments designed to test this possibility could be very informative if a connection between E_{350} and R_{236} does indeed exist.

A1.5. Predicting the function of E_{52} in RNR activity. Another residue recently implicated to be important in the overall function of *E. coli* RNR is E_{52} in β. Unpublished results from the Shao group on the human RNR indicated that the equivalent residue was essential for catalysis. Mutating E_{52} in *E. coli* β also abolished NDP reduction activity, and a few of the mutants exhibited a 10-fold weakening of the interaction between α_2 and β_2 (Lin, unpublished results), suggesting that, at least in the case of *E. coli*, E_{52} played somewhat of a dual role in structure and catalysis. E_{52} is located at the C-terminal end of a stretch of conserved residues (47-FWRPEE-52) in Ia RNR group G β subunits (**Figure A1.15A**) that sits adjacent to the β_2 dimer interface and the plausible position of Y_{356} described in the previous section. Relative to neighboring regions of β, the secondary structure of these six residues, especially R_{49} – E_{52}, appears to be more flexible, with the last four exhibiting B-factors on the order of 10 – 20 units higher than F_{47} and W_{48} in nearly all β_2 crystal structures. This flexibility ultimately results in these residues adopting an unusual random coil structure that seemingly interrupts the first α-helix of β. A comparison of all available crystal structures of *E. coli* β_2 reveals that E_{52} adopts a range of different conformations (**Figure A1.15B**) that may play a role in RNR function.

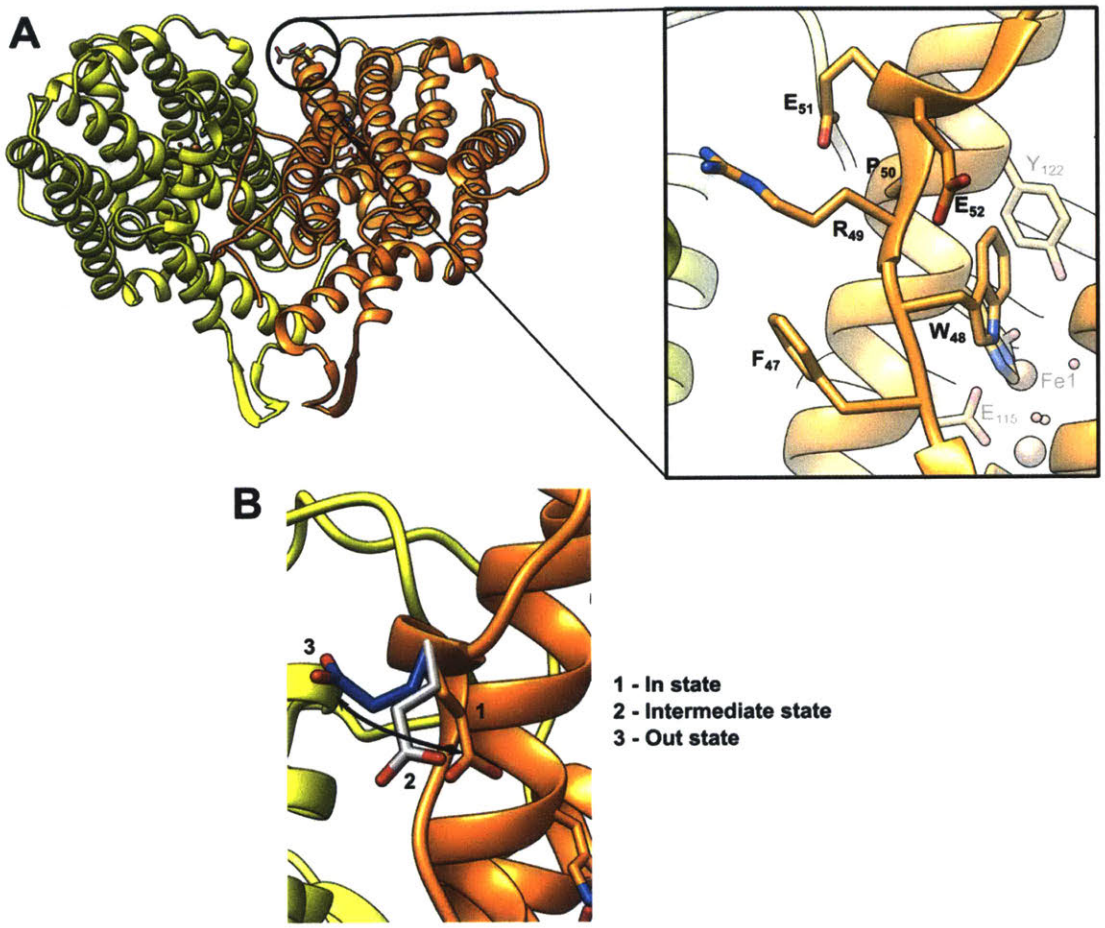


Figure A1.15. (**A**) Location of E_{52} in the crystal structure (PDB 1MXR[11]) of *E. coli* β_2. The blown-up picture shows the position of F_{47} through E_{52} relative to the $Fe(III)_2$ cluster and Y_{122} (shown in the background). These residues are conserved in the Group G class Ia RNRs and, therefore, are likely important for RNR structure and/or function. (**B**) A comparison of all available structures of β_2 reveals E_{52} adopts a range of conformations. Shown in the figure are the two extreme (orange and blue) and one intermediate conformer (gray) to illustrate the range of motion of this residue. The extreme conformers have been termed the "in" state and "out" state as indicated.

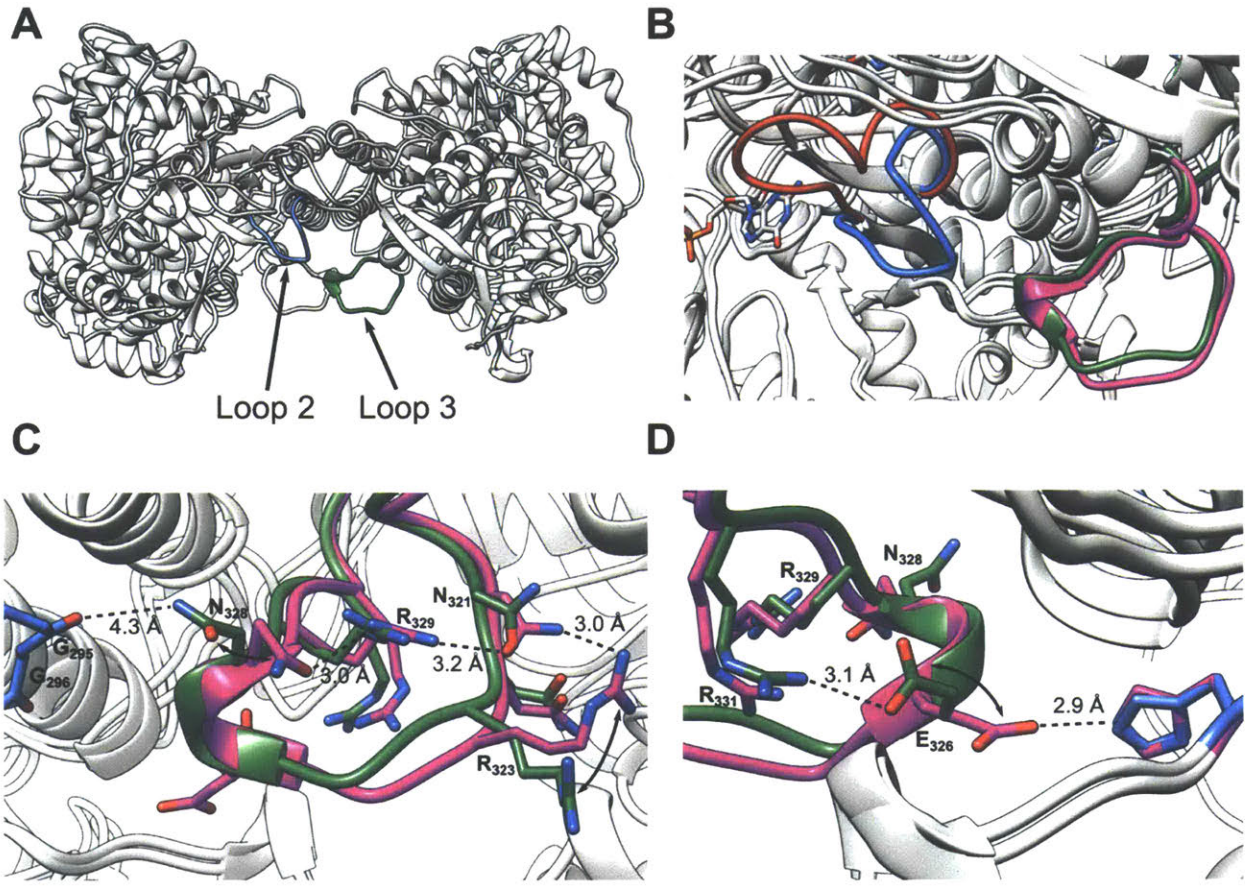


Figure A1.16. (**A**) Structural location of loop 2 and loop 3 in α_2. (**B**) Conformations of loop 2 and loop 3 in the substrate-free state (blue and dark green) and substrate-bound state (red and magenta). Binding of substrate and effector causes a drastic change in the conformation of loop 2 while loop 3 remains relatively unaffected. (**C**) Amino acid sidechain conformational changes between the substrate-free (green) and substrate-bound (magenta) states in loop 3. The atoms of G_{295} and G_{296} in the substrate-free state are shown on the left side of the panel (blue). The conformational change of loop 2 causes N_{328}, which may form a hydrogen bond to the backbone carbonyl of G_{295} in loop 2 of the opposite α protomer, to swing in towards the center of the loop to form a hydrogen bonding network with R_{329}, N_{321}, and R_{323}. (**D**) Opposite side of loop 3 from panel (**C**) showing the sidechain changes of E_{326} and R_{331}. E_{326} swings ~90° to form a hydrogen bond with H_{332} in the opposite α protomer in the substrate-bound state.

The location of E_{52} on β also appears to be a very likely a point of contact between the subunits in the $\alpha_2\beta_2$ complex. In the docking model, a loop of amino acids ($K_{320} - V_{330}$), termed loop 3,[11] from α #2 (**Figure A1.16A**) hangs down right next to E_{52} and the conserved stretch of amino acids in β #1. It has been suggested previously that the position of loop 3 is influenced by the position of two other loops that connect the specificity site to the catalytic site.[11] The most pertinent of these is loop 2 (**Figure A1.16A**), which is the primary determinant for substrate specificity of the enzyme for a particular NDP. In the substrate-free state, loop 3 from α #2 appears to be in van der Waals contact with loop 2 of α #1 (**Figure A1.16B**) with a possible hydrogen bond, given the resolution of the structure (3.2 Å), forming between the sidechain of N_{328} and the backbone carbonyl of G_{295}. Both loops exhibit the highest B-factors in these structures and are, therefore, less well-defined as compared to the surrounding regions of the protein.[11] Upon binding of substrate and effector, loop 2 flips up and makes contact with the bases of the nucleotides occupying the specificity and active sites (**Figure A1.16B**),[12] whereas loop 3 remains in a similar flexible conformation as observed in the substrate-free state. A more detailed comparison of the changes in loop 3 that occur between states reveals the reorientation of several sidechains which results in formation of new hydrogen bonding interactions within the loop as well as with α #1 (**Figure A1.16C** and **D**).

In addition to the location of critical residues at the $\alpha_2\beta_2$ interface and the mechanism of conformational gating of radical transfer as discussed above, another outstanding issue with regards to RNR activity is the mechanism of how α bound with substrate and effector nucleotide transmits a structural signal to β that indicates that the system is ready for turnover. If the sidechain reorientations observed in loop 3 really occur in solution when loop 2 flips up in the substrate bound state, then these conformational changes could be the start of this signaling process. Given

the proximity of E_{52} to loop 3, it seemed logical that the side chain reorientations in both could allow loop 3 to interact with E_{52} in β, thereby transmitting the signal into this subunit to produce the conformational change that lifts the gate and allows radical injection onto pathway. Therefore, the docking model was examined to find potential hydrogen bonding and/or salt bridge partners in loop 3 of α for E_{52} in β. Again, only residues with high conservation in the group G Ia RNRs were considered. The results of this search are shown in **Figure A1.17**. Of the four residues identified, N_{328} and R_{329} were conserved in all but a few viral sequences, in which Lys substitutions for either residue were observed. An Asp or Glu residue was found at position 326 in all but two viral sequences, but is likely not going to interact with E_{52} unless a metal ion is involved. Such an interaction would additionally require loop 3 to undergo a large conformational change. The substitutions of R_{323} were more varied than the other three, but a majority (96%) of the sequences had a positively charged residue as this position and, hence, R_{323} was included in the group of probable interactors with E_{52}.

Mutants of α with substitutions at positions 323 and 329 exhibited substantial losses of activity and, at least in the case of R_{329}, a weakening of the subunit interaction (**Table A1.1**, Lin, unpublished results). The other two residues have not been examined. These results indicate loop 3 is important for RNR function, but whether or not this is related to an inability to interact with E_{52} or other residues in β remains unclear at this point. R_{329} appears to be a major structural component in maintaining the conformation of loop 3 as Nε and both Nη atoms of its guanidinium side chain are hydrogen bonding with backbone carbonyls from N_{321}, R_{323}, and N_{328}, respectively, in both the substrate-free and substrate-bound states (**Figure A1.16C**). Therefore, a loss of tertiary structure of loop 3 may be responsible for the inactive phenotype rather than a missing interaction between R_{329} and E_{52} in β or a disruption of the hydrogen bond network that forms between N_{328},

R_{329}, N_{321}, and R_{323} (**Figure A1.16C**). Further studies are required to clarify this point and whether or not an interaction between loop 3 and E_{52} is important for transmitting the "go" signal for radical transfer in the *E. coli* class Ia RNR.

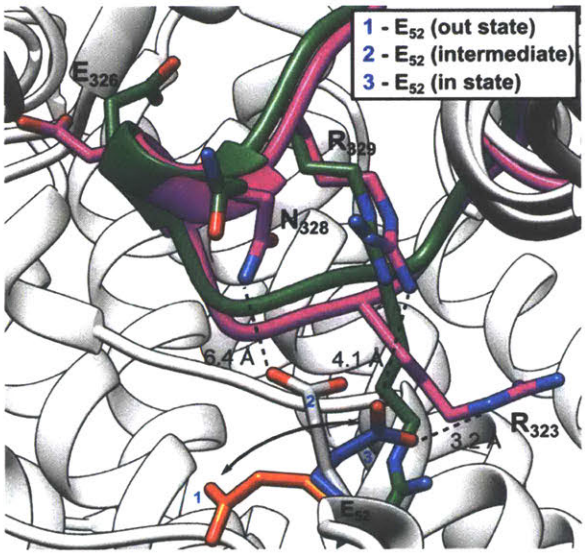


Figure A1.17. Possible hydrogen bonding/salt bridge partners in loop 3 of α for E_{52} in β of the *E. coli* class Ia RNR. The substrate-free (green) and substrate-bound (magenta) states of α have been overlaid to show sidechain reorientations that could place residues within the vicinity of E_{52} in β (distances, in Å, are indicated). The alternative conformations of E_{52} observed in β structures are also illustrated: orange = in state, gray = intermediate, blue = out state.

Table A1.1. Specific activities and K_ds for R_{323} and R_{329} mutants of *E. coli* α.[a]

Protein	K_d (μM)	SA (nmol min^{-1} mg^{-1})	% of WT activity
WT	0.18[b]	2400 ± 40	100
R_{323}K	ND	837	34
R_{329}A	4.7 ± 0.5	0	0
R_{329}E	10.6 ± 1.4	0	0
R_{329}K	2.6	2.8	0.12
R_{329}Q	2.8	0	0

[a] Qinghui Lin, Jimin Shao, JoAnne Stubbe, unpublished results.
[b] Previously reported by Climent and Sjöberg.[4]

A1.6. Conclusions. Structural observations and predictions made from a docking model of the *E. coli* class Ia RNR have led to a prediction of the location of Y_{356}, the last structurally uncharacterized redox active residue in the radical transfer pathway, in the $\alpha_2\beta_2$ complex and the identification of two hitherto uncharacterized Arg residues that are critical in the function of the enzyme. The results of these theoretical observations and subsequent experiments are beginning to reveal the intricacies of how this enzyme tightly controls radical reactivity so that it can harness the power of a free electron to complete a difficult chemical reaction. The results are also beginning to paint a picture of the structural and functional organization of the α/β interface. Continued efforts coupling experimental data with *in silico* structural interrogation will help further unmask the mysteries of this complicated enzyme and, it is hoped, lead to an opportunity to finally obtain a complete structure of the *E. coli* class Ia RNR $\alpha_2\beta_2$ complex.

A1.7. REFERENCES

1. Minnihan, E. C., Nocera, D. G., and Stubbe, J. (2013) Reversible, Long-Range Radical Transfer in *E. coli* Class la Ribonucleotide Reductase, *Accounts Chem. Res. 46*, 2524-2535.
2. Minnihan, E. C., Ando, N., Brignole, E. J., Olshansky, L., Chittuluru, J., Asturias, F. J., Drennan, C. L., Nocera, D. G., and Stubbe, J. (2013) Generation of a Stable, Aminotyrosyl Radical-Induced $\alpha_2\beta_2$ Complex of *Escherichia coli* Class Ia Ribonucleotide Reductase, *Proc. Natl. Acad. Sci. U. S. A. 110*, 3835-3840.
3. Uhlin, U., and Eklund, H. (1994) Structure of Ribonucleotide Reductase Protein R1, *Nature 370*, 533-539.
4. Climent, I., Sjöberg, B. M., and Huang, C. Y. (1991) Carboxyl-Terminal Peptides as Probes for *Escherichia coli* Ribonucleotide Reductase Subunit Interaction - Kinetic Analysis of Inhibition Studies, *Biochemistry 30*, 5164-5171.
5. Reece, S. Y., Hodgkiss, J. M., Stubbe, J., and Nocera, D. G. (2006) Proton-Coupled Electron Transfer: The Mechanistic Underpinning for Radical Transport and Catalysis In Biology, *Philos. Trans. R. Soc. B-Biol. Sci. 361*, 1351-1364.
6. Minnihan, E. C., (2012) *Mechanistic Studies of Proton-Coupled Electron Transfer in Aminotyrosine- and Fluorotyrosine-Substituted Class Ia Ribonucleotide Reductase.* Ph.D Thesis, Massachusetts Insitute of Technology.

7. Ravichandran, K., (2016) *Mechanistic Investigations of the Radical Transport Pathway in Fluorotyrosine-substituted Class Ia Ribonucleotide Reductase*. Ph.D. Thesis, Massachusets Institute of Technology.

8. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004) UCSF Chimera - A Visualization System for Exploratory Research and Analysis, *J. Comput. Chem. 25*, 1605-1612.

9. Oyala, P. H., Ravichandran, K., Funk, M. A., Stucky, P. A., Stich, T. A., Drennan, C. L., Britt, R. D., and Stubbe, J. (2016) Biophysical Characterization of Fluorotyrosine Probes Site-specifically Incorporated into Enzymes: *E. coli* Ribonucleotide Reductase as an Example, *J. Am. Chem. Soc. 138*, 7951-7964.

10. Högbom, M., Galander, M., Andersson, M., Kolberg, M., Hofbauer, W., Lassmann, G., Nordlund, P., and Lendzian, F. (2003) Displacement of the Tyrosyl Radical Cofactor in Ribonucleotide Reductase Obtained by Single-Crystal High-Field EPR and 1.4-Angstrom X-Ray Data, *Proc. Natl. Acad. Sci. U. S. A. 100*, 3209-3214.

11. Eriksson, M., Uhlin, U., Ramaswamy, S., Ekberg, M., Regnström, K., Sjöberg, B. M., and Eklund, H. (1997) Binding of Allosteric Effectors to Ribonucleotide Reductase Protein R1: Reduction of Active-Site Cysteines Promotes Substrate Binding, *Structure 5*, 1077-1092.

12. Zimanyi, C. M., Chen, P. Y. T., Kang, G., Funk, M. A., and Drennan, C. L. (2016) Molecular Basis for Allosteric Specificity Regulation in Class Ia Ribonucleotide Reductase from *Escherichia coli*, *eLife 5*, 23.

13. Lundin, D., Torrents, E., Poole, A. M., and Sjöberg, B. M. (2009) RNRdb, a Curated Database of the Universal Enzyme Family Ribonucleotide Reductase, Reveals a High Level of Misannotation in Sequences Deposited to Genbank, *BMC Genomics 10*, 8.

14. Climent, I., Sjöberg, B. M., and Huang, C. Y. (1992) Site-Directed Mutagenesis and Deletion of the Carboxyl Terminus of *Escherichia coli* Ribonucleotide Reductase Protein R2. Effects on Catalytic Activity and Subunit Interaction, *Biochemistry 31*, 4801-4807.

15. Hassan, A. Q., Wang, Y. T., Plate, L., and Stubbe, J. (2008) Methodology To Probe Subunit Interactions in Ribonucleotide Reductases, *Biochemistry 47*, 13046-13055.

16. Zhou, B. S., Su, L. L., Hu, S. Y., Hu, W. D., Yip, M. L. R., Wu, J., Gaur, S., Smith, D. L., Yuan, Y. C., Synold, T. W., Horne, D., and Yen, Y. (2013) A Small-Molecule Blocking Ribonucleotide Reductase Holoenzyme Formation Inhibits Cancer Cell Growth and Overcomes Drug Resistance, *Cancer Res. 73*, 6484-6493.

17. Nordlund, P., Sjöberg, B. M., and Eklund, H. (1990) Three-Dimensional Structure of the Free Radical Protein of Ribonucleotide Reductase, *Nature 345*, 593-598.

18. Seyedsayamdost, M. R., Chan, C. T. Y., Mugnaini, V., Stubbe, J., and Bennati, M. (2007) PELDOR Spectroscopy with DOPA-β2 and NH_2Y-α2s: Distance Measurements between Residues Involved in the Radical Propagation Pathway of *E. coli* Ribonucleotide Reductase, *J. Am. Chem. Soc. 129*, 15748-15749.

19. Yokoyama, K., Smith, A. A., Corzilius, B., Griffin, R. G., and Stubbe, J. (2011) Equilibration of Tyrosyl Radicals (Y_{356}•, Y_{731}•, Y_{730}•) in the Radical Propagation Pathway of the *Escherichia coli* Class la Ribonucleotide Reductase, *J. Am. Chem. Soc. 133*, 18420-18432.

20. Ando, N., Brignole, E. J., Zimanyi, C. M., Funk, M. A., Yokoyama, K., Asturias, F. J., Stubbe, J., and Drennan, C. L. (2011) Structural Interconversions Modulate Activity of *Escherichia coli* Ribonucleotide Reductase, *Proc. Natl. Acad. Sci. U. S. A. 108*, 21046-21051.

21. Zimanyi, C. M., Ando, N., Brignole, E. J., Asturias, F. J., Stubbe, J., and Drennan, C. L. (2012) Tangled Up in Knots: Structures of Inactivated Forms of *E. coli* Class Ia Ribonucleotide Reductase, *Structure 20*, 1374-1383.

22. Bennati, M., Weber, A., Antonic, J., Perlstein, D. L., Robblee, J., and Stubbe, J. A. (2003) Pulsed ELDOR Spectroscopy Measures the Distance Between the Two Tyrosyl Radicals in the R2 Subunit of the *E. coli* Ribonucleotide Reductase, *J. Am. Chem. Soc. 125*, 14988-14989.

23. Hoganson, C. W., Sahlin, M., Sjöberg, B. M., and Babcock, G. T. (1996) Electron Magnetic Resonance of the Tyrosyl Radical in Ribonucleotide Reductase From *Escherichia coli*, *J. Am. Chem. Soc. 118*, 4672-4679.

24. Bollinger, J. M., Jr., Tong, W. H., Ravi, N., Huynh, B. H., Edmondson, D. E., and Stubbe, J. (1994) Mechanism of Assembly of the Tyrosyl Radical-Diiron(III) Cofactor of *E. coli* Ribonucleotide Reductase. 3. Kinetics of the Limiting Fe^{2+} Reaction by Optical, EPR, and Mössbauer Spectroscopies, *J. Am. Chem. Soc. 116*, 8024-8032.

25. Nárvaez, A. J., Voevodskaya, N., Thelander, L., and Gräslund, A. (2006) The Involvement of Arg^{265} of Mouse Ribonucleotide Reductase R2 Protein in Proton Transfer and Catalysis, *J. Biol. Chem. 281*, 26022-26028.

26. Ekberg, M., Pötsch, S., Sandin, E., Thunnissen, M., Nordlund, P., Sahlin, M., and Sjöberg, B. M. (1998) Preserved Catalytic Activity in an Engineered Ribonucleotide Reductase R2 Protein wth a Nonphysiological Radical Transfer Pathway - The Importance of Hydrogen Bond Connections between the Participating Residues, *J. Biol. Chem. 273*, 21003-21008.

27. Kasanmascheff, M., Lee, W., Nick, T. U., Stubbe, J., and Bennati, M. (2016) Radical Transfer in *E. coli* Ribonucleotide Reductase: a $NH_2Y_{731}/R_{411}A$-α Mutant Unmasks a New Conformation of the Pathway Residue 731, *Chem. Sci. 7*, 2170-2178.

28. Moser, C. C., Anderson, J. L. R., and Dutton, P. L. (2010) Guidelines for Tunneling in Enzymes, *Biochim. Biophys. Acta, Bioenerg. 1797*, 1573-1586.

29. Nick, T. U., Lee, W., Koßmann, S., Neese, F., Stubbe, J., and Bennati, M. (2015) Hydrogen Bond Network between Amino Acid Radical Intermediates on the Proton-Coupled Electron Transfer Pathway of *E. coli* α2 Ribonucleotide Reductase, *J. Am. Chem. Soc. 137*, 289-298.

Appendix 1, Section 2

Class Ia RNR, Group G Sequence Alignments NrdA Proteins

E. coli K-12, subspecies MG1655 NrdA sequence highlighted in red.

* = Absolutely conserved residue
: = Conserved with strongly similar properties.
. = Conserved with weakly similar properties.

```
SalmonellaphageFelixO1_NrdAg          -------------------------------------MIKTIKKS-NGTVVSFDPERLN 21
EscherichiaphagerV5_NrdAg             -------------------------------------MITKVIKS-NGCEVAFDPERLN 21
EnterophageT5_NrdAg                   MKILGPFLSENFICQMTENKVNYFLNLDNYKKENQMSHRIEKVIKR-DGTVEDFAPEKLN 59
VibriophageKVP40_NrdAg                ---------------------------------------MLVVKS-SGITEEFEESKLM 19
AeromonasphageAeh1_NrdAg              ---------------------------------------MLVRKS-SGITENFEPKKLL 19
EnterophageRB49_NrdAg                 -------------------------------------MIKNVVKS-SGLTVPFEKEKVM 21
EnterophagePhi1_NrdAg                 -------------------------------------MIKNVVKS-SGLTVPFEKEKVM 21
EnterobacteriaphageT4_NrdAg           -------------------------------------MQLINVIKS-SGVSQSFDPQKII 22
EnterophageRB32_NrdAg                 -------------------------------------MQLINVIKS-SGVSQSFDPQKII 22
EnterophageJS98_NrdAg                 ---------------------------------------MQVQKS-SGVSQDFDAQKII 19
EnterophageRB69_NrdAg                 ---------------------------------------MQVTKS-SGVSQNFDAQKII 19
ProchlorophagePSSM2_NrdAg             ----------------------------------M-SNGTTVKKRNGRGVEPLNLEKIH 24
ProchlorococcusphagePSSM4_NrdAg       ----------------------------------MT-PKEINVIKR-DGSKAPLDLNRVH 24
Synechococcusphagesyn9_NrdAg          ----------------------------------MEA-SMTINVIKR-SGTAETLDLDKIH 25
Y_pestisMicrotus91001_NrdAg           -------------------------------------MISIVKR-NGSTEPLSEEKYN 20
Y_pestisPestoidesF_NrdAg              -------------------------------------MISIVKR-NGSTEPLSEEKYN 20
B_aphidicolaBp_NrdAg                  ----------------------------------M-NQMLFVTKR-NGKIELINLDKIH 23
B_aphidicolaSg_NrdAg                  ----------------------------------M-KNSLFVTKR-NGKKEKINLDKIH 23
C_RuthiaCm_NrdAg                      ----------------------------------M-NQDIKVIKR-NGQQESIDMEKIH 23
T_crunogenaXCL2_NrdAg                 ----------------------------------MS-TETMNVAKR-NGETEPLDLEKIH 24
M_flagellatusKT_NrdAg                 ----------MSAHEETL-------LPDWEPKMTQ-YESMKATKR-DGRQEPINLDKIH 40
N_gonorrhoeaeFA1090_NrdAg             ---------------------------------MNT-PTDLKVTKR-DGRLEAIDLDKIH 25
N_meningitidis053442_NrdAg            ---------------------------------MNT-PTDLKVTKR-DGRLEAIDLDKIH 25
N_meningitidiMC58_NrdAg               ---------------------------------MNT-PTDLKVTKR-DGRLEAIDLDKIH 25
P_atlanticaT6c_NrdAg                  ----------------------------------M-NNKLFVTKR-TGERESIDLEKIH 23
I_loihiensisL2TR_NrdAg                ----------------------------------M-NEQLYVTKR-SGEREPLNLDKIH 23
P_tunicataD2_NrdAg                    ----------------------------------M-NKQLSVSKR-DGRKELLDLDKIH 23
PsychrobacterPRwf1_NrdAg              ---------------------------------MTH-MDNIKVMKR-DGRLEPIDLDKIH 25
P_arcticus2734_NrdAg                  ---------------------------------MTH-IDKIQVTKR-DGRLEPIDLDKIH 25
P_cryohalolentisK5_NrdAg              ---------------------------------MTH-IDKIQVTKR-DGRLEPIELDKIH 25
CBlochmanniaBPEN_NrdAg                ----------------------------------M-NQTLLVTKR-NGHHEPINLDKIH 23
P_ingrahamii37_NrdAg                  ----------------------------------M-KKNLFVTKR-NGKRERIDLDKIH 23
MarinomonasMWYL1_NrdAg                -------------------------------------MSTLTVTKR-NGTTENINLEKIH 22
S_denitrificansOS217_NrdAg            ----------------------------------M-NTNMTVTKR-SGERESIDLDKIH 23
S_amazonensisSB2B_NrdAg               ---MARLF--TWICRAI----------PIYKGLSSM-NSNMTVTKR-SGEREQIDLDKIH 43
S_frigidimarina400_NrdAg              ----------------------------------M-NSNMTVTKR-SGERESIDLDKIH 23
S_oneidensisMR1_NrdAg                 ----------------------------------M-NSNMTVTKR-CGARETIDLDKIH 23
ShewanellaANA3_NrdAg                  ----------------------------------M-NSNMTVTKR-CGARETIDLDKIH 23
ShewanellaMR4_NrdAg                   ----------------------------------M-NSNMTVTKR-CGARETIDLDKIH 23
ShewanellaMR7_NrdAg                   ----------------------------------M-NSNMTVTKR-CGARETIDLDKIH 23
S_balticaOS155_NrdAg                  ----------------------------------M-NSTMTVTKR-CGDRENIDLEKIH 23
S_putrefaciensCN32_NrdAg              ----------------------------------M-NSTITVTKR-CGERENIDLEKIH 23
ShewanellaW3181_NrdAg                 ----------------------------------M-NSTITVTKR-CGERENIDLEKIH 23
S_sediminisHAWEB3_NrdAg               ----------------------------------M-NSNLKVTKR-SGERETIDLDKIH 23
S_loihicaPV4_NrdAg                    ----------------------------------M-NSNMQVTKR-SGERETIDLDKIH 23
S_pealeana700345_NrdAg                ----------------------------------M-NSNMTVTKR-SGERETIDLDKIH 23
A_succinogenes130Z_NrdAg              ----------------------------------M-NKTLMVTKR-NGKQEPLDLDKLH 23
M_succiniciproducensMBEL55E_NrdAg     -------------------------MCFSKVKTFM-NKALMVTKR-DGQVEPLDLDKIH 32
H_influenzae86028NP_NrdAg             ----------------------------------M-NKSLMVTKR-DGTQEQINLDKIH 23
H_influenzaeRdKW20_NrdAg              ---MDRFHKNTISCVLNPV-------LNNNKKAHTM-NKSLMVTKR-DGTQEQINLDKIH 48
H_somnus129PT_NrdAg                   ----------------------------------M-NKALMVTKR-DGNKEPIDLDKIH 23
P_multocidaPm70_NrdAg                 ----------------------------------M-NKALMVTKR-DGHLEPIDLDKIH 23
A_pleuropneumoniae5bL20               ----------------------------------M-QKTLMVTKR-NGKQEPFDLDKIH 23
H_ducreyi35000HP_NrdAg                ----------------------------------M-QSQLMVTKR-NGQQEPLDLEKIH 23
V_choleraeO395_NrdAg                  ----------------------------------M-NQQLTVTKR-DGRKENIDLDKIH 23
V_vulnificusCMCP6_NrdAg               ----------------------------------M-NQQLTVTKR-DGRKENIDLEKLH 23
VibrioEx25                            ----------------------------------M-NQQLTVTKR-DGRKESIDLDKIH 23
V_campbelliiBAA1116_NrdAg             ----------------------------------M-NQQLTVTKR-DGRKETIDLEKIH 23
VibrioAND4_NrdAg                      ----------------------------------M-NQQLTVTKR-DGRKETIDLEKIH 23
V_fischeriES114_NrdAg                 ----------------------------------M-TQQLTVTKR-NGSTEKINLDKIH 23
P_angustumS14_NrdAg                   ----------------------------------M-TQQLTVTKR-DGRTENIDLDKIH 23
PhotobacteriumSKA34_NrdAg             ----------------------------------M-TQQLTVTKR-DGRTESIDLDKIH 23
C_psychrerythraea34H_NrdAg            ----------------------------------M-NNNLFVTKR-NGEKETIDLEKIH 23
A_hydrophila7966_NrdAg                ------------------M-------VFRESTKGHQ-PMNLFVTKR-NGHKEPIDLDKIH 33
A_salmonicidaA449_NrdAg               -------------------------------------MNLFVTKR-NGHKEPIDLDKIH 21
B_cicadellinicolaHc_NrdAg             ----------------------------------M-SQSLLVTKR-DGRKEPINLDKIH 23
C_sakazakiiBAA894                     ----------------------------------M-NQSLLVTKR-DGSQERINLDKIH 23
K_pneumoniaeMGH78578_NrdAg            ----------------------------------M-NQSLLVTKR-DGTTERINLDKIH 23
Enterobacter638_NrdAg                 ----------------------------------M-NQSLLVTKR-DGTTERINLDKIH 23
C_koseriBAA895_NrdAg                  ----------------------------------M-NQSLLVTKR-DGRTERINLDKIH 23
S_boydiiSb227_NrdAg                   ----------------------------------M-NQNLLVTKR-DGSTERINLDKIH 23
E_coliK12ssMG1655_NrdAg               ----------------------------------M-NQNLLVTKR-DGSTERINLDKIH 23
S_flexneri58401_NrdAg                 ----------------------------------M-NQNLLVTKR-DGSTERINLDKIH 23
E_coli536_NrdAg                       ----------------------------------M-NQNLLVTKR-DGSTERINLDKIH 23
S_flexneri2a2457T_NrdAg               ----------------------------------M-NQNLLVTKR-DGSTERINLDKIH 23
STyphimuriumLT2_NrdAg                 ----------------------------------M-NQSLLVTKR-DGRTERINLDKIH 23
S_entericaTyphiTy2_NrdAg              ----------------------------------M-NQSLLVTKR-DGRTERINLDKIH 23
S_proteamaculans568_NrdAg             ----------------------------------M-NQSLLVTKR-DGRKERINLDKIH 23
Y_pestisAngola_NrdAg                  ----------------------------------M-NQSLLVTKR-DGSKERINLDKIH 23
YpseudotuberculosisIP31758_NrdAg      ----------------------------------M-NQSLLVTKR-DGSKERINLDKIH 23
                                                                          *        :  .:
```

```
SalmonellaphageFelixO1_NrdAg          KWASWADKR--GIIWSEVTMEAMKRVYEGCTTKEMHQAMIDVCVD---KQTQEYSDMAGR 76
Escherichiaphagerv5_NrdAg             KAAIFGDDG--NGNWSHIALDAYKRLYDGCTTKEVNQAIIDSCVS---RKDEEHSRMAGR 76
EnterophageT5_NrdAg                   GWAEYGCKTVD-VSWSAITMAAQKTLPKGVVDSDTLMDALIKAAESLIKDNPAYDVPAKE 118
VibriophageKVP40_NrdAg                RVLSDACKNTT-IDPHQLLARVKAQLSDGMTSKDIQKAAIKIAADSISVEEPDYQYVAAR 78
AeromonasphageAeh1_NrdAg              KILGFATAGTR-IDPNEILNGIMPSITDGMSTKNIQKLAIKFAADRISIEQSDYQYVAAR 78
EnterophageRB49_NrdAg                 KILEAAAEGHN-VDPYELFDMVEPLLTEGMSTNDIQKVLIKVAADNITVETPDYQYVASN 80
EnterophagePhi1_NrdAg                 KILEAAAEGHN-VDPYELFDMVEPLLTEGMSTNDIQKVLIKVAADNITVETPDYQYVASN 80
EnterobacteriaphageT4_NrdAg           KVLSWAAEGTS-VDPYELYENIKSYLRDGMTTDDIQTIVIKAAANSISVEEPDYQYVAAR 81
EnterophageRB32_NrdAg                 KVLSWAAEGTS-VDPYELYENIKSYLRDGMTTDDIQTIVIKAAANSISVEEPDYQYVAAR 81
EnterophageJS98_NrdAg                 KVLTWGCEGTQ-IDPYELYEEIKTHLRDGMTTKDIQNLTVKVAANLISVEEPDYQYVAAK 78
EnterophageRB69_NrdAg                 QVLTWQCEGTQ-IDPYELYEEIKTHLVDGMSTKDIQKICIKVAANSISVEEPDYQYVAAK 78
ProchlorophagePSSM2_NrdAg             VMCEEACEGLAGVSASQVEIQSGIQFYDGITTAEIQEILIRSASDLIDLDHPNYQFVAAR 84
ProchlorococcusphagePSSM4_NrdAg       HIVEHACRGLAGVSESQVEISSGLQFFDGIKTSDIQEILVRSANDLISLEAPNYQYVAAR 84
Synechococcusphagesyn9_NrdAg          TMVEHACKGLAGVSESQVEMNANLQFFDGVETKDIQEILIRSANDLISLEAPNYQFVAAR 85
Y_pestisMicrotus91001_NrdAg           RVVMWGVENIRNVSASAVAMGAAASIFDGMTTCQLHEALVKSAADLISPETPNYLQVAAR 80
Y_pestisPestoidesF_NrdAg              RVVMWGVENIRNVSASAVAMGAAASIFDGMTTCQLHEALVKSAADLISPETPNYLQVAAR 80
B_aphidicolaBp_NrdAg                  RVLNWAAKGLENISISQVALKSRIQFYNNIKTTIIHETIIKAAADLISENSPDYQYMAAR 83
B_aphidicolaSg_NrdAg                  KVLNWAAKGLENISVSQVELRSRIQFYNNISTVNIHETIIKAAADLISENTPDYQYMAAR 83
C_RuthiaCm_NrdAg                      RVVHWASQDLKGVSVSQVEINAQLAFFDGIKTEDIHETIIKSAADLISIEVPDYQYLAAR 83
T_crunogenaXCL2_NrdAg                 RVITWAAEGLENVSVSEVELRSHIQFYDGMKTSDIHETIIKSAADLISESAPDYQHLAAR 84
M_flagellatusKT_NrdAg                 RVIDWAAQGLDNVSVSQVELKSHIQFYDGIRTDVIHETIIKSAADLISEETPDYQYLAAR 100
N_gonorrhoeaeFA1090_NrdAg             RVVTWAAEGLENVSVSQVELKSHIQFYNGIRTDDIHETIIKAAADLISEDTPDYQYLAAR 85
N_meningitidis053442_NrdAg            RVVTWAADGLENVSVSQVELKSHIQFYNGIRTDDIHETIIKAAADLISEDTPDYQYLAAR 85
N_meningitidiMC58_NrdAg               RVVTWAADGLENVSVSQVELKSHIQFYNGIRTDDIHETIIKAAADLISEDTPDYQYLAAR 85
P_atlanticaT6c_NrdAg                  KVITWAAKGLNNVSVSQVEIKAHIQFYDGIKTEDIHETLIRSAADLISTDAPDYQYLAAR 83
I_loihiensisL2TR_NrdAg                RVIMWAAEGLNNVSASQVEIKSHIQFYDGIKTEDIHETIIKAAADLISEEAPDYQYMAAR 83
P_tunicataD2_NrdAg                    RVIEWAAKGLNNVSVSQVELKSHIQFYDGIRTEDIHETIIKAAADQISKESPDYQFLAAR 83
PsychrobacterPRwf1_NrdAg              KVITWAAEGLDNVSVSQVELKSHIQFYDGIKTRDIHETIIKAAADLISEDTPDYQYLAAR 85
P_arcticus2734_NrdAg                  KVIEWAAHNLENVSVSQVELKSHIQFYEGIKTRDIHETIIKSAADLISETTPDYQYLAAR 85
P_cryohalolentisK5_NrdAg              KVIEWAAHGLDNVSVSQVELKSHIQFYEGIKTRDIHETIIKSAADLISEDTPDYQYLAAR 85
CBlochmanniaBPEN_NrdAg                RVINWAAKGLEQVSVSQVELRSHIQFYNGIKTSDIQETIIKAAADLISEEEPDYQYLAAR 83
P_ingrahamii37_NrdAg                  RVITWAAEGLKNVSVSQVELNSHIQFYDGIRTDDIHETIIKAAADLISETSPDYQYLAAH 83
MarinomonasMWYL1_NrdAg                KVIAWAAEGLDNVSVSQVELKARIQFFEGIRTTDIHETLIKSAADLISESTPDYQYLAAR 82
S_denitrificansOS217_NrdAg            RVITWSTKGLKNVSVSEVELRSHLQFFDGIPTEAIHETIIKAAADLISPDSPDYQFVAAR 83
S_amazonensisSB2B_NrdAg               RVVTWAAKGLKNVSVSEVELRSHLQFFEGIPTEAIHETIIKAAADLISPESPDYQFLAAR 103
S_frigidimarina400_NrdAg              RVITWAAKGLKNVSVSEVELRSHLQFFDGIPTEAIHETIIKAAADLISPESPDYQFVAAR 83
S_oneidensisMR1_NrdAg                 RVITWAAKGLKNVSVSEVELRTHLQFFDGIPTEAIHETIIKSAADLISPESPDYQFLAAR 83
ShewanellaANA3_NrdAg                  RVITWAAKGLKNVSVSEVELRTHLQFFDGIPTEAIHETIIKSAADLISPESPDYQFLAAR 83
ShewanellaMR4_NrdAg                   RVITWAAKGLKNVSVSEVELRTHLQFFDGIPTEAIHETIIKSAADLISPESPDYQFLAAR 83
ShewanellaMR7_NrdAg                   RVITWAAKGLKNVSVSEVELRTHLQFFDGIPTEAIHETIIKSAADLISPESPDYQFLAAR 83
S_balticaOS155_NrdAg                  RVITWAAKGLKNVSVSEVELRSHLQFFDGIPTEAIHETIIKAAADLISPESPDYQFLAAR 83
S_putrefaciensCN32_NrdAg              RVITWAAKGLKNVSVSEVELRSHLQFYDGIPTETIHETIIKAAADLISPESPDYQFLAAR 83
ShewanellaW3181_NrdAg                 RVITWAAKGLKNVSVSEVELRSHLQFYDGIPTETIHETIIKAAADLISPESPDYQFLAAR 83
S_sediminisHAWEB3_NrdAg               RVITWAAKDLNNVSVSEVELRSHLQFFDGIPTEAIHETIIKAAADLISPESPDYQFLSAR 83
S_loihicaPV4_NrdAg                    RVITWAAKGLNNVSVSEVELRSHLQFYDGIPTEAIHETIIKAAADLISPESPDYQFLSAR 83
S_pealeana700345_NrdAg                RVIEWAAKGLNNVSVSEVELRSHLQFFDGIATEAIHETIIKAAADLISPESPDYQFLSAR 83
A_succinogenes130Z_NrdAg              RVITWAAEGLHNVSVSQVELRSHIQFYEGIRTADIHETIIKSAADLISKDAPDYQYLAAR 83
M_succiniciproducensMBEL55E_NrdAg     RVITWAAEGLENVSVSQVELRSHIQFYEGIRTSDIHETIIKAAADLISKDAPDYQYLAAR 92
H_influenzae86028NP_NrdAg             RVITWAAEGLDNVSVSQVELRSHIQFYEGIRTSDIHETIIKAAADLISKDSPDYQYLAAR 83
H_influenzaeRdKW20_NrdAg              RVITWAAEGLDNVSVSQVELRSHIQFYEGIRTSDIHETIIKAAADLISKDSPDYQYLAAR 108
H_somnus129PT_NrdAg                   RVITWAAEGLENVSVSLVELKSHIQFYEGIRTSDIHETIIKAAADLISKDSPDYQYLAAR 83
P_multocidaPm70_NrdAg                 RVITWAAEGLDNVSVSQVELRSHIQFYEGIRTSDIHETIIKAAADLISKDSPDYQYLAAR 83
A_pleuropneumoniae5bL20               RVITWAAEGLHNVSVSQVELSAHIQFYQGIRTADIHETIIKAAADLISTETPDYQYLAAR 83
H_ducreyi35000HP_NrdAg                RVITWAAEGLANVSVSQVELSAHIQFYQGIRTSDIHETIIKAAADLISKETPDYQYLAAR 83
V_choleraeO395_NrdAg                  RVITWAAEGLQNVSVSQVELRAHIQFYDGITTSDIHETIIKSAADLISEETPDYQYLAAR 83
V_vulnificusCMCP6_NrdAg               RVITWAAEGLDNVSVSQVELRAHIQFYDGITTSDIHETIIKSAADLISEETPDYQYLAAR 83
VibrioEx25                            RVITWAAEGLDNVSVSQVELRAHIQFYDGITTTDIHETIIKSAADLISEETPDYQYLAAR 83
V_campbelliiBAA1116_NrdAg             RVITWAAEGLNNVSVSQVELRAHIQFYDGITTSDIHETMIKSAADLISEETPDYQYLAAR 83
VibrioAND4_NrdAg                      RVITWAAEGLNNVSVSQVELRAHIQFYDGITTSDIHETIIKSAADLISEETPDYQYLAAR 83
V_fischeriES114_NrdAg                 KVITWAAEGLDNVSVSQVEMKSHIQFYDGMKTSDIHETIIKAAADLISEDTPDYQYLAAR 83
P_angustumS14_NrdAg                   RVIEWAAEGLENVSVSQVELKSHIQFYDGIRTDDIHETIIKAAADLISEDTPDYQYLAAR 83
PhotobacteriumSKA34_NrdAg             RVIEWAAEGLENVSVSQVELKSHIQFYDGIRTDDIHETIIKAAADLISEDTPDYQYLAAR 83
C_psychrerythraea34H_NrdAg            RVIAWAADGLEQVSVSQVELKSHIQFYDGIKTEDIHETIIKSAADLISEETPDYQYLAAR 83
A_hydrophila7966_NrdAg                RVLDWAAEGLDNVSVSQVELKSHIQFYDGIRTADIHETVIKAAADLISEQSPDYQYMAAR 93
A_salmonicidaA449_NrdAg               RVLDWAAEGLDNVSVSQVELKSHIQFYDGIRTADIHETVIKATADLISEQSPDYQYMAAR 81
B_cicadellinicolaHc_NrdAg             RVINWAADGLENVSVSQVELRSHIQFYDGIKTSDIHETIIKAAADLISRDAPDYQYLAAR 83
C_sakazakiiBAA894                     RVLDWAAEGLHNVSISQVELRSHIQFYDGIKTSDIHETIIKAAADLISRDAPDYQYLAAR 83
K_pneumoniaeMGH78578_NrdAg            RVLDWAAEGLNNVSISQVELRSHIQFYDGIKTADIHETIIKAAADLISRDAPDYQYLAAR 83
Enterobacter638_NrdAg                 RVLDWAAEGLNNVSISQVELRSHIQFYDGIKTSDIHETIIKAAADLISREAPDYQYMAAR 83
C_koseriBAA895_NrdAg                  RVLDWAAEGLNNVSISQVELRSHIQFYDGIKTSDIHETIIKAAADLISRDAPDYQYLAAR 83
S_boydiiSb227_NrdAg                   RVLDWAAEGLHNVSISQVELRSHIQFYDGIKTSDIHETIIKAAADLISRDAPDYQYLAAR 83
E_coliK12ssMG1655_NrdAg               RVLDWAAEGLHNVSISQVELRSHIQFYDGIKTSDIHETIIKAAADLISRDAPDYQYLAAR 83
S_flexneri58401_NrdAg                 RVLDWAAEGLHNVSISQVELRSHIQFYDGIKTSDIHETIIKAAADLISRDAPDYQYLAAR 83
E_coli536_NrdAg                       RVLDWAAEGLHNVSISQVELRSHIQFYDGIKTSDIHETIIKAAADLISRDAPDYQYLAAR 83
S_flexneri2a2457T_NrdAg               RVLDWAAEGLHNVSISQVELRSHIQFYDGIKTSDIHETIIKAAADLISRDAPDYQYLAAR 83
STyphimuriumLT2_NrdAg                 RVLDWAAEGLNNVSVSQVELRSHIQFYDGIKTSDIHETIIKAAADLISRDAPDYQYLAAR 83
S_entericaTyphiTy2_NrdAg              RVLDWAAEGLNNVSVSQVELRSHIQFYDGIKTSDIHETIIKAAADLISRDAPDYQYLAAR 83
S_proteamaculans568_NrdAg             RVIDWAAEGLNNVSVSQVELRSHIQFYDGIKTADIHETIIKAAADLISRDAPDYQYLAAR 83
Y_pestisAngola_NrdAg                  RVIDWAAEGLHNVSVSQVELRSHIQFYDGIKTSDIHETIIKAAADLISRDAPDYQYLAAR 83
YpseudotuberculosisIP31758_NrdAg      RVIDWAAEGLHNVSVSQVELRSHIQFYDGIKTSDIHETIIKAAADLISRDAPDYQYLAAR 83
                                                :       . .           :    .         :   : .
```

```
SalmonellaphageFelixO1_NrdAg          LLLGIIYKEAFGGFTK-VPTLVTFVKNMERAGLWEK---MDYSQEELEYLQGYIVHSKDI 132
Escherichiaphagerv5_NrdAg             VLIGQIYKEAFGGFTK-IPSLPSFYNTMVEKGYWED---MGYSFTELNTLDAVIDHSKDL 132
EnterophageT5_NrdAg                   LRLAQMRKRLYDSFE--PPSLRFFHDHMVSVGAWED-MSAWITDEQFEALNQVIDHDRDR 175
VibriophageKVP40_NrdAg                LAMYGLRKDVYGQFD--PIPFKQLIERNTAIEVYDSEILEKWSDSELEYLESQVKHERDF 136
AeromonasphageAeh1_NrdAg              LQMYAIRKDVYNQFD--PIHLKRHTLDLVQDGIYDKDLAIKWTDDEFNELNAYIKHDRDF 136
EnterophageRB49_NrdAg                 LAMFALRKSVYGQFQ--PIDFYSHIERNIATGKYDREILQKYSREEVEYLEKFIDHKLDF 138
EnterophagePhi1_NrdAg                 LAMFALRKSVYGQFQ--PIDFYCHIERNIATGKYDREILQKYSREEVEYLEKFIDHKLDF 138
EnterobacteriaphageT4_NrdAg           CLMFALRKHVYGQYE--PRSFIDHISYCVNAGKYDPELLSKYSAEEITFLESKIKHERDM 139
EnterophageRB32_NrdAg                 CLMFALRKHVYGQYE--PRSFIDHISYCVNEGKYDPELLSKYSAEEITFLESKIKHERDM 139
EnterophageJS98_NrdAg                 QLMFALRKEVYGQFE--PRSFIDHISYCVNEGKYDPELLSRYSAEEITFLESHIKHERDF 136
EnterophageRB69_NrdAg                 GLMFALRKDVYGQFE--PPAFIDHISYCVNEGKYDPELLSRYSAEEIIYLDSRIKHERDF 136
ProchlorophagePSSM2_NrdAg             LLLFSLRKQLFGRLRE-CPTVIDHVQKCVKKGIYDAEILDLYSEEEYNRLQSFIDHSRDF 143
ProchlorococcusphagePSSM4_NrdAg       LLLFSLRKQVYGGHPDHPTEFVEHIQSCIDRNLYDGDILKKYTNDEITELGKYIDHDRDY 144
Synechococcusphagesyn9_NrdAg          LLLFGLRKAVYNGHPDGHPPLKEHVEKCIERGVYDKTILSKYTDEEWDKLNSFMDHDRDY 145
Y_pestisMicrotus91001_NrdAg           LNLFKIRKDAFGQYD--YPNFYNHIVSNVSRGVYDEDLLKYYSLEEIAELGAYIKPKRDE 138
Y_pestisPestoidesF_NrdAg              LNLFKIRKDAFGQYD--YPNFYNHIVSNVSRGVYDEDLLKYYSLEEIAELGAYIKPKRDE 138
B_aphidicolaBp_NrdAg                  LTIFHLRKKAFGQFN--PPNLYDHVKNMIKLEKYDQHLLNDYPKQDFLIMNSFIKHRRDM 141
B_aphidicolaSg_NrdAg                  LAIFHLRKKAYGQFE--PPNLYYHVKKMVQLGKYDEKLLQNYSFEEYVEMNSFVDHRRDM 141
C_RuthiaCm_NrdAg                      LTIFHLRKKAFKSFT--PTPLFEHIKKFTDLSIYDKEILNKYSKEEIETLGAYIDHWRDM 141
T_crunogenaXCL2_NrdAg                 LAIFHLRKKAFDRFT--PPTLFEHVNKMVRNRLYDPQIQDDYTKAEFDELDNYIQHDRDM 142
M_flagellatusKT_NrdAg                 LSIFHLRKIAYGQFE--PPHLFDHVTKLTEMGKYDKHILQDYSREEFDALNDHLDHWRDM 158
N_gonorrhoeaeFA1090_NrdAg             LAIFHLRKIAYGEYE--PPHLYNHVKKLTDAGKYDRHILEDYSREEFDELNAYIDHERDM 143
N_meningitidis053442_NrdAg            LAIFHLRKIAYGEYE--PPHLYDHVKKLTDAAKYDRHILEDYSREEFDELNAYIDHERDM 143
N_meningitidiMC58_NrdAg               LAIFHLRKIAYGEYE--PPHLYDHVKKLTDAGKYDRHILEDYSREEFDELNAYIDHERDM 143
P_atlanticaT6c_NrdAg                  LAIFHLRKKAYGQFE--PPALFAHVTNMVEMQKYDQHLLSDYSEAEFQEMDDYIDHWRDM 141
I_loihiensisL2TR_NrdAg                LAIFHLRKRAYGKFE--PPRLYDQVKKMVAEHRYDEHLLEDYTEAEFDQMDEILDHWRDM 141
P_tunicataD2_NrdAg                    LAIFHLRKKAYGQFE--PPRLYDHVVKMVEDNRYDAHLLADYTQAEFEQLDEVLDHNRDL 141
PsychrobacterPRwf1_NrdAg              LAIFHLRKIAYNQFE--PPHLFDHVTKLTEAGKYDQHILQDYSREEFDELNDYIDHWRDM 143
P_arcticus2734_NrdAg                  LAIFHLRKIAYNRFT--PPHLYDHVKKLTDAGKYDAHILADYSRAEFDELEEYLDHWRDM 143
P_cryohalolentisK5_NrdAg              LAIFHLRKIAYNKFT--PPHLYDHVKKLTDAGKYDQHILADYSRAEFDELEEYLDHWRDM 143
CBlochmanniaBPEN_NrdAg                LAVFHLRKKAYGQFE--PPKLYNHVLRLIKIGKYDQHLLDDYNAEEFNQMDTFIDHWRDM 141
P_ingrahamii37_NrdAg                  LAVFHLRKKAYRRFT--PPSIFEHVSKMCSKGKYDMHLLEDYSKQEFEEMEKFIVHDRDL 141
MarinomonasMWYL1_NrdAg                LAIFHLRKKAFGDFE--PPHLYKHVKNLVEQKRYDAQLLEKYTEEDFNELNTFLDHSRDM 140
S_denitrificansOS217_NrdAg            LAVFHLRKKAYGQFE--PPSLVDHVTRLVELGKYDKNILTDYTREELEVMDTYIDHWRDM 141
S_amazonensisSB2B_NrdAg               LAVFHLRKKAFGQFE--PPKLYDHVVKLVEMGKYDRHILEDYTKEELDVLDSYIDHWRDM 161
S_frigidimarina400_NrdAg              LAVFHLRKKAFGQFE--PPTLYDHVTKLVEIGKYDMHIMADYSREELDILDTYIDHWRDL 141
S_oneidensisMR1_NrdAg                 LAVFHLRKKAFGQFE--PPKLYDHVTKLVELGKYDMHILQDYTREELDIMDTYIDHWRDM 141
ShewanellaANA3_NrdAg                  LAVFHLRKKAFGQFE--PPKLYDHVTKLVELGKYDMHILQDYTREELDIMDSYIDHWRDM 141
ShewanellaMR4_NrdAg                   LAVFHLRKKAFGQFE--PPKLYDHVTKLVELGKYDMHILQDYTREELDIMDTYIDHWRDM 141
ShewanellaMR7_NrdAg                   LAVFHLRKKAFGQFE--PPKLYDHVTKLVELGKYDMHILQDYTREELDIMDTYIDHWRDM 141
S_balticaOS155_NrdAg                  LAVFHLRKKAFGQFE--PPALYTHVTKLVELGKYDMHILQDYSREELDVMDTYIDHWRDM 141
S_putrefaciensCN32_NrdAg              LAVFHLRKKAFGQFE--PPALYDHVTKLVELGKYDMHILQDYTREELDLMDTYIDHWRDM 141
ShewanellaW3181_NrdAg                 LAVFHLRKKAFGQFE--PPALYDHVTKLVELGKYDMHILQDYTREELDLMDTYIDHWRDM 141
S_sediminisHAWEB3_NrdAg               LAIFHLRKKAFGQFE--PPKLYEHVVKLVEMGKYDAHILHDYSREELDTLDGYIDHWRDM 141
S_loihicaPV4_NrdAg                    LAIFHLRKKAFGQFE--PPSLYDHVVKLVELGKYDTHILNDYSREELDTLDSYIDHWRDM 141
S_pealeana700345_NrdAg                LAVFHLRKKAFGQFE--PPKLYDHVVKLIELGKYDQHILEDYSQEEIDTLDSYIDHWRDM 141
A_succinogenes130Z_NrdAg              LSVFHLRKKAFGQFE--PPHLYEHVKKLVEMGKYDQALLTDYGREEWEQMNDFLDHWRDM 141
M_succiniciproducensMBEL55E_NrdAg     LAIFHLRKKAYGHFD--PPRLYEHVKKLVRLGKYDESLLSDFSREEWDEMDNFLDHSRDM 150
H_influenzae86028NP_NrdAg             LAIFHLRKKAYGHFD--PPRLYDHVKKLVRMEKYDQALLDDYTREEWDTMDGFIDHWRDM 141
H_influenzaeRdKW20_NrdAg              LAIFHLRKKAYGHFD--PPRLYDHVKKLVRMEKYDQALLDDYTREEWDTMDGFIDHWRDM 166
H_somnus129PT_NrdAg                   LAVFHLRKKAYGHFE--PPKLYDHVKKLVRMGKYDQALLTDYTREEWDEMDGFLDHWRDM 141
P_multocidaPm70_NrdAg                 LAVFHLRKKAYGHFD--PPRLYDHVKKLVRMGKYDPALLVDYTREEWDEMDSYLDHWRDM 141
A_pleuropneumoniae5bL20               LAVFHLRKKAYGDFT--PPHLFTHIKKLVALGKYDKALLTDYNEQEWQLMNEFLVHSRDM 141
H_ducreyi35000HP_NrdAg                LAVFHLRKKAYGTFY--PPHLFEHIQKLVKLGKYDNLLLKQYSEEEWDIMNGFLVHERDM 141
V_choleraeO395_NrdAg                  LAVFHLRKKAYGQYE--PPTLLAHVTKMVEKGKYDKHLLEDYTREELELMDQFIDHRRDL 141
V_vulnificusCMCP6_NrdAg               LAVFHLRKKAYGQYE--PPTLYDHVARLVDMGKYDKHILEDYTKAELDELDAYLDHKRDL 141
VibrioEx25                            LAVFHLRKKAYGQYE--PPTLHDHVSRLVDMGKYDQHILEDYTKAELDELDAYIDHKRDL 141
V_campbelliiBAA1116_NrdAg             LAVFHLRKKAYGQYE--PPTLHDHVARLVDMGKYDKHILEDYTKAELDELDAYMDHKRDL 141
VibrioAND4_NrdAg                      LAVFHLRKKAYGQYE--PPTLHDHVARLVDMGKYDQHILEDYTKAELDELDAYLEHKRDL 141
V_fischeriES114_NrdAg                 LSVFHLRKKAYGQFE--PPTLFNHVLNLIELGKYDKHIVEDYTQEEFALMDQMIDHNRDM 141
P_angustumS14_NrdAg                   LAIFHLRKKAYGQFE--PPKLYEHVKNLVAKGKYDGHLLEDYTEEEFAVMDNFIDHWRDM 141
PhotobacteriumSKA34_NrdAg             LAIFHLRKKAYGQFE--PPKLYDHVKNLVEKGKYDSHLLEDYTEEEFAVMDNFIDHWRDM 141
C_psychrerythraea34H_NrdAg            LAVFHLRKKAYGQFE--PPALYDHVVKLVEAGKYDQHLLADYTKDDFEHMSSFMDHSRDL 141
A_hydrophila7966_NrdAg                LAIFHLRKKAYGQFE--PPHLYQHVAKLVDMGKYDKHLLEDYSQAEFEELNQHLDHWRDM 151
A_salmonicidaA449_NrdAg               LAIFHLRKKAYGQFE--PPHLYQHVARLVDMGKYDKHLLTDYTQAEFEELNEHLDHWRDM 139
B_cicadellinicolaHc_NrdAg             LAVLHLRKKAYGTFE--PPTLYDHVSRLVKMGKYDQHLLKDYSFEEFTQMDKFINHWRDM 141
C_sakazakiiBAA894                     LAIFHLRKKAYGQFE--PPKLYDHVVKMVELGKYDHHLLQDYSEDEFAQMDGFIDHWRDM 141
K_pneumoniaeMGH78578_NrdAg            LAIFHLRKKAFGQFE--PPALYDHVVKMVEKGKYDHHLLEDYTEEEFQQMDGFLDHWRDM 141
Enterobacter638_NrdAg                 LAIFHLRKKAYGQFE--PPALLDHVTKMVELGKYDAHLLEDYTEEEFKQMDGFIDHWRDM 141
C_koseriBAA895_NrdAg                  LAIFHLRKKAYGEFE--PPALYDHVVKMVELGKYDNHLLEDYTEEEFKQMDSFIVHDRDM 141
S_boydiiSb227_NrdAg                   LAIFHLRKKAYGQFE--PPALYDHVVKMVEMGKYDNHLLEDYTEEEFKQMDTFIDHDRDM 141
E_coliK12ssMG1655_NrdAg               LAIFHLRKKAYGQFE--PPALYDHVVKMVEMGKYDNHLLEDYTEEEFKQMDTFIDHDRDM 141
S_flexneri58401_NrdAg                 LAIFHLRKKAYGQFE--PPALYDHVVKMVEMGKYDNHLLEDYTEEEFKQMDTFIDHDRDM 141
E_coli536_NrdAg                       LAIFHLRKKAYGQFE--PPALYDHVVKMVEMGKYDNHLLEDYTEEEFKQMDTFIDHDRDM 141
S_flexneri2a2457T_NrdAg               LAIFHLRKKAYGQFE--PPALYDHVVKMVEMGKYDNHLLEDYTEEEFKQMDTFIDHDRDM 141
STyphimuriumLT2_NrdAg                 LAIFHLRKKAFGQFE--PPALYHHVVKMVELGKYDNHLLEDYTEEEFKQMDSFIVHDRDM 141
S_entericaTyphiTy2_NrdAg              LAIFHLRKKAFGQFE--PPALYHHVVKMVELGKYDNHLLEDYTEEEFKQMDSFIVHDRDM 141
S_proteamaculans568_NrdAg             LAIFHLRKKAYGQFE--PPKLLAHVTRMVEMGKYDKHLLEDYTTEEFEQMDSFIDHWRDM 141
Y_pestisAngola_NrdAg                  LAIFHLRKKAYGQFE--PPTLFAHVSKMVEMGKYDKHLLEDYTAEEFEQMDTFIDHWRDM 141
YpseudotuberculosisIP31758_NrdAg      LAIFHLRKKAYGQFE--PPTLFAHVSKMVEMGKYDKHLLEDYTAEEFEQMDTFIDHWRDM 141
                                        :  : *  :          .            ::          :   :   :    *
```

```
SalmonellaphageFelixO1_NrdAg          SYGYAVLKQFRDKYGIRDIKTGRLFESPQFMFMGMAMKAFEKQPK-----HRRLQDVIKL 187
EscherichiaphageRv5_NrdAg             GYGYAVLKQFRDKYGIKDTLKDTLFESPQMMFMGIAMAVMKNMPK-----ERRINDVVKL 187
EnterophageT5_NrdAg                   LFTSGGLKQFFDKYSRRNIATGEIYETPQFAYMGMAMAMLSQPNW-------TILDAIDL 228
VibriophageKVP40_NrdAg                DFAYAGVMQFKEKYLCKNRGTGAIYETPQYAYMLIAMCLHQEEKY------DRLKHVIDF 190
AeromonasphageAeh1_NrdAg              DLAYAGVMQFKEKYLAKDRSTDQIYETPQILYMLVAMCLHQDEKT------DRIAKVKNF 190
EnterophageRB49_NrdAg                 KMTYAGAMQLKEKYLVQDRSTKKIIETPQFAYMLIGMCLHQEETE------NRLEHVIDF 192
EnterophagePhi1_NrdAg                 KMTYAGAMQLKEKYLVQDRSTKKIIETPQFAYMLIGMCLHQEETE------NRLEHVIDF 192
EnterobacteriaphageT4_NrdAg           EFTYSGAMQLKEKYLVKDKTTGQIYETPQFAFMTIGMALHQDEPV------DRLKHVIRF 193
EnterophageRB32_NrdAg                 EFTYSGAMQLKEKYLVKDKTTGQIYETPQFAFMTIGMALHQDEPV------DRLKHVIRF 193
EnterophageJS98_NrdAg                 ELTYAGAMQLKEKYLVKDKTTGKIYETPQFAIMCIGMALHQDEPG------NRLAHIIRF 190
EnterophageRB69_NrdAg                 DLTYAGAMQLKEKYLVKDKTTGKIYETPQFAIMAIGMALHQDETH------DRLKHIIRF 190
ProchlorophagePSSM2_NrdAg             LFTYAGLRQVVDKYLVQDRSTGEVYETPQFMYLMIAATIFSKYPK-----ETRLDYVRKY 198
ProchlorococcusphagePSSM4_NrdAg       LFTYAGIRQVVDKYLVQDRSTGEVFETPQFMYMMIAATLFQDDDK-----FYRLKYVKKY 199
Synechococcusphagesyn9_NrdAg          LFTYAGIRQVVDKYLVQDRSTGEIYETPQFMYMMIAATLFQDDDP-----FYRIEYVKKY 200
Y_pestisMicrotus91001_NrdAg           YFGYAATVQLASKYLVQNRVTGAIYEGPQQLYMLVGMCLFQNWED-GCAGKTRLEMVRGF 197
Y_pestisPestoidesF_NrdAg              YFGYAATVQLASKYLVQNRVTGAIYEGPQQLYMLVGMCLFQNWED-GCAGKTRLEMVRGF 197
B_aphidicolaBp_NrdAg                  NFSYAAVKQLEGKYLIQNRVTKKIYESAQFLYILISACLFSKYPN-----ETRMKYIKNF 196
B_aphidicolaSg_NrdAg                  NFSYAAVKQLEAKYLLQNRVTGQIYESAQFLYILISACLFSKYEK-----KVRMNYIQRF 196
C_RuthiaCm_NrdAg                      NFSYAAVKQLEGKYLVQNRVTGEIYESPQLLYVLIGMYLFQEYEA-----SIRMDIVKRF 196
T_crunogenaXCL2_NrdAg                 NFSYAAVKQLEGKYLVQNRVTGKIYESPQFIYMLIPMSLFAHYPNDAEPGKTRLDYVKRF 202
M_flagellatusKT_NrdAg                 NFSYAAVKQLEGKYLVQNRVTKKIYESPQFLYMLVAMCLFARYEG-----KTRLEYIKRF 213
N_gonorrhoeaeFA1090_NrdAg             SFSYAAVKQLEGKYLVQNRVTRQIYETPQFLYVLVAMCLFSKYPK-----EARLDYVKRF 198
N_meningitidis053442_NrdAg            SFSYAAVKQLEGKYLVQNRVTRQIYETPQFLYVLVAMCLFSKYPK-----ETRLDYVKRF 198
N_meningitidiMC58_NrdAg               SFSYAAVKQLEGKYLVQNRVTRQIYETPQFLYVLVAMCLFSKYPK-----EARLGYVKRF 198
P_atlanticaT6c_NrdAg                  DFSYAAVKQLEGKYLVQNRVTGEIYESAQFLYILVAACLFADYPK-----DSRMEYIKRF 196
I_loihiensisL2TR_NrdAg                NFSYAAVKQLENKYLVQNRVTGDIYESAQFLYILVAACLFANYPK-----STRMDYVKRF 196
P_tunicataD2_NrdAg                    TFSYAAVKQLEGKYLVQNRVTGEIYESAQFLYILVAASLFSKYPQ-----SIRLDYIKRF 196
PsychrobacterPRwf1_NrdAg              NLAYAAAEQMAGKYLVQDRVTKTVYESPQFLYVLVGMCLFASYDK-----SERMHYVKRF 198
P_arcticus2734_NrdAg                  NLAYAAVEQMAGKYLVQDRVSKTVFESPQFLYMLVGMCLFASYDK-----AERLDYVKRF 198
P_cryohalolentisK5_NrdAg              NLAYAAVEQMAGKYLVQDRVSKTVFESPQFLYMLVGMCLFSRYEK-----SERMNYVKRF 198
CBlochmanniaBPEN_NrdAg                NFSYAAVKQLEGKYLVQNRVTGEIYESAQFLYILTAACLFSKYPK-----KTRMHYIQQF 196
P_ingrahamii37_NrdAg                  KFSYAAVKQLEGKYLIQNRVTGDIYESAQFLYMMVAACLFSNYEP-----AIRLSYVRRF 196
MarinomonasMWYL1_NrdAg                DFSYAAVKQLEGKYLVQNRVTGAIYESPQFIYILVAACLFAEYDK-----KERLNLIRRF 195
S_denitrificansOS217_NrdAg            NFSYAAVKQLEGKYLVQNRVTHEVYESAQFLYILVAACLFARYPA-----ETRLKYIKDF 196
S_amazonensisSB2B_NrdAg               NFSYAAVKQLEGKYLVQNRVTHEVYESAQFLYILVAACLFAGYPK-----ETRLKYVKDF 216
S_frigidimarina400_NrdAg              NFSYAAVKQLEGKYLVQNRVTHEVYESAQFLYILVAACLFARYPK-----EIRLKYIKDF 196
S_oneidensisMR1_NrdAg                 NFSYAAVKQLEGKYLVQNRVTHEIYESAQFLYILVAACLFARYPK-----ETRLQYVKDF 196
ShewanellaANA3_NrdAg                  NFSYAAVKQLEGKYLVQNRVTHEIYESAQFLYILVAACLFARYPK-----ETRLQYVKDF 196
ShewanellaMR4_NrdAg                   NFSYAAVKQLEGKYLVQNRVTHEIYESAQFLYILVAACLFARYPK-----ETRLQYVKDF 196
ShewanellaMR7_NrdAg                   NFSYAAVKQLEGKYLVQNRVTHEIYESAQFLYILVAACLFARYPK-----ETRLQYVKDF 196
S_balticaOS155_NrdAg                  DFSYAAVKQLEGKYLVQNRVTHEIYESAQFLYILVAACLFARYPK-----ETRLKYVKDF 196
S_putrefaciensCN32_NrdAg              NFSYAAVKQLEGKYLVQNRVTHEIYESAQFLYILVAACLFARYPK-----ETRLKYVKDF 196
ShewanellaW3181_NrdAg                 NFSYAAVKQLEGKYLVQNRVTHEIYESAQFLYILVAACLFARYPK-----ETRLKYVKDF 196
S_sediminisHAWEB3_NrdAg               NFSYAAVKQLEGKYLVQNRVTHEIYESAQFLYILVAACLFAKYPK-----ETRLTYIKNF 196
S_loihicaPV4_NrdAg                    NFSYAAVKQLEGKYLVQNRVTHEVYESAQFLYILVAACLFANYPR-----ETRLQYVKDF 196
S_pealeana700345_NrdAg                NFSYAAVKQLEGKYLVQNRVTHEIYESAQFLYILVAACLFAKYPK-----ETRLQYVKDF 196
A_succinogenes130Z_NrdAg              NFSYAATKQLEGKYLVQNRVTGEIYESAQFLYILVAASLFSQYPK-----ATRLDYVRRF 196
M_succiniciproducensMBEL55E_NrdAg     TFSYAAVKQLEGKYLVQNRVTGEIYESAQFLYILVAASLFSKYPA-----ETRLDYIHRF 205
H_influenzae86028NP_NrdAg             TFSYAAVKQLEGKYLVQNRVTGEIYESAQFLYLLVSASLFSKYPK-----ETRLDYVKRF 196
H_influenzaeRdKW20_NrdAg              TFSYAAVKQLEGKYLVQNRVTGEIYESAQFLYLLVSASLFSKYPK-----ETRLDYVKRF 221
H_somnus129PT_NrdAg                   TFSYAAVKQLEGKYLVQNRVTGEIYESAQFLYLLVAASLFSRYPK-----ETRLDYIHRF 196
P_multocidaPm70_NrdAg                 TFSYAAVKQLEGKYLVQNRVTGEIYESAQFLYLLVAASLFSKYPT-----ETRLSYIRRF 196
A_pleuropneumoniae5bL20               DFSYAAVKQLEGKYLVQNRVTGEIYESAQFLYLLVSACLFSQYPK-----ATRLDYVKRF 196
H_ducreyi35000HP_NrdAg                NFSYAAVKQLEGKYLVQNRVTGEIYESAQFLYILVAACLFSSYPK-----ATRLDYIHRF 196
V_choleraeO395_NrdAg                  DFSYAAVKQLEGKYFVQNRVTGQIYESAQFLYILVAACLFAKYPK-----ETRLDYIKRF 196
V_vulnificusCMCP6_NrdAg               DFSYAAVKQLEGKYFVQNRVSGQIYESAQFLYILVAACLFAKYPK-----ETRLDYVKRF 196
VibrioEx25                            NFSYAAVKQLEGKYFVQNRVTGQIYESAQFLYILVAACLFAKYPK-----ATRLDYIKRF 196
V_campbelliiBAA1116_NrdAg             DFSYAAVKQLEGKYFVQNRVSGEIYESAQFLYILVSACLFANYPK-----ETRLDYIKRF 196
VibrioAND4_NrdAg                      DFSYAAVKQLEGKYFVQNRVSGEIYESAQFLYILVSACLFANYPK-----ETRLDYIKRF 196
V_fischeriES114_NrdAg                 DFSYAAVKQLEGKYFVQNRVTGAIYESAQFLYILVAACLFAKYPK-----DTRLDYIKRF 196
P_angustumS14_NrdAg                   NFSYAAVKQLEGKYLVQNRVSGEIFESAQFLYVLIAAALFNKYPK-----ETRLDYIKRF 196
PhotobacteriumSKA34_NrdAg             NFSYAAVKQLEGKYLVQNRVSGEIFESAQFLYILIAAALFNKYPK-----ETRLDYIKRF 196
C_psychrerythraea34H_NrdAg            NFSYAAMKQLEGKYLVQNRVSGEIYESAQFLYILVSACLFAKYPV-----ESRLRYVEDF 196
A_hydrophila7966_NrdAg                DFSYAAVKQLEGKYLVQNRVTGEIYESPQFLYILVAACLFSKYPK-----DTRLDYIKRF 206
A_salmonicidaA449_NrdAg               NFSYAAVKQLEGKYLVQNRVTGEIYESPQFLYILVAACLFSKYPK-----DTRLSYIKRF 194
B_cicadellinicolaHc_NrdAg             DFAYAAVKQLEGKYLVQNRVTGEIYESAQFLYILVAACLFSSYPR-----TNRLDYIRRF 196
C_sakazakiiBAA894                     NFSYAAVKQLEGKYLVQNRVTGEIYESAQFLYILVAACLFSNYPR-----ETRLDYVKRF 196
K_pneumoniaeMGH78578_NrdAg            NFSYAAVKQLEGKYLVQNRVTGEIYESAQFLYILVAACLFSNYPR-----ETRLDYIKRF 196
Enterobacter638_NrdAg                 NFSYAAVKQLEGKYLVQNRVTGEIYESAQFLYILVAACLFSNYPR-----ETRLEYVKRF 196
C_koseriBAA895_NrdAg                  TFSYAAVKQLEGKYLVQNRVTGEIYESAQFLYILVAACLFSNYPR-----ETRLDYVKRF 196
S_boydiiSb227_NrdAg                   TFSYAAVKQLEGKYLVQNRVTGEIYESAQFLYILVAACLFSNYPR-----ETRLQYVKRF 196
E_coliK12ssMG1655_NrdAg               TFSYAAVKQLEGKYLVQNRVTGEIYESAQFLYILVAACLFSNYPR-----ETRLQYVKRF 196
S_flexneri58401_NrdAg                 TFSYAAVKQLEGKYLVQNRVTGEIYESAQFLYILVAACLFSNYPR-----ETRLQYVKRF 196
E_coli536_NrdAg                       TFSYAAVKQLEGKYLVQNRVTGEIYESAQFLYILVAACLFSNYPR-----ETRLQYVKRF 196
S_flexneri2a2457T_NrdAg               TFSYAAVKQLEGKYLVQNRVTGEIYESAQFLYILVAACLFSNYPR-----ETRLQYVKRF 196
STyphimuriumLT2_NrdAg                 TFSYAAVKQLEGKYLVQNRVTGEIYESAQFLYILVAACLFSNYPR-----ETRLDYVKRF 196
S_entericaTyphiTy2_NrdAg              TFSYAAVKQLEGKYLVQNRVTGEIYESAQFLYILVVACLFSNYPR-----ETRLDYVKRF 196
S_proteamaculans568_NrdAg             NFSYAAVKQLEGKYLVQNRVSGEIYESAQFLYILVAACLFSNYPR-----DTRLDYVKRF 196
Y_pestisAngola_NrdAg                  NFSYAAVKQLEGKYLVQNRVSGEIYESAQFLYILVSACLFSGYPR-----ETRLDYIKRF 196
YpseudotuberculosisIP31758_NrdAg      NFSYAAVKQLEGKYLVQNRVSGEIYESAQFLYILVSACLFSGYPR-----ETRLDYIKRF 196
                                          .  *.  **  ::  .  : *  *   :                    :
```

```
SalmonellaphageFelixO1_NrdAg          YTYLSDLKINAPTPYLNGLRATKSGYASCCLIKANDTAESLGIAAKVAYDMTTKQAGIGM 247
Escherichiaphagerv5_NrdAg             YTYLSDLKINAPTPYLNGLRTGKTGYASCCVIKADDTAKSIGVAREVAYTMTVNQAGIGY 247
EnterophageT5_NrdAg                   YNAMSLHKINVPTPPLVGLRSSDRGFASCCLVDSTDTLDSIDTAEHIVFKMVAARAGIGY 288
VibriophageKVP40_NrdAg                YNAVSTQKLSLPTPIMAGVRTPTRQFSSCVKIESGDSLRSINKVSEAIVDYISKRAGIGV 250
AeromonasphageAeh1_NrdAg              YDVTSLGKLSLPTPILAGVRTPTRQFSSCVKIESGDELDSINATASSIVDYISKRAGIGI 250
EnterophageRB49_NrdAg                 YNAVANKQISLPTPIMAGVRTPTRQFSSCVVIESDDSLQGINETANAIIKYISKRAGIGI 252
EnterophagePhi1_NrdAg                 YNAVANKQISLPTPIMAGVRTPTRQFSSCVVIESDDSLQGINETANAIIKYISKRAGIGI 252
EnterobacteriaphageT4_NrdAg           YEAVSTRQISLPTPIMAGCRTPTRQFSSCVVIEAGDSLKSINKASASIVEYISKRAGIGI 253
EnterophageRB32_NrdAg                 YEAVSTRQISLPTPIMAGCRTPTRQFSSCVVIEAGDSLKSINKASASIVEYISKRAGIGI 253
EnterophageJS98_NrdAg                 YDAVSTRQVSLPTPIMAGARTPTRQFSSCVVIEAGDSLKSINKTANSIIEYISKRAGIGI 250
EnterophageRB69_NrdAg                 YDAVSTHQVSLPTPIMAGARTPTRQFSSCVVIEAGDSLKSINKTSASIIEYISKRAGIGI 250
ProchlorophagePSSM2_NrdAg             YDAISRHRINIPTPIMAGVRTPIRQFASCVLVDIDDTLDSIFSSDMAIGKYVAQRAGIGI 258
ProchlorococcusphagePSSM4_NrdAg       YDAISKHKINIPTPIMGGVRTPLRQFASCVLVDVDDTLDSIFSSDMAIGKYVAQRAGIGI 259
Synechococcusphagesyn9_NrdAg          YNAISKHRLNIPTPVMAGVRTTLRQFASCVLVDVDDTLDSIFSSDMAIGKYVAQRAGIGI 260
Y_pestisMicrotus91001_NrdAg           YDVTSTFKLSLPTPIMAGVRTPTRQFSSCVLIEAADSLKGISAASSAIIDYVSRRAGIGI 257
Y_pestisPestoidesF_NrdAg              YDVTSTFKLSLPTPIMAGVRTPTRQFSSCVLIEAADSLKGISAASSAIIDYVSRRAGIGI 257
B_aphidicolaBp_NrdAg                  YDAISTFKISLPTPIMAGLRTPTRQFSSCVLIESEDNLNSINATTSAIVKYVSQRAGIGI 256
B_aphidicolaSg_NrdAg                  YNAISTFKISLPTPIMSGVRTPTRQFSSCVLIECADNLNSINATTSAIVKYVSQRAGIGI 256
C_RuthiaCm_NrdAg                      YDAVSLFKISLPTPIMAGVRTPIRQFSSCVLIEVDDDLDSISAGAGAIVKYVSQRAGIGI 256
T_crunogenaXCL2_NrdAg                 YDAASNFKLSLPTPIMSGVRTPTRQFSSCVLIECGDSLDSINATSSSIVKYVSQRAGIGV 262
M_flagellatusKT_NrdAg                 YDAVSTFKISLPTPIMSGVRTPTRQFSSCVLIECDDSLDSINATTSAIVKYVSQRAGIGV 273
N_gonorrhoeaeFA1090_NrdAg             YDAVSTFKVSLPTPIMSGVRTPTRQFSSCVLIECDDSLDSINATTSAIVKYVSQRAGIGI 258
N_meningitidis053442_NrdAg            YDAVSTFKVSLPTPIMSGVRTPTRQFSSCVLIECDDSLDSINATTSAIVKYVSQRAGIGI 258
N_meningitidiMC58_NrdAg               YDAVSTFKVSLPTPIMSGVRTPTRQFSSCVLIECDDSLDSINATTSAIVKYVSQRAGIGI 258
P_atlanticaT6c_NrdAg                  YDATSKFKISLPTPIMAGVRTPTRQFSSCVLIEAGDSLDSINATASAIVKYVSQRAGIGV 256
I_loihiensisL2TR_NrdAg                YDAASEFKISLPTPIMSGVRTPTRQFSSCVLIEAGDSLDSINATSSAIVKYVSQRAGIGI 256
P_tunicataD2_NrdAg                    YDAVSQFKISLPTPIMAGVRTPTRQFSSCVLIECGDSLDSINATSSAIVKYVSQRAGIGI 256
PsychrobacterPRwf1_NrdAg              YDAASTFKISLPTPIMSGVRTPSRQFSSCVLIECGDSLDSINATTSAIVRYVSQRAGIGI 258
P_arcticus2734_NrdAg                  YDATSEFKISLPTPIMSGVRTPSRQFSSCVLIECGDSLDSINATTSAIVRYVSQRAGIGI 258
P_cryohalolentisK5_NrdAg              YDATSEFKISLPTPIMSGVRTPSRQFSSCVLIECGDSLDSINATTSAIVRYVSQRAGIGI 258
CBlochmanniaBPEN_NrdAg                YDAISTFKISLATPIMSGVRTPTRQFSSCVLIECDDSLDSINATSSSIVKYVAQRAGIGI 256
P_ingrahamii37_NrdAg                  YDAVSQFKISLPTPIMAGVRTPTRQFSSCVLIETGDSLDSINATSSAIVKYVSQRAGIGI 256
MarinomonasMWYL1_NrdAg                YDAVSKFKISLPTPIMSGIRTPTRQFSSCVLIECGDSLDSINATSSAIVKYVSQRAGIGV 255
S_denitrificansOS217_NrdAg            YDATSNFKISLPTPIMAGVRTPTRQFSSCVLIECDDSLDSINATASSIVKYVSQRAGIGI 256
S_amazonensisSB2B_NrdAg               YDAVSTFKISLPTPIMSGVRTPTRQFSSCVLIECGDSLDSINATASSIVKYVSQRAGIGV 276
S_frigidimarina400_NrdAg              YDATSLFKISLPTPIMAGVRTPTRQFSSCVLIECDDSLDSINATASSIVKYVSQRAGIGI 256
S_oneidensisMR1_NrdAg                 YDAVSTFKISLPTPIMAGVRTPTRQFSSCVLIECGDSLDSINATASSIVKYVSQRAGIGI 256
ShewanellaANA3_NrdAg                  YDAVSTFKISLPTPIMAGVRTPTRQFSSCVLIECGDSLDSINATASSIVKYVSQRAGIGI 256
ShewanellaMR4_NrdAg                   YDAVSTFKISLPTPIMAGVRTPTRQFSSCVLIECGDSLDSINATASSIVKYVSQRAGIGI 256
ShewanellaMR7_NrdAg                   YDAVSTFKISLPTPIMAGVRTPTRQFSSCVLIECGDSLDSINATASSIVKYVSQRAGIGI 256
S_balticaOS155_NrdAg                  YDAVSTFKISLPTPIMAGVRTPTRQFSSCVLIECGDSLDSINATASSIVKYVSQRAGIGI 256
S_putrefaciensCN32_NrdAg              YDAVSTFKISLPTPIMAGVRTPTRQFSSCVLIECGDSLDSINATASSIVKYVSQRAGIGI 256
ShewanellaW3181_NrdAg                 YDAVSTFKISLPTPIMAGVRTPTRQFSSCVLIECGDSLDSINATASSIVKYVSQRAGIGI 256
S_sediminisHAWEB3_NrdAg               YDAVSTFKISLPTPIMAGVRTPTRQFSSCVLIECGDSLDSINATASSIVKYVSQRAGIGV 256
S_loihicaPV4_NrdAg                    YDATSTFKISLPTPIMSGVRTPTRQFSSCVLIECGDSLDSINATASSIVKYVSQRAGIGV 256
S_pealeana700345_NrdAg                YDATSTFKISLPTPIMAGVRTPTRQFSSCVLIECGDSLDSINATSSSIVKYVSQRAGIGV 256
A_succinogenes130Z_NrdAg              YDATSSFKISLPTPIMAGVRTPTRQFSSCVLIECGDSLDSINATASAIVKYVSQRAGIGV 256
M_succiniciproducensMBEL55E_NrdAg     YDAISTFKISLPTPIMSGVRTPTRQFSSCVLIECDDSLDSINATSSAIIKYVSQRAGIGI 265
H_influenzae86028NP_NrdAg             YDATSTFKISLPTPIMAGVRTPTRQFSSCVLIECDDSLDSINATASAIVKYVSQRAGIGI 256
H_influenzaeRdKW20_NrdAg              YDATSTFKISLPTPIMAGVRTPTRQFSSCVLIECDDSLDSINATASAIVKYVSQRAGIGI 281
H_somnus129PT_NrdAg                   YDATSTFKISLPTPIMAGVRTPTRQFSSCVLIECGDSLDSINATASAIVKYVSQRAGIGV 256
P_multocidaPm70_NrdAg                 YDATSTFKISLPTPIMAGVRTPTRQFSSCVLIECGDSLDSINATASAIVKYVSQRAGIGV 256
A_pleuropneumoniae5bL20               YDATSTFKISLPTPIMAGVRTPTRQFSSCVLIECGDSLDSINATSAAIIKYVSQRAGIGI 256
H_ducreyi35000HP_NrdAg                YDATSTFKISLPTPIMAGVRTPTRQFSSCVLIECGDSLDSINATSSAVIKYVSQRAGIGI 256
V_choleraeO395_NrdAg                  YDATSTFKISLPTPIMSGVRTPTRQFSSCVLIECGDSLDSINATASSIVRYVSQRAGIGI 256
V_vulnificusCMCP6_NrdAg               YDATSTFKISLPTPIMSGVRTPTRQFSSCVLIECGDSLDSINATASSIVRYVSQRAGIGI 256
VibrioEx25                            YDATSTFKISLPTPIMSGVRTPTRQFSSCVLIECGDSLDSINATASSIVRYVSQRAGIGI 256
V_campbelliiBAA1116_NrdAg             YDATSTFKISLPTPIMSGVRTPTRQFSSCVLIECGDSLDSINATASSIVRYVSQRAGIGI 256
VibrioAND4_NrdAg                      YDATSTFKISLPTPIMSGVRTPTRQFSSCVLIECGDSLDSINATASSIVRYVSQRAGIGI 256
V_fischeriES114_NrdAg                 YDATSQFKISLPTPIMSGVRTPTRQFSSCVLIECDDSLDSINATAGSIVRYVSQRAGIGI 256
P_angustumS14_NrdAg                   YDASSLFKISLPTPIMSGVRTPTRQFSSCVLIECDDSLDSINATASSIVRYVSQRAGIGI 256
PhotobacteriumSKA34_NrdAg             YDASSLFKISLPTPIMSGVRTPTRQFSSCVLIECDDSLDSINATASSIVRYVSQRAGIGI 256
C_psychrerythraea34H_NrdAg            YHAISRFKISLPTPIMAGVRTPTRQFSSCVLIECGDSLDSINATSSAIVKYVSQRAGIGV 256
A_hydrophila7966_NrdAg                YDAVSTFKISLPTPIMSGVRTPTRQFSSCVLIECDDSLDSINATASAIVRYVSQRAGIGI 266
A_salmonicidaA449_NrdAg               YDAVSTFKISLPTPIMSGVRTPTRQFSSCVLIECDDSLDSINATASAIVRYVSQRAGIGI 254
B_cicadellinicolaHc_NrdAg             YDAVSTFKISLPTPIMSGVRTPTRQFSSCVLIECGDSLNSINATSSAIVKYVSRRAGIGI 256
C_sakazakiiBAA894                     YDAVSTFKISLPTPIMSGVRTPTRQFSSCVLIECGDSLDSINATSSAIVKYVSQRAGIGI 256
K_pneumoniaeMGH78578_NrdAg            YDAVSTFKISLPTPIMSGVRTPTRQFSSCVLIECGDSLDSINATSSAIVKYVSQRAGIGI 256
Enterobacter638_NrdAg                 YDAVSIFKISLPTPIMSGVRTPTRQFSSCVLIECGDSLDSINATSSAIVKYVSQRAGIGI 256
C_koseriBAA895_NrdAg                  YDAVSTFKISLPTPIMSGVRTPTRQFSSCVLIECGDSLDSINATSSAIVKYVSQRAGIGI 256
S_boydiiSb227_NrdAg                   YDAVSTFKISLPTPIMSGVRTPTRQFSSCVLIECGDSLDSINATSSAIVKYVSQRAGIGI 256
E_coliK12ssMG1655_NrdAg               YDAVSTFKISLPTPIMSGVRTPTRQFSSCVLIECGDSLDSINATSSAIVKYVSQRAGIGI 256
S_flexneri58401_NrdAg                 YDAVSTFKISLPTPIMSGVRTPTRQFSSCVLIECGDSLDSINATSSAIVKYVSQRAGIGI 256
E_coli536_NrdAg                       YDAVSTFKISLPTPIMSGVRTPTRQFSSCVLIECGDSLDSINATSSAIVKYVSQRAGIGI 256
S_flexneri2a2457T_NrdAg               YDAVSTFKISLPTPIMSGVRTPTRQFSSCVLIECGDSLDSINATSSAIVKYVSQRAGIGI 256
STyphimuriumLT2_NrdAg                 YDAVSTFKISLPTPIMSGVRTPTRQFSSCVLIECGDSLDSINATSSAIVKYVSQRAGIGI 256
S_entericaTyphiTy2_NrdAg              YDAVSTFKISLPTPIMSGVRTPTRQFSSCVLIECGDSLDSINATSSAIVKYVSQRAGIGI 256
S_proteamaculans568_NrdAg             YDAISTFKISLPTPIMSGVRTPTRQFSSCVLIECGDSLDSINATSSAIVKYVSQRAGIGI 256
Y_pestisAngola_NrdAg                  YDAISTFKISLPTPIMAGVRTPTRQFSSCVLIECGDSLDSINATSSAIVKYVSQRAGIGI 256
YpseudotuberculosisIP31758_NrdAg      YDAISTFKISLPTPIMAGVRTPTRQFSSCVLIECGDSLDSINATSSAIVKYVSQRAGIGI 256
                                      *   :  ::.  ** : * *:    ::**  :. *   .:            :****
```

```
SalmonellaphageFelixO1_NrdAg          LME-TRTIGDGIRQNTIEHMGKLPYYKLVRSSVEANKQ-KSRGGSANNFYTALDPQIEDL 305
Escherichiaphagerv5_NrdAg             YLS-SRSAGDGVRGNTIKHMGKIPYYRGIDAGVKENRQ-QSRGGSATVSFLALDPQIDEL 305
EnterophageT5_NrdAg                   HLE-SRSIADPVRNGAFPHSGKLPYYRHIDRSVKANTQ-QTRGGSATVSYPYFDPEIIQL 346
VibriophageKVP40_NrdAg                NGGAIRAEGSVVRAGEIRHTGVIPFWKHFQTAVKSCSQGGVRGGAATLTYPIWHLEVEKL 310
AeromonasphageAeh1_NrdAg              NGGAIRAEGSRIRSGEVRHTGIIPFWKYFQTAVKSCSQGGVRGGAATLYYPMWHLEVEKI 310
EnterophageRB49_NrdAg                 NAGMIRAEGSKIGHGEVRHTGVIPFWKHFQTAVKSCSQGGIRGGAATLYYPMWHLEFESL 312
EnterophagePhi1_NrdAg                 NAGMIRAEGSKIGHGEVRHTGVIPFWKHFQTAVKSCSQGGIRGGAATLYYPMWHLEFESL 312
EnterobacteriaphageT4_NrdAg           NVGMIRAEGSKIGMGEVRHTGVIPFWKHFQTAVKSCSQGGIRGGAATAYYPIWHLEVENL 313
EnterophageRB32_NrdAg                 NVGMIRAEGSKIGMGEVRHTGVIPFWKHFQTAVKSCSQGGIRGGAATAYYPIWHLEVENL 313
EnterophageJS98_NrdAg                 NVGMIRAEGSRIGMGEVKHTGVIPFWKHFQTAVKSCSQGGIRGGAATAYYPIWHLEVENL 310
EnterophageRB69_NrdAg                 NVGMLRAEGSKIGMGEVRHTGVIPFWKHFQTAVKSCSQGGIRGGAATAYYPMWHLEVENL 310
ProchlorophagePSSM2_NrdAg             NAGRIRGINAKIRGGEVQHTGVVPFLKKFESTVRCCTQNGIRGGSATVHFPIWHQEISDI 318
ProchlorococcusphagePSSM4_NrdAg       NAGRIRGINSKIRGGEVQHTGVVPFLKKFEATVRCCTQNGIRGGSATVHFPIWHQEIEDI 319
Synechococcusphagesyn9_NrdAg          NAGRIRGINSKIRDGEVVHTGVVPFLKKFESTVRCCTQNGIRGGSATVHFPIWHTEIEDI 320
Y_pestisMicrotus91001_NrdAg           GFGRLRALGSEIRNGEATHTGVIPFLKHFQTAVKSCSQGGVRSGAATAFYPIWHLEVESL 317
Y_pestisPestoidesF_NrdAg              GFGRLRALGSEIRNGEATHTGVIPFLKHFQTAVKSCSQGGVRSGAATAFYPIWHLEVESL 317
B_aphidicolaBp_NrdAg                  NAGRIRALGSPIRNGDTLHTGCIPFYKHFQSAVKSCSQGGVRGGAATIFYPIWHFEIESL 316
B_aphidicolaSg_NrdAg                  NAGQIRALGSPIRNGEAFHTGCIPFYKHFQSAVKSCSQGGVRGGAATVFYPIWHLEIESL 316
C_RuthiaCm_NrdAg                      NGGKIRAIGSPIRGGEAIHTGCIPFYKHFHTAVKSCSQGGVRGGAATLFYPLWHLEVESL 316
T_crunogenaXCL2_NrdAg                 NVGQIRALGSEIRGGEAYHTGMIPFIKHFQTAVKSCSQGGVRGGAATLFFPIWHLEVESL 322
M_flagellatusKT_NrdAg                 NAGRIRALGSEIRGGEAQHTGCLPFYKLFQAAVKSCSQGGVRGGAATLFYPLWHLEVESL 333
N_gonorrhoeaeFA1090_NrdAg              NAGRIRGLDSEIRGGEARHTGCIPFFKMFQAAVKSCSQGGVRGGAATLFYPLWHIEAESL 318
N_meningitidis053442_NrdAg            NAGRIRGLDSEIRGGEARHTGCIPFFKMFQAAVKSCSQGGVRGGAATLFYPLWHIEAESL 318
N_meningitidiMC58_NrdAg               NAGRIRGLDSEIRGGEARHTGCIPFFKMFQAAVKSCSQGGVRGGAATLFYPLWHIEAESL 318
P_atlanticaT6c_NrdAg                  NAGRIRALGSPIRGGEAFHTGCIPFYKYFQTAVKSCSQGGVRGGAATLFYPLWHMEVESL 316
I_loihiensisL2TR_NrdAg                NAGRIRALGSPIRNGEAFHTGCIPFYKYFQTAVKSCSQGGVRGGAATLFYPLWHLEVESL 316
P_tunicataD2_NrdAg                    NAGRIRALGSPIRNGEAYHTGCIPFYKHFQTAVKSCSQGGVRGGAATLFYPLWHLEVENL 316
PsychrobacterPRwf1_NrdAg              NAGNIRALGSPIRGGEAQHTGCIPFYKLFQTAVKCCSQGGVRGGAATLFYPLWHLEVESL 318
P_arcticus2734_NrdAg                  NAGRIRALGSPIRGGEAQHTGCIPFYKLFQTAVKCCSQGGVRGGAATLFYPLWHLEVESL 318
P_cryohalolentisK5_NrdAg              NAGRIRALGSPIRGGEAQHTGCIPFYKLFQTAVKCCSQGGVRGGAATLFYPLWHLEVESL 318
CBlochmanniaBPEN_NrdAg                NAGRIRALGSPIRNGEAFHTGCIPFYKHFQTAVKSCSQGGVRGGAATLFYPIWHLEIENL 316
P_ingrahamii37_NrdAg                  NAGRIRALGSAIRGGEAFHTGCIPFYKHFQTAVKSCSQGGVRGGAATVFYPVWHLEVESL 316
MarinomonasMWYL1_NrdAg                NAGAIRALGSPIRGGEAFHTGCIPFYKHFQTAVKSCSQGGVRGGAATVFYPIWHLEVESL 315
S_denitrificansOS217_NrdAg            NAGRIRALGSPIRGGEAFHTGCLPFYKYFQTAVKSCSQGGVRGGAATLFYPIWHLEVESL 316
S_amazonensisSB2B_NrdAg               NAGRIRALGSPIRGGEAFHTGCLPFYKYFQTAVKSCSQGGVRGGAATLFYPIWHLEVESL 336
S_frigidimarina400_NrdAg              NAGRIRALGSPIRGGEAFHTGCLPFYKYFQTAVKSCSQGGVRGGAATLFYPLWHLEVESL 316
S_oneidensisMR1_NrdAg                 NAGRIRALGSPIRGGEAFHTGCLPFYKYFQTAVKSCSQGGVRGGAATLFYPIWHLEVESL 316
ShewanellaANA3_NrdAg                  NAGRIRALGSPIRGGEAFHTGCLPFYKYFQTAVKSCSQGGVRGGAATLFYPIWHLEVESL 316
ShewanellaMR4_NrdAg                   NAGRIRALGSPIRGGEAFHTGCLPFYKYFQTAVKSCSQGGVRGGAATLFYPIWHLEVESL 316
ShewanellaMR7_NrdAg                   NAGRIRALGSPIRGGEAFHTGCLPFYKYFQTAVKSCSQGGVRGGAATLFYPIWHLEVESL 316
S_balticaOS155_NrdAg                  NAGRIRALGSPIRGGEAFHTGCLPFYKYFQTAVKSCSQGGVRGGAATLFYPIWHLEVESL 316
S_putrefaciensCN32_NrdAg              NAGRIRALGSPIRGGEAFHTGCLPFYKYFQTAVKSCSQGGVRGGAATLFYPIWHLEVESL 316
ShewanellaW3181_NrdAg                 NAGRIRALGSPIRGGEAFHTGCLPFYKYFQTAVKSCSQGGVRGGAATLFYPIWHLEVESL 316
S_sediminisHAWEB3_NrdAg               NAGRIRALGSPIRGGEAFHTGCLPFFKYFQTAVKSCSQGGVRGGAATLFYPLWHLEVESL 316
S_loihicaPV4_NrdAg                    NAGRIRALGSPIRGGEAFHTGCLPFYKYFQTAVKSCSQGGVRGGAATLFYPMWHLEVESL 316
S_pealeana700345_NrdAg                NAGRIRALGSPIRGGEAFHTGCLPFYKYFQTAVKSCSQGGVRGGAATLFYPMWHLEVESL 316
A_succinogenes130Z_NrdAg              NAGQIRALGSPIRGGEAFHTGCIPFYKYFQTAVKSCSQGGVRGGAATVYYPIWHLEVESL 316
M_succiniciproducensMBEL55E_NrdAg     NAGAIRALGSPIRDGEAFHTGCIPFYKHFQTAVKSCSQGGVRGGAATVYFPMWHLEVESL 325
H_influenzae86028NP_NrdAg             NAGAIRALGSEIRGGEAFHTGCIPFYKYFQTAVKSCSQGGVRGGAATLYYPIWHLEAENL 316
H_influenzaeRdKW20_NrdAg              NAGAIRALGSEIRGGEAFHTGCIPFYKYFQTAVKSCSQGGVRGGAATLYYPIWHLEAENL 341
H_somnus129PT_NrdAg                   NAGAIRALGSPIRNGEAFHTGCIPFYKHFQTAVKSCSQGGVRGGAATVYYPIWHLEVESL 316
P_multocidaPm70_NrdAg                 NAGAIRALGSPIRGGEAFHTGCIPFYKYFQTAVKSCSQGGVRGGAATVYYPIWHLEVESL 316
A_pleuropneumoniae5bL20               NAGAIRALGSPIRQGEAFHTGCIPFYKHFQTAVKSCSQGGVRGGAATVFYPLWHLEAESL 316
H_ducreyi35000HP_NrdAg                NAGAIRALGSPIRNGEAFHTGCIPFYKHFQTAVKSCSQGGVRGGAATVFYPIWHLEVESL 316
V_choleraeO395_NrdAg                  NAGRIRALGSEIRGGEAFHTGCIPFYKYFQTAVKCCSQGGVRGGAATVFYPLWHGEAESL 316
V_vulnificusCMCP6_NrdAg               NAGRIRALGSEIRGGEAFHTGCIPFYKYFQTAVKCCSQGGVRGGAATVFYPMWHGEARSL 316
VibrioEx25                            NAGRIRALGSEIRGGEAFHTGCIPFYKYFQTAVKCCSQGGVRGGAATVFYPMWHGEARSL 316
V_campbelliiBAA1116_NrdAg             NAGRIRALGSEIRGGEAFHTGCIPFYKYFQTAVKCCSQGGVRGGAATVFYPLWHGEARSL 316
VibrioAND4_NrdAg                      NAGRIRALGSEIRGGEAFHTGCIPFYKYFQTAVKCCSQGGVRGGAATVFYPLWHGEARSL 316
V_fischeriES114_NrdAg                 NAGRIRALGSEIRDGEAFHTGCIPFYKYFQTAVKCCSQGGVRGGAATLFYPIWHREVESL 316
P_angustumS14_NrdAg                   NAGRIRALGSEIRDGEAFHTGCIPFYKYFQTAVKCCSQGGVRGGAATLFYPLWHREVESL 316
PhotobacteriumSKA34_NrdAg             NAGRIRALGSEIRDGEAFHTGCIPFYKYFQTAVKCCSQGGVRGGAATLFYPIWHREVESL 316
C_psychrerythraea34H_NrdAg            NAGRIRALGSTIRNGEAFHTGCIPFFKHFQTAVKSCSQGGVRGGAATLFYPLWHLEVESL 316
A_hydrophila7966_NrdAg                NAGRIRGLGSPIRGGEAFHTGCIPFYKYFQTAVKCCSQGGVRGGAATLFYPLWHTEVESL 326
A_salmonicidaA449_NrdAg               NAGRIRGLGSPIRGGEAFHTGCIPFYKYFQTAVKCCSQGGVRGGAATLFYPLWHTEVESL 314
B_cicadellinicolaHc_NrdAg             NAGRIRAIGSQIRGGEAFHTGCIPFYKHFQTAVKSCSQGGVRGGAATLFYPMWHLEVESL 316
C_sakazakiiBAA894                     NAGRIRALGSPIRGGEAFHTGCIPFYKHFQTAVKSCSQGGVRGGAATLFYPMWHLEVESL 316
K_pneumoniaeMGH78578_NrdAg            NAGRIRALGSPIRGGEAFHTGCIPFYKHFQTAVKSCSQGGVRGGAATLFYPMWHLEVESL 316
Enterobacter638_NrdAg                 NAGRIRALGSPIRGGEAFHTGCIPFYKHFQTAVKSCSQGGVRGGAATLFYPMWHLEVESL 316
C_koseriBAA895_NrdAg                  NAGRIRALGSPIRGGEAFHTGCIPFYKHFQTAVKSCSQGGVRGGAATLFYPMWHLEVESL 316
S_boydiiSb227_NrdAg                   NAGRIRALGSPIRGGEAFHTGCIPFYKHFQTAVKSCSQGGVRGGAATLFYPMWHLEVESL 316
E_coliK12ssMG1655_NrdAg               NAGRIRALGSPIRGGEAFHTGCIPFYKHFQTAVKSCSQGGVRGGAATLFYPMWHLEVESL 316
S_flexneri58401_NrdAg                 NAGRIRALGSPIRGGEAFHTGCIPFYKHFQTAVKSCSQGGVRGGAATLFYPMWHLEVESL 316
E_coli536_NrdAg                       NAGRIRALGSPIRGGEAFHTGCIPFYKHFQTAVKSCSQGGVRGGAATLFYPMWHLEVESL 316
S_flexneri2a2457T_NrdAg               NAGRIRALGSPIRGGEAFHTGCIPFYKHFQTAVKSCSQGGVRGGAATLFYPMWHLEVESL 316
STyphimuriumLT2_NrdAg                 NAGRIRALGSPIRGGEAFHTGCIPFYKHFQTAVKSCSQGGVRGGAATLFYPMWHLEVESL 316
S_entericaTyphiTy2_NrdAg              NAGRIRALGSPIRGGEAFHTGCIPFYKHFQTAVKSCSQGGVRGGAATLFYPMWHLEVESL 316
S_proteamaculans568_NrdAg             NAGRIRALGSPIRGGEAFHTGCIPFYKHFQTAVKSCSQGGVRGGAATLFYPMWHLEVESL 316
Y_pestisAngola_NrdAg                  NAGRIRALGSPIRGGEAFHTGCIPFYKHFQTAVKSCSQGGVRGGAATLFYPMWHLEVESL 316
YpseudotuberculosisIP31758_NrdAg      NAGRIRALGSPIRGGEAFHTGCIPFYKHFQTAVKSCSQGGVRGGAATLFYPMWHLEVESL 316
                                           *     :      * * :*: : .   *.   *   *.*:*.  :   . :  .:
```

```
SalmonellaphageFelixO1_NrdAg          LRLKHPTTVPSKRINEMDYSFGTNDYFWQCVQYDTDWLLFSYKDAPK-LYDMFYTASADE 364
Escherichiaphagerv5_NrdAg             IRLRNPMTVASKRINTMDYSIGINRSFINRVAKNLDWMLVSYKHAPQ-LHEGIFTMSMEE 364
EnterophageT5_NrdAg                   MQVKQQRATDENKIDKMDYSLSFNNLLLKRYLKNEDITLMSYFYAPE-VHEAFYSDDEAK 405
VibriophageKVP40_NrdAg                LVLKNNKGVEENRIRHLDYSVSINRLMMQRLVNNEYITLFSPDERHGALYGEYH-TNPEE 369
AeromonasphageAeh1_NrdAg              LVLKNNKGVEENRIRQLDYGIQINKLFFERLKNDDYITLFSPDVQNGELYDLFY-KDPVK 369
EnterophageRB49_NrdAg                 IVLKNNKGVEDNRIRHLDYGVQINNLMVERFLKRDYITLFSPDVAGGLLWDLYF-RDEEG 371
EnterophagePhi1_NrdAg                 IVLKNNKGVEDNRIRHLDYGVQINNLMVERFLKRDYITLFSPDVAGGLLWDLYF-RDEEG 371
EnterobacteriaphageT4_NrdAg           LVLKNNKGVEENRIRHMDYGVQLNDLMMERFGKNDYITLFSPHEMGGELYYSYF-KDQDR 372
EnterophageRB32_NrdAg                 LVLKNNKGVEENRIRHMDYGVQLNDLMMERFGKNDYITLFSPHEMGGELYYSYF-KDQDR 372
EnterophageJS98_NrdAg                 LVLKNNKGVDENRIRHMDYGIQLNDLMMERFGKNDYITLFSPHEMGGELYYSYF-EDQER 369
EnterophageRB69_NrdAg                 LVLKNNKGVDENRIRHMDYGIQINDLMMERLGKNDYITLFSPHEMSGELYYSFF-EDQDK 369
ProchlorophagePSSM2_NrdAg             LVLKNNKGTEDNRVRKLDYSIQLSKLFYERFIQNLPITLFSPHDVPG-LYDAFG-S--DT 374
ProchlorococcusphagePSSM4_NrdAg       IVLKNNKGTEDNRVRKLDYSVQISGLFYQRFIEDSTITLFSPHNVPD-LYDAFG-T--DA 375
Synechococcusphagesyn9_NrdAg          IVLKNNKGTEDNRVRKLDYSIQLSKIFYERFIANKDISLFSPHDVPG-LYDAFG-T--PE 376
Y_pestisMicrotus91001_NrdAg           LVLKNNRGIEENRVRHLDYGVMINRLMYRRLVHNENITLFSPHDVPG-LYDAFF-VDQDK 375
Y_pestisPestoidesF_NrdAg              LVLKNNRGIEENRVRHLDYGVMINRLMYRRLVHNENITLFSPHDVPG-LYDAFF-VDQDK 375
B_aphidicolaBp_NrdAg                  LVLKNNRGIEENRVRHMDYSVQLNKLMYQRLILGEHITLFSPSDVPN-LYDSFF-NNQEK 374
B_aphidicolaSg_NrdAg                  LVLKNNRGIEENRVRHLDYAVQINKLMYQRMLLGQKITLFSPSDVPK-LYEFFF-SDQKK 374
C_RuthiaCm_NrdAg                      LVLKNNTGTDENRIRHLDYGVQFNGLMYQRFLKDDDITLFSPHDVPG-LYDAFG-ANQEE 374
T_crunogenaXCL2_NrdAg                 IVLKNNRGVEENRARHLDYGVQVNKLMYQRMIEGGNITLFSPSDVPG-LYDAFF-EDQEE 380
M_flagellatusKT_NrdAg                 LVLKNNRGVEENRVRHLDYGVQINRLLYQRLIKGGNITLFSPHDVPG-LYDAFF-ADQDE 391
N_gonorrhoeaeFA1090_NrdAg             LVLKNNRGVEDNRIRQLDYGVQINRLLYTRLIKGGNITLFSPNEVSG-LYEAFF-ADQDE 376
N_meningitidis053442_NrdAg            LVLKNNRGVEDNRIRQLDYGVQINRLLYTRLIKGGNITLFSPNEVPG-LYEAFF-ADQDE 376
N_meningitidiMC58_NrdAg               LVLKNNRGVEDNRIRQLDYGVQINRLLYTRLIKGGNITLFSPNEVPG-LYEAFF-ADQDE 376
P_atlanticaT6c_NrdAg                  LVLKNNRGVEENRVRHLDYGVQFNKLMYQRLIKDQHISLFSPSDVPG-LYDAFF-ADQPK 374
I_loihiensisL2TR_NrdAg                LVLKNNRGVEENRVRHLDYGVQFNRTMYQRLIKDDYITLFSPSDVPG-LYDAFF-EDQDE 374
P_tunicataD2_NrdAg                    LVLKNNRGVDDNRVRHLDYGVQFNKVMYSRLIKDDYITLFSPSDVPG-LYDAFF-EDQAL 374
PsychrobacterPRwf1_NrdAg              LVLKNNRGVEDNRVRHMDYGVQINRPLYTRLIKGEDIALFSPSDVPG-LYEAFF-ADQDK 376
P_arcticus2734_NrdAg                  LVLKNNRGVEDNRVRHMDYGVQLNKTMYTRLIKGQDISLFSPGDTPG-LYDAFF-EDQDK 376
P_cryohalolentisK5_NrdAg              LVLKNNRGVEDNRVRHMDYGVQLNKTMYTRLIKGQDISLFSPGDTPG-LYDAFF-EDQDK 376
CBlochmanniaBPEN_NrdAg                LVLKNNRGIDSNRIRHLDYGVQINKLMYQRLIEGKYITLFSPSDVPG-LYEAFF-NNQNE 374
P_ingrahamii37_NrdAg                  LVLKNNRGIDSNRIRHLDYGVQLSRLFYQRLIDKKNVALFSPSDVPG-LYDAFF-EDQDE 374
MarinomonasMWYL1_NrdAg                LVLKNNRGVEENRVRHLDYGVQFNRLMYQRLIKGGNITLFSPSDVPG-LYDAFF-ADQEE 373
S_denitrificansOS217_NrdAg            LVLKNNRGVDDNRVRHLDYGVQLNKLMYQRLIKGQDISLFSPSDVPG-MYDAFF-ADQVE 374
S_amazonensisSB2B_NrdAg               LVLKNNRGVEDNRVRHLDYGVQLNKLMYQRLIKGENITLFSPSDVPG-LYDAFF-ENQDE 394
S_frigidimarina400_NrdAg              LVLKNNRGVDDNRVRHLDYGVQLNKLMYQRLIKGEYISLFSPSDVPG-LYDAFF-EDQAE 374
S_oneidensisMR1_NrdAg                 LVLKNNRGVEDNRIRHLDYGVQLNKLMYQRLIKGESISLFSPSDVPG-LYDAFF-EDQAE 374
ShewanellaANA3_NrdAg                  LVLKNNRGVEDNRIRHLDYGVQLNKLMYQRLIKGENISLFSPSDVPG-LYDAFF-EDQDE 374
ShewanellaMR4_NrdAg                   LVLKNNRGVEDNRIRHLDYGVQLNKLMYQRLIKGENISLFSPSDVPG-LYDAFF-EDQDE 374
ShewanellaMR7_NrdAg                   LVLKNNRGVEDNRIRHLDYGVQLNKLMYQRLIKGENISLFSPSDVPG-LYDAFF-EDQDE 374
S_balticaOS155_NrdAg                  LVLKNNRGVEDNRIRHLDYGVQLNKVMYQRLIKGENISLFSPSDVPG-LYDAFF-EDQNE 374
S_putrefaciensCN32_NrdAg              LVLKNNRGVEDNRIRHLDYGVQLNKVMYQRLIKGENISLFSPSDVPG-LYDAFF-EDQAE 374
ShewanellaW3181_NrdAg                 LVLKNNRGVEDNRIRHLDYGVQLNKVMYQRLIKGENISLFSPSDVPG-LYDAFF-EDQAE 374
S_sediminisHAWEB3_NrdAg               LVLKNNRGVEDNRVRQLDYGVQLNKVMYQRLIKGGMITLFSPSDVPG-LYDAFF-EDQDE 374
S_loihicaPV4_NrdAg                    LVLKNNRGVEDNRIRHLDYGVQLNKLMYQRLIQGGMITLFSPSDVPG-LYDAFF-EDQAE 374
S_pealeana700345_NrdAg                LVLKNNRGVDDNRIRHLDYGVQINKVMYQRLIQGGMISLFSPSDVPG-MYDAFF-EDQEE 374
A_succinogenes130Z_NrdAg              LVLKNNRGVEDNRVRHMDYGVQINKLMYQRLIKGGDITLFSPSDVPG-LYDAFF-ADQDK 374
M_succiniciproducensMBEL55E_NrdAg     VVLKNNRGVEENRARHMDYGVQINRTMYQRLIKGGDITLFSPSDVPG-LYEAFF-ADQAK 383
H_influenzae86028NP_NrdAg             LVLKNNRGVEDNRVRHMDYGVQLNKLMYQRLIKGGEITLFSPSDVPG-LYEAFF-ADQDK 374
H_influenzaeRdKW20_NrdAg              LVLKNNRGVEDNRVRHMDYGVQLNKLMYQRLIKGSEITLFSPSDVPG-LYEAFF-ADQDK 399
H_somnus129PT_NrdAg                   LVLKNNRGVEDNRVRHMDYGVQLNKLMYQRLIKGGEITLFSPSDVPG-LYEAFF-ADQNK 374
P_multocidaPm70_NrdAg                 LVLKNNRGVEDNRVRHMDYGVQLNKLMYQRLIKGGDITLFSPSDVPG-LYEAFF-ADQDK 374
A_pleuropneumoniae5bL20               LVLKNNRGVEDNRCRHMDYGVQLNKLMYQRLIKGGEITLFSPSDVPG-LYEDFF-ADQEK 374
H_ducreyi35000HP_NrdAg                VVLKNNRGVEDNRCRHMDYGVQINKLIYQRLIKGSDITLFSPSDVPG-LYEAFF-ADQEK 374
V_choleraeO395_NrdAg                  LVLKNNRGVEENRVRHMDYGVQLNKLMYKRLVEGGNITLFSPSDVPG-LYDAFF-QDQDE 374
V_vulnificusCMCP6_NrdAg               MVLKNNRGVEENRVRHMDYGVQLNKLMYQRLVEGGNITLFSPSDVPG-LYDAFF-ENQAE 374
VibrioEx25                            MVLKNNRGVEENRVRHMDYGVQLNKLMYQRLVEGGNITLFSPSDVPG-LYDAFF-ENQEE 374
V_campbelliiBAA1116_NrdAg             LVLKNNRGVEENRVRHMDYGVQLNKLMYQRLVEGGNITLFSPSDVPG-LYDAFF-ENQEE 374
VibrioAND4_NrdAg                      LVLKNNRGVEENRVRHMDYGVQLNKLMYQRLVEGGNITLFSPSDAPG-LYDAFF-ENQEE 374
V_fischeriES114_NrdAg                 LVLKNNRGVEENRVRHMDYGVQINKLMYTRLIKGGNISLFSPSDVPG-LYDAFF-ADQAE 374
P_angustumS14_NrdAg                   LVLKNNRGVEENRVRHMDYGVQINKLMYTRLIKGESISLFSPSDVPG-MYDAFF-ADQEE 374
PhotobacteriumSKA34_NrdAg             LVLKNNRGVEENRVRHMDYGVQINKLMYTRLIKGENISLFSPSDVPG-MYDAFF-ADQEE 374
C_psychrerythraea34H_NrdAg            LVLKNNRGVEENRVRHLDYGVQFNKLMYQRLIKGGSITLFSPSDVPG-LYDAFF-EDQEK 374
A_hydrophila7966_NrdAg                LVLKNNRGVEENRVRHMDYGVQINKLMYQRLIKGGDITLFSPSDAPG-LYDAFF-ADQDK 384
A_salmonicidaA449_NrdAg               LVLKNNRGVEENRVRHMDYGVQINKLMYQRLIKGGDITLFSPSDAPG-LYDAFF-ADQDK 372
B_cicadellinicolaHc_NrdAg             LVLKNNRGIETNRVRYMDYGVQINKLMYQRLVNNDQITLFSPSDVPG-LYDAFF-ADQNK 374
C_sakazakiiBAA894                     LVLKNNRGTEANRVRHMDYGVQINKLMYQRLLKGEDITLFSPSDVPG-LYDAFF-ADQDE 374
K_pneumoniaeMGH78578_NrdAg            LVLKNNRGTDANRVRHMDYGVQINKLMYTRLLKGGDITLFSPSDVPG-LYDAFF-ADQDE 374
Enterobacter638_NrdAg                 LVLKNNRGVEGNRVRHMDYGVQINKLMYTRLLKGQDITLFSPSDVPG-LYDAFF-ADQDE 374
C_koseriBAA895_NrdAg                  LVLKNNRGVEGNRVRHMDYGVQINKLMYTRLLKGEDITLFSPSDVPG-LYDAFF-ADQDE 374
S_boydiiSb227_NrdAg                   LVLKNNRGVEGNRVRHMDYGVQINKLMYTRLLKGEDITLFSPSDVPG-LYDAFF-ADQEE 374
E_coliK12ssMG1655_NrdAg               LVLKNNRGVEGNRVRHMDYGVQINKLMYTRLLKGEDITLFSPSDVPG-LYDAFF-ADQEE 374
S_flexneri58401_NrdAg                 LVLKNNRGVEGNRVRHMDYGVQINKLMYTRLLKGEDITLFSPSDVPG-LYDAFF-ADQEE 374
E_coli536_NrdAg                       LVLKNNRGVEGNRVRHMDYGVQINKLMYTRLLKGEDITLFSPSDVPG-LYDAFF-ADQEE 374
S_flexneri2a2457T_NrdAg               LVLKNNRGVEGNRVRHMDYGVQINKLMYTRLLKGEDITLFSPSDVPG-LYDAFF-ADQEE 374
STyphimuriumLT2_NrdAg                 LVLKNNRGVEGNRVRHMDYGVQINKLMYTRLLKGGDITLFSPSDVPG-LYDAFF-ADQDE 374
S_entericaTyphiTy2_NrdAg              LVLKNNRGVEGNRVRHMDYGIQINKLMYTRLLKGGDITLFSPSDVPG-LYDAFF-ADQDE 374
S_proteamaculans568_NrdAg             LVLKNNRGVEGNRVRHMDYGVQLNRLMYQRLIKGQDITLFSPSDVPG-LYDAFF-ADQDE 374
Y_pestisAngola_NrdAg                  LVLKNNRGVEGNRVRHMDYGVQINKLMYQRLLKGGDITLFSPSDVPG-LYDAFF-ADQDE 374
YpseudotuberculosisIP31758_NrdAg      LVLKNNRGVEGNRVRHMDYGVQINKLMYQRLLKGGDITLFSPSDVPG-LYDAFF-ADQDE 374
                                      : ::.      ::   :**..  .  :          *.*       :
```

```
SalmonellaphageFelixO1_NrdAg          FAMAVGHAVH-SGV-----------KHRRVKAREIAKLFIQQRYATGRVYPFFTNNANT 411
EscherichiaphagerV5_NrdAg             FDAEVDRIAKDDSI-----------PKTWIRARDLAVELLTQRIEVGRLFVYWTDEMNR 412
EnterophageT5_NrdAg                   FEEIYVAAEKRVASLTKIDHEGKTVPAAPKVSAKEILDTWLRIRMETGRMYAHHIGESNR 465
VibriophageKVP40_NrdAg                FERLYLELEADPKV-----------RKKRIKATELFSIFGQERSSTGRIYPIFVDNLNA 417
AeromonasphageAeh1_NrdAg              FKELYEMLENEPWC-----------RKKRIKASELFNLFAQERTGTGRIYPMFVDNVNS 417
EnterophageRB49_NrdAg                 FRQLYEKLEKDPSV-----------RKTRVKAEDLFLAVAIERFNTGRIYPTFTDNMNN 419
EnterophagePhi1_NrdAg                 FRQLYEKLEKDPSV-----------RKTRVKAEDLFLAVAIERFNTGRIYPTFTDNMNN 419
EnterobacteriaphageT4_NrdAg           FRELYEAAEKDPNI-----------RKKRIKARELFELLMTERSGTARIYVQFIDNTNN 420
EnterophageRB32_NrdAg                 FRELYEAAEKDPNI-----------RKKRIKARELFELLMTERSGTARIYVQFIDNTNN 420
EnterophageJS98_NrdAg                 FRELYEAAEKDSSI-----------RKKRIKARDLFELFMTERSGTARIYPQFVDNTNN 417
EnterophageRB69_NrdAg                 FRELYEKAEKDPTI-----------RKKRIKALELFELFMTERSGTARIYPQFVDNTNN 417
ProchlorophagePSSM2_NrdAg             FDELYTQYESDESI-----------PRTTIGAQELILDLLKERAETGRIYLMNIDHCNS 422
ProchlorococcusphagePSSM4_NrdAg       FDELYVKYENDNSI-----------PKKVVRAQELVLDLLKERAETGRLYIMNIDHCNS 423
Synechococcusphagesyn9_NrdAg          FDDLYTQYEQDSSI-----------PKKRINAQELILDLLKERAETGRLYIMNIDHCNS 424
Y_pestisMicrotus91001_NrdAg           FEELYLEYEADESI-----------RKKSIPAVDLFSTLMQERASTGRVYIANVDHMND 423
Y_pestisPestoidesF_NrdAg              FEELYLEYEADESI-----------RKKSIPAVDLFSTLMQERASTGRVYIANVDHMND 423
B_aphidicolaBp_NrdAg                  FERLYIKYENDKTI-----------RKKKVKASYLFSLMMQERTSTGRIYIQNVDHCNS 422
B_aphidicolaSg_NrdAg                  FKEIYIKYEKNRNI-----------RKKSINAIDLFCLIMRERASTGRIYIQHVDHCNS 422
C_RuthiaCm_NrdAg                      FKRLYEQYEQNDSI-----------RKQAIKARELFAKFMQERANTGRIYLQNVDHCNT 422
T_crunogenaXCL2_NrdAg                 FERLYKIYEADDTI-----------RKKTMKAVDLFTQVAQERAQTGRIYIQNVDHSNT 428
M_flagellatusKT_NrdAg                 FERLYTQYEANDSL-----------RKRIVPAVDLFSLLMQERAGTGRIYIQNVDHCNT 439
N_gonorrhoeaeFA1090_NrdAg             FERLYTKYEQDPNI-----------RKRIIPAADLFSTLMQERAGTGRIYIQNVDHCNT 424
N_meningitidis053442_NrdAg            FERLYTKYEQDPDI-----------RKRIIPAADLFSTLMQERAGTGRIYIQNVDHCNT 424
N_meningitidiMC58_NrdAg               FERLYTKYEQDPDI-----------RKRIIPAADLFSTLMQERAGTGRIYIQNVDHCNT 424
P_atlanticaT6c_NrdAg                  FEELYVKYEQDDSI-----------RKTQIKAVELFSLFMQERASTGRIYLQNVDHCNT 422
I_loihiensisL2TR_NrdAg                FERLYLQYEQDESI-----------RKKRIKAIELFGLFMQERASTGRIYVQNVDHCNT 422
P_tunicataD2_NrdAg                    FESLYVQYEQDPAI-----------RKKRIKAIELFSMFVQERASTGRIYLQNVDHCNT 422
PsychrobacterPRwf1_NrdAg              FEALYTKYEKDDNI-----------RKRHVPAADLFSLLMQERASTGRIYIQNVDHCNT 424
P_arcticus2734_NrdAg                  FEELYTKYENDPTI-----------RSRQIPAADLFSLMMQERASTGRIYIQNVDHCNT 424
P_cryohalolentisK5_NrdAg              FEELYTKYENDPAI-----------RSRQIPAADLFSLMMQERASTGRIYIQNVDHCNT 424
CBlochmanniaBPEN_NrdAg                FEHLYNIYENDHNI-----------RHRKIKAIELFSLMMQERASTGRIYIQHTDHCNT 422
P_ingrahamii37_NrdAg                  FERLYNLYEGDPDI-----------RQKSISAVELFSLLMQERASTGRIYIQNVDHCNT 422
MarinomonasMWYL1_NrdAg                FDRLYTMYEQDPSI-----------RKQTIKATELFTLFASERASTGRIYLQNVDHCNT 421
S_denitrificansOS217_NrdAg            FERLYLQYEQDTRI-----------RKKTLKAVELFSLMMQERASTGRIYIQNVDHCNT 422
S_amazonensisSB2B_NrdAg               FERLYTKYEQEPSI-----------RKKSLKAVELFSLMMQERASTGRIYVQNVDHCNT 442
S_frigidimarina400_NrdAg              FERLYLQYEQDSNI-----------RKKQIKAVELFSLMMQERASTGRIYIQNVDHCNT 422
S_oneidensisMR1_NrdAg                 FERLYLQYEQDSSI-----------RKKTLKAVDLFSLMMQERASTGRIYIQNVDHCNT 422
ShewanellaANA3_NrdAg                  FERLYLQYEQDSSI-----------RKKTLKAVELFSLMMQERASTGRIYIQNVDHCNT 422
ShewanellaMR4_NrdAg                   FERLYLQYEQDSSI-----------RKKTLKAVELFSLMMQERASTGRIYIQNVDHCNT 422
ShewanellaMR7_NrdAg                   FERLYLQYEQDSSI-----------RKKTLKAVELFSLMMQERASTGRIYIQNVDHCNT 422
S_balticaOS155_NrdAg                  FERLYVQYEQDSSI-----------RKKTIKAVELFSLMMQERASTGRIYIQNVDHCNT 422
S_putrefaciensCN32_NrdAg              FERLYLQYEQDSSI-----------RKKTIKAVELFSLMMQERASTGRIYIQNVDHCNT 422
ShewanellaW3181_NrdAg                 FERLYLQYEQDSSI-----------RKKTIKAVELFSLMMQERASTGRIYIQNVDHCNT 422
S_sediminisHAWEB3_NrdAg               FERLYHQYEQDDSV-----------RKTQVKAAELFSLLMQERASTGRIYIQNVDHCNT 422
S_loihicaPV4_NrdAg                    FERLYVKYEQDESI-----------RKKQIKAVELFSLMMQERASTGRIYIQNVDHCNT 422
S_pealeana700345_NrdAg                FERLYLKYEADDSI-----------RKKQIKAVELFSLMMQERASTGRIYIQNVDHCNT 422
A_succinogenes130Z_NrdAg              FEQLYVQYEQDPAI-----------RKRSVKAIELFSLLMQERASTGRIYIHNVDHTNT 422
M_succiniciproducensMBEL55E_NrdAg     FEELYVKYEQDPTI-----------RKRTVKAVDLFSLLMQERASTGRIYVQNVDHCNT 431
H_influenzae86028NP_NrdAg             FEELYVKYEQDPTI-----------RKRTVKAVEIFSLLMQERASTGRIYIQNVDHCNT 422
H_influenzaeRdKW20_NrdAg              FEELYVKYEQDPTI-----------RKRTVKAVEIFSLLMQERASTGRIYIQNVDHCNT 447
H_somnus129PT_NrdAg                   FETLYLKYEQDPAV-----------RKRSVKAVELFSLLMQERASTGRIYIQNVDHCNT 422
P_multocidaPm70_NrdAg                 FEQLYLQYEQDPTI-----------RQRSVKAVELFSLLMQERASTGRIYIQNVDHCNT 422
A_pleuropneumoniae5bL20               FEQLYLQYEQDPSI-----------RKRSVKAVELFSLLMQERASTGRIYIQNVDHCNT 422
H_ducreyi35000HP_NrdAg                FELLYLQYEQDPSI-----------RKKSIKAIELFSLLMQERASTGRIYIHNVDHSNT 422
V_choleraeO395_NrdAg                  FERLYVKYENDPSI-----------KKKTVKALEIFTLLMQERASTGRIYIQNVDHCNT 422
V_vulnificusCMCP6_NrdAg               FERLYVKYENDPSI-----------KKETVKALEMFTLLMQERASTGRIYIQNVDHCNT 422
VibrioEx25                            FERLYVKYENDPSI-----------KKETVKAIEMFTLLMQERASTGRIYIQNVDHCNT 422
V_campbelliiBAA1116_NrdAg             FERLYVKYENDPSI-----------KKETVKAIEMFTLLMQERASTGRIYIQNVDHCNT 422
VibrioAND4_NrdAg                      FERLYVKYENDPSV-----------KKETVKAIEMFTLLMQERASTGRIYIQNVDHCNT 422
V_fischeriES114_NrdAg                 FERLYEKYEQDTSI-----------KRETVKAIELFSLLMQERASTGRIYIQNVDHCNT 422
P_angustumS14_NrdAg                   FERLYLKYEQDSSI-----------KRETIKAIDLFSLLMQERASTGRIYIQNVDHCNT 422
PhotobacteriumSKA34_NrdAg             FERLYLKYEQDSSI-----------KRETIKAIDLFSLLMQERASTGRIYIQNVDHCNT 422
C_psychrerythraea34H_NrdAg            FERLYVQYENDDSI-----------RKKRIKAIELFTLFAQERASTGRIYLQNVDHCNT 422
A_hydrophila7966_NrdAg                FEALYVKYEQDPSI-----------RKKTLKAVDLFSLMMQERAGTGRIYIQNVDHCNT 432
A_salmonicidaA449_NrdAg               FEALYVKYEQDPAI-----------RKKTIKAVDLFSLMMQERAGTGRIYIQNVDHCNT 420
B_cicadellinicolaHc_NrdAg             FEQLYCQYEKDVSI-----------RKRSIKAVELFSLMMQERASTGRIYVQHVDHCNN 422
C_sakazakiiBAA894                     FERLYTRYEQDESI-----------RKQRIKASELFSLMMQERASTGRIYIQNVDHCNT 422
K_pneumoniaeMGH78578_NrdAg            FERLYTQYEQDDSI-----------RKQRIKAVELFSLMMQERASTGRIYIQNVDHCNT 422
Enterobacter638_NrdAg                 FERLYTQYEQDDSI-----------RKQRVKAVDLFSLMMQERASTGRIYIQNVDHCNT 422
C_koseriBAA895_NrdAg                  FERLYTKYEKDDSI-----------RKQRVKAVDLFSLMMQERASTGRIYIQNVDHCNT 422
S_boydiiSb227_NrdAg                   FERLYTKYEKDDSI-----------RKQRVKAVELFSLMMQERASTGRIYIQNVDHCNT 422
E_coliK12ssMG1655_NrdAg               FERLYTKYEKDDSI-----------RKQRVKAVELFSLMMQERASTGRIYIQNVDHCNT 422
S_flexneri58401_NrdAg                 FERLYTKYEKDDSI-----------RKQRVKAVELFSLMMQERASTGRIYIQNVDHCNT 422
E_coli536_NrdAg                       FERLYTKYEKDDSI-----------RKQRVKAVELFSLMMQERASTGRIYIQNVDHCNT 422
S_flexneri2a2457T_NrdAg               FERLYTKYEKDDSI-----------RKQRVKAVELFSLMMQERASTGRIYIQNVDHCNT 422
STyphimuriumLT2_NrdAg                 FERLYVKYEHDDSI-----------RKQRVKAVELFSLMMQERASTGRIYIQNVDHCNT 422
S_entericaTyphiTy2_NrdAg              FERLYVKYENDDSI-----------RKQRVKAVELFSLMMQERASTGRIYIQNVDHCNT 422
S_proteamaculans568_NrdAg             FERLYTQYEQDDSI-----------RQKRVKAVELFSLMMQERASTGRIYIQNVDHCNT 422
Y_pestisAngola_NrdAg                  FERLYVKYEQDSSI-----------RKQRVKAVELFSLMMQERASTGRIYIQNVDHCNT 422
YpseudotuberculosisIP31758_NrdAg      FERLYVKYEQDSSI-----------RKQRVKAVELFSLMMQERASTGRIYIQNVDHCNT 422
                                      *                          : *  :       *  ..*::     . *
```

```
SalmonellaphageFelixO1_NrdAg          HTPFK---EPLKMSNLCMEIVLPVYGFEKETDLYRD---------------------- 444
Escherichiaphagerv5_NrdAg             HTPFK---DVIYSSNLCAEIALPTKGYSDMRNIFKH---------------------- 445
EnterophageT5_NrdAg                   HGNFL---DPIRMTNLCVEITQPTRPFHHITELY------------------------ 496
VibriophageKVP40_NrdAg                QSPYI---EAIKQSNLCMEIALPTYPLGTS---------------------------- 444
AeromonasphageAeh1_NrdAg              ATPYK---EPIRMGNLCIEIALPTVPMGKKEKMLFKVKKSEVDKFKLMVATDQYKSLRYM 474
EnterophageRB49_NrdAg                 HGAFIRNVAPIKQSNLCCEIALPTKPMGRD---------------------------- 449
EnterophagePhi1_NrdAg                 HGAFIRNVAPIKQSNLCCEIALPTKPMGRD---------------------------- 449
EnterobacteriaphageT4_NrdAg           YTPFIREKAPIRQSNLCCEIAIPTNDVNSP---------------------------- 450
EnterophageRB32_NrdAg                 YTPFIREKAPIRQSNLCCEIAIPTNDVNSP---------------------------- 450
EnterophageJS98_NrdAg                 FTPFIREKAPIRQSNLCCEIAIPTTDVNSP---------------------------- 447
EnterophageRB69_NrdAg                 YTPFIREKAPIRQSNLCCEIAIPTNDVNGP---------------------------- 447
ProchlorophagePSSM2_NrdAg             HSSFLDKV---EMSNLCQEITLPTKPIQHI---------------------------- 449
ProchlorococcusphagePSSM4_NrdAg       HSSFKDKV---SMSNLCQEITLPTTPLQHI---------------------------- 450
Synechococcusphagesyn9_NrdAg          HSSFKDKV---NMSNLCQEITLPTDPLQHI---------------------------- 451
Y_pestisMicrotus91001_NrdAg           HGAFVPEIAPVHQSNLCMEITLPTKPLAFT---------------------------- 453
Y_pestisPestoidesF_NrdAg              HGAFVPEIAPVHQSNLCMEITLPTKPLAFT---------------------------- 453
B_aphidicolaBp_NrdAg                  HSAFNPKIAPIRQSNLCLEITLPTKALQDV---------------------------- 452
B_aphidicolaSg_NrdAg                  HSAFNSKLATIRQSNLCLEITLPTKPLNNI---------------------------- 452
C_RuthiaCm_NrdAg                      HSAFDAKQAPIKQSNLCMEITLPTKPLYSV---------------------------- 452
T_crunogenaXCL2_NrdAg                 HSSFDPKKAPIHQSNLCLEITLPTHPLNSV---------------------------- 458
M_flagellatusKT_NrdAg                 HSPFNPRLAPVRQSNLCMEIALPTRPLNDI---------------------------- 469
N_gonorrhoeaeFA1090_NrdAg             HSPFDPRVAPVHQSNLCMEIALPTKPLDNI---------------------------- 454
N_meningitidis053442_NrdAg            HSPFDPRVAPVHQSNLCMEIALPTKPLDNI---------------------------- 454
N_meningitidiMC58_NrdAg               HSPFDPRVAPVHQSNLCMEIALPTKPLDNI---------------------------- 454
P_atlanticaT6c_NrdAg                  HSPFDPKVAPIRQSNLCLEIALPTKPLEHV---------------------------- 452
I_loihiensisL2TR_NrdAg                HSPFDSKVAPIRQSNLCLEIALPTKPLSHV---------------------------- 452
P_tunicataD2_NrdAg                    HSPFDSKVAPIRQSNLCLEIALPTKPLKHV---------------------------- 452
PsychrobacterPRwf1_NrdAg              HSPFDPTVAPIRQSNLCMEIALPTKPLDNI---------------------------- 454
P_arcticus2734_NrdAg                  HSPFDPTVAPIRQSNLCMEIALPTKPLDNI---------------------------- 454
P_cryohalolentisK5_NrdAg              HSPFDATVAPIRQSNLCMEIALPTKPLDNI---------------------------- 454
CBlochmanniaBPEN_NrdAg                HSAFDSIAAPVRQSNLCLEITLPTKPLNNI---------------------------- 452
P_ingrahamii37_NrdAg                  HSPFDSKVAPIRQSNLCLEIALPTKPLQHI---------------------------- 452
MarinomonasMWYL1_NrdAg                HSPFDPKVAPVKQSNLCLEIALPTKPLNSI---------------------------- 451
S_denitrificansOS217_NrdAg            HSPFDPSVAPIKQSNLCLEIALPTKPLKNI---------------------------- 452
S_amazonensisSB2B_NrdAg               HSPFDSSVAPIRQSNLCLEIALPTKPLDSV---------------------------- 472
S_frigidimarina400_NrdAg              HSPFDSKVAPIKQSNLCLEIALPTKPLTNI---------------------------- 452
S_oneidensisMR1_NrdAg                 HSPFDSKVAPIRQSNLCLEIALPTKPLNNI---------------------------- 452
ShewanellaANA3_NrdAg                  HSPFDSKVAPIRQSNLCLEIALPTKPLNNI---------------------------- 452
ShewanellaMR4_NrdAg                   HSPFDSKVAPIRQSNLCLEIALPTKPLNNI---------------------------- 452
ShewanellaMR7_NrdAg                   HSPFDSKVAPIRQSNLCLEIALPTKPLNNI---------------------------- 452
S_balticaOS155_NrdAg                  HSPFDSKVAPIRQSNLCLEIALPTKPLNNI---------------------------- 452
S_putrefaciensCN32_NrdAg              HSPFDSKVAPIRQSNLCLEIALPTKPLNNI---------------------------- 452
ShewanellaW3181_NrdAg                 HSPFDSKVAPIRQSNLCLEIALPTKPLNNI---------------------------- 452
S_sediminisHAWEB3_NrdAg               HSPFDSKVAPIRQSNLCLEIALPTKPLNNI---------------------------- 452
S_loihicaPV4_NrdAg                    HSPFDAKVAPVRQSNLCLEIALPTKPLNNI---------------------------- 452
S_pealeana700345_NrdAg                HSPFDSKVAPVRQSNLCLEIALPTKPLNNI---------------------------- 452
A_succinogenes130Z_NrdAg              HSPFDPSVAPVRQSNLCLEIALPTKPLQNI---------------------------- 452
M_succiniciproducensMBEL55E_NrdAg     HSPFDPAVAPVRQSNLCLEIALPTKPLNNI---------------------------- 461
H_influenzae86028NP_NrdAg             HSPFDPQVAPVRQSNLCLEIALPTKPLQHI---------------------------- 452
H_influenzaeRdKW20_NrdAg              HSPFDPQVAPVRQSNLCLEIALPTKPLQHI---------------------------- 477
H_somnus129PT_NrdAg                   HSPFDPAVAPIRQSNLCLEIALPTKPLQHF---------------------------- 452
P_multocidaPm70_NrdAg                 HSPFDPKVAPVRQSNLCLEIALPTKPLQHF---------------------------- 452
A_pleuropneumoniae5bL20               HSPFNPQVAPIKQSNLCLEIALPTKPLNDI---------------------------- 452
H_ducreyi35000HP_NrdAg                HSPFDPSVAPIRQSNLCLEIALPTKPLNNI---------------------------- 452
V_choleraeO395_NrdAg                  HSPFDAAVAPVRQSNLCLEIALPTSPLANV---------------------------- 452
V_vulnificusCMCP6_NrdAg               HSPFDAEVAPVRQSNLCLEIALPTKPLANV---------------------------- 452
VibrioEx25                            HSPFDSEVAPVRQSNLCLEIALPTKPLTNV---------------------------- 452
V_campbelliiBAA1116_NrdAg             HSPFDSEVAPVRQSNLCLEIALPTKPLKNV---------------------------- 452
VibrioAND4_NrdAg                      HSPFDSEVAPVRQSNLCLEIALPTKPLKNV---------------------------- 452
V_fischeriES114_NrdAg                 HSPFDPSVAPVRQSNLCLEIALPTKPLTHV---------------------------- 452
P_angustumS14_NrdAg                   HSPFDPAVAPIRQSNLCLEIALPTKPLNNV---------------------------- 452
PhotobacteriumSKA34_NrdAg             HSPFDPAVAPIRQSNLCLEIALPTKPLNNV---------------------------- 452
C_psychrerythraea34H_NrdAg            HSPFDSSKAAIRQSNLCLEIALPTKPMNDI---------------------------- 452
A_hydrophila7966_NrdAg                HSPFDPSVAPVRQSNLCLEIALPTKPLNNV---------------------------- 462
A_salmonicidaA449_NrdAg               HSPFDPSVAPVRQSNLCLEIALPTKPLNNV---------------------------- 450
B_cicadellinicolaHc_NrdAg             HSPFDVNVAPIRQSNLCLEIILPTKPIEDI---------------------------- 452
C_sakazakiiBAA894                     HSPFDPVVAPVRQSNLCLEIALPTKPLDDV---------------------------- 452
K_pneumoniaeMGH78578_NrdAg            HSPFDPVVAPVRQSNLCLEIALPTKPLEDV---------------------------- 452
Enterobacter638_NrdAg                 HSPFDPAIAPVRQSNLCLEIALPTKPLYDV---------------------------- 452
C_koseriBAA895_NrdAg                  HSPFDAQVAPVRQSNLCLEIALPTKPLMDV---------------------------- 452
S_boydiiSb227_NrdAg                   HSPFDPAIAPVRQSNLCLEIALPTKPLNDV---------------------------- 452
E_coliK12ssMG1655_NrdAg               HSPFDPAIAPVRQSNLCLEIALPTKPLNDV---------------------------- 452
S_flexneri58401_NrdAg                 HSPFDPTIAPVRQSNLCLEIALPTKPLNDV---------------------------- 452
E_coli536_NrdAg                       HSPFDPAIAPVRQSNLCLEIALPTKPLNDV---------------------------- 452
S_flexneri2a2457T_NrdAg               HSPFDPTIAPVRQSNLCLEIALPTKPLNDV---------------------------- 452
STyphimuriumLT2_NrdAg                 HSPFDPVVAPVRQSNLCLEIALPTKPLNDV---------------------------- 452
S_entericaTyphiTy2_NrdAg              HSPFDPVVAPVRQSNLCLEIALPTKPLNDV---------------------------- 452
S_proteamaculans568_NrdAg             HSPFDPQIAPVRQSNLCLEIALPTKPLDDV---------------------------- 452
Y_pestisAngola_NrdAg                  HSPFDPKIAPVRQSNLCLEIALPTKPLNDI---------------------------- 452
YpseudotuberculosisIP31758_NrdAg      HSPFDPKIAPVRQSNLCLEIALPTKPLNDI---------------------------- 452
                                         :          *** **  *.
```

```
SalmonellaphageFelixO1_NrdAg          ----------------DAVKEDGEVALCFLASLVAGRIS---EDEYADVAYYALAMVD 483
EscherichiaphagerV5_NrdAg             ----------------D--VDDGEVALCFLSSLVAGRVT---PEEYEDVAYYTVLMID 482
EnterophageT5_NrdAg                   ---------KTKEQLDQMKPEDIGEVSLCNLGGVVLGRME--SLAEWEKTCYILLKFVD 544
VibriophageKVP40_NrdAg                ----------------------YEEIALCTLAAYNLGAIE--SPSEFARLAPIVIRALD 479
AeromonasphageAeh1_NrdAg              IEHERIYEVYEDEEYTTVEIEVDVSEISLCTLAAYNLGAIE--GYEDFLNVARAAVRALD 532
EnterophageRB49_NrdAg                 ----------------------DSEISLCTLAAYVLGNFDIENMEEVERIAMVMVRALD 486
EnterophagePhi1_NrdAg                 ----------------------DSEISLCTLAAYVLGNFDIENMEEVERIAMVMVRALD 486
EnterobacteriaphageT4_NrdAg           ----------------------DAEIGLCTLSAFVLDNFDWQDQDKINELAEVQVRALD 487
EnterophageRB32_NrdAg                 ----------------------DAEIGLCTLSAFVLDNFDWQDQDKINELAEVQVRALD 487
EnterophageJS98_NrdAg                 ----------------------DAEIGLCTLSAFVLDNFDWQDQDKINELAEVQVRALD 484
EnterophageRB69_NrdAg                 ----------------------DAEIGLCTLSAFVLDKFDWQDQDKINELAEVQVRALD 484
ProchlorophagePSSM2_NrdAg             --------------------DDETGEIALCILSAINIGKIR--DVSDLEGLCDLSIRSLD 487
ProchlorococcusphagePSSM4_NrdAg       --------------------DG-SGEIALCILSAINVGKLT--KLDELDDLCELAVRGLD 487
Synechococcusphagesyn9_NrdAg          --------------------DG-QGEIALCILSAINVGKLK--SLDEIDELCELAVRGLD 488
Y_pestisMicrotus91001_NrdAg           --------------------DDPNGEIALCTLSAFNLGAIR--TLDSLKDVAFYAVAALD 491
Y_pestisPestoidesF_NrdAg              --------------------DDPNGEIALCTLSAFNLGAIR--TLDSLKDVAFYAVAALD 491
B_aphidicolaBp_NrdAg                  --------------------HDKNGEISLCTLSAINLGSLN--NLNDLSKLSNLIVRALD 490
B_aphidicolaSg_NrdAg                  --------------------DDKNGEIALCTLSALNLGLIN--DLHDLKELSTLSVRALD 490
C_RuthiaCm_NrdAg                      --------------------QDEQGEVALCTLSAVNLGALD--SLNEMEALTDIIVRSLD 490
T_crunogenaXCL2_NrdAg                 --------------------DDPEGEIALCTLSAFNLGALE--DLDELDELSDLIVRALD 496
M_flagellatusKT_NrdAg                 --------------------NDENGEIALCTLSAFNLGALE--SLDELEGLADLVVRALD 507
N_gonorrhoeaeFA1090_NrdAg             --------------------NDPDGEIALCTLSAFNLGALN--SLDELEGLADLTVRALD 492
N_meningitidis053442_NrdAg            --------------------NDPNGEIALCTLSAFNLGALN--SLDELEGLADLTVRALD 492
N_meningitidiMC58_NrdAg               --------------------NDPNGEIALCTLSAFNLGALN--SLDELEGLADLTVRALD 492
P_atlanticaT6c_NrdAg                  --------------------NDENGEIALCTLSAFNLGAIE--SLEDFAEMADLAVRALD 490
I_loihiensisL2TR_NrdAg                --------------------NDEEGEIALCTLSAFNLGKIE--NLGEFENLADLAVRALD 490
P_tunicataD2_NrdAg                    --------------------MDEEGEIALCTLSAFNLGAID--SLDELEELAELAVRALD 490
PsychrobacterPRwf1_NrdAg              --------------------NDEEGEIALCTLSAVNLGEVD--SLEDIEEPAELIVRALD 492
P_arcticus2734_NrdAg                  --------------------NDEAGEIALCTLSAVNLGKVE--HVSDIEEPAELIVRALD 492
P_cryohalolentisK5_NrdAg              --------------------NDEEGEIALCTLSAVNLGKVE--NVSDIEEPAELIVRALD 492
CBlochmanniaBPEN_NrdAg                --------------------NDHNGEIALCTLSAFNLGNIE--NLDDLSELSILIVRALD 490
P_ingrahamii37_NrdAg                  --------------------MDEEGEIALCTLSAFNLGAIE--SLDELEELAELAVRALD 490
MarinomonasMWYL1_NrdAg                --------------------DDKEGEIALCTLSAFNLGALE--NLDELEELAELIVRALD 489
S_denitrificansOS217_NrdAg            --------------------DDPEGEIALCTLSALNLGAIK--DLSELEPLADLAVRALD 490
S_amazonensisSB2B_NrdAg               --------------------DDTNGEIALCTLSALNLGAIE--NLEELEGLSDLAVRALD 510
S_frigidimarina400_NrdAg              --------------------DDPDGEIALCTLSALNLGAIK--DLSELEPLADLAVRALD 490
S_oneidensisMR1_NrdAg                 --------------------DDPNGEIALCTLSALNLGTIT--DLAELEPLADLAVRALD 490
ShewanellaANA3_NrdAg                  --------------------DDPNGEIALCTLSALNLGAIK--NLAELEPLADLAVRALD 490
ShewanellaMR4_NrdAg                   --------------------DDPNGEIALCTLSALNLGAIK--NLAELEPLADLAVRALD 490
ShewanellaMR7_NrdAg                   --------------------DDPNGEIALCTLSALNLGAIK--NLAELEPLADLAVRALD 490
S_balticaOS155_NrdAg                  --------------------DDPEGEIALCTLSALNLGSIK--ELSDIEPLADLAVRALD 490
S_putrefaciensCN32_NrdAg              --------------------DDPNGEIALCTLSALNLGAIK--ELSELEPLADLAVRALD 490
ShewanellaW3181_NrdAg                 --------------------DDPNGEIALCTLSALNLGAIK--ELSELERLADLAVRALD 490
S_sediminisHAWEB3_NrdAg               --------------------DDPEGEIALCTLSALNLGAIK--DLSELEPLADLAVRALD 490
S_loihicaPV4_NrdAg                    --------------------DDPDGEIALCTLSALNLGAIK--KLDELESLADLAVRALD 490
S_pealeana700345_NrdAg                --------------------NDPDGEIALCTLSALNLGAIK--KLEELEPLADLAVRALD 490
A_succinogenes130Z_NrdAg              --------------------NDEDGEIALCTLSAFNLGKLE--NLDELDNLADLAVRALD 490
M_succiniciproducensMBEL55E_NrdAg     --------------------NDEDGEIALCTLSAFNLGKID--DLDELENLADLAVRSLD 499
H_influenzae86028NP_NrdAg             --------------------NDENGEIALCTLSAFNLGKIE--NLDELEELADLAVRSLD 490
H_influenzaeRdKW20_NrdAg              --------------------NDENGEIALCTLSAFNLGKIE--NLDELEELADLAVRSLD 515
H_somnus129PT_NrdAg                   --------------------HDENGEIALCTLSAFNLGALE--NLDELEELADLAVRALD 490
P_multocidaPm70_NrdAg                 --------------------HDEKGEIALCTLSAFNLGTLE--NLDELEGLADLAVRALD 490
A_pleuropneumoniae5bL20               --------------------NDENGEIALCTLSAFNLGTLQ--NLDEFEPLADLVVRALD 490
H_ducreyi35000HP_NrdAg                --------------------NDEQGEIALCTLSAFNLGTLQ--NLDELEPLADLAVRALD 490
V_choleraeO395_NrdAg                  --------------------EDESGEIALCTLSAFNLGAIE--SLDDFEELSDLVVRALD 490
V_vulnificusCMCP6_NrdAg               --------------------EDDSGEIALCTLSAFNLGAIN--TLDDLKELSELVVRALD 490
VibrioEx25                            --------------------EDDSGEIALCTLSAFNLGAIK--SLDDFEELSDLVVRALD 490
V_campbelliiBAA1116_NrdAg             --------------------EDDSGEIALCTLSAFNLGAIK--SLDDFEELSELVVRALD 490
VibrioAND4_NrdAg                      --------------------EDDSGEIALCTLSAFNLGAIK--SLDDFEELSELVVRALD 490
V_fischeriES114_NrdAg                 --------------------NDEEGEIALCTLSAFNLGAIE--KLEELEELAELAVRALD 490
P_angustumS14_NrdAg                   --------------------EDENGEIALCTLSAFNLGAID--SLDDLEELAELTVRALD 490
PhotobacteriumSKA34_NrdAg             --------------------EDENGEIALCTLSAFNLGAIE--SLDDLEELAELTVRALD 490
C_psychrerythraea34H_NrdAg            --------------------NDENGEIALCTLSAFNLGAID--SLDELEGLADLAVRALD 490
A_hydrophila7966_NrdAg                --------------------DDEEGEIALCTLSAFNLGAIE--KLEDLEEMADLAVRALD 500
A_salmonicidaA449_NrdAg               --------------------DDEEGEIALCTLSAFNLGAIE--KLEDLEEMADLAVRALD 488
B_cicadellinicolaHc_NrdAg             --------------------NEENGEIALCTLSALNLGTIN--SLDDFEELTMLAVRALD 490
C_sakazakiiBAA894                     --------------------NDENGEIALCTLSAFNLGAIE--SLDELEELSTLAVRALD 490
K_pneumoniaeMGH78578_NrdAg            --------------------NDENGEIALCTLSAFNLGAID--SLDELEELAVLAVRALD 490
Enterobacter638_NrdAg                 --------------------NDENGEIALCTLSAFNLGAIK--SLAELEELAVLAVRALD 490
C_koseriBAA895_NrdAg                  --------------------NDENGEIALCTLSAFNLGAIN--SLDELEELAVLAVRALD 490
S_boydiiSb227_NrdAg                   --------------------NDENGEIALCTLSAFNLGAIN--NLDELEELAILAVRALD 490
E_coliK12ssMG1655_NrdAg               --------------------NDENGEIALCTLSAFNLGAIN--NLDELEELAILAVRALD 490
S_flexneri58401_NrdAg                 --------------------NDENGEIALCTLSAFNLGAIN--SLDELEELAILAVRALD 490
E_coli536_NrdAg                       --------------------NDENGEIALCTLSAFNLGAIN--SLDELEELAILAVRALD 490
S_flexneri2a2457T_NrdAg               --------------------NDENGEIALCTLSAFNLGAIN--SLDELEELAILAVRALD 490
STyphimuriumLT2_NrdAg                 --------------------NDENGEIALCTLSAFNLGAIK--TLDELEELAILAVRALD 490
S_entericaTyphiTy2_NrdAg              --------------------NDENGEIALCTLSAFNLGAIK--TLDELEELAILAVRALD 490
S_proteamaculans568_NrdAg             --------------------NDENGEIALCTLSAFNLGAIE--SLDDLEELATLAVRALD 490
Y_pestisAngola_NrdAg                  --------------------NDENGEIALCTLSAFNLGAID--SLDDLEDLAILAVRALD 490
YpseudotuberculosisIP31758_NrdAg      --------------------NDENGEIALCTLSAFNLGAID--SLDDLEDLAILAVRALD 490
                                                            *:.** *..     .      .       :  :*
```

```
SalmonellaphageFelixO1_NrdAg          SVIDLMDYPYPSMRNHVQKRRSVGIGLTNVAHYLAKNYVNYSSRAGKTKLHELAEMHSYY 543
Escherichiaphagerv5_NrdAg             NVMEIMTYPYENMESTVKKRRSIGVGLTNVAHFIASHKVTYGSPESKQLVHDLAELHSYS 542
EnterophageT5_NrdAg                   TIIEIQDYPFPTMEYTAKKRRNVGIGLMNAAGAMAAEGLAYEGIEARNWIHREAEKLSYF 604
VibriophageKVP40_NrdAg                NLLDYQDYPVEAAEK-AKYRRALGIGITNFAYFLAKNFVGYDSS-AKYLVHEWMEAMSYY 537
AeromonasphageAeh1_NrdAg              NLLDYQDYPVVQALA-AKKRRALGMGVTNFAYFLAKNFVNYSSGNANKLVHRTMEMMQFA 591
EnterophageRB49_NrdAg                 NLLDYQDYPVEQALT-AKKRRSLGVGVTNFASYLADNFFTYED--ANEFTHKTFERLQYA 543
EnterophagePhi1_NrdAg                 NLLDYQDYPVEQALT-AKKRRSLGVGVTNFASYLADNFFTYED--ANEFTHKTFERLQYA 543
EnterobacteriaphageT4_NrdAg           NLLDYQGYPVPEAEK-AKKRRNLGVGVTNYAAWLASNFASYED--ANDLTHELFERLQYG 544
EnterophageRB32_NrdAg                 NLLDYQGYPVPEAEK-AKKRRNLGVGVTNYAAWLASNFASYED--ANDLTHELFERLQYG 544
EnterophageJS98_NrdAg                 NLLDYQDYPVPEALK-AKKRRNLGVGVTNYAAWLASNFASYED--ANDLTHELFERLQYG 541
EnterophageRB69_NrdAg                 NLLDYQDYPVPEALK-AKKRRNLGVGVTNYAAWLASNFASYED--ANYLTHELFERLQYG 541
ProchlorophagePSSM2_NrdAg             ELIDFQGYPVRAAEIATKARRSLGVGFIGLAHYLAKQGVKYEDPEAWKLVHDLTEVFQYN 547
ProchlorococcusphagePSSM4_NrdAg       ALIDYQQYPVTAAEQSTKNRRSLGIGYIGLAHYLAKNGAKYDSQKAFDLVHKLTERFQFA 547
Synechococcusphagesyn9_NrdAg          SLIDYQGYPIKAAEISTLNRRSLGIGFIGLAHYLAKHGAKYNSPKAYKLVHDLTERFQYA 548
Y_pestisMicrotus91001_NrdAg           SLLDYQDYPMAAAEIPAKARRSLGIGVTNFAYYLAKNGFNYSDAAGNQLVHETFEAIQYY 551
Y_pestisPestoidesF_NrdAg              SLLDYQDYPMAAAEIPAKARRSLGIGVTNFAYYLAKNGFNYSDAAGNQLVHETFEAIQYY 551
B_aphidicolaBp_NrdAg                  SVLDYQEYPILSAKKSALSRRTLGIGVINFAYYLAKNGVRYSDGSANNLTHRTFEAIQYH 550
B_aphidicolaSg_NrdAg                  EILEYQNYPVVCAKKSAISRRSLGIGVINFAYYLAKNKVRYSDGSAKNLTHKTFEAIQYY 550
C_RuthiaCm_NrdAg                      CLLDYQDYPIKAAELSSKSRRTLGIGVTNLAYYLAKNGVKYSDGSGNQLIHKTFEALQYY 550
T_crunogenaXCL2_NrdAg                 SLLDYQDYPVEAARKASMGRRTLGVGVTNLAYYLAKNHVKYSDGSANNLVHKAFEAIQYY 556
M_flagellatusKT_NrdAg                 ALLDYQDYPVKAAFNATQKRRSLGVGVINYAYYLAKNGTGYTDDAALGLTHRTFEAIQYY 567
N_gonorrhoeaeFA1090_NrdAg              ALLDYQGYPVEAARTSTMDRRSLGIGVINYAYYLAKNGVRYSDGSALGLTHRTFEAIQYY 552
N_meningitidis053442_NrdAg            ALLDYQGYPVEAARTSTMGRRSLGIGVINYAYYLAKNGVRYSDGSALGLTHRTFEAIQYY 552
N_meningitidiMC58_NrdAg               ALLDYQGYPVEAARTSTMGRRSLGIGVINYAYYLAKNGVRYSDGSALGLTHRTFEAIQYY 552
P_atlanticaT6c_NrdAg                  NLLDYQDYPVPAAYNATMGRRTLGIGVINFAYYLAKNGVKYSDGSANGLIHRTFEAMQYH 550
I_loihiensisL2TR_NrdAg                SLLDYQEYPVPAAYNATMNRRTLGIGVINFAYYLAKHGVKYSDGSANGLVHQTFEAIQYY 550
P_tunicataD2_NrdAg                    NLLDYQDYPVPAAKNATMNRRTLGVGVVNYAYYLAKNGKKYSDGSANALTHQTFEAIQYY 550
PsychrobacterPRwf1_NrdAg              ALLDYQDYPVKAAENGSMNRRTLGVGVINYAYYLAKNGTRYSDGSALGLTHQTFEALQYY 552
P_arcticus2734_NrdAg                  ALLDYQDYPVKAAQNGSMRRRTLGVGVINYAYYLAKNGVRYSDDSALGLTHQTFEALQFY 552
P_cryohalolentisK5_NrdAg              ALLDYQDYPVKAAQNGSMRRRTLGVGVINYAYYLAKNGVRYSDDSAIGLTHQTFEALQFY 552
CBlochmanniaBPEN_NrdAg                ALLDYQHYPIPAAQRSAVRRRSLGIGVINYAYYLAKHGVRYSDGSANKLTHRTFEAIQYH 550
P_ingrahamii37_NrdAg                  SLLDYQDYPVKAAENSSMARRTLGIGVINFAYYLAKNGVKYSDGSANNLTHKTFEAIQYY 550
MarinomonasMWYL1_NrdAg                SLLDYQNYPIPAAQRATELRRTLGVGVINYAYYLAKNGVKYSDGSANELTHKTFEAIQYF 549
S_denitrificansOS217_NrdAg            NLLDYQDYPIISAQLGSMNRRTLGIGVINFANYLAKNGKKYSDGSANNLTHRTFEAIQFY 550
S_amazonensisSB2B_NrdAg               SLLDYQDYPIKAAKLGSMNRRTLGIGVINFANYLAKNGKKYSDGSANNLTHKTFEAIQFN 570
S_frigidimarina400_NrdAg              NLLDYQDYPIKAAQTSSMNRRTLGIGVINFANYLAKNGKKYSDGSANNLTHKTFEAIQFY 550
S_oneidensisMR1_NrdAg                 NLLDYQDYPIKAAQISSMNRRTLGIGVINFANYLAKNGMKYSDGSANNLTHKTFEAIQFY 550
ShewanellaANA3_NrdAg                  NLLDYQDYPIKAAEISSMNRRTLGIGVINFANYLAKNGMKYSDGSANNLTHKTFEAIQFY 550
ShewanellaMR4_NrdAg                   NLLDYQDYPIKAAEISSMNRRTLGIGVINFANYLAKNGMKYSDGSANNLTHKTFEAIQFY 550
ShewanellaMR7_NrdAg                   NLLDYQDYPIKAAEISSMNRRTLGIGVINFANYLAKNGMKYSDGSANNLTHKTFEAIQFY 550
S_balticaOS155_NrdAg                  NLLDYQDYPIKAAQIASMNRRTLGIGVINFANYLAKNGMKYSDGSANELTHRTFEAIQYY 550
S_putrefaciensCN32_NrdAg              NLLDYQDYPIKAAQIASMNRRTLGIGVINFANYLAKNGMKYSDGSANNLTHRTFEAIQFY 550
ShewanellaW3181_NrdAg                 NLLDYQDYPIKAAQIASMNRRTLGIGVINFANYLAKNGMRYSDGSANNLTHRTFEAIQFY 550
S_sediminisHAWEB3_NrdAg               NLLDYQDYPITAAQKSSMNRRTLGIGVINFANYLAKEGVRYSDGSANNITHKTFEAIQYY 550
S_loihicaPV4_NrdAg                    SLLDYQDYPIVSAHKGSMNRRTLGIGVINFANYLAKEGVRYSDGSANGITHKTFEAIQYY 550
S_pealeana700345_NrdAg                NLLDYQDYPIISAQKASMNRRTLGIGVINFANYLAKEGVRYSDGSANGITHKTFEAIQFY 550
A_succinogenes130Z_NrdAg              ALLDYQDYPVAAARRSSLGRRSLGIGVINYAYYLAKNNMRYSNGSANALTHRTFEAIQYY 550
M_succiniciproducensMBEL55E_NrdAg     ALLDYQDYPVPAAKRSSLGRRALGIGVINYAYYLAKNGVRYSDGSANNLTHRTFEAIQYY 559
H_influenzae86028NP_NrdAg             ALLDYQDYPVVAAKRSSLARRSLGIGVINYAYYLAKNGVRYSDGSANDLTHRTFEAIQYY 550
H_influenzaeRdKW20_NrdAg              ALLDYQDYPVVAAKRSSLARRSLGIGVINYAYYLAKNGVRYSDGSANDLTHRTFEAIQYY 575
H_somnus129PT_NrdAg                   ALLDYQDYPVAAAKRSSLARRSLGIGVINYAYYLAKNNVRYSDGSANDLTHRTFEAIQYY 550
P_multocidaPm70_NrdAg                 ALLDYQDYPVIAAKKSSLARRSLGIGVINYAYYLAKNGVRYSDSSANDLTHRTFEAIQYY 550
A_pleuropneumoniae5bL20               ALLDYQDYPVKAAKHAALARRSLGIGVINYAYYLAKNGVRYSDGSANALTHRTFEAMQYY 550
H_ducreyi35000HP_NrdAg                ALLDYQNYPVKAAEKGAIARRSLGIGVINYAYYLAKHGVRYSDGSANDLTHRTFEAVQYY 550
V_choleraeO395_NrdAg                  ALLDYQDYPLPAARRSSMNRRTLGVGVINYAYYLAKNGVKYSDGSANGLTHRTFEAMQYH 550
V_vulnificusCMCP6_NrdAg               ALLDYQDYPLPAAYKSTMNRRTLGVGVINYAYYLAKHGVKYSDGSANGLTHRTFEAIQYY 550
VibrioEx25                            ALLDYQDYPLPAAYKSTMNRRTLGVGVINFAYYLAKHGVKYSDGSANGLTHRTFEAMQYY 550
V_campbelliiBAA1116_NrdAg             ALLDYQDYPLPAAYKSTMNRRTLGVGVINFAYYLAKNGVKYSDGSANGLTHRTFEAIQYY 550
VibrioAND4_NrdAg                      ALLDYQDYPLPAAYKSTMNRRTLGVGVINFAYYLAKNGVKYSDGSANGLTHRTFEAIQYY 550
V_fischeriES114_NrdAg                 ALLDYQDYPLPAAYKSTMNRRTLGVGVINYAYYLAKNGVKYSDGSANGLTHKTFEAIQYY 550
P_angustumS14_NrdAg                   ALLDYQDYPLPAARRSTMNRRTLGVGVINFAYYLAKNGVRYSDGSANNLTHEAFEAIQYY 550
PhotobacteriumSKA34_NrdAg             ALLDYQDYPLPAARRSTMNRRTLGVGVINFAYYLAKHGVRYSDGSANNLTHEAFEAIQYY 550
C_psychrerythraea34H_NrdAg            SLLDYQDYPIAAAQSASMARRTLGIGVINYAYYLAKNGLYYSNGSANNLTHRTFEAIQFY 550
A_hydrophila7966_NrdAg                ALLDYQDYPLKAAMKGSMNRRTLGIGVINYAYYLAKNGARYSDGSGLALTHRTFEAIQYY 560
A_salmonicidaA449_NrdAg               ALLDYQDYPLKAAQKGSMNRRTLGIGVINYAYYLAKNGARYSDGSGLALTHRTFEAIQFY 548
B_cicadellinicolaHc_NrdAg             ALLDYQSYPLTSAQRGALNRRPLGIGVINYAYYLAKNGVRYSDGSANNLTHRTFEAIQYY 550
C_sakazakiiBAA894                     ALLDYQDYPIPAAKRGAMGRRTLGIGVINFAYYLAKHGVRYSDGSANNLTHKTFEAIQYY 550
K_pneumoniaeMGH78578_NrdAg            ALLDYQDYPIPAAKRGAMGRRTLGIGVINFAYYLAKHGKRYSDGSANNLTHKTFEAIQYY 550
Enterobacter638_NrdAg                 ALLDYQDYPIPAAKRGAMGRRTLGIGVINFAYWLAKNGKRYSDGSANNLTHQTFEAIQYY 550
C_koseriBAA895_NrdAg                  ALLDYQDYPIPAAKRGAMGRRTLGIGVINYAYYLAKHGVRYSDGSANNLTHKTFEAIQYY 550
S_boydiiSb227_NrdAg                   ALLDYQDYPIPAAKRGAMGRRTLGIGVINFAYYLAKHGKRYSDGSANNLTHKTFEAIQYY 550
E_coliK12ssMG1655_NrdAg               ALLDYQDYPIPAAKRGAMGRRTLGIGVINFAYYLAKHGKRYSDGSANNLTHKTFEAIQYY 550
S_flexneri58401_NrdAg                 ALLDYQDYPIPAAKRGAMGRRTLGIGVINFAYYLAKHGKRYSDVSANNLTHKTFEAIQYY 550
E_coli536_NrdAg                       ALLDYQDYPIPAAKRGAMGRRTLGIGVINFAYYLAKHGKRYSDGSANNLTHKTFEAIQYY 550
S_flexneri2a2457T_NrdAg               ALLDYQDYPIPAAKRGAMGRRTLGIGVINFAYYLAKHGKRYSDGSANNLTHKTFEAIQYY 550
STyphimuriumLT2_NrdAg                 ALLDYQDYPIPAAKRGAMGRRTLGIGVINFAYWLAKNGKRYSDGSANNLTHKTFEAIQYY 550
S_entericaTyphiTy2_NrdAg              ALLDYQDYPIPAAKRGAMGRRTLGIGVINFAYWLAKNGKRYSDGSANNLTHKTFEAIQYY 550
S_proteamaculans568_NrdAg             ALLDYQDYPIKAAHRGAMGRRTLGIGVINYAYYLAKNGVRYSDGSANALTHKTFEAIQYY 550
Y_pestisAngola_NrdAg                  ALLDYQDYPIAAAQRGAMGRRTLGIGVINFAYYLAKNGVRYSDGSANNLTHRTFEAIQYY 550
YpseudotuberculosisIP31758_NrdAg      ALLDYQDYPIAAAQRGAMGRRTLGIGVINFAYYLAKNGVRYSDGSANNLTHRTFEAIQYY 550
                                       ::: **          ** :*:*    * :* .  *    .   *  * .:
```

```
SalmonellaphageFelixO1_NrdAg            LHEASLRLAKERGVPEYMKFTKYPEGWVPPKTANKKIDEK-H-DAKLRYDWDDLAQRIKE 601
Escherichiaphagerv5_NrdAg               LHKASLRLAKERGVCEWMHKTKYPEGWLPIDTYNKEVDSVVK-DPSLKQDWEGLRKEIIE 601
EnterophageT5_NrdAg                     LHKASVRLAKEQGACEWFNRTKPSKGILVIDTYKKTVDELVS--VGLEMDWESLRADILK 662
VibriophageKVP40_NrdAg                  NIQASMMLAKERGANKWHRKTRWAKGQLPIDWYNKNVDKLVK--PNYNLDWEALRADVIR 595
AeromonasphageAeh1_NrdAg                NLVASMELAKERGANEWHHKTKWADGLLPIDWYNKNVDNIVP--NDLKMDWEWLRAEIKQ 649
EnterophageRB49_NrdAg                   LIKASVQLAKEKGACELMHETRYGRGELPIDWYKKTVDELVD--PVYEMDWEELRKDLKK 601
EnterophagePhi1_NrdAg                   LIKASVRLAKEKGACELMHETRYGRGELPIDWYKKTVDELVD--PVYEMDWEELREDLKK 601
EnterobacteriaphageT4_NrdAg             LIKASIKLAKEKGPSEYYSDTRWSRGELPIDWYNKKIDQIAA--PKYVCDWSALREDLKL 602
EnterophageRB32_NrdAg                   LIKASIKLAKEKGPCEYYSDTRWSRGELPIDWYNKKIDQIAA--PKYVCDWSALREDLKL 602
EnterophageJS98_NrdAg                   LIRASIKLAEEKGPCEYYSDTKWSRGELPIDWYNKKIDQIAA--PNYVCDWSALREDLAK 599
EnterophageRB69_NrdAg                   LIKASIKLAKEKGPCEYYSDTKWSRGELPIDWYNKKIDQIAA--PNYVCDWKNFELNSKE 599
ProchlorophagePSSM2_NrdAg               LIKSSVNLAKEKGACTYSDRTKYSQGILPIDTYKQDVDELVP--NNLSCDWESLREDVKQ 605
ProchlorococcusphagePSSM4_NrdAg         LLSASNQLAMEKGPCGYFGKTKYADGILPIDTYKNDVDEIIP--NDLSCDWEFLRGRILE 605
Synechococcusphagesyn9_NrdAg            LLIASNRMAMEKGPCGYFDKTKYSDGILPIDTYKKEIDDIVS--NDLSCDWEFLRERIKQ 606
Y_pestisMicrotus91001_NrdAg             LLDASCRLAEAKGACDWFSHTKYAQGQLPVDHYRKTLDTNPETSFELKMPWEELRGRIRE 611
Y_pestisPestoidesF_NrdAg                LLDASCRLAEAKGACDWFSHTKYAQGQLPVDHYRKTLDTNPETSFELKMPWEELRGRIRE 611
B_aphidicolaBp_NrdAg                    LLNASCVLAQEKGACHWFSHTNYYKGILPIDTYKKNVDEICN--EPLHLDWELLRKKIKK 608
B_aphidicolaSg_NrdAg                    LLKASCELAKEKGSCSLFNQTNYYLGKFPIDTYKKDIDSICN--EPLHLNWNLLRKKIKK 608
C_RuthiaCm_NrdAg                        SLKASNTLAKELGVCPLFSQTQYSQGIMPIDSYKKDIDAFCD--CKLELDWDTLRQNIKS 608
T_crunogenaXCL2_NrdAg                   LLKASLNLSKEYGKCDFFEDTKYAQGILPIDTYKKELDEFCT--EPLHYDWETLRSEIQK 614
M_flagellatusKT_NrdAg                   LLKASVQLAREFGPCGAFNETTYSQGILPIDTYKKDLDAVCQ--EALQLDWEQLRKDIVA 625
N_gonorrhoeaeFA1090_NrdAg               LLKASANLAKEYGACTLFNQTVYSQGKLPIDTYKKDLDAVCG--EPLHYDWESLRADIVK 610
N_meningitidis053442_NrdAg              LLKASANLAKEYGACTLFNQTVYSQGNLPIDTYKKDLDAVCN--EPLHYDWESLRAEIVK 610
N_meningitidiMC58_NrdAg                 LLKASANLAKEYGACTLFNQTVYSQGKLPIDTYKKDLDAVCG--EPLHYDWESLRAEIVK 610
P_atlanticaT6c_NrdAg                    LMRASCNLAKEKGACPKFDETTYSQGIMPIDTYKKDLDKICN--EPLHCDWDSLRADIKQ 608
I_loihiensisL2TR_NrdAg                  LMQASMNLAKERGACPKFNETTYSQGIMPTDTYKKDLDKVCD--EPLHLDWDKLRADIKQ 608
P_tunicataD2_NrdAg                      LMKASNELAKEKGACPKFHETTYSQGLMPIDTYKKDLDQICS--EPLRLDWDGLRQSIKQ 608
PsychrobacterPRwf1_NrdAg                LLKASNKLAKEKGACPAFNETTYSQGILPIDTYKADLDKVCN--EPLRLDWEALRQEIKT 610
P_arcticus2734_NrdAg                    LLKASNKLAKEQGACPAFNETTYSQGILPIDTYKKDLDKICQ--EPLHLDWETLRTEITT 610
P_cryohalolentisK5_NrdAg                LLKASNKLAKEQGACPAFNETTYSQGILPIDTYKKDLDKICQ--EPLHLDWETLRGEITE 610
CBlochmanniaBPEN_NrdAg                  LLRASNNLAQEKGSCEWFNDTSYSKGILPIDTYKTDLNSFIN--EPLHYNWEDLRVKIKK 608
P_ingrahamii37_NrdAg                    LLKASNKLAFERGACPKFNETRYSQGILPLDTYKKSLDNICD--EQYHLDWETLRSDIVK 608
MarinomonasMWYL1_NrdAg                  LLKASNKLAKEFGPCLAFNETTYSKGILPIDSYKKEIDNIVS--NELFYDWETLRNDIVT 607
S_denitrificansOS217_NrdAg              LLKASMNLAKEQGACPLFNETTYSKGILPIDTYKRDLDAICS--EPLHMDWEALRKDILE 608
S_amazonensisSB2B_NrdAg                 LLKASMNLAKEKGACPAFHETTYSQGILPIDTYKRDLDKICD--EPLHLDWEGLRADIKH 628
S_frigidimarina400_NrdAg                LLKASMNLAKEQGACPLFHETTYSKGILPIDTYKRDLDKICD--EPLHMDWESLRSDIKQ 608
S_oneidensisMR1_NrdAg                   LLKASMNLAKEKGACPLFNETTYAQGILPIDTYKRDLDKICS--EPLHLDWEALRSDIKE 608
ShewanellaANA3_NrdAg                    LLKASMNLAKEKGACPLFNETTYAQGILPIDTYKRDLDKICS--EPLHLDWETLRQDIKQ 608
ShewanellaMR4_NrdAg                     LLKASMNLAKEQGACPLFNETTYAQGILPIDTYKRDLDKICS--EPLHLDWETLRQDIKQ 608
ShewanellaMR7_NrdAg                     LLKASMNLAKEQDACPLFNETTYAQGILPIDTYKRDLDKICS--EPLHLDWETLRQDIKQ 608
S_balticaOS155_NrdAg                    LLKASMNLAKEQGACPLFHETTYSQGILPIDTYKRDLDKICA--EPLHLDWESLRSDIKQ 608
S_putrefaciensCN32_NrdAg                LLKASMNLAKEQGACPLFNETTYSQGILPIDTYKRDLDKICN--EPLHLDWESLRSDIKQ 608
ShewanellaW3181_NrdAg                   LLKASMNLAKEQGACPLFNETTYSQGILPIDTYKRDLDKICN--EPLHLDWESLRSDIKQ 608
S_sediminisHAWEB3_NrdAg                 LLKASMNLAKEQGACPSFHETSYSKGILPIDTYKRHLDQICD--EPLHMDWETLRQDIVQ 608
S_loihicaPV4_NrdAg                      LLKASMNLAKEQGACPAFHETNYAKGILPIDTYKRDLDKICD--EPLHMDWETLRGEIKT 608
S_pealeana700345_NrdAg                  LLKASMKLAKEQGACPAFHETTYSKGILPIDTYKRDLDKICD--EPLHLDWETLREEIKT 608
A_succinogenes130Z_NrdAg                LLKASMNLAKEQGKCELFHETRYAQGILPIDTYKKDVDGLTD--EPLHYDWETLRADIQK 608
M_succiniciproducensMBEL55E_NrdAg       LLKASMNLAKELGACEYFNETTYAKGILPIDTYKKDVDNLTS--EPLHYDWEQLRTEILE 617
H_influenzae86028NP_NrdAg               LLKASMNLAKEQGACEYFNETTYAKGILPIDTYKKDLDSLTQ--EPLHYDWESLRKDIQE 608
H_influenzaeRdKW20_NrdAg                LLKASMNLAKEQGACEYFNETTYAKGILPIDTYKKDLDSLTQ--EPLHYDWESLRKDIQE 633
H_somnus129PT_NrdAg                     LLKASMNLAKEIGACEYFHETTYAQGILPIDTYKKDIDNLTK--EPLHYDWEALRKEIQE 608
P_multocidaPm70_NrdAg                   LLKASMNLAKEIGPCEYFNETTYSQGILPIDTYKKDIDNLTK--EPLHYDWETLRKEIQE 608
A_pleuropneumoniae5bL20                 LLKASLNLAKEQGACKLFDETSYAQGILPIDSYKKELDNLTT--EPLHYDWENLRAEIKQ 608
H_ducreyi35000HP_NrdAg                  LLKASMNLAKELGPCKGFNETTYAKGILPIDSYKKDIDQLTQ--EPSHYDWEALRADIKQ 608
V_choleraeO395_NrdAg                    LLRASMKLAKEQGRCPLFHETNYAKGLMPIDTYKKDIDLVCA--EPLHYDWDKLRHEIME 608
V_vulnificusCMCP6_NrdAg                 LLKASVELAKEQGKCPLFHETNYAKGLMPIDTYKKDIDLICD--EPLHYDWDALRAEIME 608
VibrioEx25                              LLKASVELAKEQGKCPSFDETNYAKGLLPIDTYKKDIDLICD--EALHYDWDSLRKEIME 608
V_campbelliiBAA1116_NrdAg               LLKASVALAKEQGKCPSFHETNYAKGLLPIDTYKKDVDLICE--EELHYDWDSLRQEIME 608
VibrioAND4_NrdAg                        LLKASVMLAKEQGKCPLFDETNYAKGLLPIDTYKKDVDLICD--EKLHYDWGTLRQEIME 608
V_fischeriES114_NrdAg                   LLKASVGLAKEQGACPAFNETTYAQGILPIDSYKKDIDAICN--EELHLDWETLREEIKT 608
P_angustumS14_NrdAg                     LLKASVNLAKEQGACPAFNETTYSQGILPIDTYKKDINNICS--AELRYDWETLREEIKT 608
PhotobacteriumSKA34_NrdAg               LLKASVNLAKEQGACPAFNETTYSQGILPIDTYKKDINNICS--AELRYDWETLREEIKT 608
C_psychrerythraea34H_NrdAg              LLKATNELAKEQGACPKFDETLLSKGILPIDTYKKTIDDITG--EPLHLDWESLRESIKK 608
A_hydrophila7966_NrdAg                  LLKASVQLAREFGPCPAFNETTYAQGILPIDTYKKDLDNLCN--EPLHLDWEALREEIRT 618
A_salmonicidaA449_NrdAg                 LLKASVQLAREFGPCPAFNETTYAQGILPIDTYKKDLDNLCN--EPLHLDWETLREEIKT 606
B_cicadellinicolaHc_NrdAg               LLKASNKLAKERGACPWFHETTYAQGILPIDTYKKDLDSIVS--EPLHYDWESLRKDIKK 608
C_sakazakiiBAA894                       LLKASNELAKEQGACPWFNETTYSKGILPIDTYKKDLDAIAS--EPLHYDWESLRESIKT 608
K_pneumoniaeMGH78578_NrdAg              LLKASNELAIEQGACPWFNETTYAQGILPIDTYKKDLDGIVS--EPLHYDWEALRESIKT 608
Enterobacter638_NrdAg                   LMKASNELAKEQGACPWFNETTYAKGILPIDTYKKDLDAIVN--EPLHLDWETLRESIKT 608
C_koseriBAA895_NrdAg                    LLKASNELAKEQGACPWFNETTYAQGILPIDTYKKDLDAIAN--EALHYNWEALRESIKT 608
S_boydiiSb227_NrdAg                     LLKASNELAKEQGECPWFNETTYAKGILPIDTYKKDLDAIAN--EPLHYDWEALRESIKT 608
E_coliK12ssMG1655_NrdAg                 LLKASNELAKEQGACPWFNETTYAKGILPIDTYKKDLDTIAN--EPLHYDWEALRESIKT 608
S_flexneri58401_NrdAg                   LLKASNELAKEQGACPWFNETTYAKGILPIDTYKKDLDAIAN--EPLHYDWEALRESIKT 608
E_coli536_NrdAg                         LLKASNELAKEQGACPWFNETTYAKGILPIDTYKKDLDAIAN--EPLHYDWEALRESIKT 608
S_flexneri2a2457T_NrdAg                 LLKASNELAKEQGACPWFNETTYAKGILPIDTYKKDLDAIAN--EPLHYDWEALRESIKT 608
STyphimuriumLT2_NrdAg                   LLKASNELAKEQGACPWFNETTYAKGILPIDTYKKDLDAIVN--EPLHYDWEQLRESIKT 608
S_entericaTyphiTy2_NrdAg                LLKASNELAKEQGACPWFNETTYAKGILPIDTYKKDLDAIVN--EPLHYDWEQLRESIKT 608
S_proteamaculans568_NrdAg               LLKASNRLAQEQGACPWFNETTYSQGILPIDTYKKDLDVICQ--EPLHYDWETLRKEIKE 608
Y_pestisAngola_NrdAg                    LLKASNALAREQGACPWFNETTYSQGILPIDTYKKDLDAISN--EPLHYDWETLRKEIQT 608
YpseudotuberculosisIP31758_NrdAg        LLKASNALAREQGACPWFNETTYSQGILPIDTYKKDLDAISN--EPLHYDWETLRKEIQT 608
                                           :: ::          *   * . . . ::          * :
```

```
SalmonellaphageFelixO1_NrdAg          NGGIRNSVLEAYMPNESSSLATNTTNGLYPIRDFILTKKSATGNVLFIVPDYEELK--YV 659
Escherichiaphagerv5_NrdAg             NGGLRNSVLEAYMPNESSSLATNTTNGVYPVREHIIFKKSPQGQVLFIVPDYDELK--EY 659
EnterophageT5_NrdAg                   Y-GMRNSVLTAQMPGESSSVLLGVTNSIEPPRKIVSIKGSAVNKVIAIAPGATDWETLMS 721
VibriophageKVP40_NrdAg                Y-GMRNDTTSAFMPCESSSQISNSTNGIEPPRGLVSTKQSKDGVFNQVVPEVDLLK--DE 652
AeromonasphageAeh1_NrdAg              H-GLRNDTLSAFMPCESSSQITNSTNGIEPPRGLVSTKQSKDGVFNQLVPEVDMLG--SE 706
EnterophageRB49_NrdAg                 Y-GIRNSTLSALMPCESSSQVSNSTNGIEPPRGIVSVKGSKEGQFNQVIPNVENNA--DV 658
EnterophagePhi1_NrdAg                 Y-GIRNSTLSALMPCESSSQVSNSTNGIEPPRGIVSVKGSKEGQFNQVIPNVENNA--DV 658
EnterobacteriaphageT4_NrdAg           F-GIRNSTLSALMPCESSSQVSNSTNGYEPPRGPVSVKESKEGSFNQVVPNIEHNI--DL 659
EnterophageRB32_NrdAg                 F-GIRNSTLSALMPCESSSQVSNSTNGIEPPRGPVSVKESKEGSFNQVVPNIEHNI--DL 659
EnterophageJS98_NrdAg                 H-GIRNSTLSALMPCESSSQVSNSTNGIEPPRGPVSIKESKEGNFRQVVPNVEHNQ--GL 656
EnterophageRB69_NrdAg                 H-GIRNSTLSALMPCESSSQVSNSTNGIEPPRGPVSIKESKEGNFRQVVPNIEHNM--SL 656
ProchlorophagePSSM2_NrdAg             Y-GIRNSTLSAQMPSESSSVVCNATNGIEPPRGYLSIKKSKKGPLKQIVPSYNTLK--NN 662
ProchlorococcusphagePSSM4_NrdAg       Y-GLRHSTLSAQMPSESSSIVSNATNGIEPPRDYLSIKKSKKGPLKQIVPSYSTLK--NN 662
Synechococcusphagesyn9_NrdAg          Y-GLRHSTLSAQMPSESSSVVSNATNGIEPPRDYLSIKKSKKGPLKQIVPQFNSLK--AN 663
Y_pestisMicrotus91001_NrdAg           H-GLRNSTLTAQMPCETSSQITNSTNGIEPPRGPVSVKSSKDGIVKMVVPDFAQLK--DQ 668
Y_pestisPestoidesF_NrdAg              H-GLRNSTLTAQMPCETSSQITNSTNGIEPPRGPVSVKSSKDGIVKMVVPDFAQLK--DQ 668
B_aphidicolaBp_NrdAg                  Y-GLRNSTLSALMPSETSSQISNATNGIEPPRGLISIKASKDGMLRQVVPEYKRLK--NN 665
B_aphidicolaSg_NrdAg                  Y-GLRNSTLSALMPSETSSQISNATNGIEPPRGFISIKSSKDGMLRQVVPEYKKLK--SE 665
C_RuthiaCm_NrdAg                      T-GLRNSTLTALMPCESSSQISNSTNGIEPPRGFVSIKQSKDGILKQIVPEYERLK--NQ 665
T_crunogenaXCL2_NrdAg                 H-GLRNSTMTALMPCETSSQITNSTNGIEPPRGYVSVKASKDGILKQVVPGFKEFR--DQ 671
M_flagellatusKT_NrdAg                 H-GLRNSTLTALMPSETSSQIANATNGIEPPRGLVSVKQSKDGILRQVVPEIERLR--NQ 682
N_gonorrhoeaeFA1090_NrdAg             Y-GLRNSTLTALMPSETSSQIANATNGIEPPRGLVTVKASKDGILKQVVPEFETLK--NA 667
N_meningitidis053442_NrdAg            Y-GLRNSTLTALMPSETSSQIANATNGIEPPRGLVTVKASKDGILKQVVPEFETLK--NA 667
N_meningitidiMC58_NrdAg               Y-GLRNSTLTALMPSETSSQIANATNGIEPPRGLVTVKASKDGILKQVVPEFETLK--NA 667
P_atlanticaT6c_NrdAg                  Y-GLRNSTLSALMPSETSSQISNATNGIEPPRGHISVKSSKDGVLKQVVPEYELLK--DK 665
I_loihiensisL2TR_NrdAg                H-GMRNSTLSALMPSETSSQISNATNGIEPPRGFVSIKSSKDGVTKQVVPDYENLK--DK 665
P_tunicataD2_NrdAg                    Y-GMRNSTLSALMPSETSSQISNATNGIEPPRGHVSVKASKDGILKQVVPEYERLK--DQ 665
PsychrobacterPRwf1_NrdAg              H-GLRNSTLSALMPSETSSQIANATNGIEPPRGLVSVKASKDGILKQVVPDILNLK--GN 667
P_arcticus2734_NrdAg                  H-GLRNSTLTALMPSETSSQIANATNGIEPPRGLVSIKASKDGILKQVVPEIDKLK--DQ 667
P_cryohalolentisK5_NrdAg              H-GLRNSTLTALMPSETSSQIANATNGIEPPRGLVSIKASKDGILKQVVPEINKLK--NQ 667
CBlochmanniaBPEN_NrdAg                Y-GLRNSTLSALMPSETSSQISNATNGIEPPRGYISIKTSKNGILRQVVPEFIKLK--SD 665
P_ingrahamii37_NrdAg                  H-GLRNSTLSALMPSETSSQISNATNGIEPPRGFVSVKASKDGILKQVVPEYEKYK--QN 665
MarinomonasMWYL1_NrdAg                H-GLRNSTLTALMPSETSSQISNATNGIEPPRGFVSIKASKDGILKQVVPEFERLK--EN 664
S_denitrificansOS217_NrdAg            H-GLRNSTLSALMPSETSSQISNATNGIEPPRGLISVKASKDGQLKQVVPDFDELR--DS 665
S_amazonensisSB2B_NrdAg               Y-GLRNSTLSALMPSETSSQISNATNGIEPPRGLISVKASKDGQLKQVVPDFDELR--NA 685
S_frigidimarina400_NrdAg              Y-GLRNSTLSALMPSETSSQISNATNGIEPPRGLISVKASKDGQLKQVVPDFELLK--DK 665
S_oneidensisMR1_NrdAg                 Y-GLRNSTLSALMPSETSSQISNATNGIEPPRGLISVKASKDGQLKQVVPDFDTLK--NN 665
ShewanellaANA3_NrdAg                  H-GLRNSTLSALMPSETSSQISNATNGIEPPRGLISVKASKDGQLKQVVPDFDELK--YN 665
ShewanellaMR4_NrdAg                   H-GLRNSTLSALMPSETSSQISNATNGIEPPRGLISVKASKDGQLKQVVPDFDELK--YN 665
ShewanellaMR7_NrdAg                   H-GLRNSTLSALMPSETSSQISNATNGIEPPRGLISVKASKDGQLKQVVPDFDELK--YN 665
S_balticaOS155_NrdAg                  Y-GLRNSTLSALMPSETSSQISNATNGIEPPRGLISVKSSKDGQLKQVVPDFDTLK--DS 665
S_putrefaciensCN32_NrdAg              Y-GLRNSTLSALMPSETSSQISNATNGIEPPRGLISVKASKDGQLKQVVPDFDTLK--DN 665
ShewanellaW3181_NrdAg                 Y-GLRNSTLSALMPSETSSQISNATNGIEPPRGLISVKASKDGQLKQVVPDFDTLK--DN 665
S_sediminisHAWEB3_NrdAg               H-GLRNSTLSALMPSETSSQISNATNGIEPPRGLISVKSSKDGQLKQVVPDFDKYQ--YS 665
S_loihicaPV4_NrdAg                    H-GLRNSTLSALMPSETSSQISNATNGIEPPRGLISVKASKDGQLKQVVPDFEKYQ--YS 665
S_pealeana700345_NrdAg                H-GLRNSTLSALMPSETSSQISNATNGIEPPRGLISVKASKDGQLKQVVPDFEKYQ--YS 665
A_succinogenes130Z_NrdAg              Y-GLRNSTVTALMPSETSSQISNATNGIEPPRGHISVKASKDGILKQVVPDYENLG--EN 665
M_succiniciproducensMBEL55E_NrdAg     F-GLRNSTLTALMPSETSSQISNATNGIEPPRGHISVKASKDGILKQVVPDYENLS--DK 674
H_influenzae86028NP_NrdAg             F-GLRNSTLTALMPSETSSQISNATNGIEPPRGHVSIKASKDGILKQVVPEYENLM--DN 665
H_influenzaeRdKW20_NrdAg              F-GLRNSTLTALMPSETSSQISNATNGIEPPRGHVSIKASKDGILKQVVPEYENLM--DN 690
H_somnus129PT_NrdAg                   F-GLRNSTLTALMPSETSSQISNATNGIEPPRGHVSIKASKDGILKQVVPDYENLG--EN 665
P_multocidaPm70_NrdAg                 F-GLRNSTLTALMPSETSSQISNATNGIEPPRGHVSVKASKDGILKQVVPEYETLG--EN 665
A_pleuropneumoniae5bL20               F-GLRNSTLSSLMPSETSSQISNATNGIEPPRGYVSIKASKDGILKQVVPEYEKLR--SQ 665
H_ducreyi35000HP_NrdAg                F-GLRNSTLTALMPSETSSQISNATNGIEPPRGYVSIKASKDGILKQVVPEYETLK--DQ 665
V_choleraeO395_NrdAg                  H-GLRNSTLTALMPSETSSQISNATNGIEPPRGYVSVKASKDGILKQVVPEFTRLK--EN 665
V_vulnificusCMCP6_NrdAg               H-GLRNSTLTALMPSETSSQISNATNGIEPPRGYVSVKASKDGILKQVVPDFIELK--NN 665
VibrioEx25                            H-GLRNSTLTALMPSETSSQISNATNGIEPPRGYVSVKASKDGILKQVVPEFTELK--DN 665
V_campbelliiBAA1116_NrdAg             H-GLRNSTLTALMPSETSSQISNATNGIEPPRGFVSVKASKDGILKQVVPDFAELK--DN 665
VibrioAND4_NrdAg                      H-GLRNSTLTALMPSETSSQISNATNGIEPPRGFVSVKASKDGILKQVVPDFAELK--ND 665
V_fischeriES114_NrdAg                 H-GLRNSTLSALMPSETSSQISNATNGIEPPRGFVSVKASKDGILKQVVPDFANLK--NN 665
P_angustumS14_NrdAg                   H-GLRNSTLSALMPSETSSQISNATNGIEPPRGFVSVKASKDGILKQVVPEYAKYK--DN 665
PhotobacteriumSKA34_NrdAg             H-GLRNSTLSALMPSETSSQISNATNGIEPPRGFVSVKASKDGILKQVVPEYAKYK--EN 665
C_psychrerythraea34H_NrdAg            H-GVRNSTVSALMPSETSSQISNATNGIEPPRGLISIKASKDGVLKQVVPEYEKLK--GN 665
A_hydrophila7966_NrdAg                V-GLRNSTLTALMPSETSSQISNATNGIEPPRGLVSVKASKDGILKQVVPEYDRLK--DQ 675
A_salmonicidaA449_NrdAg               V-GLRNSTLTALMPSETSSQISNSTNGIEPPRGLVSVKASKDGILKQVVPEYHRLK--DQ 663
B_cicadellinicolaHc_NrdAg             S-GLRNSTLSAMMPSETSSQISNATNGIEPPRGHVSIKASKNGILRQVVPESVRLH--NA 665
C_sakazakiiBAA894                     H-GLRNSTLSALMPSETSSQISNATNGIEPPRGHISIKASKDGILRQVVPDYEQLK--DQ 665
K_pneumoniaeMGH78578_NrdAg            H-GLRNSTLSALMPSETSSQISNATNGIEPPRGHVSIKASKDGILRQVVPDYENLQ--NA 665
Enterobacter638_NrdAg                 H-GLRNSTLSALMPSETSSQISNATNGIEPPRGHVSIKASKDGVLRQVVPDYELLK--DS 665
C_koseriBAA895_NrdAg                  H-GLRNSTLSALMPSETSSQISNATNGIEPPRGHVSIKASKDGVLRQVVPDYEGLK--DR 665
S_boydiiSb227_NrdAg                   H-GLRNSTLSALMPSETSSQISNATNGIEPPRGYVSIKASKDGILRQVVPDYEHLH--DA 665
E_coliK12ssMG1655_NrdAg               H-GLRNSTLSALMPSETSSQISNATNGIEPPRGYVSIKASKDGILRQVVPDYEHLH--DA 665
S_flexneri58401_NrdAg                 H-GLRNSTLSALMPSETSSQISNATNGIEPPRGYVSIKASKDGILRQVVPDYEHLH--DA 665
E_coli536_NrdAg                       H-GLRNSTLSALMPSETSSQISNATNGIEPPRGYVSIKASKDGILRQVVPDYEHLH--DA 665
S_flexneri2a2457T_NrdAg               H-GLRNSTLSALMPSETSSQISNATNGIEPPRGYVSIKASKDGILRQVVPDYEHLH--DA 665
STyphimuriumLT2_NrdAg                 H-GLRNSTLSALMPSETSSQISNATNGIEPPRGYVSIKASKDGILRQVVPDYEHLK--DA 665
S_entericaTyphiTy2_NrdAg              H-GLRNSTLSALMPSETSSQISNATNGIEPPRGYVSIKASKDGILRQVVPDYEHLK--DA 665
S_proteamaculans568_NrdAg             T-GLRNSTLSALMPSETSSQISNATNGIEPPRGHISIKASKDGILRQVVPEYERLK--ND 665
Y_pestisAngola_NrdAg                  H-GLRNSTLSALMPSETSSQISNATNGIEPPRGYVSIKASKDGILRQVVPEYERLK--GA 665
YpseudotuberculosisIP31758_NrdAg      H-GLRNSTLSALMPSETSSQISNATNGIEPPRGYVSIKASKDGILRQVVPEYERLK--GA 665
                                        *:*...  : ** *:**     **.  * *  :  * *       : *
```

```
SalmonellaphageFelixO1_NrdAg                YEIAWDID---TFDMIDCYAIVQKFTGQAISSDFYVDYAKSK--KVSLAQALKYMIYANS 714
EscherichiaphageRv5_NrdAg                   YTSAWDVD---TNDMIDMYAVIQKFTGQAISADLYIDYTKLKDGKISMKEQLGYLIRATK 716
EnterophageT5_NrdAg                         YKLAYDV---DRIEWIKWVATMQKFFSQSISTNMYYDYTKFENEIIPGPVVVRDFMTAVK 778
VibriophageKVP40_NrdAg                      YGLAWKVCEKDNFGYLTLVAIMQKFVDQSISANQYYDPSKFPNGKIPMERIFEELAFCSH 712
AeromonasphageAeh1_NrdAg                    YETAWQMTKHGMKGYLDIVAIMQKFVDQMISANTYYDPAVYPGEKIPMKVVLKDISRAQK 766
EnterophageRB49_NrdAg                       IDLLWTLIKKGNKPYLTLAAIMQKFVDQAISSNTSYDPSTYKTGKVDIADVLDDMLFAQY 718
EnterophagePhi1_NrdAg                       IDLLWTLIKKGNKPYLTLAAIMQKFVDQAISSNTSYDPSTYKTGKVDIADVLDDMLFAQY 718
EnterobacteriaphageT4_NrdAg                 YDYTWKLAKKGNKPYLTQVAIMLKWVCQSASANTYYDPQIFPKGKVPMSIMIDDMLYGWY 719
EnterophageRB32_NrdAg                       YDYTWKLAKKGNKPYLTQVAIMLKWVCQSASANTYYDPQIFPKGKVPMSIMIDDMLYGWY 719
EnterophageJS98_NrdAg                       YDYAWSLAKKGNKPYLTQVAIMLKWVCQSASANTYYDPQNYEKGKVPMSVMLDDLLYFWY 716
EnterophageRB69_NrdAg                       YDYAWKLAKKGNKPYLTQVAIMLKWVCQSASANTYYDPAVFEKGKVPMSVMLDDLLYFWY 716
ProchlorophagePSSM2_NrdAg                   YTLLWDMP--NNTGYINIVAVMQKFFDQAISGNWSYNPQHFDEGEVPTSVMAQDLLTTYK 720
ProchlorococcusphagePSSM4_NrdAg             YTLLWDMH--NNDGYIKVTAIMQKFFDQAISGNWSYNPENYPNNEVPMQQMAMDWLNTYK 720
Synechococcusphagesyn9_NrdAg                YDLLWDMP--SNEGYIKIVSVMQKFFDQAISGNWAYNPENYPNNEVPVSEMAKDLLTTYK 721
Y_pestisMicrotus91001_NrdAg                 YEYLWDMP--DNRGYLTKVAIIQKFFDQAISANTNYDPTRFLAIRFPMMKLLEDLLFAYQ 726
Y_pestisPestoidesF_NrdAg                    YEYLWDMP--DNRGYLTKVAIIQKFFDQAISANTNYDPTRFPGDKVPMMKLLEDLLFAYQ 726
B_aphidicolaBp_NrdAg                        YELLWNIP--NNTGYLQLASIMQKFVDQSISANTNYDPKLFPNNKIPMQQLILDLLTAYK 723
B_aphidicolaSg_NrdAg                        YELLWEIP--NNIGYLELVAIMQKFVDQSISANTNYDPKRFLNEKIPMKQLIYDLLVAYK 723
C_RuthiaCm_NrdAg                            YELLWDIK--DNEGYLQLCGIMQKFVDQSISTNMHYDPSQFEGNRVPMKLFLMDLFRAYK 723
T_crunogenaXCL2_NrdAg                       YELLWQIP--DNKGYLQLVGIMQKFVDQAISANTNYDPAKFEDGRVPIKLVLQDLLYAYK 729
M_flagellatusKT_NrdAg                       YQLLWSLP--NNDGYLKLVGIMQKFVDQSISSNTNYDPKRFEGGKVPMKQLLKDLLQAYK 740
N_gonorrhoeaeFA1090_NrdAg                   YETLWQLP--GNEGYLKLVGVMQKFVDQAISANTAYDPGKFEGNKVSMKQMLKDLLTAYK 725
N_meningitidis053442_NrdAg                  YETLWQLP--GNEGYLKLVGVMQKFVDQSISANTAYDPGKFEGGKVSMKQMLKDLLTAYK 725
N_meningitidiMC58_NrdAg                     YETLWQLP--GNEGYLKLVGVMQKFVDQSISANTAYDPGKFEGGKVSMKQMLKDLLTAYK 725
P_atlanticaT6c_NrdAg                        YELLWDLP--SNDGYLQLTGIMQKFVDQTISANTSYDPNKYDGGKVPMKLLLKDLL-TYQ 722
I_loihiensisL2TR_NrdAg                      YELLWQIP--NNKGYLELVGIMQKFVDQTISANTNYDPSKFQAGKVPMKQLLQDLLTAYK 723
P_tunicataD2_NrdAg                          YELLWDIP--SNDGYLQLVGIMQKFIDQTVSANTNYDPGKFEGGKVPMKLLLKDLLTAYK 723
PsychrobacterPRwf1_NrdAg                    YELLWQMP--NNDGYLKLVAVMQKFIDQSISANTNYDPTRFEGNRVPMKVLLKDLLTAYK 725
P_arcticus2734_NrdAg                        YELLWQMP--NNDGYLKLVAVMQKFVDQSISANTNYDPVRYEGGRVPMKILLKDLLNAYK 725
P_cryohalolentisK5_NrdAg                    YELLWQMP--NNDGYLKLVAVMQKFVDQSISANTNYDPVRYEGGRVPMKILLKDLLNAYK 725
CBlochmanniaBPEN_NrdAg                      YELLWQIP--DNRGYLQLVGIMQKFIDQSISTNTNYDPSRFPENKVPMKQLLSDLLYAYR 723
P_ingrahamii37_NrdAg                        YELLWDIP--NNQGYLELVGIMQKFIDQAISANTNYDPKRFSEGRVPMTNLLKDLLVAYK 723
MarinomonasMWYL1_NrdAg                      YELLWTIP--SNDGYLQLVGIMQKFVDQAISANTNYDPQKFESQKVPMKVILKDLLTAYK 722
S_denitrificansOS217_NrdAg                  YELLWNMP--NNDGYLQLVGIMQKFVDQAISANTNYDPSRFEGRKVPMQTLLKDLLGAYK 723
S_amazonensisSB2B_NrdAg                     YELLWQMP--SNDGYLQLVGIMQKFVDQAISANTNYDPAKFEGRKVPMQLLLKDLLTAYK 743
S_frigidimarina400_NrdAg                    YELLWNMP--NNDGYIQLVGLMQKFIDQAISANTNYDPSRFEGGRVPMQTLLKDLLSVYK 723
S_oneidensisMR1_NrdAg                       YELLWQMP--SNEGYLQLVGIMQKFVDQAISANTNYDPARFEGRKVPMQTLLKDLLNAYK 723
ShewanellaANA3_NrdAg                        YELLWQMP--SNEGYLQLVGIMQKFVDQAISANTNYDPTRFEGRKVPMQTLLKDLLNAYK 723
ShewanellaMR4_NrdAg                         YELLWQMP--SNEGYLQLVGIMQKFVDQAISANTNYDPTRFEGRKVPMQTLLKDLLNAYK 723
ShewanellaMR7_NrdAg                         YELLWQMP--SNEGYLQLVGIMQKFVDQAISANTNYDPTRFEGRKVPMQTLLKDLLNAYK 723
S_balticaOS155_NrdAg                        YELLWKMP--SNDGYLQLVGIMQKFVDQAISANTNYDPSRFEGRKVPMQTLLKDLLNAYK 723
S_putrefaciensCN32_NrdAg                    YELLWKMP--SNDGYLQLVGIMQKFVDQAISANTNYDPSRFEGRKVPMQTLLKDLLNAYK 723
ShewanellaW3181_NrdAg                       YELLWKMP--SNDGYLQLVGIMQKFVDQAISANTNYDPSRFEGRKVPMQTLLKDLLNAYK 723
S_sediminisHAWEB3_NrdAg                     YELLWQMP--NNEGYLQLVGLMQKFVDQAISANTNYDPSRFEGQKVPMQVLLKDLLNAYK 723
S_loihicaPV4_NrdAg                          YELLWQMP--GNEGYLQLVGLMQKFVDQSISANTNYDPSRFPGAKVPMQVLLKDLLTAYK 723
S_pealeana700345_NrdAg                      YELLWQMP--NNDGYLQLVGLMQKFVDQSISANTNYDPSRFEGGRVPMKVLLKDLLTAYK 723
A_succinogenes130Z_NrdAg                    YELLWDMP--TMDGYLHLVGIMQKFVDQAISANTNYDPQRFEDGKVPMKVLLKDLLTAYK 723
M_succiniciproducensMBEL55E_NrdAg           YELLWDMP--NMDGYLHLVGIMQKFVDQSISANTNYDPKRFEDDKVPMKILLKDLLTAYK 732
H_influenzae86028NP_NrdAg                   YELLWDIP--SNDGYLHLVGIMQKFVDQAISANTNYDPKRFEDGKVPMKVLLKDLLTAYK 723
H_influenzaeRdKW20_NrdAg                    YELLWDIP--SNDGYLHLVGIMQKFVDQAISANTNYDPKRFEDGKVPMKVLLKDLLTAYK 748
H_somnus129PT_NrdAg                         YELLWDIP--SNDGYLHLVGIMQKFIDQAISANTNYDPKRFEDGKVPMKVLLKDLLTAYK 723
P_multocidaPm70_NrdAg                       YELLWDLP--NNDGYLHLVGIMQKFVDQAISANTNYDPQRFEDGKVPMKRLLGDLLTAYK 723
A_pleuropneumoniae5bL20                     YELLWDIP--DMSGYLQLVGIMQKFIDQAISANTNYDPQKFEDGKVPMKRLLADLLTAYK 723
H_ducreyi35000HP_NrdAg                      YELLWDIP--NMDGYLHLVGIMQKFVDQAISANTNYDPKRFEDDKVPMKILLKDLLTAYK 723
V_choleraeO395_NrdAg                        YELLWNIG--SNDGYLHLVGIMQKFVDQAISANTNYDPSNFDSGKVPMKKLLQDLLTAYK 723
V_vulnificusCMCP6_NrdAg                     YELLWNIG--SNDGYLHLVGIMQKFVDQAISANTNYDPARYESGKVPMKKLLQDLLTAYK 723
VibrioEx25                                  YELLWNIG--SNDGYLHLVGIMQKFVDQAISANTNYDPSRYETGKVPMKKLLQDLLTAYK 723
V_campbelliiBAA1116_NrdAg                   YELLWNIG--SNDGYLHLVGIMQKFVDQAISANTNYDPSGYESGKVPMKKLLQDLLTAYK 723
VibrioAND4_NrdAg                            YELLWNIG--SNEGYLHLVGIMQKFVDQAISANTNYDPSGYDSGKVPMKKLLQDLLTAYK 723
V_fischeriES114_NrdAg                       YELLWNIG--SNDGYLQLVGIMQKFVDQAISANTNYDPNIQPNGKVPMKLLLKDLLTAYK 723
P_angustumS14_NrdAg                         YELLWNIG--NNDGYLQLVGIMQKFIDQAISANTNYDPSKQAGGKVPMKLLLKDLLNAYK 723
PhotobacteriumSKA34_NrdAg                   YELLWNIG--SNDGYLQLVGIMQKFIDQAISANTNYDPSKQAGGKVPMKLLLKDLLNAYK 723
C_psychrerythraea34H_NrdAg                  YELLWNIP--DNQGYLELVGIMQKFVDQAISANTNYDPNKFEGGKVPVKQILKDILTAYK 723
A_hydrophila7966_NrdAg                      YELLWKQP--SVDGYLQLVGIMQKFVDQAISANTNYDPARFEGNKVPMKQLLKDLLTAYK 733
A_salmonicidaA449_NrdAg                     YELLWKQP--SVDGYLQLVGIMQKFVDQAISANTNYDPARFEGNKVPMKQLLKDLLTAYK 721
B_cicadellinicolaHc_NrdAg                   YEFLWDMP--NNNGYLHLVGVMQKFVDQSISANTNYDPSRYPYGKVPMKQLLKDLLTAYK 723
C_sakazakiiBAA894                           YELLWEMP--GNDGYLQLVGIMQKFIDQSISANTNYDPTRFPSGKVPMQQLLKDLLTAYK 723
K_pneumoniaeMGH78578_NrdAg                  YELLWEMP--NNDGYLQLVGIMQKFIDQSISANTNYDPTRFPSGKVPMQQLLKDLLNAYK 723
Enterobacter638_NrdAg                       YELLWEMP--NNDGYLQLVGIMQKFIDQSISANTNYDPTRFPSGKVPMQQLLKDLLTAYK 723
C_koseriBAA895_NrdAg                        YELLWEMP--SNDGYLQLVGIMQKFIDQAISANTNYDPTRFPSGKVPMQQLLKDLLTAYK 723
S_boydiiSb227_NrdAg                         YELLWEMP--GNDGYLQLVGIMQKFIDQSISANTNYDPSRFPSGKVPMQQLLKDLLTAYK 723
E_coliK12ssMG1655_NrdAg                     YELLWEMP--GNDGYLQLVGIMQKFIDQSISANTNYDPSRFPSGKVPMQQLLKDLLTAYK 723
S_flexneri58401_NrdAg                       YELLWEMP--GNDGYLQLVGIMQKFIDQSISANTNYDPSRFPSGKVPMQQLLKDLLTAYK 723
E_coli536_NrdAg                             YELLWEMP--GNDGYLQLVGIMQKFIDQSISANTNYDPSRFPSGKVPMQQLLKDLLTAYK 723
S_flexneri2a2457T_NrdAg                     YELLWEMP--GNDGYLQLVGIMQKFIDQSISANTNYDPSRFPSGKVPMQQLLKDLLTAYK 723
STyphimuriumLT2_NrdAg                       YELLWEMP--NNDGYLQLVGIMQKFIDQSISANTNYDPSRFPSGKVPMQQLLKDLLTAYK 723
S_entericaTyphiTy2_NrdAg                    YELLWEMP--NNDGYLQLVGIMQKFIDQSISANTNYDPSRFPSGKVPMQQLLKDLLTAYK 723
S_proteamaculans568_NrdAg                   YELLWEMP--NNDGYLQLVGLMQKFIDQSISANTNYDPTRFPGGKVPMKQLLKDLLTTYK 723
Y_pestisAngola_NrdAg                        YELLWEMP--NNDGYLQLVGLMQKFVDQAISANTNYDPTRFPSGKVPMKQLLKDLLTTYK 723
YpseudotuberculosisIP31758_NrdAg            YELLWEMP--NNDGYLQLVGLMQKFVDQAISANTNYDPTRFPSGKVPMKQLLKDLLTTYK 723
                                                :          :   . : *:  *  * :   :         .
```

```
SalmonellaphageFelixO1_NrdAg          VGMKTMYYLNSRIGVGKSA-------------------------LQDAYCE--GCGV 744
EscherichiaphagerV5_NrdAg             FGMKTWYYLNSRVGAGDSLAEELHAKKQAED--ETRKNYPLIADSSVEDPYCE--SCAL 771
EnterophageT5_NrdAg                   YGWKTWYYANFNTANGGGAG------------------------EEAAGCASGGCQL 811
VibriophageKVP40_NrdAg                YGIKTLYYSNTRDGADGTE-------------------------DG-EDCD--GCKV 741
AeromonasphageAeh1_NrdAg              HGVKSLYYSNTRDGADGTE-------------------------EQDSSCD--VCKV 796
EnterophageRB49_NrdAg                 YGLKTLYYANVRDGADGT--------------------------EDASECS--SCKL 747
EnterophagePhi1_NrdAg                 YGLKTLYYANVRDGADGT--------------------------EDSSECS--SCKL 747
EnterobacteriaphageT4_NrdAg           YGIKNFYYHNTRDGSGTDDYEIE---------------------TPKADDCA--ACKL 754
EnterophageRB32_NrdAg                 YGIKNFYYHNTRDGSGTDDYEIE---------------------TPKAEDCS--SCKL 754
EnterophageJS98_NrdAg                 FGGKNLYYHNTRDGSGTDDYEIE---------------------TPKADDCA--ACKL 751
EnterophageRB69_NrdAg                 FGGKNLYYHNTRDGSGTDDYELE---------------------TPKAEDCS--ACKL 751
ProchlorophagePSSM2_NrdAg             YGWKTSYYQNTYDVKTDEIDLSVSNPNDVGIQGSTRLENLISEL--ENQEECE--SCAI 775
ProchlorococcusphagePSSM4_NrdAg       YGWKTSYYQNTYDAKKDGDDD---------LAPPVPTENLINDILTLEEENCE--SCTI 768
Synechococcusphagesyn9_NrdAg          YGWKTSYYQNTYDAKKDAPIE-----------ESGDVDKLINEILQSEEDDCD--SCKV 767
Y_pestisMicrotus91001_NrdAg           QGVKTLYYHNTRDGAGKRDDDDQVSE-------------PSAAEVVEPEDECDGACKI 771
Y_pestisPestoidesF_NrdAg              QGVKTLYYHNTRDGAGKRDDDDQVSE-------------PSAAEVVEPEDECDGACKI 771
B_aphidicolaBp_NrdAg                  LGLKTLYYQNTRDGASDHQTE--------------------DIQFSKNENCNHGACKI 761
B_aphidicolaSg_NrdAg                  LGIKTLYYQNTRDGAEDNQNLNK------------------SIENIKEDNCTSGSCTI 763
C_RuthiaCm_NrdAg                      YGIKTLYYHTTRDGGDDWQEKE-----------------------EDNACADGVCKL 757
T_crunogenaXCL2_NrdAg                 AGLKTLYYHNTRDGADDSHE-------------------------DDGCAGGACKL 760
M_flagellatusKT_NrdAg                 LGIKTLYYHNTRDGASEGGAEAG----------------------DDGCESGACKI 774
N_gonorrhoeaeFA1090_NrdAg             YGVKTLYYHNTRDGADDTQTDIQ----------------------DDGCAGGACKI 759
N_meningitidis053442_NrdAg            YGVKTLYYHNTRDGADDTQTDIQ----------------------DDGCAGGACKI 759
N_meningitidiMC58_NrdAg               YGVKTLYYHNTRDGADDTQTDIQ----------------------DDGCAGGACKI 759
P_atlanticaT6c_NrdAg                  LGVKTLYYHNTRDGAADTSQQ-----E------LKSTQFVPQEAVVEEDDDCAGGACKI 770
I_loihiensisL2TR_NrdAg                LGVKTLYYHNTRDGAADNQSDIAARKG------AAASQAQPTEVIIEEDDDCAGGACKI 776
P_tunicataD2_NrdAg                    FGVKTLYYHNTRDGASDAQNE--------------------IVADADDDCAGGACKI 760
PsychrobacterPRwf1_NrdAg              LGVKTLYYHNTRDGANDAQADI-----------------------DDGCASGACKL 758
P_arcticus2734_NrdAg                  MGVKTLYYHNTRDGANDAQADM-----------------------EDDCAGGACKI 758
P_cryohalolentisK5_NrdAg              LGVKTLYYHNTRDGANDAQADM-----------------------EDDCAGGACKI 758
CBlochmanniaBPEN_NrdAg                LGIKTLYYHNTRDGAQDVQKTLP------------------H--NQSNICTSDACTI 760
P_ingrahamii37_NrdAg                  MGVKTLYYHNTRDGADDENI-------------------------IDEADDCAGGACKI 757
MarinomonasMWYL1_NrdAg                LGVKTLYYHNTRDGADDKQEDM-----------------------DDCAGGACKI 754
S_denitrificansOS217_NrdAg            LGVKTLYYHNTRDGASDQFDEIT------------------AVIEEDDDCAGGACKI 762
S_amazonensisSB2B_NrdAg               LGLKTLYYHNTRDGATDQHDDIA------------------VVEKEDDGCAGGACKI 782
S_frigidimarina400_NrdAg              LGIKTLYYHNTRDGASDQFDDII------------------AVEEEDDGCGGGACKI 762
S_oneidensisMR1_NrdAg                 LGVKTLYYHNTRDGASDQHDDIT------------------NIEKEDDGCAGGACKI 762
ShewanellaANA3_NrdAg                  LGVKTLYYHNTRDGASDQHDDIT------------------NIEKEDDGCAGGACKI 762
ShewanellaMR4_NrdAg                   LGVKTLYYHNTRDGASDQHDDIT------------------NIEKEDDGCAGGACKI 762
ShewanellaMR7_NrdAg                   LGVKTLYYHNTRDGASDQHDDIT------------------NIEKEDDGCAGGACKI 762
S_balticaOS155_NrdAg                  LGVKTLYYHNTRDGASDSHDDIT------------------NLEKEDDGCAGGACKI 762
S_putrefaciensCN32_NrdAg              LGVKTLYYHNTRDGASDQHDDIT------------------NLEKEDDGCAGGACKI 762
ShewanellaW3181_NrdAg                 LGVKTLYYHNTRDGASDQHDDIT------------------NLEKEDDGCAGGACKI 762
S_sediminisHAWEB3_NrdAg               LGVKTLYYHNTRDGATDSHSDIV------------------SVEKEDEGCEGGACKI 762
S_loihicaPV4_NrdAg                    YGVKTLYYHNTRDGASDKYDDIT------------------SIEKEDDSCAGGACKI 762
S_pealeana700345_NrdAg                YGVKTLYYHNTRDGASDSHDDIT------------------AIEKEDDSCAGGACKI 762
A_succinogenes130Z_NrdAg              YGLKTLYYQNTRDGAEDAQEDM-----------------------DDGCAGGACKI 756
M_succiniciproducensMBEL55E_NrdAg     YGLKTLYYQNTRDGADDAQEDL-----------------------DDGCAGGACKI 765
H_influenzae86028NP_NrdAg             YGLKTLYYQNTRDGAEDVQEDL-----------------------DDGCAGGACKI 756
H_influenzaeRdKW20_NrdAg              YGLKTLYYQNTRDGAEDVQEDL-----------------------DDGCAGGACKI 781
H_somnus129PT_NrdAg                   YGLKTLYYQNTRDGAEDMQEDL-----------------------DDACAGGACKI 756
P_multocidaPm70_NrdAg                 YGLKTLYYQNTRDGAEDVQEDI-----------------------DDGCSGGACKI 756
A_pleuropneumoniae5bL20               YGLKTLYYQNTRDGAEDAQEDL-----------------------NDGCAGGACKI 756
H_ducreyi35000HP_NrdAg                YGLKTLYYQNTRDGAEDSQEDL-----------------------DANCADGACRI 756
V_choleraeO395_NrdAg                  FGVKTLYYHNTRDGAKDEQTDAT------------------VK-VQEDDCASGACKI 761
V_vulnificusCMCP6_NrdAg               FGVKTLYYHNTRDGAKDEQKDA-------------------VQ-PQDDDCAGGGCKI 760
VibrioEx25                            FGVKTLYYHNTRDGAKDEQKDA-------------------VQ-PQDDDCAGGGCKI 760
V_campbelliiBAA1116_NrdAg             FGVKTLYYHNTRDGAKDDQKDA-------------------VQ-PQDDDCAGGGCKI 760
VibrioAND4_NrdAg                      FGVKTLYYHNTRDGAKDDQKDA-------------------VQ-PQDDDCAGGGCKI 760
V_fischeriES114_NrdAg                 LGLKTLYYHNTRDGASDSQDDL-----------------------GQDDDCAGGGCKI 758
P_angustumS14_NrdAg                   LGVKTLYYHNTRDGASDSQGDM-----------------------AADD-CT--SCKI 755
PhotobacteriumSKA34_NrdAg             LGVKTLYYHNTRDGASDSQGDM-----------------------AADD-CT--SCKI 755
C_psychrerythraea34H_NrdAg            LGVKTLYYHNTRDGASDDIE-------------------------IEDDSCAGGACKI 756
A_hydrophila7966_NrdAg                YGLKTLYYHNTRDGADDAQTDL-----------------------QDDGCAGGACKI 767
A_salmonicidaA449_NrdAg               YGLKTLYYHNTRDGADDAQADM-----------------------QDDGCAGGACKI 755
B_cicadellinicolaHc_NrdAg             FGIKTLYYQNTRDGAKETQEDIQ--------------------V--TQQEDCESEACKI 760
C_sakazakiiBAA894                     LGVKTLYYQNTRDGAEDAQDDLA--------------------PSI-QDDGCESGACKI 761
K_pneumoniaeMGH78578_NrdAg            FGVKTLYYHNTRDGAEDAQDDLA--------------------PSI-QDDGCESGACKI 761
Enterobacter638_NrdAg                 FGVKTLYYHNTRDGAEDAQDDMA--------------------PSI-QDDGCESGACKI 761
C_koseriBAA895_NrdAg                  FGVKTLYYHNTRDGAEDAQDDLV--------------------PSI-QDDGCESGACKI 761
S_boydiiSb227_NrdAg                   FGVKTLYYQNTRDGAEDAQDDLV--------------------PSI-QDDGCESGACKI 761
E_coliK12ssMG1655_NrdAg               FGVKTLYYQNTRDGAEDAQDDLV--------------------PSI-QDDGCESGACKI 761
S_flexneri58401_NrdAg                 FGVKTLYYQNTRDGAEDAQDDLV--------------------PSI-QDDGCESGACKI 761
E_coli536_NrdAg                       FGVKTLYYQNTRDGAEDAQDDLV--------------------PSI-QDDGCESGACKI 761
S_flexneri2a2457T_NrdAg               FGVKTLYYQNTRDGAEDAQDDLV--------------------PSI-QDDGCESGACKI 761
STyphimuriumLT2_NrdAg                 FGVKTLYYQNTRDGAEDAQDDLA--------------------PSI-QDDGCESGACKI 761
S_entericaTyphiTy2_NrdAg              FGVKTLYYQNTRDGAEDAQDDLA--------------------PSI-QDDGCESGACKI 761
S_proteamaculans568_NrdAg             FGVKTLYYQNTRDGADDIQEDLL--------------------PAKGGDDDCESGACKI 762
Y_pestisAngola_NrdAg                  FGVKTLYYQNTRDGADDVQEDIQ--------------------SSLPGDDDCEGGACKI 762
YpseudotuberculosisIP31758_NrdAg      FGVKTLYYQNTRDGADDVQEDIQ--------------------SSLPGDDDCEGGACKI 762
                                       * *. ** .                                  * :
```

Appendix 1, Section 3

Class Ia RNR, Group G Sequence Alignments NrdB Proteins

E. coli K-12, subspecies MG1655 NrdB sequence highlighted in red.

* = Absolutely conserved residue
: = Conserved with strongly similar properties.
. = Conserved with weakly similar properties.

```
VibriophageKVP40_NrdBg                          ---------- 0
AeromonasphageAeh1_NrdBg                        ---------- 0
EnterophageRB49_NrdBg                           ---------- 0
EnterophagePhi1_NrdBg                           ---------- 0
EnterophageRB69_NrdBg                           ---------- 0
EnterophageJS98_NrdBg                           ---------- 0
EnterophageT4_NrdBg                             ---------- 0
EnterophageRB32_NrdBg                           ---------- 0
ProchlorococcusphagePSSM2_NrdBg                 ---------- 0
ProchlorococcusphagePSSM4_NrdBg                 ---------- 0
Y_pestisAngola_NrdBg_1                          ---------- 0
Y_pestisMicrotus91001_NrdBg_1                   ---------- 0
Y_pestisPestoidesF_NrdBg_2                      ---------- 0
C_RuthiamagnificaCm_NrdBg                       ---------- 0
C_psychrerythraea34H_NrdBg                      ---------- 0
S_sediminisHAWEB3_NrdBg                         ---------- 0
I_loihiensisL2TR_NrdBg                          ---------- 0
P_atlanticaT6c_NrdBg                            ---------- 0
B_aphidicolaBp_NrdBg                            ---------- 0
B_aphidicolaSg_NrdBg                            ---------- 0
P_ingrahamii37_NrdBg                            ---------- 0
MarinomonasMWYL1_NrdBg                          ---------- 0
P_tunicataD2_NrdBg                              ---------- 0
S_loihicaPV4_NrdBg                              ---------- 0
S_pealeana700345_NrdBg                          ---------- 0
S_denitrificansOS217_NrdBg                      ---------- 0
S_frigidimarina400_NrdBg                        ---------- 0
S_amazonensisSB2B_NrdBg                         ---------- 0
S_oneidensisMR1_NrdBg                           ---------- 0
ShewanellaANA3_NrdBg                            ---------- 0
ShewanellaMR4_NrdBg                             ---------- 0
ShewanellaMR7_NrdBg                             ---------- 0
S_balticaOS155_NrdBg                            ---------- 0
S_putrefaciensCN32_NrdBg                        ---------- 0
ShewanellaW3181_NrdBg                           ---------- 0
PsychrobacterPRwf1_NrdBg                        ---------- 0
P_arcticus2734_NrdBg                            ---------- 0
P_cryohalolentisK5_NrdBg                        ---------- 0
CBlochmanniaBPEN_NrdBg                          ---------- 0
A_hydrophila7966_NrdBg                          ---------- 0
A_salmonicidaA449_NrdBg                         ---------- 0
P_multocidaPm70_NrdBg                           ---------- 0
H_somnus129PT_NrdBg                             ---------- 0
H_influenzae86028NP_NrdBg                       ---------- 0
H_influenzaeRdKW20_NrdBg                        ---------- 0
A_succinogenes130Z_NrdBg                        ---------- 0
M_succiniciproducensMBEL55E_NrdBg               ---------- 0
A_pleuropneumoniae5bL20_NrdBg                   ---------- 0
H_ducreyi35000HP_NrdBg                          ---------- 0
B_cicadellinicolaHc_NrdBg                       ---------- 0
C_sakazakiiBAA894_NrdBg                         ---------- 0
Enterobacter638_NrdBg                           ---------- 0
C_koseriBAA895_NrdBg                            ---------- 0
K_pneumoniaeMGH78578_NrdBg                      ---------- 0
E_coli536_NrdBg                                 ---------- 0
S_flexneri2a2457T_NrdBg                         ---------- 0
S_flexneri58401_NrdBg                           MRCRMRRKRLIRPTARFVGLIRRASVASGSGCRMQRERLIRPTARICRPDKTRQRRIRHR 60
E_coliK12MG1655_NrdBg                           ---------- 0
S_boydiiSb227_NrdBg                             ---------- 0
S_TyphimuriumLT2_NrdBg                          ---------- 0
S_entericaTy2_NrdBg                             ---------- 0
S_proteamaculans568_NrdBg                       ---------- 0
Y_pestisAngolaNrdBg_2                           ---------- 0
Y_pestisMicrotus91001_NrdBg_2                   ---------- 0
Y_pestisPestoidesF_NrdBg_1                      ---------- 0
Y_pseudotuberculosisIP31758_NrdBg               ---------- 0
M_flagellatusKT_NrdBg                           ---------- 0
N_gonorrhoeaeFA1090_NrdBg                       ---------- 0
N_meningitidis053442_NrdBg                      ---------- 0
N_meningitidisMC58_NrdBg                        ---------- 0
V_choleraeO395_NrdBg                            ---------- 0
V_vulnificusCMCP6_NrdBg                         ---------- 0
V_campbelliiBAA1116_NrdBg                       ---------- 0
VibrioAND4_NrdBg                                ---------- 0
V_fischeriES114_NrdBg                           ---------- 0
P_angustumS14_NrdBg                             ---------- 0
PhotobacteriumSKA34_NrdBg                       ---------- 0
```

```
VibriophageKVP40_NrdBg                 ----------------MSKTIFNKNKVDHLGQPLFLGEDPNISRFETVKHGVFEQLTEK 43
AeromonasphageAeh1_NrdBg               ----------------MDTVFNKNVVDHMKAPLFLGEAMGLARYETVKHPAIDKLTEK 42
EnterophageRB49_NrdBg                  ----------------MSVFNREHVDIMNEPMFLGSGLGIARYDIQRHQVFEELIEK 41
EnterophagePhi1_NrdBg                  ----------------MSVFNREHVDIMNEPMFLGSGLGIARYDIQRHQVFEELIEK 41
EnterophageRB69_NrdBg                  ----------------MSTVFNTQQVDVLNEPMFFGSGLGLARYDIQRHKVFEELIE- 41
EnterophageJS98_NrdBg                  ----------------MSTVFNTTPVDVLVEPMFFGSGLGIARYDIQRHKVFEELCEK 42
EnterophageT4_NrdBg                    ----------------MSTVFNTNPVDVLNEPMFFGSGLGLARYDIQRHRVFEELIER 42
EnterophageRB32_NrdBg                  ----------------MSTVFNTNPVDVLKEPMFFGSGLGIARYDIQRHRVFEELTEK 42
ProchlorococcusphagePSSM2_NrdBg        ------------MAEVKGMTVFNTEEVNTKKQPMFFGKPLGVQRYDNFKYPSFENLTKQ 47
ProchlorococcusphagePSSM4_NrdBg        --------MSPAQSKPVESMTVFNTEEVDTKKQPMFFGKPLGVQRYDSYKYPVFDKLTQQ 52
Y_pestisAngola_NrdBg_1                 ----------------MTYSTFRLGANDATKEPMFLGQSVNVARYDQQKYRDFEKLIEK 43
Y_pestisMicrotus91001_NrdBg_1          ----------------MTYSTFRLGANDATKEPMFLGQSVNVARYDQQKYRDFEKLIEK 43
Y_pestisPestoidesF_NrdBg_2             ----------------MTYSTFRLGANDATKEPMFLGQSVNVARYDQQKYRDFEKLIEK 43
C_RuthiamagnificaCm_NrdBg              ----------------MSYSIFNKKISNTLIQPMFFGDSMNVTRFDKQKFEIFEKLTEK 43
C_psychrerythraea34H_NrdBg             ----------------MTYTTFNQVPNNALLEPMFLGNSVNVSRYDQQRYIAFEKLIEK 43
S_sediminisHAWEB3_NrdBg                ----------------MTYTTFNQVKNDQLLEPMFLGNSVNVSRYDQQKHVSFEKLIEK 43
I_loihiensisL2TR_NrdBg                 ----------------MSYSTFNQEQNNPLTEPMFLGNSVNVARYDQQKHSIFEKLIEK 43
P_atlanticaT6c_NrdBg                   ----------------MKYTTFNQINNDPLQEPMFFGNPVNVARYDQQKHSIFEKLIEK 43
B_aphidicolaBp_NrdBg                   ----------------MTYTTFSKNKNNQLYEPMFFGQSVNISRYDQQKYEIFEKLIEK 43
B_aphidicolaSg_NrdBg                   ----------------MSYTTFSKKKNNQLKEPMFFGQSVNIARYDQQKYNIFEKLIEK 43
P_ingrahamii37_NrdBg                   ----------------MAYTTFSTIPNDAMKEPMFFGNPVNVARYDQQRFAVFEMLIEK 43
MarinomonasMWYL1_NrdBg                 ----------------MSYSTFNRKHFDSTKEPMFFGENVNVARYDQQKHPIFEKLIEK 43
P_tunicataD2_NrdBg                     ----------------MSYSTFSRNHNNQLKEPMFFGQSVNVSRYDQQKHPIFEKLIEK 43
S_loihicaPV4_NrdBg                     ----------------MAYSTFCQTPNNALLEPMFLGQSVNVARYDIQKYEVFEKLIEK 43
S_pealeana700345_NrdBg                 ----------------MAYSTFCQTPNNAMLEPMFLGQSVNVARYDQQKYEVFEKLIEK 43
S_denitrificansOS217_NrdBg             ----------------MAYSTFCQTPNDATKEPMFFGQSVNVARYDQQKYDIFEKLIEK 43
S_frigidimarina400_NrdBg               ----------------MAYSTFCQTPNDSTKEPMFFGQSVNVARYDQQKYEVFEKLIEK 43
S_amazonensisSB2B_NrdBg                --MLTGGFAPFYRNREIMAYTTFCQTPNDATREPMFFGQSVNVARYDQQRYEVFEKLIEK 58
S_oneidensisMR1_NrdBg                  ----------------MTYSTFCQTPNDATREPMFFGQSVNVARYDQQKYEVFEKLIEK 43
ShewanellaANA3_NrdBg                   --MRDPFGIPTYWNTRTMTYSTFCQTPNDATREPMFFGQSVNVARYDQQKYEVFEKLIEK 58
ShewanellaMR4_NrdBg                    ----------------MTYSTFCQTPNDATREPMFFGQSVNVARYDQQKYEVFEKLIEK 43
ShewanellaMR7_NrdBg                    ----------------MTYSTFCQTPNDATREPMFFGQSVNVARYDQQKYEVFEKLIEK 43
S_balticaOS155_NrdBg                   ----------------MAYSTFCQTPNDATREPMFLGQSVNVARYDQQKYEVFEKLIEK 43
S_putrefaciensCN32_NrdBg               ----------------MAYSTFCQTPNDATREPMFLGQSVNVARYDQQKYEVFEKLIEK 43
ShewanellaW3181_NrdBg                  ----------------MAYSTFCQTPNDATREPMFLGQSVNVARYDQQKYEVFEKLIEK 43
PsychrobacterPRwf1_NrdBg               ----------------MTYSIFSQTKNDALKEPMFFGQPVNVARYDQQKHPIFEQLIEK 43
P_arcticus2734_NrdBg                   ----------------MTYSIFSQTPNNALTEPMFFGNPVNVARYDQQKHPIFEQLIEK 43
P_cryohalolentisK5_NrdBg               ----------------MTYSIFSQTPNNALTEPMFFGNPVNVARYDQQKHPIFEQLIEK 43
CBlochmanniaBPEN_NrdBg                 ----------------MIYTTFSQNKNNQLLEPMFLGQPVNIARFDQQKHSIFEKLIEK 43
A_hydrophila7966_NrdBg                 ----------------MAYSTFCQTQNDQLKEPMFFGQSVNVARYDQQKHEVFERLIEK 43
A_salmonicidaA449_NrdBg                ----------------MAYSTFCQTQNDQLKEPMFFGQSVNVARYDQQKHEVFERLIEK 43
P_multocidaPm70_NrdBg                  ----------------MAYTTFSQNKNDQLKEPMFFGQNVNVARYDQQKYETFEKLIEK 43
H_somnus129PT_NrdBg                    ----------------MAYTTFSQTKNNQLLEPMFFGQNVNVARYDQQKYETFEKLIEK 43
H_influenzae86028NP_NrdBg              ----------------MAYTTFSQNKNDQLKEPMFFGQNVNVARYDQQKYETFEKLIEK 43
H_influenzaeRdKW20_NrdBg               ----------------MAYTTFSQNKNDQLKEPMFFGQNVNVARYDQQKYETFEKLIEK 43
A_succinogenes130Z_NrdBg               ----------------MAYTTFSQNKNDQLLEPMFFGQNVNVARYDQQKYETFEKLIEK 43
M_succiniciproducensMBEL55E_NrdBg      ----------------MAYTTFSQNKNDQLKEPMFFGQNVNVSRYDQQKYETFEKLIEK 43
A_pleuropneumoniae5bL20_NrdBg          ----------------MAYTTFSQNKNDQLKEPMFFGQNVNVARYDQQKYELFEKLIEK 43
H_ducreyi35000HP_NrdBg                 ----------------MAYTTFSQNKNDQLKEPMFFGQNVNVARYDQQKYELFEKLIEK 43
B_cicadellinicolaHc_NrdBg              ----------------MVYTTFSPKKNNQLLEPMFLGQSVNVVRFDQQKYESFEKLIEK 43
C_sakazakiiBAA894_NrdBg                ----------------MAYTTFSQTKNDQLLEPMFFGQPVNVARYDQQKYEIFEKLIEK 43
Enterobacter638_NrdBg                  ----------------MAYTTFSQTKNDQLKEPMFFGQPVNVARYDQQKYDIFEKLIEK 43
C_koseriBAA895_NrdBg                   ----------------MAYTTFSQTKNDQLKEPMFFGQPVNVARYDQQKYDIFEKLIEK 43
K_pneumoniaeMGH78578_NrdBg             ----------------MAYTTFSQTKNDQLLEPMFFGQPVNVARYDQQKYDIFEKLIEK 43
E_coli536_NrdBg                        ----------------MAYTTFSQTKNDQLKEPMFFGQPVNVARYDQQKYDIFEKLIEK 43
S_flexneri2a2457T_NrdBg                ----MPYPALNSQQDTLMAYTTFSQTKNDQLKEPMFFGQPVNVARYDQQKYDIFEKLIEK 56
S_flexneri58401_NrdBg                  MRRKMPYPALNSQQDTLMAYTTFSQTKNDQLKEPMFFGQPVNVARYDQQKYDIFEKLIEK 120
E_coliK12MG1655_NrdBg                  ----------------MAYTTFSQTKNDQLKEPMFFGQPVNVARYDQQKYDIFEKLIEK 43
S_boydiiSb227_NrdBg                    ----------------MAYTTFSQTKNDQLKEPMFFGQPVNVARYDQQKYDIFEKLIEK 43
S_TyphimuriumLT2_NrdBg                 ----------------MAYTTFSQTKNDQLKEPMFFGQPVNVARYDQQKYDIFEKLIEK 43
S_entericaTy2_NrdBg                    ----------------MAYTTFSQTKNDQLKEPMFFGQPVNVARYDQQKYDIFEKLIEK 43
S_proteamaculans568_NrdBg              ----------------MAYTTFSQNKNNQLLEPMFFGQSVNVARFDQQKHEIFEKLIEK 43
Y_pestisAngolaNrdBg_2                  ----------------MAYTTFSQNKNNQLLEPMFFGQSVNVARFDQQKHAIFEKLIEK 43
Y_pestisMicrotus91001_NrdBg_2          --MPITLFVLCTTPEETMAYTTFSQNKNNQLLEPMFFGQSVNVARFDQQKHAIFEKLIEK 58
Y_pestisPestoidesF_NrdBg_1             ----------------MAYTTFSQNKNNQLLEPMFFGQSVNVARFDQQKHAIFEKLIEK 43
Y_pseudotuberculosisIP31758_NrdBg      ----------------MAYTTFSQNKNNQLLEPMFFGQSVNVARFDQQKHAIFEKLIEK 43
M_flagellatusKT_NrdBg                  ----------------MAYSIFSKVDNEQLKEPMFFGQPVNVARYDQQKYPVFEKLIEK 43
N_gonorrhoeaeFA1090_NrdBg              ----------MSCEHLVMSYSTFPKTKNDALKEPMFFGQPVNVARYDQQKYEVFEKLIEK 50
N_meningitidis053442_NrdBg             ----------MSCEHLTMSYSTFPKTKNDALNEPMFFGQPVNVARYDQQKYEVFEKLIEK 50
N_meningitidisMC58_NrdBg               ----------MSCEHLTMSYSTFPKTKNDALNEPMFFGQPVNVARYDQQKYEVFEKLIEK 50
V_choleraeO395_NrdBg                   -----MLRQEAKRMKATMAYSTFTQKKNDQLKEPMFLGQPVNVARYDQQKYEIFEKLIEK 55
V_vulnificusCMCP6_NrdBg                ----------------MAYSTFNQTKNDQLKEPMFLGQSVNVARYDQQKFEIFEKLIEK 43
V_campbelliiBAA1116_NrdBg              ----------------MAYSTFNQNKNDQLKEPMFLGQSVNVARYDQQKFEIFEKLIEK 43
VibrioAND4_NrdBg                       ----------------MAYSTFNQNKNDQLKEPMFLGQSVNVARYDQQKFEIFEKLIEK 43
V_fischeriES114_NrdBg                  ----------------MAYSTFRQEKNDQLKEPMFLGNSVNVARYDQQKFQIFEKLIEK 43
P_angustumS14_NrdBg                    ----------------MAYSTFCQTRNDQLKEPMFFGQSVNVARYDQQKFEIFEKLIEK 43
PhotobacteriumSKA34_NrdBg              ----------------MAYSTFCQTRNDQLKEPMFFGQSVNVARYDQQKFEIFEKLIEK 43
                                                        : *     :    *:*:*.   : *::  :.  :: * :
```

```
VibriophageKVP40_NrdBg                  QLSFFWRPEEVDLTLDSQEFD-RMPEHWKHVFTQNLKYQSLLDSVQGRAPNIAFLPICSD 102
AeromonasphageAeh1_NrdBg                QLSFFWRPEEVDLSKDAIDYN-QLSEAEQHVFTQNLKYQSLLDSVQGRAPNLVFLPIVSD 101
EnterophageRB49_NrdBg                   QLSFFWRPEEISLATDRIQYL-NMGDVEKRFFTENLKYQTLLDSIQGRAPSAAFGPIVSD 100
EnterophagePhi1_NrdBg                   QLSFFWRPEEISLATDRIQYL-NMGDVEKRFFTENLKYQTLLDSIQGRAPSAAFGPIVSD 100
EnterophageRB69_NrdBg                   ILSFFWRPEEVNLMLDSAQFN-KLPQFQQDICTNNLKYQSLLDSIQGRAPSAVLMALISD 100
EnterophageJS98_NrdBg                   QLSFFWRPEEVNLMMDAAQFN-KLPQFQQNIFTNNLKYQSLLDSIQGRAPSAALQALISD 101
EnterophageT4_NrdBg                     QISFFWRPEEVNLMMDAAQFN-KLPQYQQNIFTNNLKYQSLLDSIQGRAPSAVLMSLISD 101
EnterophageRB32_NrdBg                   SLSFFWRPEEVNLMMDAAQFN-KLPQYQQNIFTNNLKYQSLLDSIQGRAPSAVLMSLISD 101
ProchlorococcusphagePSSM2_NrdBg         QLGYFWRPEEVSLQKDRGDYQ-TLRPEQKHIYTSNLKYQIMLDSVQGRAPGMAFLPYCSL 106
ProchlorococcusphagePSSM4_NrdBg         QLGYFWRPEEVSLQKDRGDYQ-QLTPEQKHIFTSNLKYQIMLDSVQGRAPGMAFAPYCSL 111
Y_pestisAngola_NrdBg_1                  QLSFFWRPEEVDITTDRIDFNTKLQEHERHIFLSNLRYQTLLDSVQGRSPNATLLPLISI 103
Y_pestisMicrotus91001_NrdBg_1           QLSFFWRPEEVDITTDRIDFNTKLQEHERHIFLSNLRYQTLLDSVQGRSPNATLLPLISI 103
Y_pestisPestoidesF_NrdBg_2              QLSFFWRPEEVDITTDRIDFNTKLQEHERHIFLSNLRYQTLLDSVQGRSPNATLLPLISI 103
C_RuthiamagnificaCm_NrdBg               QLSFFWRPEEIDVSKDKIDFS-KLLPNEKHIFVSNLQYQILLDSVQGRAPNIAFLPIVSL 102
C_psychrerythraea34H_NrdBg              QLSFFWRPEEVDVSKDRADWQ-SLTDSEKHIFISNLKYQTLLDSMAARSVNAVLLPLVSL 102
S_sediminisHAWEB3_NrdBg                 QLSFFWRPEEIDISKDRADWQ-RLTESEKHIFISNLKYQTLLDSMAARSVNAVLLPLVSL 102
I_loihiensisL2TR_NrdBg                  QISFFWRPEEIDVSRDRIDFN-KLTASEKHVFISNLKYQTLLDSIQGRSPNIALLPIVSL 102
P_atlanticaT6c_NrdBg                    QISFFWRPEEVDVSKDRVDFQ-KLSESEKHIFISNLKYQTLLDSIQGRSPNIAFLPIISI 102
B_aphidicolaBp_NrdBg                    QLSFFWRPEEVDLSRDYIDFQ-NLPQHEKHIFVSNLKYQTLLDSIQGRSPNIALLPIVSL 102
B_aphidicolaSg_NrdBg                    QLSFFWRPEEIDLSKDRIDFD-NLPSNEKHIFISNLKYQTLLDSIQGRSPNIAFLPIVSI 102
P_ingrahamii37_NrdBg                    QLSFFWRPEEIDVSRDRIDFL-NLPDHEQHIFISNLKYQTLLDSIQGRSPNIALLPIVSL 102
MarinomonasMWYL1_NrdBg                  QLSFFWRPEEVDLSTDRKDFQ-RLEEHEQHIFLSNLKYQTLLDSVQGRSPNVALLPIVSL 102
P_tunicataD2_NrdBg                      QLSFFWRPEEVDVSKDRVDFE-NLPDHEKHIFLSNLKYQTLLDSIQGRSPNVALLPIVSL 102
S_loihicaPV4_NrdBg                      QLSFFWRPEEVDVSKDKIDYA-ALPDHEKHIFISNLKYQTLLDSIQGRSPNVAFLPLVSL 102
S_pealeana700345_NrdBg                  QLSFFWRPEEVDVSRDKIDYA-ALPEHEKHIFISNLKYQTLLDSIQGRSPNVAFLPLVSL 102
S_denitrificansOS217_NrdBg              QLSFFWRPEEVDVSRDKIDFT-ALPEHEKHIFISNLKYQTLLDSIQGRSPNVAFLPLVSL 102
S_frigidimarina400_NrdBg                QLSFFWRPEEVDVSRDKIDFG-NLPEHEQHIFLSNLKYQTLLDSIQGRSPNVAFLPLVSL 102
S_amazonensisSB2B_NrdBg                 QLSFFWRPEEVDVSRDKIDFG-KLPDHERHIFISNLKYQTLLDSIQGRSPNVAFLPLVSL 117
S_oneidensisMR1_NrdBg                   QLSFFWRPEEVDVSRDKIDYN-SLPDHEKHIFISNLKYQTLLDSIQGRSPNVAFLPLVSL 102
ShewanellaANA3_NrdBg                    QLSFFWRPEEVDVSRDKIDYN-SLPDHEKHIFISNLKYQTLLDSIQGRSPNVAFLPLVSL 117
ShewanellaMR4_NrdBg                     QLSFFWRPEEVDVSRDKIDYN-SLPDHEKHIFISNLKYQTLLDSIQGRSPNVAFLPLVSL 102
ShewanellaMR7_NrdBg                     QLSFFWRPEEVDVSRDKIDYN-SLPDHEKHIFISNLKYQTLLDSIQGRSPNVAFLPLVSL 102
S_balticaOS155_NrdBg                    QLSFFWRPEEVDVSRDKIDYG-ALPDHEKHIFISNLKYQTLLDSIQGRSPNVAFLPLVSL 102
S_putrefaciensCN32_NrdBg                QLSFFWRPEEVDVSRDKIDYN-DLPDHEKHIFISNLKYQTLLDSIQGRSPNVAFLPLVSL 102
ShewanellaW3181_NrdBg                   QLSFFWRPEEVDVSRDKIDYN-DLPDHEKHIFISNLKYQTLLDSIQGRSPNVAFLPLVSL 102
PsychrobacterPRwf1_NrdBg                QLSFFWRPEEVDVSKDRIDYG-NLSTHEQHIFISNLKYQTLLDSIQGRSPNVVLLPLVSI 102
P_arcticus2734_NrdBg                    QLSFFWRPEEVDVSRDRMDYG-NLSSHEQHIFISNLKYQTLLDSIQGRSPNVVLLPLVSI 102
P_cryohalolentisK5_NrdBg                QLSFFWRPEEVDVSKDRMDFG-NLSSHEQHIFLSNLKYQTLLDSIQGRSPNVVLLPLVSI 102
CBlochmanniaBPEN_NrdBg                  QLSFFWRPEEIDISKDRIDFQ-SLPDHEKHIFISNLKYQTLLDSIQGRSPNIALLPLISI 102
A_hydrophila7966_NrdBg                  QLSFFWRPEEVDVSHDRIDFQ-ALPPHEQHIFISNLKYQTLLDSIQGRSPNVALLPLVSI 102
A_salmonicidaA449_NrdBg                 QLSFFWRPEEVDVSHDRIDYQ-ALPPHEQHIFISNLKYQTLLDSIQGRSPNVALLPLVSI 102
P_multocidaPm70_NrdBg                   QLSFFWRPEEVDVSQDRIDYQ-ALPEHEKHIFISNLKYQTLLDSIQGRSPNVALLPLVSI 102
H_somnus129PT_NrdBg                     QLSFFWRPEEVDVSQDRIDYQ-ALPEHEKHIFISNLKYQTLLDSIQGRSPNVALLPLVSI 102
H_influenzae86028NP_NrdBg               QLSFFWRPEEVDVSQDRIDYA-ALPEHEKHIFISNLKYQTLLDSIQGRSPNVALLPLVSI 102
H_influenzaeRdKW20_NrdBg                QLSFFWRPEEVDVSQDRIDYA-ALPEHEKHIFISNLKYQTLLDSIQGRSPNVALLPLVSI 102
A_succinogenes130Z_NrdBg                QLSFFWRPEEVDVSQDRIDYA-ALPDHEKHIFISNLKYQTLLDSIQGRSPNVALLPLVSI 102
M_succiniciproducensMBEL55E_NrdBg       QLSFFWRPEEVDVAQDRIDYA-ALPEHEKHIFISNLKYQTLLDSIQGRSPNVALLPIVSL 102
A_pleuropneumoniae5bL20_NrdBg           QLSFFWRPEEVDVSQDRIDYQ-SLPEHEKHIFISNLKYQTLLDSIQGRSPNVAFLPLVSI 102
H_ducreyi35000HP_NrdBg                  QLSFFWRPEEVDVSQDRIDYQ-ALPEHEKHIFISNLKYQTLLDSIQGRSPNVAFLPLVSI 102
B_cicadellinicolaHc_NrdBg               QLSFFWRPEEVDISLDRIDYQ-ALPQHEKHIFISNIKYQTLLDSIQGRSPNVALLPLISI 102
C_sakazakiiBAA894_NrdBg                 QLSFFWRPEEVDVSRDRIDFQ-ALPEHEKHIFISNLKYQTLLDSIQGRSPNVALLPLISI 102
Enterobacter638_NrdBg                   QLSFFWRPEEVDVSRDRIDFQ-ALPDHEKHIFLSNLKYQTLLDSIQGRSPNVALLPLISI 102
C_koseriBAA895_NrdBg                    QLSFFWRPEEVDVSRDRIDFQ-SLPEHEKHIFISNLKYQTLLDSIQGRSPNVALLPLISI 102
K_pneumoniaeMGH78578_NrdBg              QLSFFWRPEEVDVSRDRIDYQ-ALPEHEKHIFISNLKYQTLLDSIQGRSPNVALLPLISI 102
E_coli536_NrdBg                         QLSFFWRPEEVDVSRDRIDYQ-ALPEHEKHIFISNLKYQTLLDSIQGRSPNVALLPLISI 102
S_flexneri2a2457T_NrdBg                 QLSFFWRPEEVDVSRDRIDYQ-ALPEHEKHIFISNLKYQTLLDSIQGRSPNVALLPLISI 115
S_flexneri58401_NrdBg                   QLSFFWRPEEVDVSRDRIDYQ-ALPEHEKHIFISNLKYQTLLDSIQGRSPNVALLPLISI 179
E_coliK12MG1655_NrdBg                   QLSFFWRPEEVDVSRDRIDYQ-ALPEHEKHIFISNLKYQTLLDSIQGRSPNVALLPLISI 102
S_boydiiSb227_NrdBg                     QLSFFWRPEEVDVSRDRIDYQ-ALPEHEKHIFISNLKYQTLLDSIQGRSPNVALLPLISI 102
S_TyphimuriumLT2_NrdBg                  QLSFFWRPEEVDVSRDRIDYQ-ALPEHEKHIFISNLKYQTLLDSIQGRSPNVALLPLISI 102
S_entericaTy2_NrdBg                     QLSFFWRPEEVDVSRDRIDYQ-ALPEHEKHIFISNLKYQTLLDSIQGRSPNVALLPLISI 102
S_proteamaculans568_NrdBg               QLSFFWRPEEVDVSRDRIDYQ-ALPEHEKHIFISNLKYQTLLDSIQGRSPNVALLPLISI 102
Y_pestisAngolaNrdBg_2                   QLSFFWRPEEIDVSRDRIDYN-ALPDHEKHIFISNLKYQTLLDSIQGRSPNVALLPLISI 102
Y_pestisMicrotus91001_NrdBg_2           QLSFFWRPEEIDVSRDRIDYN-ALPDHEKHIFISNLKYQTLLDSIQGRSPNVALLPLISI 117
Y_pestisPestoidesF_NrdBg_1              QLSFFWRPEEIDVSRDRIDYN-ALPDHEKHIFISNLKYQTLLDSIQGRSPNVALLPLISI 102
Y_pseudotuberculosisIP31758_NrdBg       QLSFFWRPEEIDVSRDRIDYN-ALPDHEKHIFISNLKYQTLLDSIQGRSPNVALLPLISI 102
M_flagellatusKT_NrdBg                   QLSFFWRPEEVDVSQDRIDYQ-GLPEHEKHIFISNLKYQTLLDSIQGRSPNVALLPLVSL 102
N_gonorrhoeaeFA1090_NrdBg               QLSFFWRPEEIDVSRDRIDYA-NLPEHEKHIFISNLKYQTLLDSIQGRSPNVALLPLVSI 109
N_meningitidis053442_NrdBg              QLSFFWRPEEIDVSRDRIDYA-NLPEHEKHIFISNLKYQTLLDSIQGRSPNVALLPLVSI 109
N_meningitidisMC58_NrdBg                QLSFFWRPEEIDVSRDRIDYA-NLPEHEKHIFISNLKYQTLLDSIQGRSPNVALLPLVSI 109
V_choleraeO395_NrdBg                    QLSFFWRPEEVDVSSDRIDYG-KLPDHEKHIFISNLKYQTLLDSIQGRSPNVALLPLVSL 114
V_vulnificusCMCP6_NrdBg                 QLSFFWRPEEVDVSSDRIDYN-KLPDHEKHIFISNLKYQTLLDSIQGRSPNVALLPLVSL 102
V_campbelliiBAA1116_NrdBg               QLSFFWRPEEVDVSSDRIDYN-KLPDHEKHIFISNLKYQTLLDSIQGRSPNVALLPLVSL 102
VibrioAND4_NrdBg                        QLSFFWRPEEVDVSSDRIDYN-KLPDHEKHIFISNLKYQTLLDSIQGRSPNVALLPIVSL 102
V_fischeriES114_NrdBg                   QLSFFWRPEEVDVSGDRIDYN-NLPEHEQHIFISNLKYQTLLDSIQGRSPNVALLPLVSI 102
P_angustumS14_NrdBg                     QLSFFWRPEEVDVSGDRINYN-NLPDHEKHIFISNLKYQTLLDSIQGRSPNVALLPIVSL 102
PhotobacteriumSKA34_NrdBg               QLSFFWRPEEVDVSGDRINYN-NLPDHEKHIFISNLKYQTLLDSIQGRSPNVALLPIVSL 102
                                         :.:******:.:  *  ::   :    : .  .*::** :***: .*:   .:    *
```

```
VibriophageKVP40_NrdBg                NSLENWITTWDFSETIHSRSYTHIMRNLYSNPSAVFDEIVLDKSIMARAEAVTEHYDKLI 162
AeromonasphageAeh1_NrdBg              VALETWITTWSFSETIHSRSYTHIMRNLYVDPSAVFDRILLDSKIMARATSVTQYYDNLI 161
EnterophageRB49_NrdBg                 PTLDTWLQTWTFSETIHSRSYTHIMRNCFIDPASEFDQIIDNPAIMKRAESIGKYYDDVL 160
EnterophagePhi1_NrdBg                 PTLDTWLQTWTFSETIHSRSYTHIMRNCFIDPASEFDQIIDNPAIMKRAESIGQYYDDVL 160
EnterophageRB69_NrdBg                 PSLDTWVATWTFSETIHSRSYTHIMRNLYTDPSKVFDEIVLDEAIMKRAESIGKYYDDVL 160
EnterophageJS98_NrdBg                 PSLDTWVATWTFSETIHSRSYTHIMRNLYTDPSKVFDEILLDEAIMKRAESIGRYYDDVI 161
EnterophageT4_NrdBg                   PSLDTWVATWTFSETIHSRSYTHIMRNLYTDPSKVFDEIVLDEAIMKRAESIGRYYDDVL 161
EnterophageRB32_NrdBg                 PSLDTWVATWTFSETIHSRSYTHIMRNLYTDPSKVFDEIVLDEAIMKRAESIGRYYDDVL 161
ProchlorococcusphagePSSM2_NrdBg       PELEACMEVWSFMEMIHSRSYTYVIKNVYADPSEVFDKILSDDRILSRAASVTESYDDFI 166
ProchlorococcusphagePSSM4_NrdBg       PELEGCINVWQFMEMIHSRSYTYIIKNVYSDPSEVFDTIIKDTNILQRAESVTKSYDDFV 171
Y_pestisAngola_NrdBg_1                PELETWVETWSFSETIHSRSYTHIIRGMVDDPSIVFDGIVTDEEIISRAVSISSEYDRLY 163
Y_pestisMicrotus91001_NrdBg_1         PELETWVETWSFSETIHSRSYTHIIRGMVDDPSIVFDGIVTDEEIISRAVSISSEYDRLY 163
Y_pestisPestoidesF_NrdBg_2            PELETWVETWSFSETIHSRSYTHIIRGMVDDPSIVFDGIVTDEEIISRAVSISSEYDRLY 163
C_RuthiamagnificaCm_NrdBg             PEVENWIETWSFSETIHSRSYTHIIRAIINEPGVIFDDIMKTDEIIQRAESASKHYDGLI 162
C_psychrerythraea34H_NrdBg            PEVETWVETWAFSETIHSRSYTHILRNLFTEPGEIFDDIIVNPAILKRASSIAKFFDDVI 162
S_sediminisHAWEB3_NrdBg               PEVETWVETWAFSETIHSRSYTHILRNLFTEPGEVFDDIVVNPAILKRATSIAKYFDDVI 162
I_loihiensisL2TR_NrdBg                PELETWIETWSFFETIHSRSYTHILRNLFSDPSEVFEDIVINEQIKIRANDISKYYDDLI 162
P_atlanticaT6c_NrdBg                  PELETWTETWSFFETIHSRSYTHILRNLFGDPSSIFSDIVENEQIKKRAADISKYYDDLI 162
B_aphidicolaBp_NrdBg                  PELETWIETWSFSETIHSRSYTHIIRNIINTPSLIFDDIIDNKNIANRAKDIAKYYDDLI 162
B_aphidicolaSg_NrdBg                  PELETWIETWSFSETIHSRSYTHIIRNIINNPSIVFDDIISNKNINDRAQDISIYYDELI 162
P_ingrahamii37_NrdBg                  PELETWIETWSFSETIHSRSYTHILRNLFTEPAEIFEDIVTNVEIQKRATDISGYYDALI 162
MarinomonasMWYL1_NrdBg                PELETWIETWSFSETIHSRSYTHIIRNIVNDPTKIFDDIVTNEEITKRADSVSKYYDRLI 162
P_tunicataD2_NrdBg                    PELETWIETWSFSETIHSRSYTHIIRNVTKQPELIFDDIVDNDKIKERADEVTRYYDDLI 162
S_loihicaPV4_NrdBg                    PELETWIETWSFSETIHSRSYTHIIRNIVNDPSVVFDDIVQNEEILKRASDIADYYDNLI 162
S_pealeana700345_NrdBg                PELETWIETWSFSETIHSRSYTHIIRNIVNDPSIVFNDIVENEEILKRASDIANYYDHLI 162
S_denitrificansOS217_NrdBg            PELETWIETWSFSETIHSRSYTHIIRNIVNDPSVIFDDIVVNDEILKRAGDIAKYYDDLI 162
S_frigidimarina400_NrdBg              PELETWIETWSFSETIHSRSYTHIIRNIVNNPSLVFDDIVQNDEILKRAGDIAAYYDDLI 162
S_amazonensisSB2B_NrdBg               PELETWIETWSFSETIHSRSYTHIIRNIVNDPSSVFDDIVVNEEILKRAGDISAYYDELI 177
S_oneidensisMR1_NrdBg                 PELETWIETWSFSETIHSRSYTHIIRNIVNDPSVIFDDIVVNHEILRRATDIAEFYDHLI 162
ShewanellaANA3_NrdBg                  PELETWIETWSFSETIHSRSYTHIIRNIVNDPSVIFDDIVVNQEILRRATDIAEYYDHLI 177
ShewanellaMR4_NrdBg                   PELETWIETWSFSETIHSRSYTHIIRNIVNDPSVIFDDIVVNQEILRRATDIAEYYDHLI 162
ShewanellaMR7_NrdBg                   PELETWIETWSFSETIHSRSYTHIIRNIVNDPSVIFDDIVVNQEILRRATDIAEYYDHLI 162
S_balticaOS155_NrdBg                  PELETWIETWSFSETIHSRSYTHIIRNIVNDPSVIFDDIVVNKEILSRATDIAAYYDHLI 162
S_putrefaciensCN32_NrdBg              PELETWIETWSFSETIHSRSYTHIIRNIVNDPSVIFDDIVVNKEILRRATDIAEYYDHLI 162
ShewanellaW3181_NrdBg                 PELETWIETWSFSETIHSRSYTHIIRNIVNDPSVIFDDIVVNKEILRRATDIAEYYDHLI 162
PsychrobacterPRwf1_NrdBg              PELETWIETWSFSETIHSRSYTHIIRNIVNDPAVVFDDIMENEHILQRASDIAYYYDDLY 162
P_arcticus2734_NrdBg                  PELETWIETWTFSETIHSRSYTHIIRNIINDPAIVFDDIMQNEHILGRAADIAKYYDDLY 162
P_cryohalolentisK5_NrdBg              PELETWIETWTFSETIHSRSYTHIIRNIINDPAIVFDDIMQNEHILGRAADIAKYYDDLY 162
CBlochmanniaBPEN_NrdBg                PELEIWVETWSFFETIHSRSYTHIIRNIVNHPSLIFDDIISNKEILKRAKDISTYYDNLI 162
A_hydrophila7966_NrdBg                PELETWIETWAFSETIHSRSYTHIIRNIVNTPAQVFDDIVTNEQILKRAGSISHFYDDLI 162
A_salmonicidaA449_NrdBg               PELETWIETWAFSETIHSRSYTHIIRNIVNNPGQVFDDIVTNEQILKRAGSISCFYDDLI 162
P_multocidaPm70_NrdBg                 PELETWIETWTFSETIHSRSYTHIIRNIVNDPSVIFDDIVTNEEIIKRAKDISAYYDDLI 162
H_somnus129PT_NrdBg                   PELETWIETWTFSETIHSRSYTHIIRNIVNDPSIVFDDIVTNEEIIKRARDISSYYDDLI 162
H_influenzae86028NP_NrdBg             PELETWIETWTFSETIHSRSYTHIIRNIVNDPSIVFDDIVTNEEIIKRAQDISSYYDDLI 162
H_influenzaeRdKW20_NrdBg              PELETWIETWTFSETIHSRSYTHIIRNIVNDPSIVFDDIVTNEEIIKRAQDISSYYDDLI 162
A_succinogenes130Z_NrdBg              PELETWIETWTFSETIHSRSYTHIIRNIVNDPSIVFDDIVTNEEIIKRAKDISAYYDDLI 162
M_succiniciproducensMBEL55E_NrdBg     PELETWIETWTFSETIHSRSYTHIIRNIVNDPSVVFDDIVTNEEIIKRAKDISAYYDDLI 162
A_pleuropneumoniae5bL20_NrdBg         PELETWIETWTFSETIHSRSYTHIIRNIVNDPSIVFDDIVTNEEIIKRAKDISCYYDDLI 162
H_ducreyi35000HP_NrdBg                PELETWIETWTFSETIHSRSYTHIIRNIVNDPSIVFDDIVTNEEIIKRARDISCYYDDLI 162
B_cicadellinicolaHc_NrdBg             PELETWVETWAFSETIHSRSYTHIIRNIVNDPSLVFDDIVTNTEILKRAKDIAAFYDDLI 162
C_sakazakiiBAA894_NrdBg               PELETWVETWAFSETIHSRSYTHIIRNIVNDPAVVFDDIVTNEQILKRAEGISHYYDDLI 162
Enterobacter638_NrdBg                 PELETWVETWAFSETIHSRSYTHIIRNIVNDPAVVFDDIVTNEQIQKRAEGISHYYDELI 162
C_koseriBAA895_NrdBg                  PELETWVETWAFSETIHSRSYTHIIRNIVNDPAVVFDDIVTNEQIQKRAEGISHYYDELI 162
K_pneumoniaeMGH78578_NrdBg            PELETWVETWAFSETIHSRSYTHIIRNIVNDPATVFDDIVTNEQIQKRAEGISHYYDSLI 162
E_coli536_NrdBg                       PELETWVETWAFSETIHSRSYTHIIRNIVNDPSVVFDDIVTNEQIQKRAEGISSYYDELI 162
S_flexneri2a2457T_NrdBg               PELETWVETWAFSETIHSRSYTHIIRNIVNDPSVVFDDIVTNEQIQKRAEGISSYYDELI 175
S_flexneri58401_NrdBg                 PELETWVETWAFSETIHSRSYTHIIRNIVNDPSVVFDDIVTNEQIQKRAEGISSYYDELI 239
E_coliK12MG1655_NrdBg                 PELETWVETWAFSETIHSRSYTHIIRNIVNDPSVVFDDIVTNEQIQKRAEGISSYYDELI 162
S_boydiiSb227_NrdBg                   PELETWVETWAFSETIHSRSYTHIIRNIVNDPSVVFDDIVTNEQIQKRAEGISSYYDELI 162
S_TyphimuriumLT2_NrdBg                PELETWVETWAFSETIHSRSYTHIIRNIVNDPAVVFDDIVTNEQIQKRAEGISAYYDELI 162
S_entericaTy2_NrdBg                   PELETWVETWAFSETIHSRSYTHIIRNIVNDPAIVFDDIVTNEQIQKRAEGISAYYDELI 162
S_proteamaculans568_NrdBg             PELETWIETWSFSETIHSRSYTHIIRNIVNDPAIVFDDIVTNEEILKRAKDISGYYDDLI 162
Y_pestisAngolaNrdBg_2                 PELETWVETWSFSETIHSRSYTHIIRNIVNDPSVVFDDIVTNEEILKRAKDISAYYDDLI 162
Y_pestisMicrotus91001_NrdBg_2         PELETWVETWSFSETIHSRSYTHIIRNIVNDPSVVFDDIVTNEEILKRAKDISAYYDDLI 177
Y_pestisPestoidesF_NrdBg_1            PELETWVETWSFSETIHSRSYTHIIRNIVNDPSVVFDDIVTNEEILKRAKDISAYYDDLI 162
Y_pseudotuberculosisIP31758_NrdBg     PELETWVETWSFSETIHSRSYTHIIRNIVNDPSVVFDDIVTNEEILKRAKDISAYYDDLI 162
M_flagellatusKT_NrdBg                 PELETWVETWAFSETIHSRSYTHIIRNIVSDPALIFDDIVRNENIVARASDISAYYDDLI 162
N_gonorrhoeaeFA1090_NrdBg             PELETWVETWSFSETIHSRSYTHIIRNIVNDPSVVFDDIVENEYITARAEDIACYYDDLI 169
N_meningitidis053442_NrdBg            PELETWIETWSFSETIHSRSYTHIIRNIVNDPSVVFDDIVENEYITARAEDIACYYDDLI 169
N_meningitidisMC58_NrdBg              PELETWIETWSFSETIHSRSYTHIIRNIVNDPSVVFDDIVENEYITARAEDIACYYDDLI 169
V_choleraeO395_NrdBg                  PELETWIETWSFSETIHSRSYTHIIRNIVNDPAVVFDDIVENEHIIKRAKDIAHYYDDLI 174
V_vulnificusCMCP6_NrdBg               PELETWIETWSFSETIHSRSYTHIIRNIVNDPSVVFDDIVENEHILKRAKDIAHYYDDLI 162
V_campbelliiBAA1116_NrdBg             PELETWIETWSFSETIHSRSYTHIIRNIVNDPSVVFDDIVENEHILKRAKDIAHYYDDLI 162
VibrioAND4_NrdBg                      PELETWIETWSFSETIHSRSYTHIIRNIVNDPSVVFDDIVENEHILKRAKDIAHYYDDLI 162
V_fischeriES114_NrdBg                 PELETWIETWSFSETIHSRSYTHIIRNIVNNPSVIFDDIVENEEILKRAKDIAHYYDNLI 162
P_angustumS14_NrdBg                   PELETWIETWSFSETIHSRSYTHIIRNIVNNPSVVFDDIVENEHIIKRAGDISKYYDALI 162
PhotobacteriumSKA34_NrdBg             PELETWIETWSFSETIHSRSYTHIIRNIVNNPSVVFDDIVENEHIVKRAGDISKYYDDLI 162
                                        ::    .* * * *******::::     *   *. *:    *  **      :* .
```

```
VibriophageKVP40_NrdBg                AETQKLNQLKTMNLD------------E--AHILMQESAVKEALYLCMHAVNALEAIRF 207
AeromonasphageAeh1_NrdBg              KELKTYDLLTQSVFS--.----------P--ELRVAQLTKVKEALYLAMHAVNALEAIRF 206
EnterophageRB49_NrdBg                 YWIRKYENAKEDLRYWEERDDQTA--LDDARRFFEDTKRDLMASIYLCAHCVNALEAIRF 218
EnterophagePhi1_NrdBg                 YWIRKYENAKEDLRYWEERDDHTA--LDDARRFFEDTKRDLMASIYLCAHCVNALEAIRF 218
EnterophageRB69_NrdBg                 VKTRNWENAKEVYECSIGSCQQNDIEMY--EEIMFNAKRDLMKSLYLCLHVINALEAIRF 218
EnterophageJS98_NrdBg                 EKTRLYQNALEEYHEFG------------EGVKQAKYELMRSLYLCLHVINALEAIRF 207
EnterophageT4_NrdBg                   VKTREWENAKDMVEYYKDQG-LI----LADKDVEQRAKRDLMKSLYLCLHVINALEAIRF 216
EnterophageRB32_NrdBg                 VKTREWENAKEFVELAKESP-DADFRLNRAIKQEAEAKCALMKSLYLCLHVINALEAIRF 220
ProchlorococcusphagePSSM2_NrdBg       NEAQGWGQSNLWRDMDKS--------LD--TSLPVLEMKEVKRKLYRAVTNVNILEGIRF 216
ProchlorococcusphagePSSM4_NrdBg       NDAHEYDSGNAWKFAREG--------H----PAGTLDRRELKRKLYRAIANVNILEGIRF 219
Y_pestisAngola_NrdBg_1                EMTCARQHLGEDEFERLY------------VSEFDGKPYPLQRQLFRTLVSVNALEAIRF 211
Y_pestisMicrotus91001_NrdBg_1         EMTCARQHLGEDEFERLY------------VSEFDGKPYPLQRQLFRTLVSVNALEAIRF 211
Y_pestisPestoidesF_NrdBg_2            EMTCARQHLGEDEFERLY------------VSEFDGKPYPLQRQLFRTLVSVNALEAIRF 211
C_RuthiamagnificaCm_NrdBg             KYTQAYLLHGTGKLKIQE------------EEVSIDLYCLKKQLYLTIMSVNILEAVRF 209
C_psychrerythraea34H_NrdBg            LTTQLLQSQGEGTYEVEG------------K-TVEVSTRKLKERLFLAVCSVNALEAIRF 209
S_sediminisHAWEB3_NrdBg               LTTQLLQSQGEGSYEVNG------------E-TIEVSARKLKERLYLAVCSVNALEAIRF 209
I_loihiensisL2TR_NrdBg                FYTQLLQTHGEGDVVVEG------------K-THTVTQREIKKKLFLCINSVNALEAIRF 209
P_atlanticaT6c_NrdBg                  FYTQLWQTLGEGTHTING------------E-EHVVTMRTLKIKLFLCMNSVNALEAIRF 209
B_aphidicolaBp_NrdBg                  ELTSYWHLFGEGKHTING------------K-KIKIDLHELKKKLYLCLISVNALEAIRF 209
B_aphidicolaSg_NrdBg                  RITAYWHLLGEGNHKINE------------K-KVEINLKLLKRRLYLCLISVNVLEAIRF 209
P_ingrahamii37_NrdBg                  HATQFMQLHGLEQTDIDYIN------LE--NQVEKLNLRAIKKKLYLCICSVNALEAIRF 214
MarinomonasMWYL1_NrdBg                EMVNLYNTMGMGTFDVPG------------RGSIEISMPKLKKALYLTIASVNVLEAIRF 210
P_tunicataD2_NrdBg                    EYTSIYNLYGEGKHQIND------------Q-TVVINLFELKKKLYRAIMSVNILEAIRF 209
S_loihicaPV4_NrdBg                    ELSQIYHLMGEGTHEING------------K-TVEVTTRAIKKALYLCMMSVNVLEAIRF 209
S_pealeana700345_NrdBg                KLSQAYHLLGEGSHEIDG------------E-IIEVTKREIKKALYLCMVSVNVLEAIRF 209
S_denitrificansOS217_NrdBg            HLTQIFNLHGAGTHLIEG------------K-QVTVNLRVLKKSLYLCMMSVNALEAIRF 209
S_frigidimarina400_NrdBg              NLTQINNLLGEGTHIING------------K-EITVNSRILKKSLYLCMMSVNALEAIRF 209
S_amazonensisSB2B_NrdBg               KLTQIYHLHGEGTHTVDG------------H-VVDVTVRALKKALYLCMMSVNALEAIRF 224
S_oneidensisMR1_NrdBg                 KLTQIYHLLGEGTHVIDG------------E-PVEVNTRTLKKALYLCMMSVNALEAIRF 209
ShewanellaANA3_NrdBg                  KLTQVYHLLGEGTHVIDG------------E-PIEVNTRTLKKALYLCMMSVNALEAIRF 224
ShewanellaMR4_NrdBg                   KLTQVYHLLGEGTHVIDG------------E-PIEVNTRTLKKALYLCMMSVNALEAIRF 209
ShewanellaMR7_NrdBg                   KLTQVYHLLGEGTHVIDG------------E-PIEVNTRTLKKALYLCMMSVNALEAIRF 209
S_balticaOS155_NrdBg                  QLTQVFHLLGAGTHVIDG------------K-EVIVNTRTLKKALYLCMMSVNALEAIRF 209
S_putrefaciensCN32_NrdBg              KLTQVFNLFGAGTHLIDG------------E-ELVVNTRTLKKALYLCMMSVNALEAIRF 209
ShewanellaW3181_NrdBg                 KLTQVFNLFGAGTHLIDG------------E-ELVVNTRTLKKALYLCMMSVNALEAIRF 209
PsychrobacterPRwf1_NrdBg              QHSQLYSLFGPGTHIVNG------------E-EKYVDLRKLKKQLYLCLMAVNVLEAIRF 209
P_arcticus2734_NrdBg                  YSSSLYNLYGAGTHTING------------K-QITVDLKSLKKQLYLCLMAVNVLEAIRF 209
P_cryohalolentisK5_NrdBg              YSSSLYNLYGAGTHTING------------K-QITIDLKSLKKQLYLCLMAVNVLEAIRF 209
CBlochmanniaBPEN_NrdBg                ELTSYYHILGPGIHQINN------------N-TVAVNLHELKKKLYLCLISVNVLEAIRF 209
A_hydrophila7966_NrdBg                EATALYNLHGEGTHRVAG------------R-DVTITLRELKKKLYLCLMSVNALEAIRF 209
A_salmonicidaA449_NrdBg               EATALYNLHGEGTHSVAG------------Q-MRTITLRELKKKLYLCLMSVNALEAIRF 209
P_multocidaPm70_NrdBg                 RDSQLYTLYGEGTYNVDG------------Q-DCVVTLRNLKRQLYLCLMSVNALEAIRF 209
H_somnus129PT_NrdBg                   RDSQLYGLYGEGTYTVDG------------K-ACKVTLRNLKRQLYLCLMSVNALEAIRF 209
H_influenzae86028NP_NrdBg             RDSQLYGLYGEGTYTVDG------------K-ECVVTLRSLKKQLYLCLMSVNALEAIRF 209
H_influenzaeRdKW20_NrdBg              RDSQLYGLYGEGTYTVDG------------K-ECVVTLRSLKKQLYLCLMSVNALEAIRF 209
A_succinogenes130Z_NrdBg              RDSQLYSLYGEGTYQVDG------------K-ECAVTLRNLKKRLYLCLMSVNALEAIRF 209
M_succiniciproducensMBEL55E_NrdBg     RDSQLYNLFGEGTYKVEG------------K-ECKVTLRNLKKQLYLCLMSVNALEAIRF 209
A_pleuropneumoniae5bL20_NrdBg         RDSQLYTLYGEGCYTVDG------------K-QYQVTLRNLKRQLYLCLMSVNVLEAIRF 209
H_ducreyi35000HP_NrdBg                RDSQLYTLYGEGTYQVDG------------K-ACAVTLRNLKRQLYLCLMSVNVLEAIRF 209
B_cicadellinicolaHc_NrdBg             QMTSYFHLFGEGNHSING------------K-LVIVNQHELKKKLYLCLMSINALEAIRF 209
C_sakazakiiBAA894_NrdBg               EMTSYWHLLGEGTHSVNG------------K-TVTVNLHELKKKLYLCLMSVNALEAIRF 209
Enterobacter638_NrdBg                 EMTSYWHLLGEGTHSVNG------------K-TVTVNLRALKKQLYLCLMSVNALEAIRF 209
C_koseriBAA895_NrdBg                  EMTSYWHLLGEGTHNVNG------------K-TVTVNLRALKKQLYLCLMSVNALEAIRF 209
K_pneumoniaeMGH78578_NrdBg            EMTSYWHLLGEGTHTVNG------------K-TVTVNLRELKKKLYLCLMSVNALEAIRF 209
E_coli536_NrdBg                       EMTSYWHLLGEGTHTVNG------------K-TVTVSLRELKKKLYLCLMSVNALEAIRF 209
S_flexneri2a2457T_NrdBg               EMTSYWHLLGEGTHTVNG------------K-TVTVSLRELKKKLYLCLMSVNALEAIRF 222
S_flexneri58401_NrdBg                 EMTSYWHLLGEGTHTVNG------------K-TVTVSLRELKKKLYLCLMSVNALEAIRF 286
E_coliK12MG1655_NrdBg                 EMTSYWHLLGEGTHTVNG------------K-TVTVSLRELKKKLYLCLMSVNALEAIRF 209
S_boydiiSb227_NrdBg                   EMTSYWHLLGEGTHTVNG------------K-TVTVSLRELKKKLYLCLMSVNALEAIRF 209
S_TyphimuriumLT2_NrdBg                EMTSYWHLLGEGTHTVNG------------K-TVVVNLRELKKKLYLCLMSVNALEAIRF 209
S_entericaTy2_NrdBg                   EMTSYWHLLGEGTHTVNG------------K-TVVVNLRELKKKLYLCLMSVNALEAIRF 209
S_proteamaculans568_NrdBg             EMTGYYHLLGEGTHQVNG------------K-TVTVTLRALKKQLYLCLMSVNALEAIRF 209
Y_pestisAngolaNrdBg_2                 EMTSYYHLLGEGTHQVNG------------K-TVVVKLRDLKKQLYLCLMSVNALEAIRF 209
Y_pestisMicrotus91001_NrdBg_2         EMTSYYHLLGEGTHQVNG------------K-TVVVKLRDLKKQLYLCLMSVNALEAIRF 224
Y_pestisPestoidesF_NrdBg_1            EMTSYYHLLGEGTHQVNG------------K-TVVVKLRDLKKQLYLCLMSVNALEAIRF 209
Y_pseudotuberculosisIP31758_NrdBg     EMTSYYHLLGEGTHQVNG------------K-TVVVKLRDLKKQLYLCLMSVNALEAIRF 209
M_flagellatusKT_NrdBg                 HYTMLYSQLGKGDHVING------------E-TKTVNLRELKKLLYLCLMSVNILEAIRF 209
N_gonorrhoeaeFA1090_NrdBg             EYTQYYNLLGEGVHNVGG------------K-PVTVSLRGLKKKLYLCLMCVNVLEAIRF 216
N_meningitidis053442_NrdBg            EYTQYYNLLGEGMHNVGG------------K-PVTVSLRGLKKKLYLCLMCVNVLEAIRF 216
N_meningitidisMC58_NrdBg              EYTQYYNLLGEGVHNVGG------------K-PVTVSLRGLKKKLYLCLMCVNVLEAIRF 216
V_choleraeO395_NrdBg                  ELTNDYHRYGEGQHQING------------E-TITVSLKELKKKLYLCLMSVNALEAIRF 221
V_vulnificusCMCP6_NrdBg               QLTNDYHRFGEGEHVVNG------------E-TIKVSLYELKKKLYLCLMSVNALEAIRF 209
V_campbelliiBAA1116_NrdBg             QATNDYHRYGEGEHVLNG------------E-TVKVSLYDLKKKLYICLMSVNALEAIRF 209
VibrioAND4_NrdBg                      QATNDYHRYGEGEHLLNG------------E-TVKVSLYDLKKKLYICLMSVNALEAIRF 209
V_fischeriES114_NrdBg                 QATNDYHRFGEGTHVLNG------------Q-EITISKRELKKKLYLCLMSVNALEAIRF 209
P_angustumS14_NrdBg                   SATNDYHRYGEGSHSANG------------V-PFTVNLHELKKKLYLCLMSVNALEAIRF 209
PhotobacteriumSKA34_NrdBg             TVTNDYNRFGEGNHSANG------------I-PFTVSLRELKKKLYLCLMSVNALEAIRF 209
                                                                        :  ::    :* **.:**
```

```
VibriophageKVP40_NrdBg                YVSFACTFNFAE-QGIMEGNAKVMKLIARDEQLHLKGTQALISIWQNYKDDPEMAEISLR 266
AeromonasphageAeh1_NrdBg              YVSFACTFSFGE-RGIMEGNAKIMRLIARDEQLHLKGTQYLIRAWQNAEDDSEMYEIAKR 265
EnterophageRB49_NrdBg                 YVSFACTFSFAE-RGLLEGNAKIMKFIARDEQLHLKGTQYIIRQWQTGVDGDEWKQIAEE 277
EnterophagePhi1_NrdBg                 YVSFACTFSFAE-RGLLEGNAKIMKFIARDEQLHLKGTQYIIRQWQTGADGDEWKQIAEE 277
EnterophageRB69_NrdBg                 YVSFACTFNFHKNMEIMEGNAKIMKFIARDEQLHLRVLSTLSVNFNLVLMAMSGFKLLKN 278
EnterophageJS98_NrdBg                 YVSFACTFNFHKNMEIMEGNAKIMKFIARDEQLHLKGTQYIIRQLQLGTDGDEWVKIAQE 267
EnterophageT4_NrdBg                   YVSFACTFNFHKNMEIMEGNAKIMKFIARDEQLHLKGTQYIIRQLQSGTDGDEWVKIAQE 276
EnterophageRB32_NrdBg                 YVSFACTFNFHKNMEIMEGNAKIMKFIARDEQLHLKGTQYIIRQLQSGTDGDEWVKIAQE 280
ProchlorococcusphagePSSM2_NrdBg       YVSFACSFAFGE-LKVMEGSAKIISLIARDENQHLAITQNILNNWRKG-DDPEMVEIMKE 274
ProchlorococcusphagePSSM4_NrdBg       YVSFACSFAFGE-NKLMEGSAKILSLISRDESQHLVITQQILKKWADG-DDPEMQEIAKE 277
Y_pestisAngola_NrdBg_1                YVSFACTFAFGE-RKLLEGNTKIMRFIARDEALHCEGTERMIRFMRTGREGLMWKMIAAE 270
Y_pestisMicrotus91001_NrdBg_1         YVSFACTFAFGE-RKLLEGNTKIMRFIARDEALHCEGTERMIRFMRTGREGLMWKMIAAE 270
Y_pestisPestoidesF_NrdBg_2            YVSFACTFAFGE-RKLLEGNTKIMRFIARDEALHCEGTERMIRFMRTGREGLMWKMIAAE 270
C_RuthiamagnificaCm_NrdBg             YVSFACSFAFAE-RKVMEGNAKIIKMIARDEALHLTGTQHMLNLMSSGKDDLDMQKISKE 268
C_psychrerythraea34H_NrdBg            YVSFACSFAFAE-RELLEGNAKIIKLIARDEALHLTGTQHILNNWRSGKDDPEMKIISEE 268
S_sediminisHAWEB3_NrdBg               YVSFACSFAFAE-RELLEGNAKIIKLIARDEALHLSGTQHILKLWANGNDDPEMVEVCQE 268
I_loihiensisL2TR_NrdBg                YVSFACTFAFAE-RELMEGNAKIIKLIARDENVHLTSTQHILNLWQYGEDDPEMKEIAEE 268
P_atlanticaT6c_NrdBg                  YVSFACTFAFAE-RELMEGNSKIIRLISRDENLHLTSTQHILNLWAKGKDDPEMAEIALE 268
B_aphidicolaBp_NrdBg                  YVSFACSFAFAE-REKMEGNAKIIRLIARDEALHLTSTQHILNILHNDKNNEGMSDISKE 268
B_aphidicolaSg_NrdBg                  YVSFACSFAFAE-REIMEGNAKIIRLIARDEALHLTGTQHILNILNNEKNNENMKNTVLE 268
P_ingrahamii37_NrdBg                  YVSFACSFAFAE-RDLLEGNAKIIKLIARDEALHLNSTQQILNLWADGKDDPEMAEIAIE 273
MarinomonasMWYL1_NrdBg                YVSFACSFAFAE-RTLMEGNAKIIKLIARDEALHLTGTQHMLNLMASGKDDPEMAVIAAE 269
P_tunicataD2_NrdBg                    YVSFACSFAFAE-RELMEGNAKIIKLIARDEALHLSGTQHMLNIMQDGQDDPEMAIVAAQ 268
S_loihicaPV4_NrdBg                    YVSFACSFAFAE-RKLMEGNAKIIRLIARDEALHLNSTQHILKLMQGGKDDPEMAEIAKE 268
S_pealeana700345_NrdBg                YVSFACSFAFAE-RKVMEGNAKIIRLIARDEALHLNSTQHILKIMHAGKDDPEMAEIAKE 268
S_denitrificansOS217_NrdBg            YVSFACSFAFAE-RKLMEGNAKIIRLIARDEALHLNSTQHILALMQGGKDDPEMAEIAIE 268
S_frigidimarina400_NrdBg              YVSFACSFAFAE-RRVMEGNAKIIRLIARDEALHLNSTQHILALMQGGKDDPEMGEIAAE 268
S_amazonensisSB2B_NrdBg               YVSFACSFAFAE-RKLMEGNAKIIRLIARDEALHLNSTQHILNIMQGGKDDAEMAEIAIE 283
S_oneidensisMR1_NrdBg                 YVSFACSFAFAE-RKLMEGNAKIIRLIARDEALHLNSTQHILTIMQGGKDDPEMGEIAIE 268
ShewanellaANA3_NrdBg                  YVSFACSFAFAE-RKLMEGNAKIIRLIARDEALHLNSTQHILTIMQGGKDDPEMAEIAAE 283
ShewanellaMR4_NrdBg                   YVSFACSFAFAE-RKLMEGNAKIIRLIARDEALHLNSTQHILTIMQAGKDDPEMGEIAAE 268
ShewanellaMR7_NrdBg                   YVSFACSFAFAE-RKLMEGNAKIIRLIARDEALHLNSTQHILTIMQAGKDDPEMGEIAAE 268
S_balticaOS155_NrdBg                  YVSFACSFAFAE-RKLMEGNAKIIRLIARDEALHLNSTQHILSIMQGGKDDPEMGEIAAE 268
S_putrefaciensCN32_NrdBg              YVSFACSFAFAE-RKLMEGNAKIIRLIARDEALHLNSTQHILNIMQGGKDDPEMGEIAAQ 268
ShewanellaW3181_NrdBg                 YVSFACSFAFAE-RKLMEGNAKIIRLIARDEALHLNSTQHILNIMQGGKDDPEMGEIAAQ 268
PsychrobacterPRwf1_NrdBg              YVSFACSFAFAE-RKLMEGNAKIIKLIARDEALHLTSTQHMINIMASGKDDPEMAEISAQ 268
P_arcticus2734_NrdBg                  YVSFACSFAFAE-RKLMEGNAKIIKLIARDEALHLTGTQHILNLMRIGRDDPEMVDIAKE 268
P_cryohalolentisK5_NrdBg              YVSFACSFAFAE-RKLMEGNAKIIKLIARDEALHLTGTQHILNLMRIGRDDPEMVDIAKE 268
CBlochmanniaBPEN_NrdBg                YVSFACSFAFAE-RELMEGNAKIIRLIARDEALHLNATQHILNLMRSGSDDPEMADISTE 268
A_hydrophila7966_NrdBg                YVSFACSFAFAE-RELMEGNAKIIKMIARDEALHLTGTQHMLNLMRSGQDDPEMAEIAAE 268
A_salmonicidaA449_NrdBg               YVSFACSFAFAE-RELMEGNAKIIKMIARDEALHLTGTQHMLNLMRTGQDDPEMAEISAE 268
P_multocidaPm70_NrdBg                 YVSFACSFAFAE-RKLMEGNAKIIKFIARDEALHLTGTQHILNIMASGQDDPEMAEIVEE 268
H_somnus129PT_NrdBg                   YVSFACSFAFAE-RKLMEGNAKIIKFIARDEALHLTGTQHILNIMAAGQDDPEMAEIAEE 268
H_influenzae86028NP_NrdBg             YVSFACSFAFAE-RRLMEGNAKIIKFIARDEALHLTGTQHILNIMAAGQDDPEMAEIAEE 268
H_influenzaeRdKW20_NrdBg              YVSFACSFAFAE-RRLMEGNAKIIKFIARDEALHLTGTQHILNIMAAGQDDPEMAEIAEE 268
A_succinogenes130Z_NrdBg              YVSFACSFAFAE-RQLMEGNAKIIKFIARDEALHLTGTQHILNIMAAGQDDPEMAEIAEE 268
M_succiniciproducensMBEL55E_NrdBg     YVSFACSFAFAE-RQLMEGNAKIIKFIARDEALHLTGTQHILNIMAAGQDDPEMAEIAEE 268
A_pleuropneumoniae5bL20_NrdBg         YVSFACSFAFAE-RQLMEGNAKIIKFIARDEYLHLTGTQNMLNIIASGQDDPEMAEIAEE 268
H_ducreyi35000HP_NrdBg                YVSFACSFAFAE-RQLMEGNAKIIKFIARDEYLHLAGTQQMLNIIASGQDDPEMAEIAIE 268
B_cicadellinicolaHc_NrdBg             YVSFACSFAFAE-RKLMEGNAKIIRLIARDEALHLAGTQHMLNLMRLGKDDPEMADIALE 268
C_sakazakiiBAA894_NrdBg               YVSFACSFAFAE-RELMEGNAKIIRLIARDEALHLTGTQHMLNLMRSGEDDPEMAAIAEE 268
Enterobacter638_NrdBg                 YVSFACSFAFAE-RKLMEGNAKIIRLIARDEALHLTGTQHMLNLLRSGVDDPEMAEIAEE 268
C_koseriBAA895_NrdBg                  YVSFACSFAFAE-RKLMEGNAKIIRLIARDEALHLTGTQHMLNLLRSGIDDPEMAEIAEE 268
K_pneumoniaeMGH78578_NrdBg            YVSFACSFAFAE-RKLMEGNAKIIRLIARDEALHLTGTQHMLNLLRSGSDDPEMAEIAEE 268
E_coli536_NrdBg                       YVSFACSFAFAE-RELMEGNAKIIRLIARDEALHLTGTQHMLNLLRSGADDPEMAEIAEE 268
S_flexneri2a2457T_NrdBg               YVSFACSFAFAE-RELMEGNAKIIRLIARDEALHLTGTQHMLNLLRSGADDPEMAEIAEE 281
S_flexneri58401_NrdBg                 YVSFACSFAFAE-RELMEGNAKIIRLIARDEALHLTGTQHMLNLLRSGADDPEMAEIAEE 345
E_coliK12MG1655_NrdBg                 YVSFACSFAFAE-RELMEGNAKIIRLIARDEALHLTGTQHMLNLLRSGADDPEMAEIAEE 268
S_boydiiSb227_NrdBg                   YVSFACSFAFAE-RELMEGNAKIIRLIARDEALHLTGTQHMLNLLRSGADDPEMAEIAEE 268
S_TyphimuriumLT2_NrdBg                YVSFACSFAFAE-RELMEGNAKIIRLIARDEALHLTGTQHMLNLLRSGVDDPEMAEIAEE 268
S_entericaTy2_NrdBg                   YVSFACSFAFAE-RELMEGNAKIIRLIARDEALHLTGTQHMLNLLRSGVDDPEMAEIAEE 268
S_proteamaculans568_NrdBg             YVSFACSFAFAE-RELMEGNAKIIKLIARDEALHLTGTQHMLNLMRSGADDPEMAEIAVE 268
Y_pestisAngolaNrdBg_2                 YVSFACSFAFAE-RELMEGNAKIIKMIARDEALHLTGTQHILNLMRLGEDDPEMAEIAEE 268
Y_pestisMicrotus91001_NrdBg_2         YVSFACSFAFAE-RELMEGNAKIIKMIARDEALHLTGTQHILNLMRLGEDDPEMAEIAEE 283
Y_pestisPestoidesF_NrdBg_1            YVSFACSFAFAE-RELMEGNAKIIKMIARDEALHLTGTQHILNLMRLGEDDPEMAEIAEE 268
Y_pseudotuberculosisIP31758_NrdBg     YVSFACSFAFAE-RELMEGNAKIIKMIARDEALHLTGTQHILNLMRLGEDDPEMAEIAEE 268
M_flagellatusKT_NrdBg                 YVSFACSFAFAE-RKVMEGNAKIIRLIARDEALHLTGTQHMLNIMRSGQDDPEFAGIASE 268
N_gonorrhoeaeFA1090_NrdBg             YVSFACSFAFAE-RELMEGNAKIIKLIARDEALHLTGTQHMLNLMRSGVDDPEMAEIAAE 275
N_meningitidis053442_NrdBg            YVSFACSFAFAE-RELMEGNAKIIKLIARDEALHLTGTQHMLNLMRSGVDDSEMAEIAAE 275
N_meningitidisMC58_NrdBg              YVSFACSFAFAE-RELMEGNAKIIKLIARDEALHLTGTQHMLNLMRSGVDDSEMAEIAAE 275
V_choleraeO395_NrdBg                  YVSFACSFAFAE-RELMEGNAKIIKLIARDEALHLTGTQHMLNLLRNGMDDFTFLQVAEE 280
V_vulnificusCMCP6_NrdBg               YVSFACSFAFAE-RELMEGNAKIIKLIARDEALHLNGTQHMINLLRNGQDDFSFMQIAEE 268
V_campbelliiBAA1116_NrdBg             YVSFACSFAFAE-RELMEGNAKIIKLIARDEALHLNGTQHMINLLRNGQDDFLFMQIAEE 268
VibrioAND4_NrdBg                      YVSFACSFAFAE-RELMEGNAKIIKLIARDESLHLNGTQHMINLLRNGQDDFIFMQIAEE 268
V_fischeriES114_NrdBg                 YVSFACSFAFAE-RELMEGNAKIIKLIARDESLHLTSTQHMLNLLRTGQDDPEFVEIAAE 268
P_angustumS14_NrdBg                   YVSFACSFAFAE-RELMEGNAKIIKLIARDEALHLTGTQHILNLLRNGHDDPEMAEIAAE 268
PhotobacteriumSKA34_NrdBg             YVSFACSFAFAE-RELMEGNAKIIKLIARDEALHLTGTQHILNLLRNGHDDPEMVEIATE 268
                                      ******:* * :    :**.:*:: :*:***  *    . :                  .
```

```
VibriophageKVP40_NrdBg                LQPEAEKIFMQVVDQEKEWAHSLFSQGTVPGLNAQILCEYVEYLADSRMRSVGLNSPYP- 325
AeromonasphageAeh1_NrdBg              LEHKAVQIFLEVAEQEKEWARDLLGNVTIPGLSLSSMLDYIDYLTVSRMRSAGLPAPEEL 325
EnterophageRB49_NrdBg                 LEGEATKIFLDVVEQEKEWVDHLFAMGAIPNLSPEILKTYVEYLADSRMRAAGLKSPFS- 336
EnterophagePhi1_NrdBg                 LEGEATKIFLDVVEQEKEWVDHLFAMGAIPNLSPEILKTYVEYLADSRMRAAGLKSPFS- 336
EnterophageRB69_NrdBg                 VNKKQLTFLWKLIVKYKEWAAHLFKDGTCPGINTQSMCAFVDYLTVSRMKQCGLPCPITD 338
EnterophageJS98_NrdBg                 CEQEAVDIFMEVNRQEKEWAAHLFKDGTCPGLNTQNMGAFVDYLTVSRMKQCGLPCPITD 327
EnterophageT4_NrdBg                   CEQEAVDIFMEVNRQEKDWAVHLFKDGDVPGLNTNSMWSFIDYLTVSRMKQCGLPCPITD 336
EnterophageRB32_NrdBg                 CEQEAVDIFMEVNRQEKDWAVHLFKDGDVPGLNTNSMWSFIDYLTVSRMKQCGLPCPITD 340
ProchlorococcusphagePSSM2_NrdBg       EEQWTYQMFEKCVNEEKMWAEYLFKDGSMIGLNDKLLHQYVEWVANRRMRSIGMKPVYDI 334
ProchlorococcusphagePSSM4_NrdBg       EKSFVKDMFRNTVNEEKAWADYLFKEGSMIGLNARLLNNYVEWIANRRMKAIGLDPIYDQ 337
Y_pestisAngola_NrdBg_1                EESFIYQTMMDVAGQEMRWADYLFKDGSMIGLNADILKSYVKYRTNLAMRRLGLRPLFPE 330
Y_pestisMicrotus91001_NrdBg_1         EESFIYQTMMDVAGQEMRWADYLFKDGSMIGLNADILKSYVKYRTNLAMRRLGLRPLFPE 330
Y_pestisPestoidesF_NrdBg_2            EESFIYQTMMDVAGQEMRWADYLFKDGSMIGLNADILKSYVKYRTNLAMRRLGLRPLFPE 330
C_RuthiamagnificaCm_NrdBg             CQDEVITMFREACEQEKDWAEYLFRDGSMIGLNAQILKQYLEYITNVRMKALNLSLIFD- 327
C_psychrerythraea34H_NrdBg            LRDEGLQIFLDVVEQEKEWAQYLFKDGSMIGLNADILNQYIEYISNQRLTAIGFGAPFD- 327
S_sediminisHAWEB3_NrdBg               LREAGRQVFLDVVEQEKEWADYLFKDGSMIGLNAEILKQYIEYISNNRMEAIGYEAPFA- 327
I_loihiensisL2TR_NrdBg                LRDDAYEIFMTAVKQEKQWAHYLFKDGSMIGLNESMLCQYVEYIANLRMQAIGFDAPFEN 328
P_atlanticaT6c_NrdBg                  YKEEATNIFLDAVNQEKEWAEYLFQNGSMIGLNKDILCQYVEFIANHRMSAIGLDQPFDI 328
B_aphidicolaBp_NrdBg                  CYEECYHLFIRVSDQEKIWAKYLFQNGSMLGLNQEILSQYIEYITNIRMKAIGLPSPFK- 327
B_aphidicolaSg_NrdBg                  CREEAIKIFISASEQEKKWAKYLFQDGSMLGLNKDILCQYIEYITNIRMHAIGFKMPFE- 327
P_ingrahamii37_NrdBg                  CFAEGREIFLQAAQQEKDWAKYLFKDGSMIGLNEQILSQYVEYIANLRMQAIGYAPEFE- 332
MarinomonasMWYL1_NrdBg                CQDEVRQIFIEAAEQEKEWATYLFKDGSMIGLNAQILGQYVEYITNVRMQAVGLETCFA- 328
P_tunicataD2_NrdBg                    LREESIEMFRTAAQQEKEWAEYLFKDGSMIGLNKDILCQYVEYITNVRMSAVGLPQQFS- 327
S_loihicaPV4_NrdBg                    CEQTAIDIFVKAAEQEKEWAKYLFKDGSMIGLNEQILCQYVEYITNERMKSVNLVSPYA- 327
S_pealeana700345_NrdBg                CEQTAIDIFIKAAEQEKEWAKYLFKDGSMIGLNETILCQYVEYITNERMKAVNFASPYA- 327
S_denitrificansOS217_NrdBg            CQQQAYDLFYKAAEQEKEWAKYLFKDGSMIGLNEQILCQYVEYITNQRMKSVHLPQPYA- 327
S_frigidimarina400_NrdBg              CKQEAYDLFVKAAEQEKEWAKYLFKDGSMIGLNEQILCQYVEFITNQRMKSVHLPQPYD- 327
S_amazonensisSB2B_NrdBg               CQQEAYDLFLRAAEQEKEWAKYLFKDGSMIGLNEQILCQYVEYITNQRMKSVNLPLPYP- 342
S_oneidensisMR1_NrdBg                 CQQEAYNLFLKAAEQEKEWAKYLFKDGSMIGLNEQILCQYVEYITNQRMKSVNLPQPYA- 327
ShewanellaANA3_NrdBg                  CQQEAYNLFLKAAEQEKEWAKYLFKDGSMIGLNEQILCQYVEYITNQRMKSVNLPLPYA- 342
ShewanellaMR4_NrdBg                   CQQEAYNLFLKAAEQEKEWAKYLFKDGSMIGLNEQILCQYVEYITNQRMKSVNLPQPYA- 327
ShewanellaMR7_NrdBg                   CQQEAYNLFLKAAEQEKEWAKYLFKDGSMIGLNEQILCQYVEYITNQRMKSVNLPQPYA- 327
S_balticaOS155_NrdBg                  CQQEAFDLFFKAAEQEKEWAKYLFKDGSMIGLNEQILCQYVEFITNQRMKSVNLPQPYL- 327
S_putrefaciensCN32_NrdBg              CQQEAYDLFLKAAEQEKEWAKYLFKDGSMIGLNEQILCQYVEFITNQRMKSVNLPQPYP- 327
ShewanellaW3181_NrdBg                 CQQEAYDLFLKAAEQEKEWAKYLFKDGSMIGLNEQILCQYVEFITNQRMKSVNLPQPYP- 327
PsychrobacterPRwf1_NrdBg              CHNESIEIFRKAAEQEKEWAGYLFKDGSMIGLNKDILCQYVEYITNLRMEAVGLPGIFPE 328
P_arcticus2734_NrdBg                  CYEESIEIFRKAAEQEKEWAGYLFKDGSMIGLNKDILCQYIEYITNLRMEAVGLPVAFPK 328
P_cryohalolentisK5_NrdBg              CYEESIEIFRKAAEQEKEWAGYLFKDGSMIGLNKDILCQYIEYITNLRMEAVGLPVAFPN 328
CBlochmanniaBPEN_NrdBg                CQQESYALFLMAAEQEKNWANYLFKDGSMIGLNKDILCQYIEYITNIRMQAIGLRTPFN- 327
A_hydrophila7966_NrdBg                CEQACFELFKEAAIQEKEWAEYLFQDGSMIGLNKDILCQYVEYITNLRMNAVGLQPAFAG 328
A_salmonicidaA449_NrdBg               CEQACFEIFKEAAIQEKEWAEYLFQDGSMIGLNKDILCQYVEYITNLRMNAVGLQPAFAG 328
P_multocidaPm70_NrdBg                 CKQEAYDLFVAAAEQEKEWADYLFKDGSMIGLNRDILVQYVEYITNIRMQAVGLPLPFQ- 327
H_somnus129PT_NrdBg                   CKQEAYELFVAAAEQEKEWADYLFKDGSMIGLNKDILVQYVEYITNIRMQAVGLPLPFQ- 327
H_influenzae86028NP_NrdBg             CKQEAYDLFVAAAEQEKAWADYLFKDGSMIGLNRDILVQYVEYITNIRMQAVGLPLPFQ- 327
H_influenzaeRdKW20_NrdBg              CKQEAYDLFVAAAEQEKAWADYLFKDGSMIGLNRDILVQYVEYITNIRMQAVGLPLPFQ- 327
A_succinogenes130Z_NrdBg              CKQEAYDLFVAAAEQEKAWADYLFKDGSMIGLNKDILVQYVEYITNIRMQAVGLPLPFE- 327
M_succiniciproducensMBEL55E_NrdBg     CKQEAYDLFLAAAEQEKAWADYLFKDGSMIGLNKDILVQYVEYITNIRMQAVGLPLPFQ- 327
A_pleuropneumoniae5bL20_NrdBg         CKQEAYDLFVAAAEQEKEWAAYLFKDGSMIGLNEDILVKYVEYITNIRMQAVGLPRPFA- 327
H_ducreyi35000HP_NrdBg                CQQEAYDLFVAAAEQEKEWAEYLFKDGSMIGLNKEILVKYVEYITNIRMQSVGLPLPFA- 327
B_cicadellinicolaHc_NrdBg             CQHDSYHLFLXAANQEKDWASYLFRDGSMIGLNEEILCQYIEYITNIRMQAVLLDLPFA- 327
C_sakazakiiBAA894_NrdBg               CKQECYDLFVLAAQQEKEWAEYLFQGGSMIGLNKDILCQYVEYITNIRMQAVGLDLPFK- 327
Enterobacter638_NrdBg                 CKQECYDLFVLAALQEKEWAEYLFRDGSMIGLNKDILCQYIEYITNIRMQAVGLELPFK- 327
C_koseriBAA895_NrdBg                  SKQACYDLFVQAAQQEKEWADYLFRDGSMIGLNKDILCQYVEYITNIRMQAVGLELPFQ- 327
K_pneumoniaeMGH78578_NrdBg            CKQECYDLFVLAAQQEKEWADYLFRDGSMIGLNKDILCQYVEYITNIRMQAVGLDLPFQ- 327
E_coli536_NrdBg                       CKQECYDLFVQAAQQEKDWADYLFRDGSMIGLNKDILCQYVEYITNIRMQAVGLDLPFQ- 327
S_flexneri2a2457T_NrdBg               CKQECYDLFVQAAQQEKDWADYLFRDGSMIGLNKDILCQYVEYITNIRMQAVGLDLPFQ- 340
S_flexneri58401_NrdBg                 CKQECYDLFVQAAQQEKDWADYLFRDGSMIGLNKDILCQYVEYITNIRMQAVGLDLPFQ- 404
E_coliK12MG1655_NrdBg                 CKQECYDLFVQAAQQEKDWADYLFRDGSMIGLNKDILCQYVEYITNIRMQAVGLDLPFQ- 327
S_boydiiSb227_NrdBg                   CKQECYDLFVQAAQQEKDWADYLFRDGSMIGLNKDILCQYVEYITNIRMQAVGLDLPFQ- 327
S_TyphimuriumLT2_NrdBg                CKQECYDLFVQAAQQEKEWADYLFRDGSMIGLNKDILCQYVEYITNIRMQAVGLDLPFQ- 327
S_entericaTy2_NrdBg                   CKQECYDLFVQAAQQEKEWADYLFRDGSMIGLNKDILCQYVEYITNIRMQAVGLDLPFQ- 327
S_proteamaculans568_NrdBg             CQQQCYDLFVLAAEQEKEWAEYLFRDGSMIGLNKDILCQYVEYITNIRMQAVGLGLPFA- 327
Y_pestisAngolaNrdBg_2                 CQQQCYDLFVLAAQQEKEWAEYLFRDGSMIGLNKEILCQYVEYITNIRMQAVGLGIPFS- 327
Y_pestisMicrotus91001_NrdBg_2         CQQQCYDLFVLAAQQEKEWAEYLFRDGSMIGLNKEILCQYVEYITNIRMQAVGLGIPFS- 342
Y_pestisPestoidesF_NrdBg_1            CQQQCYDLFVLAAQQEKEWAEYLFRDGSMIGLNKEILCQYVEYITNIRMQAVGLGIPFS- 327
Y_pseudotuberculosisIP31758_NrdBg     CQQQCYDLFVLAAQQEKEWAEYLFRDGSMIGLNKEILCQYVEYITNIRMQAVGLGIPFS- 327
M_flagellatusKT_NrdBg                 LQDECFEMFKRAAEQEKKWADYLFKDGSMIGLNRDILCQYVEYITNNRMASVGLSAAFPN 328
N_gonorrhoeaeFA1090_NrdBg             LQDECFQLFKKAAEQEKEWAAYLFKDGSMIGLNKEILSQYVEYITNLRMQAVGLPAGFEG 335
N_meningitidis053442_NrdBg            LQDECFQLFKKAAEQEKEWAAYLFKDGSMIGLNKEILSQYVEYITNLRMQAVGLPAGFEG 335
N_meningitidisMC58_NrdBg              LQDECFQLFKKAAEQEKEWAAYLFKDGSMIGLNKEILSQYVEYITNLRMQAVGLPAGFEG 335
V_choleraeO395_NrdBg                  CKQESFDLFKEAAEQEKEWAEYLFKDGSMIGLNKDILCQYVEYITNIRMQAVGLEPAYPA 340
V_vulnificusCMCP6_NrdBg               AKQQCFDLFKEAAEQEKEWAEYLFKDGSMIGLNKDILCQYVEYITNIRMQAVGLPVAYPN 328
V_campbelliiBAA1116_NrdBg             EKQECFDLFKEAAEQEKEWAEYLFKDGSMIGLNKDILSQYVEYITNIRMQAVGLPAAYPE 328
VibrioAND4_NrdBg                      VKQECFDLFKEAAEQEKEWAEYLFKDGSMIGLNKDILSQYVEYITNIRMQAVGLPKAYPE 328
V_fischeriES114_NrdBg                 CEQECFDLFKAAAEQEKQWANYLFKDGSMIGLNKDILCQYVEYITNLRMKAVGLNSAYEG 328
P_angustumS14_NrdBg                   CEQECFDIFKAAAEQEKEWADYLFKDGSMIGLNKQILCQYVEYITNLRMVAVGLRPAYDG 328
PhotobacteriumSKA34_NrdBg             CEQDCFDLFKAAAEQEKEWADYLFKDGSMIGLNKQILCQYVEYITNLRMVAVGLRPAYEG 328
                                             :    :   *.  *:       :.   :  ::.: :   :
```

```
VibriophageKVP40_NrdBg                  --RRKHPLPWIRKWLNSDSVQVAAQEAELSSYLVGQIDSTVRP--EFLKTLEV--  374
AeromonasphageAeh1_NrdBg                NIGMKHPLPWIRKWLNSDSVQVAAQEAELSSYLIGQVDSTLDE--TFFEELEV--  376
EnterophageRB49_NrdBg                   --VRRHPLPWMRKWLNSDEVQNAAQEVEI-MYLVSQVNTDLTD--EDYNRFGKYL  386
EnterophagePhi1_NrdBg                   --VRRHPLPWMRKWLNSDEVQNAAQEVEI-MYLVSQVNTDLTD--DDYNRFGKYL  386
EnterophageRB69_NrdBg                   -APTKHPYPWIREYLNSDLVQAAPQEVEISSYLVAQIDNDVDQ--NVINSYKKYF  390
EnterophageJS98_NrdBg                   -APTRHPYPWIREYLNSDAVQAAPQEVEISSYLVAQIDNDVDE--KVMASYKKYF  379
EnterophageT4_NrdBg                     -APVKHPYPWIREYLNSDNVQSAPQEVELSSYLVAQIDNDVDD--KVMMSFKKYF  388
EnterophageRB32_NrdBg                   -APVKHPYPWIREYLNSDNVQSAPQEVELSSYLVAQIDNDVDD--KVMMSFKKYF  392
ProchlorococcusphagePSSM2_NrdBg         -AAKNNPLPWTEHWISSKGLQVAPQETEVESYIVGGIKQDVKK--DTFAGFKL--  384
ProchlorococcusphagePSSM4_NrdBg         -PLRNNPLPWTEHWLNSKGQQNAPQETEIESYVIGGIKQDVKG--DSFAGFKL--  387
Y_pestisAngola_NrdBg_1                  --IKDDPLVWMNKWLLSDTLQIAPQEAEQSNYLVGQIDSAVDR--SGLSQFADL-  380
Y_pestisMicrotus91001_NrdBg_1           --IKDDPLVWMNKWLLSDTLQIAPQEAEQSNYLVGQIDSAVDR--SGLSQFADL-  380
Y_pestisPestoidesF_NrdBg_2              --IKDDPLVWMNKWLLSDTLQIAPQEAEQSNYLVGQIDSAVDR--SGLSQFADL-  380
C_RuthiamagnificaCm_NrdBg               --ECTNPLPWINHWLDSDNVQVAPQETEITSYLVSAIDNTLDEQDSDFNDFEL--  378
C_psychrerythraea34H_NrdBg              --IKSNPLPWMNAYLVSDNVQVAPQETEISSYLVGQVDSSVDA--DDFDDFDL--  376
S_sediminisHAWEB3_NrdBg                 --IKNNPLPWINTYLVSDNVQVAPQESEITSYLVGQVDATVSD--DDFDDFEL--  376
I_loihiensisL2TR_NrdBg                  --QRSNPLPWMNKYLVSDNVQVAPQESEITSYLVGQIDSEVNA--SDLGDFDL--  377
P_atlanticaT6c_NrdBg                    ---KNNPLPWMNNYLNSDNVQVAPQESEISSYLVGQIDSEVGS--NDFDEFEL--  376
B_aphidicolaBp_NrdBg                    --ILSNPIPWINSWLISDNVQVAPQEISVSSYLIGQINSEIND--SDFEKFKL--  376
B_aphidicolaSg_NrdBg                    --KTSNPIPWINSWLNSDYIQAAPQETEISSYLV-QIESEISN--QEFKNFKL--  375
P_ingrahamii37_NrdBg                    --VKSNPIPWINNWLVSDNVQVAPQETEISSYLVGQIDSKVDS--EDLSTFEL--  381
MarinomonasMWYL1_NrdBg                  --TKHNPIPWINAWLVSDNVQVAPQEAEISSYLVGQTDNSLDD--SDFDDFDL--  377
P_tunicataD2_NrdBg                      --ITSNPIPWINSWLVSDNVQVAPQEAEISSYLVGQIDSQVEA--SDFGDFDL--  376
S_loihicaPV4_NrdBg                      --DQSNPLPWMKNWLESDSVQVAPQEVEVSSYLVGQIDASVDE--SEFADFDL--  376
S_pealeana700345_NrdBg                  --EQTNPLPWMKNWLESDSVQVAPQEVEVSSYLVGQIDSSIDD--NEFSDFDL--  376
S_denitrificansOS217_NrdBg              --EISNPLPWMKNWLESDAVQVAPQEVEVSSYLVGQIDSSVNS--DEFLDFDL--  376
S_frigidimarina400_NrdBg                --EQSNPLPWMKNWLESDAVQVAPQEVEVSSYLVGQIDSSVDG--DEFLDFDL--  376
S_amazonensisSB2B_NrdBg                 --ESTNPLPWMKNWLESDAVQVAPQEVEVSSYLVGQIDASVDE--NEFSDFDL--  391
S_oneidensisMR1_NrdBg                   --EQTNPLPWMKNWLESDAVQVAPQEVEVSSYLVGQIDAAVDS--DEFLDFDL--  376
ShewanellaANA3_NrdBg                    --ELSNPLPWMKNWLESDAVQVAPQEVEVSSYLVGQIDSAVDS--DEFLDFDL--  391
ShewanellaMR4_NrdBg                     --EQSNPLPWMKNWLESDAVQVAPQEVEVSSYLVGQIDSAVDS--DEFLDFDL--  376
ShewanellaMR7_NrdBg                     --EQSNPLPWMKNWLESDAVQVAPQEVEVSSYLVGQIDSAVDS--DEFLDFDL--  376
S_balticaOS155_NrdBg                    --EITNPLPWMKNWLESDAVQVAPQEVEVSSYLVGQIDSAVDS--DEFLDFDL--  376
S_putrefaciensCN32_NrdBg                --ESTNPLPWMKNWLESDAVQVAPQEVEVSSYLVGQIDSAVDG--DEFLDFDL--  376
ShewanellaW3181_NrdBg                   --ESTNPLPWMKNWLESDAVQVAPQEVEVSSYLVGQIDSAVDG--DEFLDFDL--  376
PsychrobacterPRwf1_NrdBg                --SKSNPIPWINTWLSSDNVQVAPQETEISSYLVGQIDSDLSD--SDFDDFEL--  377
P_arcticus2734_NrdBg                    --SKSNPIPWINTWLSSDNVQVAPQETEISSYLVGQIDSDLSD--SDFDDFEL--  377
P_cryohalolentisK5_NrdBg                --SKSNPIPWINTWLSSDNVQVAPQETEISSYLVGQIDSDLSD--SDFDDFEL--  377
CBlochmanniaBPEN_NrdBg                  --IKSNVIPWINSWLSSDNVQVAPQEVEASSYLVGQIDPNIDE--NEFHDFQL--  376
A_hydrophila7966_NrdBg                  --AKQNPIPWINTWLSSDNVQVAPQEVEVSSYLVGQIDSEVNA--EDLGDFEL--  377
A_salmonicidaA449_NrdBg                 --AKQNPIPWINTWLSSDNVQVAPQEVEVSSYLVGQIDSEVHA--DDLGDFEL--  377
P_multocidaPm70_NrdBg                   --ARSNPIPWINAWLVSDNVQVAPQEVEVSSYLVGQIDSKVDT--KDFGDFDL--  376
H_somnus129PT_NrdBg                     --ARSNPIPWINAWLVSDNVQVAPQEVEVSSYLVGQIDAKVDT--NDFGDFDL--  376
H_influenzae86028NP_NrdBg               --TRSNPIPWINAWLVSDNVQVAPQEVEVSSYLVGQIDSKVDT--NDFDDFSL--  376
H_influenzaeRdKW20_NrdBg                --TRSNPIPWINAWLVSDNVQVAPQEVEVSSYLVGQIDSKVDT--NDFDDFSL--  376
A_succinogenes130Z_NrdBg                --ARSNPIPWINAWLVSDNVQVAPQEVEVSSYLVGQIDSKVDA--DDFGDFDL--  376
M_succiniciproducensMBEL55E_NrdBg       --ARSNPIPWINAWLVSDNVQVAPQEVEISSYLIGQIDSKVDT--SDFGDFDL--  376
A_pleuropneumoniae5bL20_NrdBg           --ASSNPIPWINAWLVSDNVQVAPQEVEVSSYLVGQINAKVDT--NDFDSFEL--  376
H_ducreyi35000HP_NrdBg                  --TQSNPIPWINAWLVSDNVQVAPQEVEVSSYLVGQIDSKVDS--NDFDNFEL--  376
B_cicadellinicolaHc_NrdBg               --TRSNPIPWINSWLVSDNVQVAPQEVEVSSYLVGQINSEVNS--DDFSDFQL--  376
C_sakazakiiBAA894_NrdBg                 --TRSNPIPWINTWLVSDNVQVAPQEVEVSSYLVGQIDSEIDH--DDLSSFQL--  376
Enterobacter638_NrdBg                   --TRSNPIPWINTWLVSDNVQVAPQEVEVSSYLVGQIDSEVNT--DDLSSFEL--  376
C_koseriBAA895_NrdBg                    --TRSNPIPWINTWLVSDNVQVAPQEVEVSSYLVGQIDSEVDT--DDLSNFQL--  376
K_pneumoniaeMGH78578_NrdBg              --TRSNPIPWINTWLVSDNVQVAPQEVEVSSYLVGQIDSEVDA--DDLSNFQL--  376
E_coli536_NrdBg                         --TRSNPIPWINTWLVSDNVQVAPQEVEVSSYLVGQIDAEVDT--DDLSNFQL--  376
S_flexneri2a2457T_NrdBg                 --TRSNPIPWINTWLVSDNVQVAPQEVEVSSYLVGQIDAEVDT--DDLSNFQL--  389
S_flexneri58401_NrdBg                   --TRSNPIPWINTWLVSDNVQVAPQEVEVSSYLVGQIDAEVDT--DDLSNFQL--  453
E_coliK12MG1655_NrdBg                   --TRSNPIPWINTWLVSDNVQVAPQEVEVSSYLVGQIDSEVDT--DDLSNFQL--  376
S_boydiiSb227_NrdBg                     --TRSNPIPWINTWLVSDNVQVAPQEVEVSSYLVGQIDSEVDT--DDLSNFQL--  376
S_TyphimuriumLT2_NrdBg                  --TRSNPIPWINTWLVSDNVQVAPQEVEVSSYLVGQIDSEVDT--DDLSNFQL--  376
S_entericaTy2_NrdBg                     --TRSNPIPWINTWLVSDNVQVAPQEVEVSSYLVGQIDSEVDT--DDLSNFQL--  376
S_proteamaculans568_NrdBg               --TRSNPIPWINAWLVSDNVQVAPQEVEVSSYLVGQIDSEVHA--DDLSDFEL--  376
Y_pestisAngolaNrdBg_2                   --TRTNPIPWINSWLVSDNVQVAPQEVEVSSYLVGQIDSEVSA--DDLSDFEL--  376
Y_pestisMicrotus91001_NrdBg_2           --TRTNPIPWINSWLVSDNVQVAPQEVEVSSYLVGQIDSEVSA--DDLSDFEL--  391
Y_pestisPestoidesF_NrdBg_1              --TRTNPIPWINSWLVSDNVQVAPQEVEVSSYLVGQIDSEVSA--DDLSDFEL--  376
Y_pseudotuberculosisIP31758_NrdBg       --TRTNPIPWINSWLVSDNVQVAPQEVEVSSYLVGQIDSEVSA--DDLSDFEL--  376
M_flagellatusKT_NrdBg                   --AKNNPIPWINTWLASDTVQVAPQEVELSSYLIGQIDAEVSS--EDLGDFEL--  377
N_gonorrhoeaeFA1090_NrdBg               --ANQNPIPWINAWLSSDNVQVAPQEVEISSYLIGQIDSEVNT--DDLGDFEL--  384
N_meningitidis053442_NrdBg              --ANQNPIPWINAWLSSDNVQVAPQEVEISSYLIGQIDSEVNT--DDLGDFEL--  384
N_meningitidisMC58_NrdBg                --ANQNPIPWINAWLSSDNVQVAPQEVEISSYLIGQIDSEVNT--DDLGDFEL--  384
V_choleraeO395_NrdBg                    --ATNNPIPWINAWLSSDNVQVAPQEAEISSYLVGQIDNEVSA--DDFEGFEL--  389
V_vulnificusCMCP6_NrdBg                 --ATSNPIPWINAWLSSDNVQVAPQEAEISSYLVGQIDNEVHS--DDFEGFEL--  377
V_campbelliiBAA1116_NrdBg               --ATSNPIPWINAWLSSDNVQVAPQEAEISSYLVGQIDNEVSA--DDFEGFEL--  377
VibrioAND4_NrdBg                        --ATTNPIPWINAWLSSDNVQVAPQEAEISSYLVGQIDNEVSS--GDFEGFEL--  377
V_fischeriES114_NrdBg                   --ATQNPIPWINSWLSSDNVQVAPQETEISSYLVGQIDNDVDA--DDLGDFEL--  377
P_angustumS14_NrdBg                     --ATQNPIPWINSWLSSDNVQVAPQEAEISSYLVGQIDNDVSA--DDLGDFEL--  377
PhotobacteriumSKA34_NrdBg               --ATQNPIPWINSWLSSDNVQVAPQEAEISSYLVGQIDNDVSA--DDLGDFEL--  377
                                              .   * . :: *.  * * ** .   *::   .  :
```

Appendix 2

Supplementary Data for Chapter 2

Supplementary results and discussion of SV-AUC data collected for *B. subtilis* Ib RNR.

Figure A2.1. SV-AUC results for $Fe(III)_2$-Y• and $Mn(III)_2$-Y• His_6-NrdF.

Figure A2.2. SV-AUC results for His_6-NrdE.

Figure A2.3. SV-AUC results for 1:1 mixtures of His_6-NrdE:$Fe(III)_2$-Y• His_6-NrdF

Figure A2.4. Anion exchange chromatogram of His_6-NrdE run under identical conditions to those in **Figure 2.7A**.

Table A2.1. Comparison of activities reported in the literature for $Me(III)_2$-Y• NrdF proteins (Me = Fe or Mn) isolated from a variety of sources and organisms.

Table A2.2. *In vivo* concentrations of *B. subtilis* class Ib RNR subunits as reported in Muntel *et al.*, 2014.[1]

Table A2.3. *In vivo* concentrations of *B. subtilis* NrdE during transition to glucose starvation.[2]

Table A2.4. *In vivo* concentrations of *B. subtilis* class Ib RNR subunits and YmaB during heat stress.[2]

Table A2.5. *In vivo* concentrations of *B. subtilis* class Ib RNR subunits and YmaB as reported in Maaß *et al.*, 2011.[3]

Table A2.6. Molecular weights of *B. subtilis* RNR subunits and complexes determined by SEC.

A2.1. SUPPLEMENTARY RESULTS

A2.1.1. SV-AUC results for holo-Me(III)$_2$-Y• NrdF (Me = Fe or Mn). Over the entire concentration range tested (~1 – 12 μM), the $s_{20,w}$ distributions for both Fe(III)$_2$-Y•- and Mn(III)$_2$-Y•-loaded His$_6$-NrdF displayed two peaks with average weighted sedimentation coefficients ($s_{<Wt>}$[20,w]) of 3.3 S and 4.8 S (**Figure A2.1A** and **C**). The signal contributions of the two peaks remained constant over the entire concentration range and the weight averaged $s_{20,w}$ isotherms (**Figure A2.1B** and **D**) remained fairly constant over the concentration range tested, indicating two non-interacting species were present in solution. Globally fitting all data sets in Sedphat using the "hybrid local continuous/global discrete species model" yielded $s_{20,w}$ values of 3.7 S, 68.3% confidence interval (CI) [3.3,3.9] (~20% of the total signal), and 4.8 S, 68.3% CI[4.6, 5.0] (~80% of the total signal), and M_W values of 45.3 kDa, 68.3% CI[38.3, 53.7] and 89.8 kDa, 68.3% CI[78.2,108.8] (**Table 2.6**) for the series of experiments using Mn(III)$_2$-Y• His$_6$-NrdF. The results obtained for Fe(III)$_2$-Y• His$_6$-NrdF were similar ($s_{20,w}$ = 3.5 S, 68.3% CI[3.0, 3.8] and 4.8 S, 68.3% CI[4.7, 5.0]; M_W = 40.7 kDa, 68.3% CI[35.8, 45.6] and 83.5 kDa, 68.3% CI[77.4, 110]). These values, along with the frictional ratio (f/f_o) and R_s determined in Sednterp with the global $s_{20,w}$ values, clearly indicate that the larger species corresponds to His$_6$-NrdF dimer and the smaller species to His$_6$-NrdF monomer (likely apo- or mismetallated) when compared with the SEC results (**Table 2.6**) and the predictions made by HYDROPRO[4] (**Table 2.7**). Large deviations between the experimentally measured and predicted hydrodynamic properties can be expected since the crystal structure of His$_6$-NrdF used to make the predictions is incomplete (missing the N-terminal His-tag and the C-terminal tails of each β unit). Additionally, apo-His$_6$-NrdF may not have the same structure as holo-protein, and since a structure of the apo-form is not available, deviations in the predicted and measured parameters for this protein can be expected.

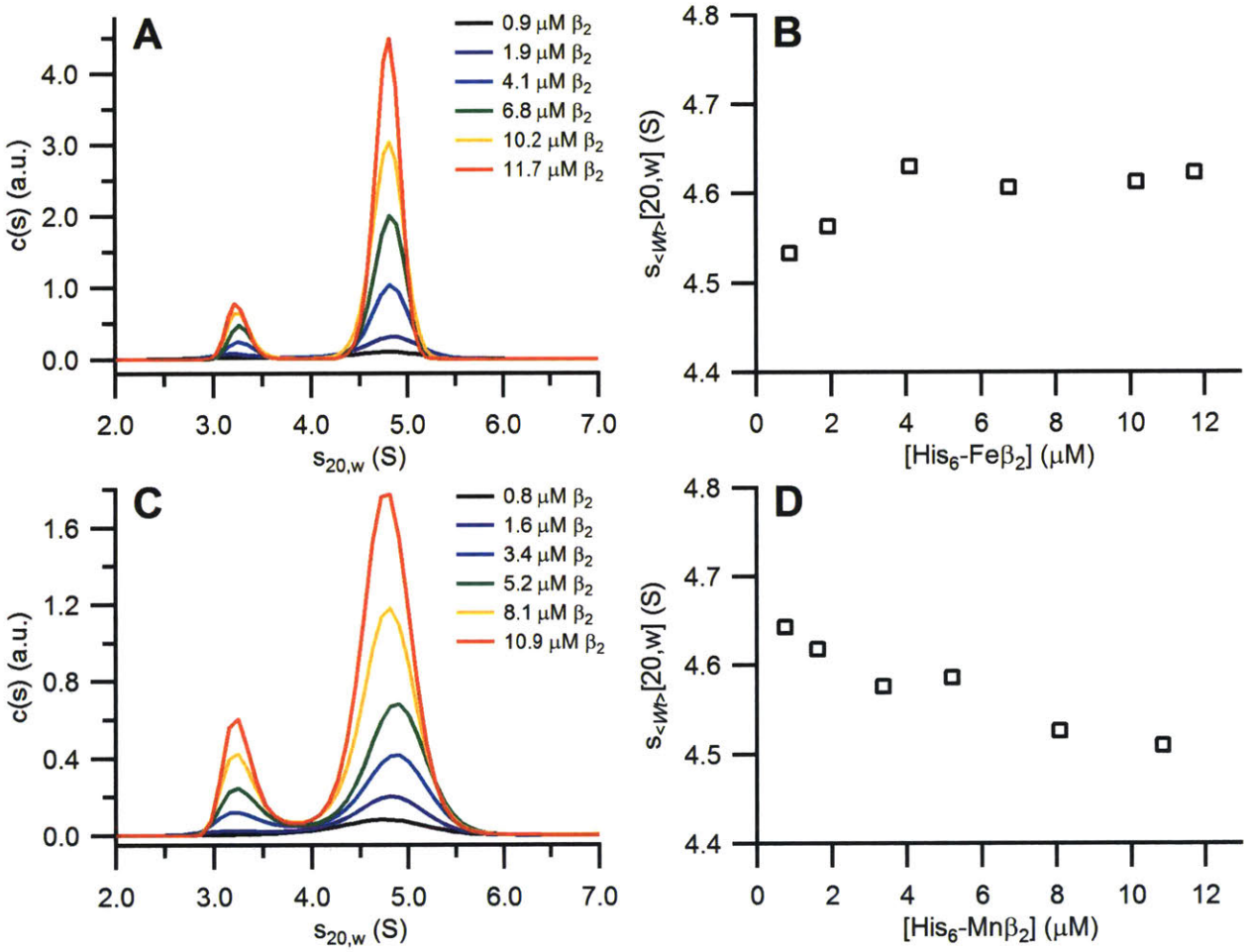


Figure A2.1. SV-AUC results for $Me(III)_2$-Y• His_6-NrdF. Sedimentation coefficient distributions show two peaks whose ratio is maintained over the entire concentration range examined. The weight averaged $s_{20,w}$ isotherms show that, despite slight upward or downward trends with increasing concentration, $s_{<Wt>}[20,w]$ remains relatively constant, indicating the two species are discrete, non-interacting species. (**A**) $s_{20,w}$ distributions for samples of $Fe(III)_2$-Y• His_6-NrdF. (**B**) Weight averaged $s_{20,w}$ isotherm for the $Fe(III)_2$-Y• His_6-NrdF samples. (**C**) Sedimentation coefficient distributions for samples of $Mn(III)_2$-Y• His_6-NrdF. (**D**) Weight averaged $s_{20,w}$ isotherm for the $Mn(III)_2$-Y• His_6-NrdF samples.

A2.1.2. SV-AUC results for NrdE. Using A_{280} as the detection wavelength, the detectable concentration range of His_6-NrdE was 1 – 12 μM (**Table 2.2**). At physiological concentrations (~ 1 μM), His_6-NrdE, in the absence of nucleotides, sedimented as a single peak ($s_{w,20}$ = 5.6, **Figure A2.2A**) with hydrodynamic properties consistent with His_6-NrdE monomer (**Table 2.6**). As the concentration increased, the peak broadened and eventually evolved into a reaction boundary at ~9 μM His_6-NrdE consisting of two peaks with $s_{20,w}$ = 5.6 S and 7.4 S (**Figure A2.2A**). These values and the associated hydrodynamic properties and M_W are consistent with His_6-NrdE monomer and dimer (**Table 2.7**). To probe the self-association of His_6-NrdE at higher concentrations, A_{250} detection was used, and an estimation of protein concentration was obtained using the A_{280}:A_{250} ratio (2.5 ± 0.2). The range detectable using this wavelength was 3 – 40 μM. In this concentration regime, the peak corresponding to α_2 shifted to ~9.7 S and broadened into a second reaction boundary at the highest concentrations tested (27 – 32 μM His_6-NrdE) with the larger sedimenting species appearing at ~12 S (**Figure A2.2B**). The isotherm of the weight averaged $s_{20,w}$ plotted against the entire concentration range analyzed (**Figure A2.2C**) never reaches an upper asymptote, appearing to increase exponentially from a lower asymptote of ~5.8 S. The α subunits of the class I RNRs have long been known to have a propensity to form aggregates, and the results reported here, while not surprising, provide the first systematic study of this phenomenon. The weight averaged f/f_o decreased from ~1.25 at 1 μM His_6-NrdE to ~0.90 at 32 μM (**Figure A2.2D**). Given that the physically possible minimum f/f_o is 1.0 for a perfect sphere, these unrealistic f/f_o indicate extra boundary broadening during sedimentation that is not attributable to diffusion alone. These values instead indicate NrdE self-association reactions are fast relative to the timescale of sedimentation.[5]

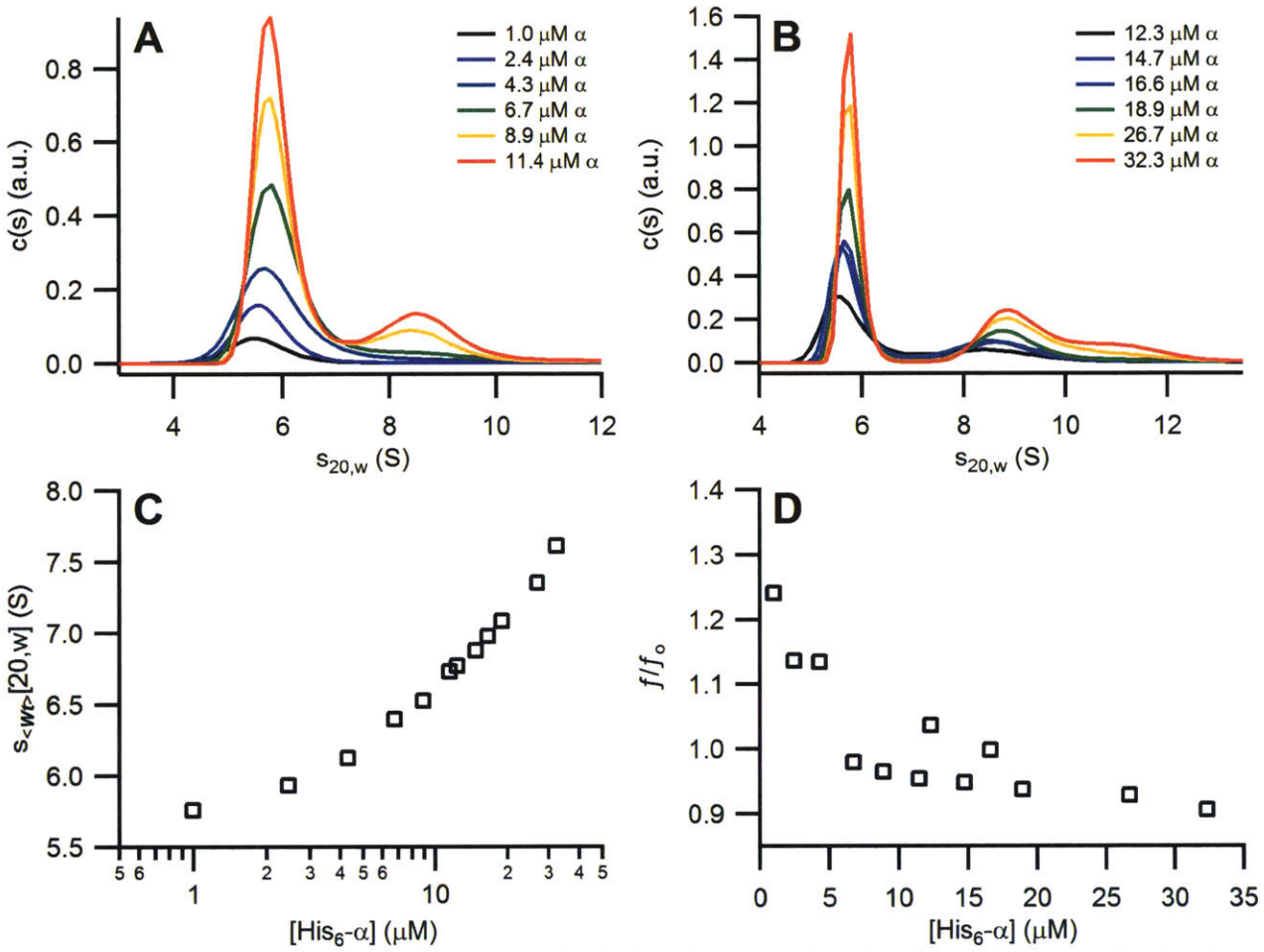


Figure A2.2. SV-AUC results for His_6-NrdE in the absence of nucleotides. (**A**) $s_{20,w}$ distributions for the samples monitored at A_{280}. The protein exists mainly as a monomer, but self-associates to produce a second species consistent with dimer at concentrations as low as 4 μM. (**B**) $s_{20,w}$ distributions for the samples monitored at A_{250}. Additional self-association becomes evident when the concentration of His_6-NrdE exceeds ~25 μM. (**C**) Weight averaged $s_{20,w}$ isotherm constructed from all data sets at both detection wavelengths. In the concentration range tested, His_6-NrdE produces no temporally stable complexes, as evidenced by the absence of an upper asymptote. (**D**) The weight average f/f_o decreases as His_6-NrdE concentrations increase, falling below 1, the value expected for a perfect sphere of the same M_W, at about 6 μM monomer. This indicates that the His_6-NrdE quaternary structures are interconverting at rates faster than the timescale of sedimentation.[5]

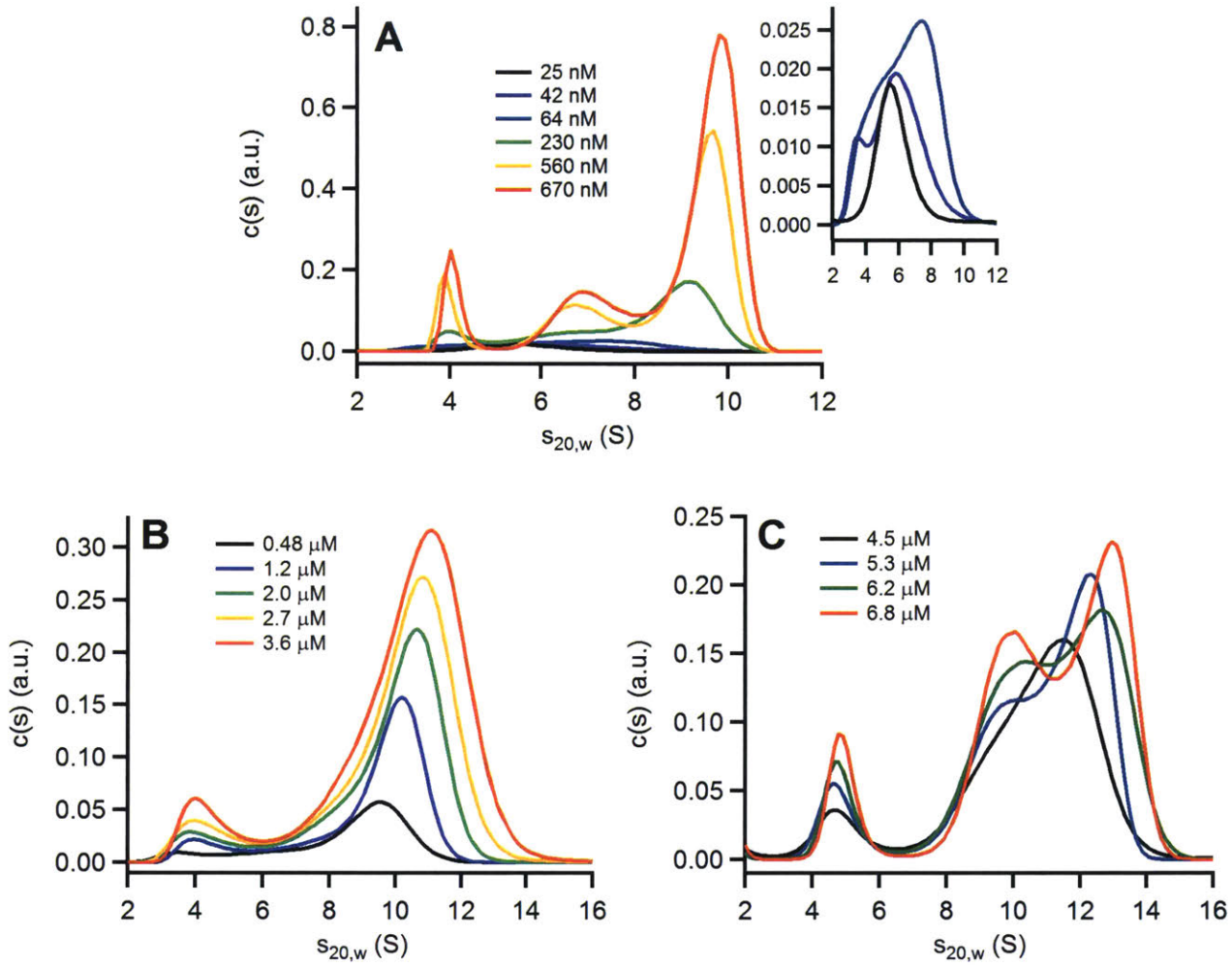


Figure A2.3. SV-AUC results for a 1:1 mixture of His_6-α_2:Fe(III)$_2$-Y• His_6-β_2 in the absence of nucleotides at the indicated concentrations. (**A**) $s_{20,w}$ distributions for samples monitored at A_{230}. Inset: a blow up of the three lowest concentration distributions. These results show an initial peak at roughly 4.4 S that broadens out and develops into a reaction boundary consisting of a mixture of low (~6.5 S) and high (~9.4 S) molecular weight species. Both peaks shift to higher $s_{20,w}$ values as the concentration of subunits increases. Also present is a peak corresponding to His_6-NrdF monomer (~3.9 S) as expected from the results of this subunit run alone (**Figure A2.1**). (**B**) $s_{20,w}$ distributions for samples monitored at A_{280}. The low and high M_W species merge forming one very broad peak that continues to shift to higher $s_{20,w}$ values as the subunit concentration increases (~11.2 S at 3.6 μM subunits). (**C**) $s_{20,w}$ distributions for samples monitored at A_{250}. New larger molecular weight oligomeric form(s) (~12.5 S) are observed as the subunit concentrations increase above 5.3 μM. The peaks corresponding to this species begins to separate from a peak at ~10 S as the concentration increases to ~7 μM. These data illustrate that complexes of *B. subtilis* Ib RNR form fairly easily, but are temporally unstable and interconvert despite the "tight" subunit affinity measured by activity assays.

A2.1.3. SV-AUC analysis of 1:1 NrdE:Fe(III)$_2$-Y• NrdF mixtures. As mentioned in Chapter 2, a reaction boundary with apparent peaks at 4.4 S and 9.7 S was observed at 1.2 μM (**Figure 2.10B**). These values are not necessarily the true $s_{20,w}$ of these entities as each protein exhibits a different f/f_o, but the c(s) model only employs a single, weight averaged f/f_o to fit the data. Protein interactions and their kinetics relative to the timescale of sedimentation influence the fitting of this ratio, typically leading to under or over estimation of the $s_{20,w}$ for the actual species corresponding to the peaks displayed in the distributions. A plot of the best fit f/f_o from the c(s) models versus loading concentration (**Figure 2.10C**) for the entire collection of SV experiments with 1:1 His$_6$-NrdE:His$_6$-NrdF mixtures shows a sharp decrease in value as the protein concentration increases. As described previously for His$_6$-NrdE, this observation likely indicates extra sedimentation boundary broadening resulting from protein interactions that are rapid relative to the timescale of sedimentation.[5] To try to get a better idea of the molecular weights and hydrodynamic properties of the entities, the data for the 1.2 μM sample was fit for $s_{20,w}$ and M_W (**Table 2.6**) in Sedphat by treating the peaks as discrete species with individual f/f_o that are surrounded by segments of continuous distribution that can be fitted for their own ratios.[5] In order to determine the more precise f/f_o (and R_s) for each species, the best fit $s_{20,w}$ were entered into the program Sednterp, using the M_Ws expected from amino acid sequence, as those determined by solely fitting SV-AUC data in Sedphat are typically ± 10% of the actual value. The calculated hydrodynamic properties of the smaller species (**Table 2.6**), in comparison with the SEC data and HYDROPRO predictions (**Table 2.7**), strongly suggest that it corresponds to His$_6$-NrdF monomer. The results for the larger species ($s_{20,w}$ = 9.6 S) are similar to those of the large species observed in the SEC data (**Table 2.6**), but are smaller than predicted for an $\alpha_2\beta_2$ complex (**Table 2.7**). The formation of larger complexes (**Figure A2.3C**) prevent rigorous treatment of the data to determine the kinetics and

affinities of the binding partners, which can be estimated by plotting the weight averaged $s_{20,w}$ versus loading concentration to generate a binding isotherm. As shown in **Figure 2.10D**, the weight averaged $s_{20,w}$ isotherm is complicated and we currently do not have a model that would satisfactorily explain the data.

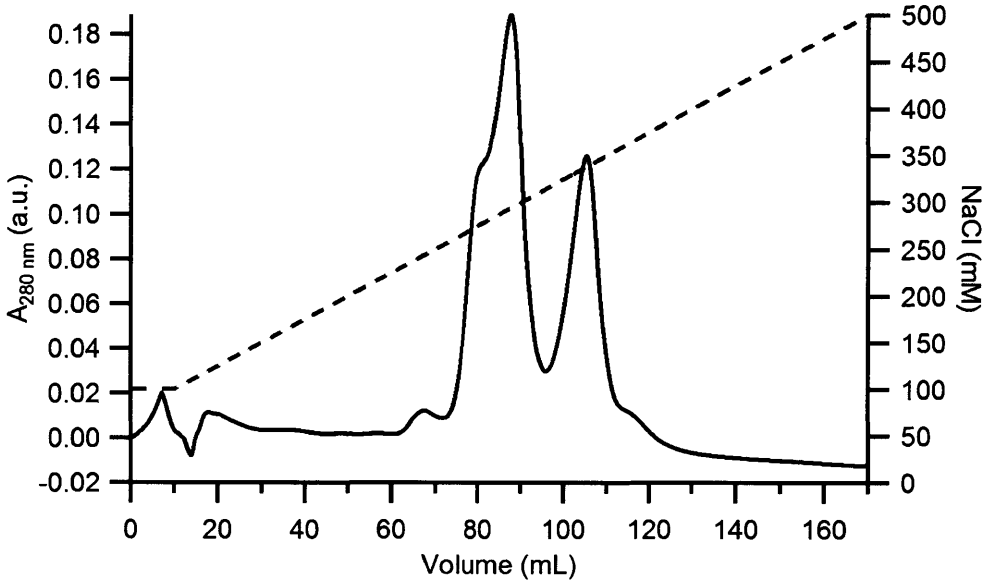


Figure A2.4. Anion exchange chromatogram of His_6-NrdE (~270 μM, 50 μL) run on a FPLC using a PorosHQ/20 column. The buffer used was 50 mM Tris, pH 7.6, 5% (w/v) glycerol. The column was washed with 0.5 CVs of Tris buffer supplemented with 100 mM NaCl before being developed with a 160 mL linear gradient from 100 – 500 mM NaCl in Tris buffer at a flow rate of 1 mL min^{-1}.

Table A2.1. Activities of $Me(III)_2$-Y• NrdF (Me = Fe or Mn) from various sources.

Source	Metal	Y•/β_2	Reductant	Equiv. α	SA (U mg^{-1})	Reference
Bacillus anthracis						
			NrdH	4 – 8	7.2 – 9.6	6
Reconstituted	Fe	0.57	TR/TRR/NADPH	~0.4	~6.5	7
			DTT	4 – 8	8.4 – 8.9	6
			NrdH	4 – 8	56 – 70	6
Reconstituted	Mn	0.3 – 0.4	TR/TRR/NADPH	~0.4	56	7
			DTT	4 – 8	27 – 40	6
Bacillus subtilis						
As-isolated[a]	Mn	0.4 – 0.5	DTT	10	160	8
Reconstituted	Fe	1.0 – 1.2	TR/TRR/NADPH	1	156	9
			DTT	1	52	9
Reconstituted	Mn	1.1	TR/TRR/NADPH	1	1475	9
			DTT	1	146	9
Corynebacterium ammoniagenes						
As-isolated[b]	Fe	0.05	DTT	1 – 4	36	10
As-isolated[a]	Mn	0.36	DTT	2	69000[c]	11
Reconstituted	Fe	0.2	DTT	1 – 4	48	10
Reconstituted	Mn	0.03	DTT	1 – 4	9	10
Escherichia coli						
As-isolated	Mn	0.2	DTT	5	720	12
Reconstituted	Fe	0.7	DTT	5	300	13
Reconstituted	Mn	0.25	DTT	5	600	13
Lactococcus lactis						
As-isolated	N/R[d]	N/R[d]	DTT	N/R[d]	198	14

Mycobacterium tuberculosis						
As-isolated	Fe	0.3 – 0.4	DTT	1 – 32	120	[15]
Staphylococcus aureus						
Reconstituted	N/R[d]	N/R[d]	NrdH	2	48	[16]
			TR/TRR/NADPH	2	61	[16]
Streptococcus pyogenes						
As-isolated	Fe	1	DTT	N/R[d]	169	[17]
Streptococcus sanguinis						
Reconstituted	Fe	1.2	NrdH/TRR/NADPH	2	1500	[18]
			DTT	10	170	[18]
Reconstituted	Mn	0.9	NrdH/TRR/NADPH	2	5000	[18]
			DTT	10	260	[18]
Salmonella typhimurium						
As-isolated	Fe	1	DTT	0.05 – 3	830	[19]
Reconstituted	Fe	0.4	DTT	1 – 4	325	[10]

[a] Isolated from the endogenous species. [b] Isolated from *E. coli* overexpression strains. [c] This reported activity is dubious as it is substantially higher than any other class I RNR activity reported. The strain of *C. ammoniagenes* genetically manipulated to overexpress NrdF is no longer available, thus this activity cannot be confirmed. [d] N/R = not reported.

The data presented in **Tables A2.2 – A2.5** were reported in the original references as number of molecules per cell. In order to convert the number of protein molecules per cell into a molar concentration, the cell dimensions (average cell length: 3.34 ± 0.82 μm, average cell width: 1.03 ± 0.08 μm) reported by Henriques *et al.*[20] were used to calculate a cell volume of (2.58 ± 1.06) x 10^{-15} L, assuming a *B. subtilis* cell is rod shaped with hemi-spherical cell poles.

Table A2.2. *In vivo* concentrations of *B. subtilis* class Ib RNR subunits and reductants as reported in Muntel *et al.*, 2014[1].[a]

Protein	Minimal medium[b]		Complex medium[c]	
	# molecules per cell	[Protein] (μM)	# molecules per cell	[Protein] (μM)
NrdE	268	0.2 ± 0.1	508	0.4 ± 0.2
NrdF	288	0.2 ± 0.1	446	0.3 ± 0.1
TrxA	7758	6.0 ± 3.0	7948	6.0 ± 3.0
TrxB	434	0.3 ± 0.1	644	0.5 ± 0.2
YtnI	-	-	97	0.10 ± 0.03

[a] Cells were grown at 37 °C and harvested in exponential growth phase.
[b] Minimal medium: 2 g L^{-1} $(NH_4)_2SO_4$, 1.4 g L^{-1} K_2HPO_4, 0.6 g L^{-1} KH_2PO_4, 0.1 g L^{-1} sodium citrate, 0.2 g L^{-1} $MgSO_4$, 0.001 g L^{-1} $MnSO_4$, 5 g L^{-1} glucose, 0.001 g L^{-1} $FeCl_3$.
[c] Complex medium: 10 g L^{-1} Casein hydrolysate, 3.9 g L^{-1} L-glutamic acid, 1.3 g L^{-1} L-alanine, 1.5 g L^{-1} L-asparagine, 1.4 g L^{-1} KH_2PO_4, 1.4 g L^{-1} NH_4Cl, 0.1 g L^{-1} Na_2SO_4, 0.1 g L^{-1} NH_4NO_3, 0.1 g L^{-1} $MgSO_4$, 0.02 g L^{-1} $CaCl_2$, 0.075 g L^{-1} $MnSO_4$, 0.001 g L^{-1} $FeCl_3$.

Table A2.3. *In vivo* concentrations of *B. subtilis* NrdE during transition to glucose starvation[2].[a,b]

Time point	# molecules per cell	[Protein] (μM)
Exponential phase	707	0.5 ± 0.2
Transition phase	733	0.5 ± 0.2
Maximal OD	748	0.5 ± 0.2
Glucose starvation, 60 min	529	0.4 ± 0.2
Glucose starvation, 120 min	430	0.3 ± 0.1
Glucose starvation, 180 min	356	0.3 ± 0.1
Glucose starvation, 240 min	476	0.3 ± 0.1

[a] Cells were grown at 37 °C and culture aliquots were harvested in a time course covering exponential to late stationary phase with glucose starvation.
[b] M9 minimal medium: 0.1% (w/v) glucose, 0.1% (w/v) *L*-malate, 0.015 g L^{-1} $CaCl_2 \cdot 2H_2O$, 0.246 g L^{-1} $MgSO_4 \cdot 7H_2O$, 8.5 g L^{-1} $Na_2HPO_4 \cdot 2H_2O$, 3 g L^{-1} KH_2PO_4, 0.5 g L^{-1} NaCl, g L^{-1} NH_4Cl, and trace elements (10 ml stock solution for 1 L medium). Trace elements consisted of (per liter): $ZnCl_2$ (0.170 g), $CuCl_2 \cdot 2H_2O$ (0.043 g), $MnCl_2 \cdot 4H_2O$ (0.100 g), $CoCl_2 \cdot 6H_2O$ (0.060 g), and $Na_2MoO_4 \cdot 2H_2O$ (0.060 g).

Table A2.4. *In vivo* concentrations of *B. subtilis* class Ib RNR subunits, reductants, and YmaB during heat stress[2].[a,b]

Time point	Protein	# molecules per cell	[Protein] (μM)
Exponential phase	NrdE	1153	0.8 ± 0.34
	NrdF	1434	1.1 ± 0.5
	YmaB	999	0.7 ± 0.3
	TrxA	9490	6.9 ± 3.1
	TrxB	7310	5.3 ± 2.4
10 min Heat stress	NrdE	1185	0.9 ± 0.4
	NrdF	1087	0.8 ± 0.4
	YmaB	516	0.4 ± 0.2
	TrxA	13545	9.9 ± 4.4
	TrxB	6290	4.6 ± 2.1
30 min Heat stress	NrdE	1011	0.7 ± 0.3
	NrdF	1007	0.7 ± 0.3
	YmaB	893	0.7 ± 0.3
	TrxA	19205	14.0 ± 6.3
	TrxB	6037	4.4 ± 2.0
60 min Heat stress	NrdE	1018	0.7 ± 0.3
	NrdF	1152	0.8 ± 0.4
	YmaB	1535	1.1 ± 0.5
	TrxA	18531	13.5 ± 6.1
	TrxB	7853	5.7 ± 2.6

[a] Cells were grown at 37 °C to mid-exponential phase and then subjected to a temperature jump to 52 °C.
[b] Belitsky minimal medium: 50 mM Tris, pH 7.5, 15 mM $(NH_4)_2SO_4$, 8 mM $MgSO_4$, 27 mM KCl, 7 mM sodium citrate, 2 mM $CaCl_2$, 1 μM $FeSO_4$, 10 μM $MnSO_4$, 4.5 mM potassium glutamate.

Table A2.5. *In vivo* concentrations of *B. subtilis* class Ib RNR subunits, reductants, and YmaB as reported in Maaß *et al.*, 2011[3].[a,b]

Time point	Protein	# molecules per cell	[Protein] (μM)
Exponential phase	NrdE	494	0.4 ± 0.2
	NrdF	1945	1.4 ± 0.6
	YmaB	915	0.7 ± 0.3
	TrxA	35338	5.6 ± 1.9
	TrxB	768	1.2 ± 0.4
Early stationary phase after glucose exhaustion	NrdE	385	0.3 ± 0.1
	NrdF	977	0.7 ± 0.3
	YmaB	508	0.4 ± 0.2
	TrxA	16009	5.6 ± 3.7
	TrxB	492	1.7 ± 1.2
Late stationary phase after glucose exhaustion	NrdE	209	0.2 ± 0.1
	NrdF	664	0.5 ± 0.2
	YmaB	1277	0.9 ± 0.4
	TrxA	-	-
	TrxB	337	0.8 ± 0.4

[a] Cells were grown at 37 °C in an aerated fermenter (dissolved O_2 > 50% saturation) and culture aliquots were collected between exponential and late stationary phase with glucose exhaustion.

[b] Batch minimal medium composition: 8.8 g L^{-1} glucose·H_2O, 2.0 g L^{-1} $Na_2SO_4 \cdot 10H_2O$, 2.68 g L^{-1} $(NH_4)_2SO_4$, 1.0 g L^{-1} NH_4Cl, 14.6 g L^{-1} K_2HPO_4, 4.02 g L^{-1} $NaH_2PO_4 \cdot 2H_2O$, 7% (w/v) $CaCl_2 \cdot 2H_2O$, 2 mM $MgSO_4 \cdot 7H_2O$; 3 ml L^{-1} of trace element solution (16.7 g L^{-1} $FeCl_3 \cdot 6H_2O$, 20.1 g L^{-1} Na_2EDTA, 0.18 g L^{-1} $ZnSO_4 \cdot 7H_2O$, 0.1 g L^{-1} $MnSO_4 \cdot H_2O$, 0.16 g L^{-1} $CuSO_4 \cdot 5H_2O$, 0.18 g L^{-1} $CoCl_2 \cdot 6H_2O$, 0.6 mg L^{-1} Na_2MoO_4).

Table A2.6. Molecular weights of *B. subtilis* RNR subunits and complexes determined by SEC.[a]

Experiment	Corresponding chromatogram in **Figure 2.9**	Average V_e (mL)	M.W. (kDa)	Possible species	Expected M.W. (kDa)
23 μM His_6-F$e\beta_2$	A*(i)*	11.4	219	?	-
		13.3	95	β_2	81.2
		14.8	50	β	40.6
18 μM His_6-M$n\beta_2$	A*(ii)*	13.4	93	β_2	81.2
		14.7	51	β	40.6
2.7 μM His_6-α_2	B*(i)*	14.1	67	α	83.0
27 μM His_6-α_2	B*(ii)*	9.4	537	α_7	581
		10.2	381	α_5	415
		12.2	155	α_2	166
		13.9	74	α	83.0
1 μM His_6-$\alpha_2\beta_2$	C *(i)*	10.1	417	?	-
		11.8	192	$\alpha_2\beta_2$	247
		13.1	110	?	-
		14.4	59	?	-
4 μM His_6-$\alpha_2\beta_2$	C*(ii)*	9.9	431	?	-
		11.1	253	$\alpha_2\beta_2$	247
		13.3	95	β_2 or α	81.2 or 83.0
		14.2	65	α (?)	83.0
8 μM His_6-$\alpha_2\beta_2$	C*(iii)*	10.1	386	?	-
		11.3	227	$\alpha_2\beta_2$	247
		13.7	79	β_2 or α	81.2 or 83.0
23 μM His_6-$\alpha_2\beta_2$	C*(iv)*	10.1	398	?	-
		11.2	255	$\alpha_2\beta_2$	247
		13.1	110	?	-

[a] Molecular weights were calculated from plots of K_{av} versus log(M_W), where K_{av} is the partition function $K_{av} = (V_e - V_c)/(V_o - V_c)$ in which V_e is the retention volume of the species of interest, V_c is the volume of the column (24 mL for a Superdex 200 GL10/300 column), and V_o is the void volume of the column. This analysis was used because $s_{20,w}$ values for each species were not determined, thus the method of Siegel and Monty[21] could not be used.

A2.2. REFERENCES

1. Muntel, J., Fromion, V., Goelzer, A., Maaβ, S., Mäder, U., Büttner, K., Hecker, M., and Becher, D. (2014) Comprehensive Absolute Quantification of the Cytosolic Proteome of *Bacillus subtilis* by Data Independent, Parallel Fragmentation in Liquid Chromatography/Mass Spectrometry (LC/MS^E), *Mol. Cell. Proteomics 13*, 1008-1019.

2. Maaβ, S., Wachlin, G., Bernhardt, J., Eymann, C., Fromion, V., Riedel, K., Becher, D., and Hecker, M. (2014) Highly Precise Quantification of Protein Molecules per Cell During Stress and Starvation Responses in *Bacillus subtilis*, *Mol. Cell. Proteomics 13*, 2260-2276.

3. Maaβ, S., Sievers, S., Zühlke, D., Kuzinski, J., Sappa, P. K., Muntel, J., Hessling, B., Bernhardt, J., Sietmann, R., Völker, U., Hecker, M., and Becher, D. (2011) Efficient, Global-Scale Quantification of Absolute Protein Amounts by Integration of Targeted Mass Spectrometry and Two-Dimensional Gel-Based Proteomics, *Anal. Chem. 83*, 2677-2684.

4. Ortega, A., Amorós, D., and de la Torre, J. G. (2011) Prediction of Hydrodynamic and Other Solution Properties of Rigid Proteins from Atomic- and Residue-Level Models, *Biophys. J. 101*, 892-898.

5. Balbo, A., and Schuck, P. (2005) Analytical Ultracentrifugation in the Study of Protein Self-Association and Heterogeneous Protein-Protein Interactions, In *Protein-Protein Interactions: A Molecular Cloning Manual* (Golemis, E. A., and Adams, P. D., Eds.), pp 253-277, Cold Spring Harbor Laboratory Press, New York.

6. Crona, M., Torrents, E., Røhr, A. K., Hofer, A., Furrer, E., Tomter, A. B., Andersson, K. K., Sahlin, M., and Sjöberg, B. M. (2011) NrdH-Redoxin Protein Mediates High Enzyme Activity in Manganese-Reconstituted Ribonucleotide Reductase from *Bacillus anthracis*, *J. Biol. Chem. 286*, 33053-33060.

7. Gustafsson, T. N., Sahlin, M., Lu, J., Sjöberg, B. M., and Holmgren, A. (2012) *Bacillus anthracis* Thioredoxin Systems, Characterization and Role as Electron Donors for Ribonucleotide Reductase, *J. Biol. Chem. 287*, 39686-39697.

8. Zhang, Y., and Stubbe, J. (2011) *Bacillus subtilis* Class Ib Ribonucleotide Reductase is a Dimanganese(III)-Tyrosyl Radical Enzyme, *Biochemistry 50*, 5615-5623.

9. Parker, M. J., Zhu, X. L., and Stubbe, J. (2014) *Bacillus subtilis* Class Ib Ribonucleotide Reductase: High Activity and Dynamic Subunit Interactions, *Biochemistry 53*, 766-776.

10. Huque, Y., Fieschi, F., Torrents, E., Gibert, I., Eliasson, R., Reichard, P., Sahlin, M., and Sjöberg, B. M. (2000) The Active Form of the R2F Protein of Class Ib Ribonucleotide Reductase from *Corynebacterium ammoniagenes* is a Diferric Protein, *J. Biol. Chem. 275*, 25365-25371.

11. Stolle, P., Barckhausen, O., Oehlmann, W., Knobbe, N., Vogt, C., Pierik, A. J., Cox, N., Schmidt, P. P., Reijerse, E. J., Lubitz, W., and Auling, G. (2010) Homologous Expression of the *nrdF* Gene of *Corynebacterium ammoniagenes* Strain ATCC 6872 Generates a Manganese-metallocofactor (R2F) and a Stable Tyrosyl Radical (Y•) Involved in Ribonucleotide Reduction, *FEBS J. 277*, 4849-4862.

12. Cotruvo, J. A., and Stubbe, J. (2011) *Escherichia coli* Class Ib Ribonucleotide Reductase Contains a Dimanganese(III)-Tyrosyl Radical Cofactor *in vivo*, *Biochemistry 50*, 1672-1681.

13. Cotruvo, J. A., and Stubbe, J. (2010) An Active Dimanganese(III)-Tyrosyl Radical Cofactor in *Escherichia coli* Class Ib Ribonucleotide Reductase, *Biochemistry 49*, 1297-1309.

14. Jordan, A., Pontis, E., Åslund, F., Hellman, U., Gibert, I., and Reichard, P. (1996) The Ribonucleotide Reductase System of *Lactococcus lactis* - Characterization of an NrdEF Enzyme and a New Electron Transport Protein, *J. Biol. Chem. 271*, 8779-8785.

15. Yang, F. D., Curran, S. C., Li, L. S., Avarbock, D., Graf, J. D., Chua, M. M., Lui, G. Z., Salem, J., and Rubin, H. (1997) Characterization of Two Genes Encoding the *Mycobacterium tuberculosis* Ribonucleotide Reductase Small Subunit, *J. Bacteriol. 179*, 6408-6415.

16. Rabinovitch, I., Yanku, M., Yeheskel, A., Cohen, G., Borovok, I., and Aharonowitz, Y. (2010) *Staphylococcus aureus* NrdH Redoxin Is a Reductant of the Class Ib Ribonucleotide Reductase, *J. Bacteriol. 192*, 4963-4972.

17. Roca, I., Torrents, E., Sahlin, M., Gibert, I., and Sjöberg, B. M. (2008) NrdI Essentiality for Class Ib Ribonucleotide Reduction in *Streptococcus pyogenes*, *J. Bacteriol. 190*, 4849-4858.

18. Makhlynets, O., Boal, A. K., Rhodes, D. V., Kitten, T., Rosenzweig, A. C., and Stubbe, J. (2014) *Streptococcus sanguinis* Class Ib Ribonucleotide Reductase: High Activity with Both Iron and Manganese Cofactors and Structural Insights, *J. Biol. Chem. 289*, 6259-6272.

19. Jordan, A., Pontis, E., Atta, M., Krook, M., Gibert, I., Barbé, J., and Reichard, P. (1994) A Second Class I Ribonucleotide Reductase in *Enterobacteriaceae* - Characterization of the *Salmonella typhimurium* Enzyme, *Proc. Natl. Acad. Sci. U. S. A. 91*, 12892-12896.

20. Henriques, A. O., Glaser, P., Piggot, P. J., and Moran, C. P. (1998) Control of Cell Shape and Elongation by the *rodA* Gene in *Bacillus subtilis*, *Mol. Microbiol. 28*, 235-247.

21. Siegel, L. M., and Monty, K. J. (1966) Determination of Molecular Weights and Frictional Ratios of Proteins in Impure Systems by Use of Gel Filtration and Density Gradient Centrifugation. Application to Crude Preparations of Sulfite and Hydroxylamine Reductases, *Biochim. Biophys. Acta 112*, 346-362.

Appendix 3

Supplementary Data for Chapter 3

Table A3.1. Apparent K_m values for effectors binding to class I RNRs from various sources.

Table A3.2. Temperature dependence of effector binding to the *E. coli* class Ia RNR α_2 subunit.

Table A3.3. Apparent K_ds or approximations for effector binding to *P. aeruginosa* Ia and *S.* Typhimurium Ib RNR α subunits.

Table A3.4. Apparent K_m values for substrate binding to class I RNRs from various sources.

Table A3.5. Temperature dependence of substrate binding to the *E. coli* class Ia RNR α_2 subunit.

Table A3.6. Catalytic efficiencies of class I RNRs from different sources.

Table A3.7. Estimated *in vivo* nucleotide concentrations in *B. subtilis* grown on glucose.

Table A3.8. Estimated *in vivo* nucleotide concentrations in *B. subtilis* grown on glucose and malate.

Table A3.9. Estimated *in vivo* nucleotide concentrations in *B. subtilis* grown on glucose and fumarate.

Table A3.10. Estimated *in vivo* nucleotide concentrations in *B. subtilis* grown on glucose and citrate.

Table A3.11. Estimated *in vivo* nucleotide concentrations in *B. subtilis* from various sources.

Table A3.12. Predicted hydrodynamic properties of the tagless *B. subtilis* Ib RNR subunits and complexes.

Table A3.13. NDP reduction activity by the *B. subtilis* Ib RNR stimulated by non-canonical effector nucleotides.

Figure A3.1. Stereo images of the *E. coli* Ia RNR ATP-cone domain depicting how two nucleotides can bind to it.

Supplementary discussion of SV-AUC data fitting and results with NrdE in the presence of dATP.

Table A3.1. Apparent K_m values for effectors binding to class I RNRs from various sources.

Varied nucleotide	Constant nucleotide[a]	Organism	Subclass	Apparent K_m (μM)	Reference
ATP	CDP (50 μM)	*E. coli*	Ia	180	1
ATP	CDP (500 μM)	*E. coli*	Ia	70	1
TTP	CDP (500 μM)	*E. coli*	Ia	1.5	1
ATP	CDP (200 μM)	*M. tuberculosis*	Ib	100	2
ATP	CDP	*S.* Typhimurium	Ib	170	3
ATP	CDP (1 mM)	*S. pyogenes*	Ib	300	4
dATP	CDP (500 μM)	*S.* Typhimurium	Ib	3.0	5
dATP	CDP	*S.* Typhimurium	Ib	2.5	3
dATP	CDP (1 mM)	*S. pyogenes*	Ib	1.1	4
dATP	CDP (500 μM)	*S. sanguinis*	Ib	3.4[b]	6
ATP	CDP	T4 phage	Ia	70	7
ATP	UDP	T4 phage	Ia	60	7
dATP	CDP	T4 phage	Ia	0.2	7
dATP	UDP	T4 phage	Ia	0.4	7
dCTP	CDP	T4 phage	Ia	20	7
dCTP	UDP	T4 phage	Ia	20	7
dGTP	GDP	T4 phage	Ia	0.5	7
dGTP	ADP	T4 phage	Ia	0.7	7
TTP	CDP	T4 phage	Ia	2	7
TTP	UDP	T4 phage	Ia	2	7
TTP	GDP	T4 phage	Ia	4	7
TTP	ADP	T4 phage	Ia	4	7

[a] The numbers in brackets, where shown, are the reported final concentrations of the constant nucleotide.
[b] This apparent K_m is the only previously reported value for Ib RNRs containing the $Mn(III)_2$-Y• cofactor.

Table A3.2. Temperature dependence of effector binding to the *E. coli* class Ia RNR α_2 subunit.

Temperature		2 °C				4 °C		25 °C					
Method[a]		RD/GF[b]		NB[c]		UF[d]		RD/GF[b]		UF[d]		SPR[e]	
E	β_2 present?	K_d (μM)	*n*	K_d (μM)	*n*	K_d (μM)	*n*	K_d (μM)	*n*	K_d (μM)	*n*	K_d (μM)	*n*
ATP	N	~100	~4			13	3.3			80 158* 120*,f	3.7 6.8* 3.8*,f	153 221[g]	NR[i] NR[i]
dATP													
h	N	0.03	1.8	0.06	1 – 2	0.11	1.9	0.3	~2	0.43	1.8	0.15	NR[i]
l	N	0.5	1.85	0.8	3 – 4[h]	1	3.6[h]			6	3.7[h]	19.3 8.3[g]	NR[i] NR[i]
l	Y							1.0	~2				
dGTP	N	0.08	~2			0.21	2			0.77	1.9		
TTP	N	0.3	~2	0.4	1 – 2	0.5	1.7	1 – 5	~2	1.9 4.2*	1.8 2.1*	2.8	NR[i]

[a] RD/GF = rapid dialysis/gel filtration, NB = nitrocellulose binding assay, UF = ultrafiltration, SPR = surface plasmon resonance. [b] Data reproduced from Brown and Reichard, 1969.[8] [c] Data reproduced from Söderman and Reichard, 1986.[9] [d] Data reproduced from Ormö and Sjöberg, 1990,[10] unless marked by * which were taken from C.M. Zimanyi Ph.D. Thesis.[11] [e] Data reproduced from Crona *et al.*, 2010.[12] [f] Assays additionally contained 100 or 500 μM dGTP to occupy the specificity sites. [g] Assays include 1 mM TTP to occupy the specificity sites. [i] Not reported.

Table A3.3. Apparent K_ds or approximations for effector binding to *P. aeruginosa* Ia and *S.* Typhimurium Ib RNR α subunits.

Effector	Temperature (°C)	Apparent K_d (μM)	n^a	Reference
***Pseudomonas aeruginosa* Ia RNR**				
dATP	4	1.51	2.62[b]	[13]
dGTP	4	1.15	~1.3	[13]
***Salmonella* Typhimurium Ib RNR[c]**				
ATP	4	~500	~2	[3]
dATP	4	~1	~1.9	[3]
dGTP	4	~0.12	~2.6	[3]
TTP	4	~3	~1.8	[3]

[a] Number of binding sites (*n*) reported per monomer for *P. aeruginosa* and per dimer for *S.* Typhimurium.
[b] Includes binding sites in two ATP-cone domains encode on the N-terminus of this protein.
[c] Scatchard plots were curved concave up, suggesting positive cooperativity in effector binding. The values reported in the apparent K_d column are concentrations at which binding sites were half occupied.

Table A3.4. Apparent K_m values for substrate binding to class I RNRs from various sources.

Variable nucleotide	Constant nucleotide[a]	Organism	Subclass	Apparent K_m (μM)	Reference
ADP	dGTP (500 μM) + ATP (2 mM)	*B. taurus*	Ia	28 – 50	14
GDP	TTP (500 μM) + ATP (2 mM)	*B. taurus*	Ia	43 – 50	14
CDP	ATP (5 mM)	*B. taurus*	Ia	30 – 32	14
UDP	ATP (5 mM)	*B. taurus*	Ia	100	14
ADP	dGTP (100 μM)	*E. coli*	Ia	30	15
GDP	TTP (20 μM)	*E. coli*	Ia	25	15
CDP	ATP (700 μM)	*E. coli*	Ia	37	16
UDP	ATP (1 – 2 mM)	*E. coli*	Ia	220	1
ADP	dGTP (330 μM)	*M. musculus*	Ia	12	17
GDP	TTP (330 μM)	*M. musculus*	Ia	4.9	17
CDP	ATP (3 mM)	*M. musculus*	Ia	2.0	18
UDP	ATP (3 mM)	*M. musculus*	Ia	6.4	18
ADP	dGTP (7 μM)	T4 phage	Ia	48	7
GDP	TTP (40 μM)	T4 phage	Ia	45	7
CDP	ATP (500 μM)	T4 phage	Ia	56	7
CDP	dATP (2 μM)	T4 phage	Ia	43	7
CDP	TTP (20 μM)	T4 phage	Ia	57	7
UDP	ATP (300 μM)	T4 phage	Ia	100	7
ADP	dGTP (500 μM) + ATP (1.6 mM)	*T. brucei*	Ia	33	19
GDP	TTP (500 μM) + ATP (1.6 mM)	*T. brucei*	Ia	10	19
CDP	ATP (1.6 mM)	*T. brucei*	Ia	3.2	19
CDP	dATP (100 μM)	*T. brucei*	Ia	4.0	19
UDP	ATP (1.6 mM)	*T. brucei*	Ia	72	19
UDP	dATP (100 μM)	*T. brucei*	Ia	9.6	19
CDP	ATP (5 mM)	*P. aeruginosa*	Ia	27	20
CDP	dATP (200 μM)	*S. pyogenes*	Ib	340	4
CDP	dATP (300 μM)	*S.* Typhimurium	Ib	16	5
CDP	dATP	*S.* Typhimurium	Ib	20	3

[a] The numbers in brackets, where shown, are the reported final concentrations of the constant nucleotide.

Table A2.5. Temperature dependence of substrate binding to the *E. coli* class Ia RNR α_2 subunit.

Temperature		0 °C				4 °C		20 °C		25 °C			
Method		ED[d]		RD[d]		UF[e]		ED[d]		UF[e]		SPR[f]	
S	E	K_d (μM)	n	K_d (μM)	n	K_d (μM)	n	K_d (μM)	n	K_d (μM)	n	K_d (μM)	n
ADP	none			38	1.9			420	3.4				
	dATP(β,γ-CH_2)[b]							320	2.5				
	dATP(β,γ-CH_2)[c]							580	2.5				
	dGTP(α-S)			28	1.9			70	1.7				
	TTP(β,γ-CH_2)			27	1.6			130	1.8				
CDP	none	130	1.4					1000	1.6			333	NR[g]
	dATP(β,γ-CH_2)[b]	130	1.5					88	1.2				
	dATP(β,γ-CH_2)[c]	170	1.5										
	TTP											136	NR[g]
	TTP(β,γ-CH_2)	260	1.5					250	1.1				
GDP	none	30	1.5	26	1.4			110	1.3			80	NR[g]
	dATP(β,γ-CH_2)[b]							110	1.2				
	dATP(β,γ-CH_2)[c]	43	1.6					190	1.7				
	dGTP(α-S)	32	1.7	24	1.4			56	1.1				
	TTP					26	2.0			24	2.0	22	NR[g]
	TTP(β,γ-CH_2)	26	1.8	13	1.3			22	1.3				

[a] ED = equilibrium dialysis, RD = rapid dialysis, UF = ultrafiltration, SPR = surface plasmon resonance. [b] Low concentrations that stimulate pyrimidine NDP reduction. [c] High concentrations that inhibit RNR. [d] Data reproduced from von Döbelin and Reichard, 1976 [21]. [e] Data reproduced from Ormö and Sjöberg, 1990 [10]. [f] Data reproduced from Crona *et al.*, 2010 [12]. [g] Not reported.

Table A3.6. Catalytic efficiencies of class I RNRs from different sources.

Substrate	Effector(s)	$k_{cat} K_m^{-1}$ ($M^{-1} s^{-1}$)	Reference
E. coli[a, b]			
ADP	100 μM dGTP	5.6×10^3	*15*
CDP	700 μM ATP	1.3×10^3	*16*
GDP	20 μM TTP	7.0×10^3	*15*
UDP	1 – 2 mM ATP	500	*1*
***Bos taurus* (calf thymus)**[c]			
ADP	500 μM dGTP + 2 mM ATP	1.4×10^5	*14*
CDP	5 mM ATP	2.0×10^5	*14*
GDP	500 μM TTP + 2 mM ATP	1.1×10^5	*14*
UDP	5 mM ATP	4.0×10^4	*14*
Trypanosoma brucei[b]			
ADP	500 μM dGTP + 1.6 mM ATP	5.0×10^3	*19*
CDP	1.6 mM ATP	1.1×10^4	*19*
	100 μM dATP	2.1×10^4	*19*
GDP	500 μM TTP + 1.6 mM ATP	1.0×10^4	*19*
UDP	100 μM dATP	8.0×10^3	*19*
Mus musculus[b]			
ADP	330 μM dGTP	1.5×10^4	*17*
CDP	3 mM ATP	1.5×10^5	*18*
GDP	330 μM TTP	5.7×10^4	*17*
UDP	3 mM ATP	3.9×10^4	*18*

[a] No modern k_{cat}, K_m, or $k_{cat} K_m^{-1}$ values have been reported, so the catalytic efficiencies were estimated from older data collected when *E. coli* Ia RNR activities were not optimized.
[b] Catalytic efficiency reported relative to α_2.
[c] Catalytic efficiency reported relative to $\alpha_2\beta_2$.

In vivo nucleotide concentrations

The estimates of the *in vivo* concentrations of nucleotides in *B. subtilis* reported over the next five pages were pulled from a variety of sources. In most of the reports, the concentrations were given as nmol nucleotide per mg cell dry weight (CDW). To convert these concentrations into estimates in molar units, the following conversion factors were used.

1. CDW was converted into a cell hydrated weight (CHW) assuming 51.5% of the mass of a hydrated cell corresponded to CDW.[22]

2. The CHW was converted to number of cells assuming the mass of a single hydrated *B. subtilis* cell was 150 fg.[23]

The data in **Tables A3.7 – A3.10** were from a report in which the phase-dependent changes in metabolite concentrations were monitored from mid-exponential to mid- or late stationary phase in *B. subtilis* cells grown on glucose alone or with another carbon source.[24] In a separate study, the phase-dependent cell volume changes for *B. subtilis* were reported for a similar time course that monitored *in vivo* protein concentrations.[25] Both studies used the same strain of *B. subtilis* (168 trp^+) with similar genotypes (the bacteria used in the metabolite study[24] had *ctc::lacZ* gene fusion and *gsiB::gfp* reporter inserted in the *amyE* locus) and similar media (M9 supplemented with 0.1% (w/v) glucose and 0.1% malate) which resulted in very similar growth curves between the two studies. Thus, the curves were compared to find similar time points between the two studies and the corresponding cell volumes were used to convert the number of cells calculated in 2 above into volume expressed as liters. For the other reports listed in **Table A3.11** requiring conversion of nucleotide amounts to molar concentrations, cultures were typically grown for short periods of time before harvest, resulting in only one nucleotide concentration determination. A cell volume of 1.31 μm^3 was used in these cases.[25]

Table A3.7. Estimated *in vivo* nucleotide concentrations in *B. subtilis* grown on glucose.[a]

Time point[b]	210 min	270 min	300 min	360 min	420 min
Phase[c]	ME	T	MO	60S	120S
Cell volume (μm³)[c]	1.31	1.15	0.98	0.82	0.86
Nucleotide	**Estimated *in vivo* concentration (μM)**				
ATP	270	360	200	30	30
ADP	100	170	120	60	80
AMP	30	70	60	140	160
CTP	30	40	20	9	20
CDP	10	20	10	10	20
CMP	4	10	10	40	100
GTP	70	100	50	10	30
GDP	20	30	20	20	40
GMP	5	10	10	20	100
UTP	90	120	70	10	20
UDP	40	70	50	20	30
UMP	20	30	30	20	50

[a] Growth at 37 °C with 300 rpm shaking in M9 minimal medium supplemented with 0.1% (w/v) glucose.[24]
[b] Time after inoculation as reported in Meyer *et al.*, 2014.[24]
[c] From Maaß *et al.*, 2014.[25] Abbreviations: ME = mid-exponential, T = transient, MO = maximum optical density, 60S = 60 min into stationary, 120S = 120 min into stationary.

Table A3.8. Estimated *in vivo* nucleotide concentrations in *B. subtilis* grown on glucose and malate.[a]

Time point[b]	150 min	240 min	300 min	390 min	450 min
Phase[c]	ME	T	MO	[d]	[d]
Cell volume (μm³)[c]	1.31	1.15	0.98	0.82	0.86
Nucleotide	**Estimated *in vivo* concentration (μM)**				
ATP	600	580	580	170	150
ADP	100	230	200	170	160
AMP	10	80	80	190	170
CTP	220	300	180	3	7
CDP	60	50	60	30	40
CMP	80	8	70	7	ND[e]
GTP	210	280	290	370	350
GDP	2	10	7	40	60
GMP	230	170	130	70	70
UTP	200	160	120	30	40
UDP	30	60	40	50	40
UMP	10	50	40	30	40

[a] Growth at 37 °C with 300 rpm shaking in M9 minimal medium supplemented with 0.1% (w/v) glucose and 0.1% (w/v) malate.[24]

[b] Time after inoculation as reported in Meyer *et al.*, 2014.[24]

[c] From Maaß *et al.*, 2014.[25] Abbreviations: ME = mid-exponential, T = transient, MO = maximum optical density.

[d] Similar time points were not collected by Maaß *et al.*,[25] but the cell volumes measured by them at 60 min and 120 min into stationary phase were still used as approximations.

[e] Not detected.

Table A3.9. Estimated *in vivo* nucleotide concentrations in *B. subtilis* grown on glucose and fumarate.[a]

Time point[b]	210 min	300 min	330 min	360 min	450 min
Phase[c]	ME	T	MO	[d]	120S
Cell volume (μm³)[c]	1.31	1.15	0.98	0.82	0.86
Nucleotide	**Estimated *in vivo* concentration (μM)**				
ATP	410	70	90	130	290
ADP	130	120	110	130	250
AMP	50	250	130	120	210
CTP	30	10	10	20	40
CDP	5	20	20	20	50
CMP	ND[e]	100	2	ND[e]	20
GTP	130	20	30	40	80
GDP	20	10	20	30	20
GMP	6	30	30	60	40
UTP	130	10	20	30	60
UDP	30	20	20	30	50
UMP	50	90	70	30	90

[a] Growth at 37 °C with 300 rpm shaking in M9 minimal medium supplemented with 0.1% (w/v) glucose and 0.1% (w/v) fumarate.[24]

[b] Time after inoculation as reported in Meyer *et al.*, 2014.[24]

[c] From Maaß *et al.*, 2014.[25] Abbreviations: ME = mid-exponential, T = transient, MO = maximum optical density, 120S = 120 min into stationary.

[d] A similar time point was not collected by Maaß *et al.*,[25] but the cell volumes measured by them at 60 min into stationary phase was still used as an approximation.

[e] Not detected.

Table A3.10. Estimated *in vivo* nucleotide concentrations in *B. subtilis* grown on glucose and citrate.[a]

Time point[b]	210 min	300 min	330 min	390 min	450 min
Phase[c]	ME	T	MO	60S	120S
Cell volume (μm^3)[c]	1.31	1.15	0.98	0.82	0.86
Nucleotide	**Estimated *in vivo* concentration (μM)**				
ATP	380	280	160	30	110
ADP	140	140	100	70	180
AMP	50	70	60	120	230
CTP	30	30	20	10	50
CDP	20	20	10	20	70
CMP	5	10	8	50	130
GTP	100	80	50	20	70
GDP	30	20	20	30	70
GMP	8	10	20	50	130
UTP	120	90	50	20	70
UDP	50	50	30	20	60
UMP	30	30	10	30	60

[a] Growth at 37 °C with 300 rpm shaking in M9 minimal medium supplemented with 0.1% (w/v) glucose and 0.1% (w/v) citrate.[24]

[b] Time after inoculation as reported in Meyer *et al.*, 2014.[24]

[c] From Maaß *et al.*, 2014.[25] Abbreviations: ME = mid-exponential, T = transient, MO = maximum optical density, 60S = 60 min into stationary, 120S = 120 min into stationary.

Table A3.11. Estimated *in vivo* nucleotide concentrations in *B. subtilis* from various sources.

Nucleotide	Strain	Genotype	Medium	Concentration	Ref.
ATP	SB19E	*pyr2 cdd-1 lys*	BisTris minimal medium	200 μM	26
	168	-	BisTris minimal medium	150 μM	27
	SB19E	*pyr2 cdd-1 lys*	BisTris minimal medium	160 μM	27
	QB-917	*trpC2 hisA1 thr-5*	MOPS minimal medium	200 μM	28
	168	*trpC2*	2x SG medium	630 μM	29
	168	*trpC2*	C minimal medium w/ 0.5% (w/v) glucose	80 μM	30
	168	*trpC2*	C minimal medium w/ 0.5% (w/v) malate	80 μM	30
ADP	168	*trpC2*	2x SG medium	550 μM	29
AMP	168	*trpC2*	2x SG medium	990 μM	29
dATP	168	-	BisTris minimal medium	7 μM	27
	SB19E	*pyr2 cdd-1 lys*	BisTris minimal medium	7 μM	27
CTP	SB19E	*pyr2 cdd-1 lys*	BisTris minimal medium	40 μM	26
	168	-	BisTris minimal medium	50 μM	27
	SB19E	*pyr2 cdd-1 lys*	BisTris minimal medium	50 μM	27
	168	*trpC2*	2x SG medium	115 μM	29
CDP	168	*trpC2*	2x SG medium	50 μM	29
CMP	SB19E	*pyr2 cdd-1 lys*	BisTris minimal medium	2 μM	26
	168	*trpC2*	2x SG medium	220 μM	29
dCTP	QB16	*ilvA1 sacA78*	BisTris minimal medium, 37 °C	270 μM	31
	QB16	*ilvA1 sacA78*	BisTris minimal medium, 46 °C	150 μM	31
	SB19E	*pyr2 cdd-1 lys*	BisTris minimal medium	20 μM	26
	168	-	BisTris minimal medium	30 μM	27
	SB19E	*pyr2 cdd-1 lys*	BisTris minimal medium	30 μM	27
dCMP	SB19E	*pyr2 cdd-1 lys*	BisTris minimal medium	0.5 μM	26

GTP	168	-	BisTris minimal medium	90 μM	27
	SB19E	*pyr2 cdd-1 lys*	BisTris minimal medium	90 μM	27
	QB-917	*trpC2 hisA1 thr-5*	MOPS minimal medium	60 μM	28
	168	*trpC2*	2x SG medium	90 μM	29
GDP	168	*trpC2*	2x SG medium	40 μM	29
GMP	168	*trpC2*	2x SG medium	110 μM	29
dGTP	168	-	BisTris minimal medium	4 μM	27
TTP	SB19E	*pyr2 cdd-1 lys*	BisTris minimal medium	15 μM	26
	SB19E	*pyr2 cdd-1 lys*	BisTris minimal medium	4 μM	27
	QB16	*ilvA1 sacA78*	BisTris minimal medium, 37 °C	340 μM	31
	QB16	*ilvA1 sacA78*	BisTris minimal medium, 46 °C	170 μM	31
	168	-	BisTris minimal medium	10 μM	27
	SB19E	*pyr2 cdd-1 lys*	BisTris minimal medium	10 μM	27
TMP	SB19E	*pyr2 cdd-1 lys*	BisTris minimal medium	1 μM	26
UTP	QB16	*ilvA1 sacA78*	BisTris minimal medium, 37 °C	2.5 mM	31
	QB16	*ilvA1 sacA78*	BisTris minimal medium, 46 °C	1.8 mM	31
	SB19E	*pyr2 cdd-1 lys*	BisTris minimal medium	90 μM	26
	168	-	BisTris minimal medium	80 μM	27
	SB19E	*pyr2 cdd-1 lys*	BisTris minimal medium	90 μM	27
dUMP	QB16	*ilvA1 sacA78*	BisTris minimal medium, 37 °C	< 15 μM	31
	QB16	*ilvA1 sacA78*	BisTris minimal medium, 46 °C	< 15 μM	31
	SB19E	*pyr2 cdd-1 lys*	BisTris minimal medium	0.5 μM	26

Table A3.12. Predicted hydrodynamic properties of the tagless *B. subtilis* Ib RNR subunits and complexes.

Protein	R_s (Å)	$s_{20,w}$ (S)[a]	f x 10^{-7} (gs^{-1})[a]	s_{max} (S)[b]	f_o x 10^{-7} (gs^{-1})[b]	f/f_o
α	37.0	5.2	0.693	6.7	0.540	1.28
β	27.3	3.3	0.516	4.0	0.423	1.22
α_2	48.0	8.0	0.906	10.6	0.681	1.33
β_2	34.5	5.2	0.651	6.3	0.533	1.22
Symmetric $\alpha_2\beta_2$[c]	51.2	10.9	0.966	13.6	0.776	1.25
Asymmetric $\alpha_2\beta_2$[d]	54.4	10.3	1.027	13.6	0.776	1.32

[a] Predicted from crystal structures in HYDROPRO using $\rho = 0.998$ g cm^{-3} and $\eta = 0.01$ poise.
[b] Maximum sedimentation and frictional coefficients expected for a perfect sphere of the same molecular weight (**Table 3.2**).
[c] Based on alignment of *B. subtilis* Ib RNR with the *E. coli* Ia docking model.[32]
[d] Based on alignment of *B. subtilis* Ib RNR with the *S.* Typhimurium holo-RNR complex.[33]

Table A3.13. NDP reduction activity by the *B. subtilis* Ib RNR stimulated by non-canonical effector nucleotides.[a]

Substrate		Effector					
		None	ATP	dATP	dCTP	dGTP	TTP
	[S/E]	-	2.9 mM	2.5 μM	50 μM	20 μM	15 μM
ADP	1 mM	1.7 ± 1.5	11 ± 2	8 ± 4	5 ± 1	420 ± 20	40 ± 3
GDP	250 μM	ND[b]	85 ± 5	12 ± 6	30 ± 3	12 ± 4	245 ± 10
CDP	500 μM	13 ± 3	385 ± 20	185 ± 6	20 ± 1	7 ± 6	245 ± 5
UDP	1 mM	18 ± 6	435 ± 12	215 ± 13	30 ± 7	30 ± 10	260 ± 6

[a] Activities are reported in nmol min^{-1} mg^{-1} $\alpha_2\beta_2$.
[b] No activity detected above background.

Activity expressed as a percentage of the maximum for a given substrate

Substrate		Effector					
		None	ATP	dATP	dCTP	dGTP	TTP
	[S/E]	-	2.9 mM	2.5 μM	50 μM	20 μM	15 μM
ADP	1 mM	0.4	3	2	1	100	10
GDP	250 μM	0	35	5	12	5	100
CDP	500 μM	3	100	48	5	2	64
UDP	1 mM	4	100	49	7	7	60

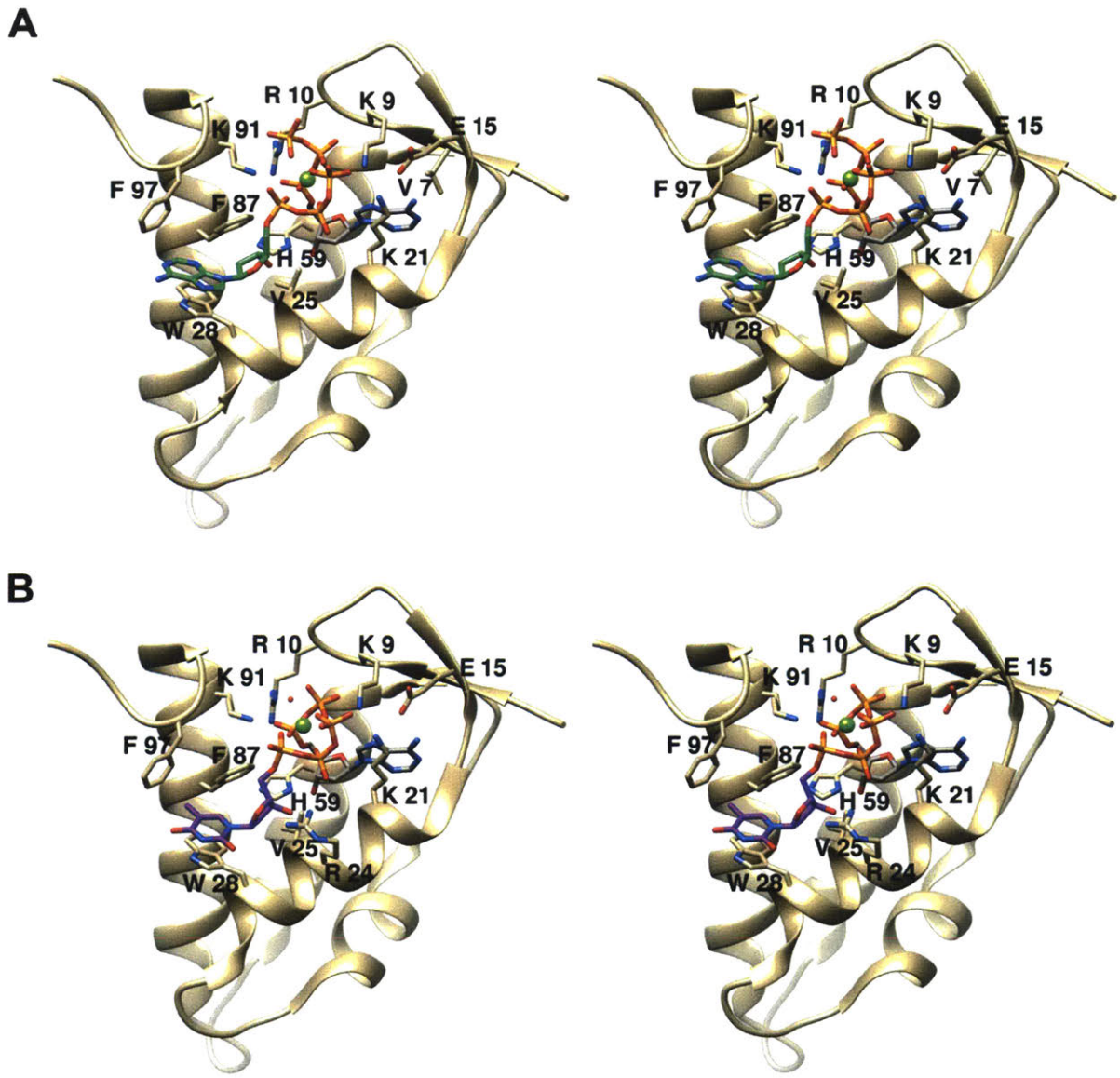


Figure A3.1. Stereo images depicting the way that the ATP-cone domain on the *E. coli* class Ia RNR α subunit can bind two nucleotides. (**A**) ATP-cone containing two dATP molecules (carbons colored gray and green) from PDB 5CNT.[34] (B) ATP-cone containing dATP (gray carbons) and TTP (purple carbons) from PDB 5CNV.[34] The binding of the second nucleotide is mediated by (1) completing the first-sphere coordination shell of a Mg^{2+} ion (green sphere) bound to dATP by the 5′-triphosphate group of the second nucleotide, and (2) placement of the nucleobase into a hydrophobic cleft composted of W_{28}, F_{87}, and F_{97}.

A3.1. SUPPLEMENTARY DISCUSSION

A3.1.1. The buffer mismatch model of Sedfit. Since the nucleotides used in the SV-AUC experiments with tagless NrdE and NrdF have some absorption at 280 nm, it was necessary to switch to using the centrifuge's Rayleigh interference optical detection system to monitor protein sedimentation. This detection system is extremely sensitive to even the smallest of differences in solvent co-solutes, like buffer molecules, salts, osmolytes (i.e. glycerol), reductants (e.g. DTT and TCEP), and nucleotides, between the reference and sample sectors of a cell. If the solutions in the respective sectors are not perfectly matched in terms of geometry (i.e. menisci matched) and composition, co-solute redistribution and/or gradient formation during sedimentation will be detected by the interference optics, leading to signal offsets that introduce artifacts and errors into the fits of the data. Since it is difficult in practice to get exact buffer and menisci matches, a buffer sedimentation correction module[35] has been added to Sedfit to be implemented, if necessary, for fitting data sets collected with the interference optical system.

The buffer sedimentation correction module introduces parameters for buffer mismatch (signal, buoyant molecular weight, and sedimentation coefficient) and menisci mismatch (radial difference, sample:reference sector signal ratio). These parameters can be fixed or floated and fit, depending on knowledge of the source of the differences between the reference and sample sectors. For fitting the NrdE and NrdF data sets collected with interference monitoring, the position of the reference sector meniscus was visually determined and the distance difference between the sample and reference menisci fixed. The other four parameters were floated and fit concomitantly with the c(s) modeling of the protein sedimentation.

A3.1.2. Technicalities of collecting interference and A_{280} data on the analytical ultracentrifuge. The biggest factor that limits the time resolution of monitoring sedimentation by

A_{280} is that the instrument is required to scan the entire length of the solution column in order to record its concentration profile in the reference and sample sectors. The scan speed is limited to ~0.3 mm s^{-1}, and the solution columns typically obtained in the experiments reported for NrdE and NrdF were on the order of 12 mm, thus each cell required ~36 s to scan the entire length. During this time the analyte is still sedimenting, meaning the concentration profile of the sector changes during the course of the scan. In contrast, a laser pulse is used to generate the radiation (λ = 675 nm) required to read the refractive index of the solution column when monitoring sedimentation by interference optics. No scanning is required, thus the refractive index of the entire length of the solution column is read nearly instantaneously.

One other reason why the NrdE dimer peak was missed in monitoring sedimentation by A_{280} was likely that the amount of dimer was too low to be detected above background noise. The instrument requires protein to be present in sufficient amounts such that the optical density is at least 0.1 a.u. whereas the amounts of NrdE dimer detected by interference in samples with a loading concentration of 2 μM or less would have an $A_{280} \approx 0.03$ a.u. (i.e. ~0.4 μM, assuming 3.3 fringes corresponds to a protein concentration of 1 mg mL^{-1}).

A3.1.3. Further discussion of the SV-AUC results for NrdE in the presence of dATP. The observed sedimentation behavior of NrdE in the presence of dATP can be rationalized using effective particle theory.[36] With little to no dATP present, the rate of NrdE dimer dissociation appears to be fast relative to the timescale of sedimentation, thus any NrdE dimer formed will exist for a very short period of time before dissociating into monomers. This leads to the detection of an effective dimeric species with a time-averaged $s_{<w_t>}[20,w]$ that is smaller than expected (i.e. $s_{<w_t>}[20,w]$ = 6.2 S for peak **2** in the absence of nucleotides, **Figure 3.12A**).[36] Binding of dATP stabilizes the dimer and, therefore, slows the rate of dimer dissociation and/or increases the rate of

monomer association relative to sedimentation, resulting in a shift of the transport properties of the system to be more like NrdE dimer. This is manifested in the shifting of the peaks in the $s_{20,w}$ distributions in **Figure 3.12A** to higher values with increasing dATP concentrations and the displacement of the total amount of protein from monomer (peak **1**) to dimer (peak **2**). At 50 μM dATP, the situation is reversed as NrdE dimer represents the most prominent species. Upon dissociation, NrdE monomer is present for fleeting amounts of time before it re-associates and binds dATP, leading to an effective monomeric species with a time-averaged $s_{<Wt>}[20,w]$ that is larger than expected (i.e. $s_{<Wt>}[20,w]$ = 5.9 S for peak **1** in the 50 μM dATP sample, **Figure 3.12A**).[36]

A3.2. REFERENCES

1. Larsson, A., and Reichard, P. (1966) Enzymatic Synthesis of Deoxyribonucleotides. IX. Allosteric Effects in the Reduction of Pyrimidine Ribonucleotides by the Ribonucleoside Diphosphate Reductase System of *Escherichia coli*, *J. Biol. Chem. 241*, 2533-2539.
2. Yang, F. D., Curran, S. C., Li, L. S., Avarbock, D., Graf, J. D., Chua, M. M., Lui, G. Z., Salem, J., and Rubin, H. (1997) Characterization of Two Genes Encoding the *Mycobacterium tuberculosis* Ribonucleotide Reductase Small Subunit, *J. Bacteriol. 179*, 6408-6415.
3. Eliasson, R., Pontis, E., Jordan, A., and Reichard, P. (1996) Allosteric Regulation of the Third Ribonucleotide Reductase (NrdEF Enzyme) from Enterobacteriaceae, *J. Biol. Chem. 271*, 26582-26587.
4. Roca, I., Torrents, E., Sahlin, M., Gibert, I., and Sjöberg, B. M. (2008) NrdI Essentiality for Class Ib Ribonucleotide Reduction in *Streptococcus pyogenes*, *J. Bacteriol. 190*, 4849-4858.
5. Jordan, A., Pontis, E., Atta, M., Krook, M., Gibert, I., Barbé, J., and Reichard, P. (1994) A Second Class I Ribonucleotide Reductase in *Enterobacteriaceae* - Characterization of the *Salmonella typhimurium* Enzyme, *Proc. Natl. Acad. Sci. U. S. A. 91*, 12892-12896.
6. Makhlynets, O., Boal, A. K., Rhodes, D. V., Kitten, T., Rosenzweig, A. C., and Stubbe, J. (2014) *Streptococcus sanguinis* Class Ib Ribonucleotide Reductase: High Activity with Both Iron and Manganese Cofactors and Structural Insights, *J. Biol. Chem. 289*, 6259-6272.
7. Berglund, O. (1972) Ribonucleoside Diphosphate Reductase Induced by Bacteriophage T4. II. Allosteric Regulation of Substrate Specificity and Catalytic Activity, *J. Biol. Chem. 247*, 7276-7281.

8. Brown, N. C., and Reichard, P. (1969) Role of Effector Binding in Allosteric Control of Ribonucleoside Diphosphate Reductase, *J. Mol. Biol. 46*, 39-55.

9. Söderman, K., and Reichard, P. (1986) A Nitrocellulose Filter Binding Assay for Ribonucleotide Reductase, *Anal. Biochem. 152*, 89-93.

10. Ormö, M., and Sjöberg, B. M. (1990) An Ultrafiltration Assay for Nucleotide Binding to Ribonucleotide Reductase, *Anal. Biochem. 189*, 138-141.

11. Zimanyi, C. M., (2013) *Structural Studies of Allosteric Regulation in the Class Ia Ribonucleotide Reductase from Escherichia coli*. Ph.D. Thesis, Massachusetts Institute of Technology.

12. Crona, M., Furrer, E., Torrents, E., Edgell, D. R., and Sjöberg, B. M. (2010) Subunit and Small-Molecule Interaction of Ribonucleotide Reductases Via Surface Plasmon Resonance Biosensor Analyses, *Protein Eng. Des. Sel. 23*, 633-641.

13. Jonna, V. R., Crona, M., Rofougaran, R., Lundin, D., Johansson, S., Brännström, K., Sjöberg, B. M., and Hofer, A. (2015) Diversity in Overall Activity Regulation of Ribonucleotide Reductase, *J. Biol. Chem. 290*, 17339-17348.

14. Eriksson, S., Thelander, L., and Åkerman, M. (1979) Allosteric Regulation of Calf Thymus Ribonucleoside Diphosphate Reductase, *Biochemistry 18*, 2948-2952.

15. Larsson, A., and Reichard, P. (1966) Enzymatic Synthesis of Deoxyribonucleotides. X. Reduction of Purine Ribonucleotides; Allosteric Behavior and Substrate Specificity of the Enzyme System from *Escherichia coli* B, *J. Biol. Chem. 241*, 2540-2549.

16. Holmgren, A., Reichard, P., and Thelander, L. (1965) Enzymatic Synthesis of Deoxyribonucleotides, VIII. Effects of ATP and dATP in the CDP Reductase System from *E. coli*, *Proc. Natl. Acad. Sci. U. S. A. 54*, 830-836.

17. Scott, C. P., Kashlan, O. B., Lear, J. D., and Cooperman, B. S. (2001) A Quantitative Model for Allosteric Control of Purine Reduction by Murine Ribonucleotide Reductase, *Biochemistry 40*, 1651-1661.

18. Kashlan, O. B., Scott, C. P., Lear, J. D., and Cooperman, B. S. (2002) A Comprehensive Model for the Allosteric Regulation of Mammalian Ribonucleotide Reductase. Functional Consequences of ATP- and dATP-Induced Oligomerization of the Large Subunit, *Biochemistry 41*, 462-474.

19. Hofer, A., Ekanem, J. T., and Thelander, L. (1998) Allosteric Regulation of *Trypanosoma brucei* Ribonucleotide Reductase Studied *in vitro* and *in vivo*, *J. Biol. Chem. 273*, 34098-34104.

20. Torrents, E., Westman, M., Sahlin, M., and Sjöberg, B. M. (2006) Ribonucleotide Reductase Modularity - Atypical Duplication of the ATP-cone Domain in *Pseudomonas aeruginosa*, *J. Biol. Chem. 281*, 25287-25296.

21. von Döbeln, U., and Reichard, P. (1976) Binding of Substrates to *Escherichia coli* Ribonucleotide Reductase, *J. Biol. Chem. 251*, 3616-3622.

22. Bratbak, G., and Dundas, I. (1984) Bacterial Dry Matter Content and Biomass Estimations, *Appl. Environ. Microbiol. 48*, 755-757.

23. Burg, T. P., Godin, M., Knudsen, S. M., Shen, W., Carlson, G., Foster, J. S., Babcock, K., and Manalis, S. R. (2007) Weighing of Biomolecules, Single Cells and Single Nanoparticles in Fluid, *Nature 446*, 1066-1069.

24. Meyer, H., Weidmann, H., Mäder, U., Hecker, M., Völker, U., and Lalk, M. (2014) A Time Resolved Metabolomics Study: The Influence of Different Carbon Sources During Growth and Starvation of *Bacillus subtilis*, *Mol. Biosyst. 10*, 1812-1823.

25. Maaβ, S., Wachlin, G., Bernhardt, J., Eymann, C., Fromion, V., Riedel, K., Becher, D., and Hecker, M. (2014) Highly Precise Quantification of Protein Molecules per Cell During Stress and Starvation Responses in *Bacillus subtilis*, *Mol. Cell. Proteomics 13*, 2260-2276.

26. Møllgaard, H., and Neuhard, J. (1978) Deoxycytidylate Deaminase from *Bacillus subtilis* - Purification, Characterization, and Physiological Function, *J. Biol. Chem. 253*, 3536-3542.

27. Møllgaard, H. (1980) Deoxyadenosine/Deoxycytidine Kinase from *Bacillus subtilis*: Purification, Characterization, and Physiological Function, *J. Biol. Chem. 255*, 8216-8220.

28. Saxild, H. H., and Nygaard, P. (1991) Regulation of Levels of Purine Biosynthetic Enzymes in *Bacillus subtilis* - Effects of Changing Purine Nucleotide Pools, *J. Gen. Microbiol. 137*, 2387-2394.

29. Soga, T., Ohashi, Y., Ueno, Y., Naraoka, H., Tomita, M., and Nishioka, T. (2003) Quantitative Metabolome Analysis Using Capillary Electrophoresis Mass Spectrometry, *J. Proteome Res. 2*, 488-494.

30. Meyer, F. M., and Stülke, J. (2013) Malate Metabolism in *Bacillus subtilis*: Distinct Roles for Three Classes of Malate-Oxidizing Enzymes, *FEMS Microbiol. Lett. 339*, 17-22.

31. Neuhard, J., Price, A. R., Schack, L., and Thomassen, E. (1978) Two Thymidylate Synthetases in *Bacillus subtilis*, *Proc. Natl. Acad. Sci. U. S. A. 75*, 1194-1198.

32. Uhlin, U., and Eklund, H. (1994) Structure of Ribonucleotide Reductase Protein R1, *Nature 370*, 533-539.

33. Uppsten, M., Färnegårdh, M., Jordan, A., Eliasson, R., Eklund, H., and Uhlin, U. (2003) Structure of the Large Subunit of Class Ib Ribonucleotide Reductase from *Salmonella typhimurium* and Its Complexes with Allosteric Effectors, *J. Mol. Biol. 330*, 87-97.

34. Zimanyi, C. M., Chen, P. Y. T., Kang, G., Funk, M. A., and Drennan, C. L. (2016) Molecular Basis for Allosteric Specificity Regulation in Class Ia Ribonucleotide Reductase from *Escherichia coli*, *eLife 5*, 23.

35. Zhao, H. Y., Brown, P. H., Balbo, A., Fernandez-Alonso, M. D., Polishchuck, N., Chaudhry, C., Mayer, M. L., Ghirlando, R., and Schuck, P. (2010) Accounting for Solvent Signal Offsets in the Analysis of Interferometric Sedimentation Velocity Data, *Macromol. Biosci. 10*, 736-745.

36. Schuck, P. (2010) Sedimentation Patterns of Rapidly Reversible Protein Interactions, *Biophys. J. 98*, 2005-2013.

Appendix 4

Supplemental Data for Chapter 4

^{1}H-NMR Spectra of Nucleotides Isolated from NrdE

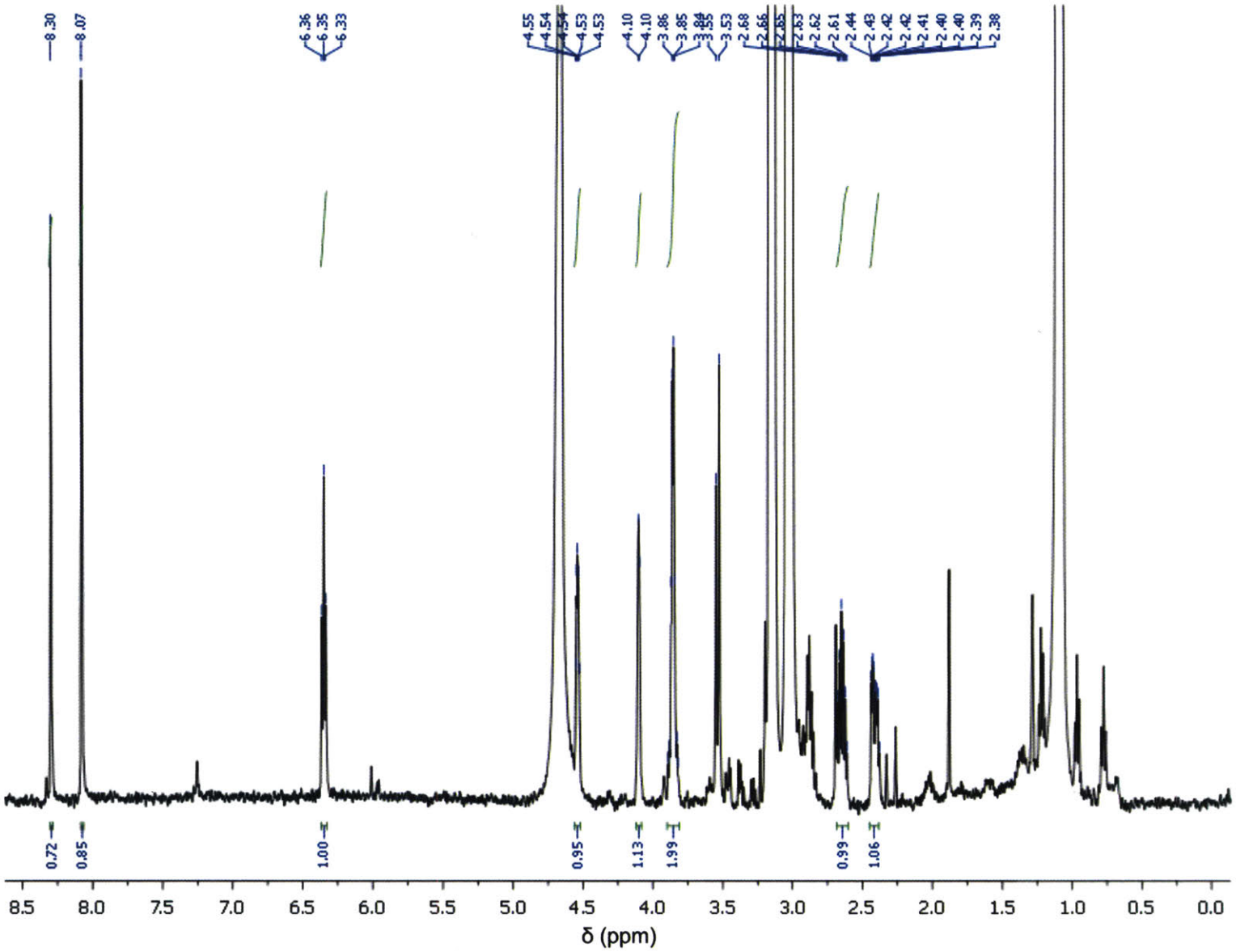


Figure A4.1. Full ^{1}H-NMR spectrum of nucleoside monophosphate isolated from NrdE and eluting off of the DEAE A25 column at ~230 mM TEABC. Peak chemical shifts are shown at the top of the spectrum and the integration at the bottom.

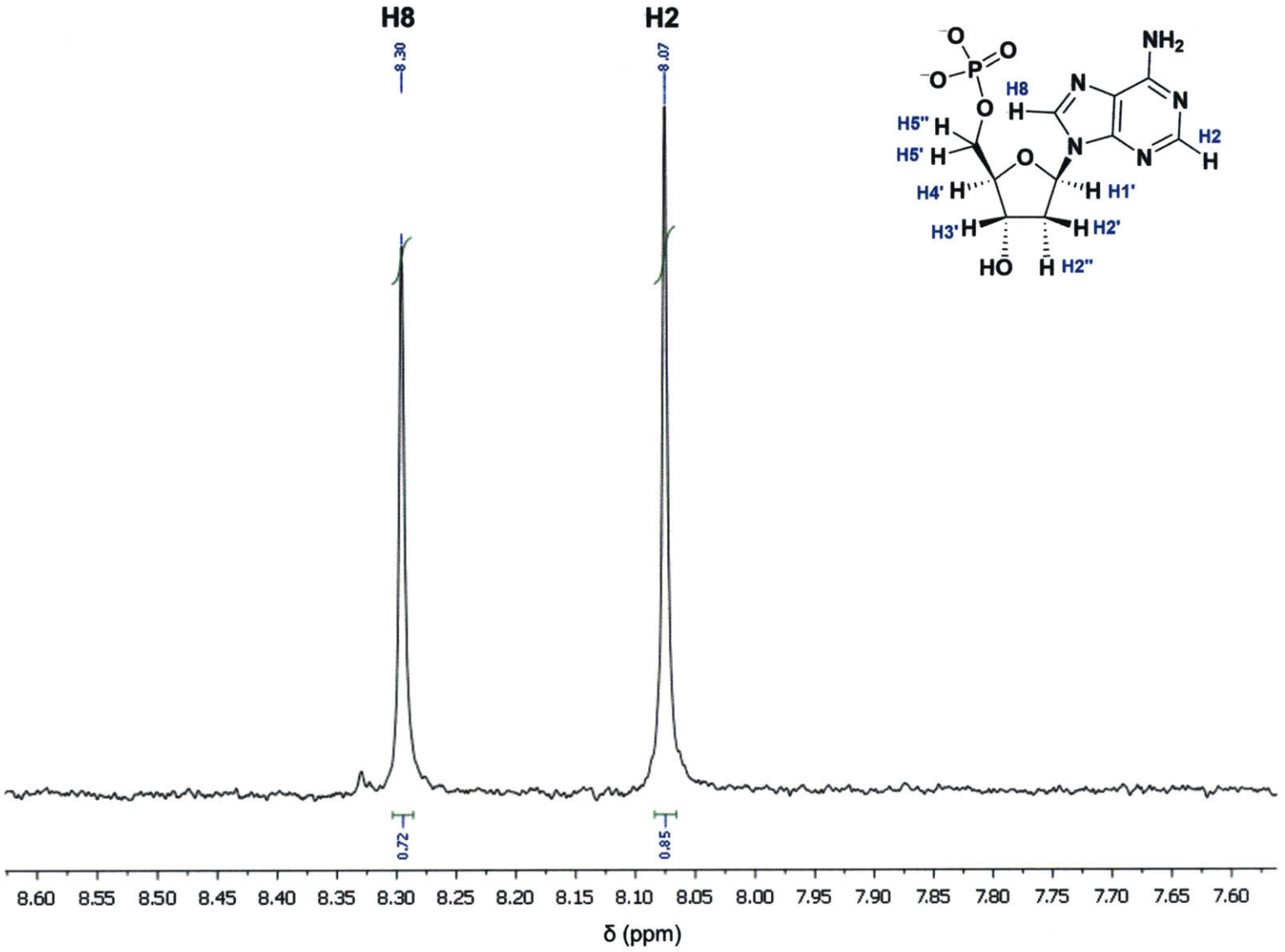


Figure A4.2. Aromatic region of the ^{1}H-NMR spectrum of the nucleoside monophosphate isolated from NrdE.

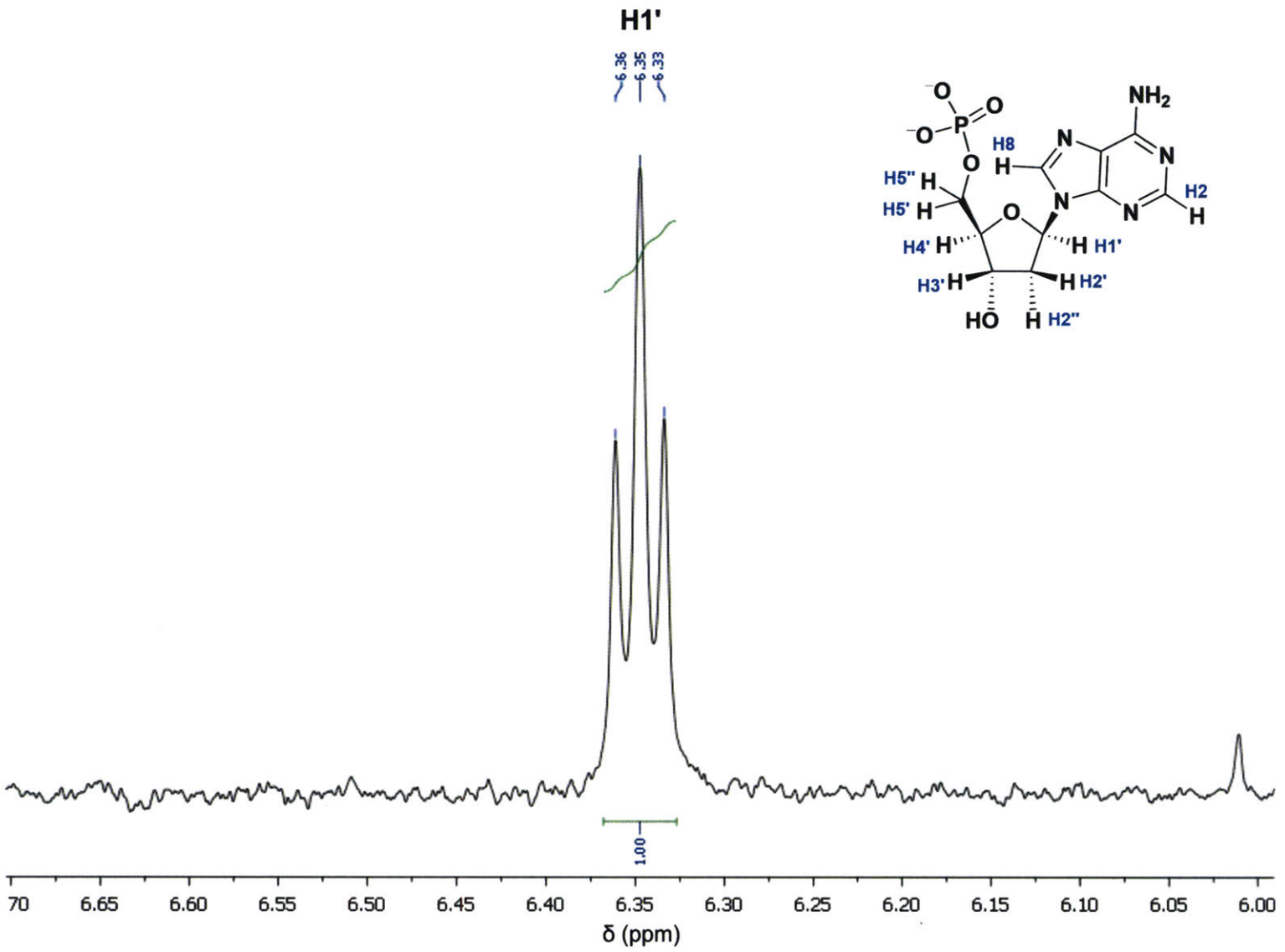


Figure A4.3. Blow-up of the ^{1}H-NMR spectrum of the nucleoside monophosphate isolated from NrdE in the region of H1′.

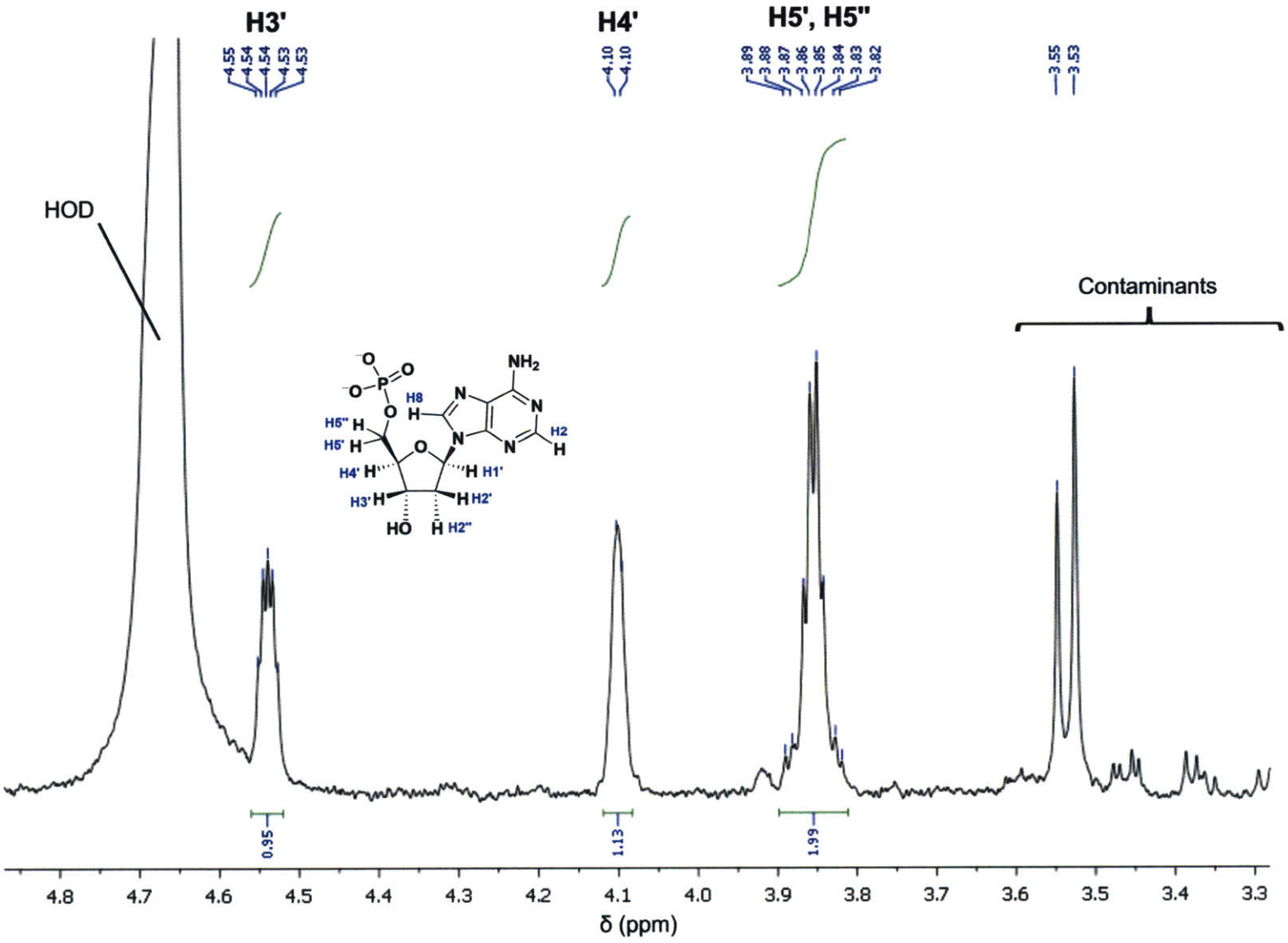


Figure A4.4. Blow up of the 1H-NMR spectrum of the nucleoside monophosphate isolated from NrdE in the region of the 3′, 4′, and 5′ hydrogen atoms.

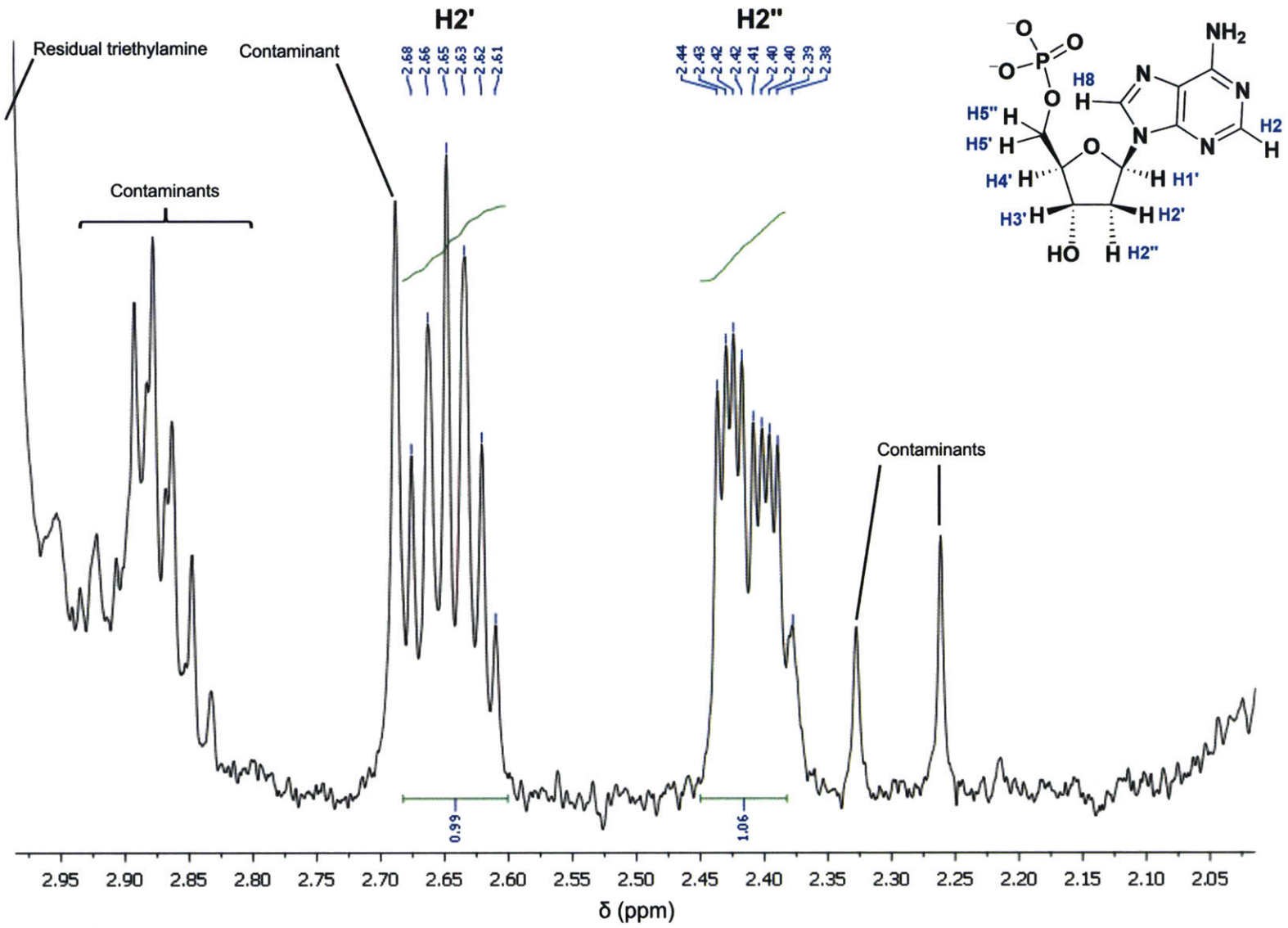


Figure A4.5. Blow up of the ^{1}H-NMR spectrum of the nucleoside monophosphate isolated from NrdE in the region of the 2′ hydrogen atoms.

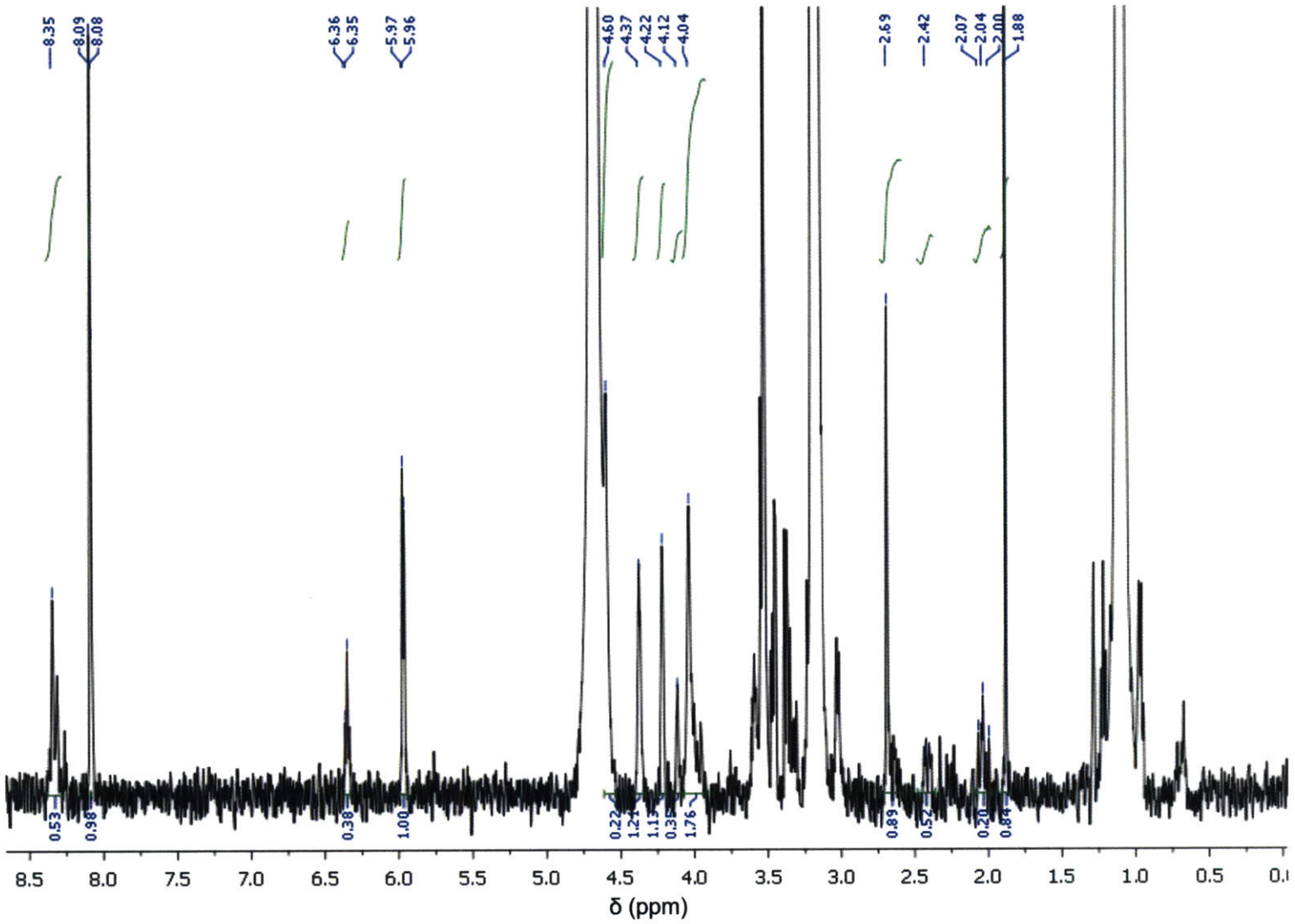


Figure A4.6. Full ^{1}H-NMR spectrum of the compound isolated from NrdE and eluting off of the DEAE A25 column at ~390 mM TEABC. Peak chemical shifts are shown at the top of the spectrum and the integration at the bottom.

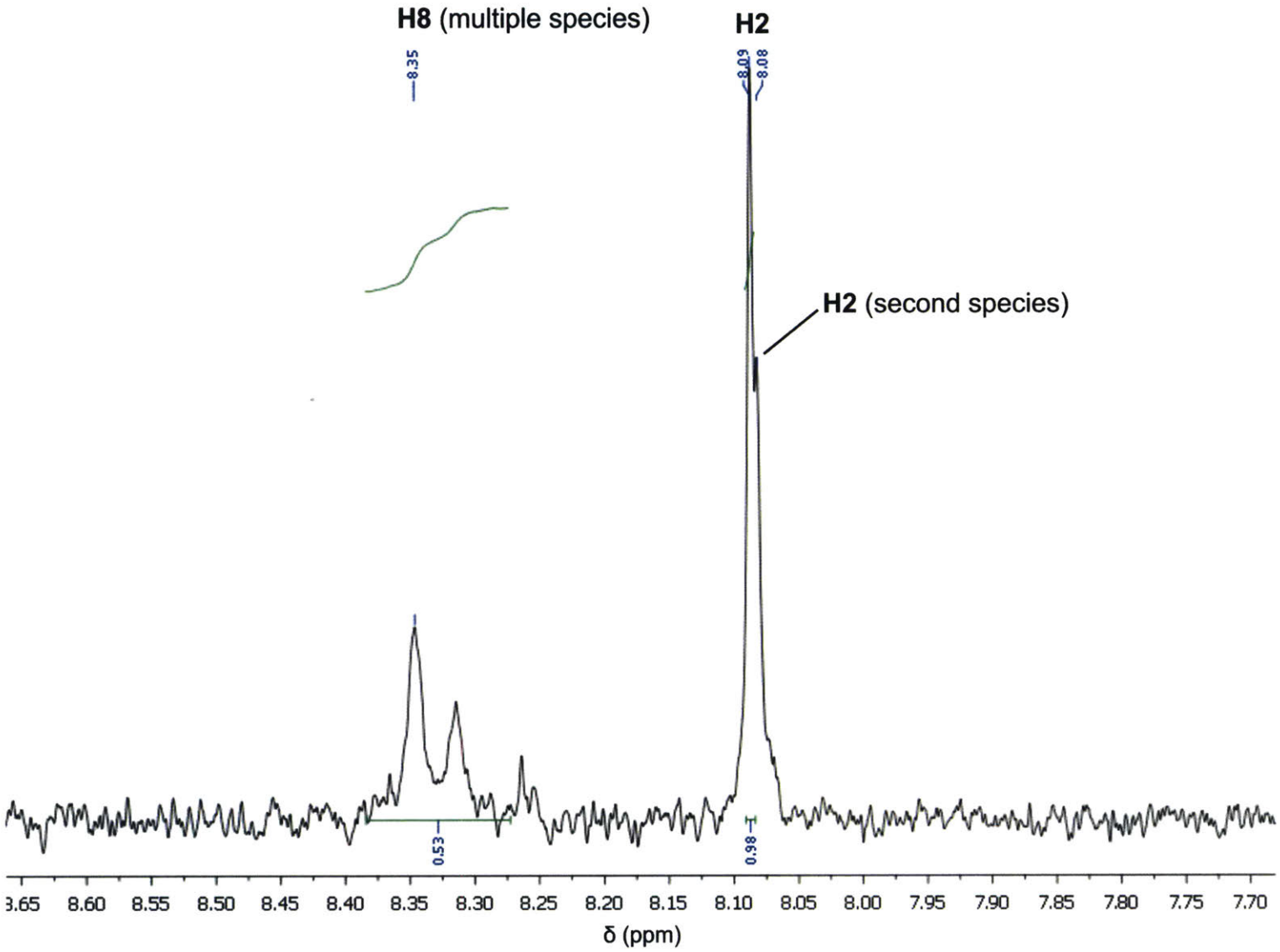


Figure A4.7. Aromatic region of the ^{1}H-NMR spectrum of the compound isolated from NrdE that eluted off of the DEAE A25 column at ~390 mM TEABC.

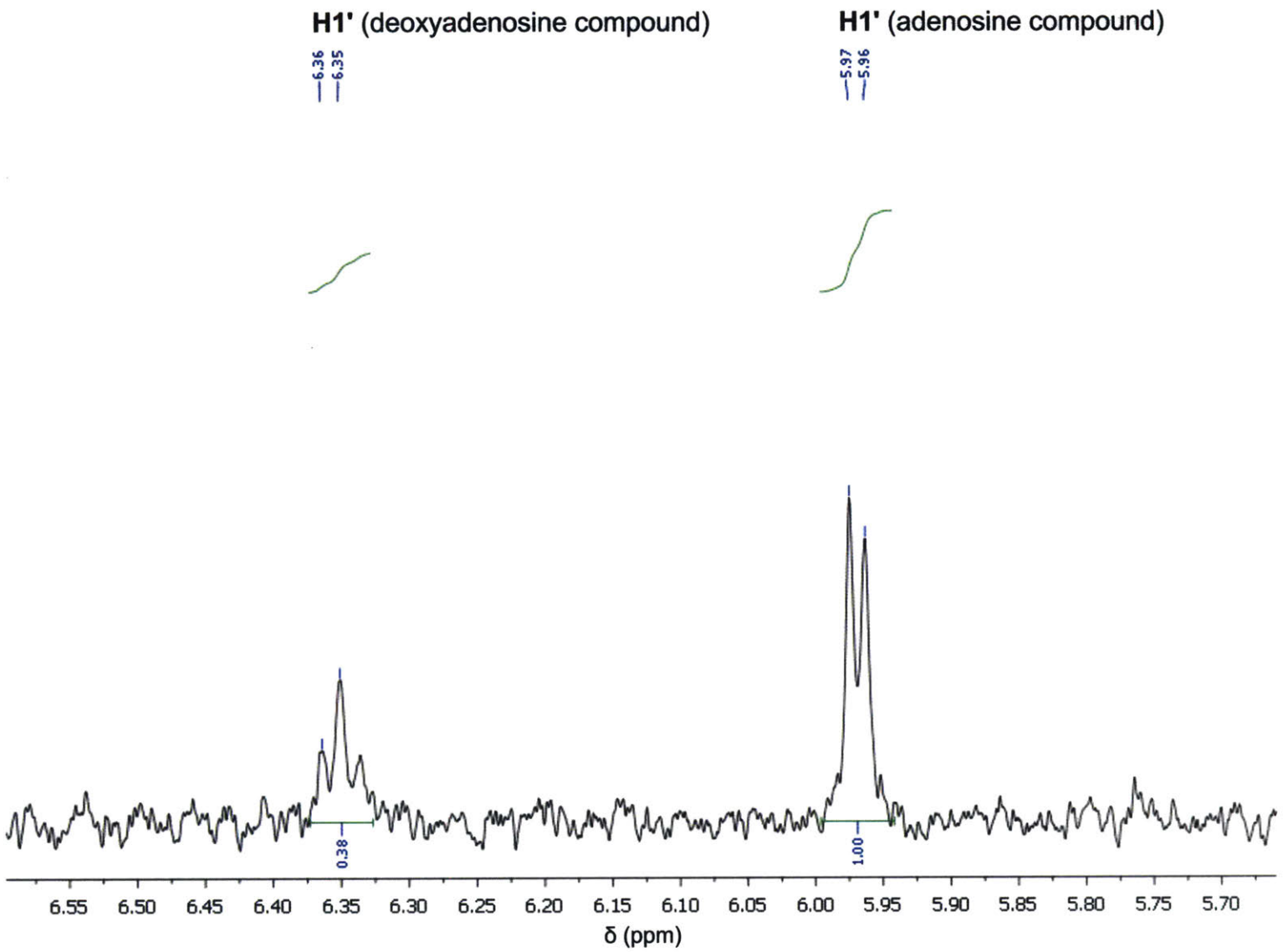


Figure A4.8. Blow up of the ^{1}H-NMR spectrum of the compound isolated from NrdE that eluted at ~390 mM TEABC from the DEAE A25 column. The region shows the 1′ hydrogen atom signals from an adenosine and deoxyadenosine analog.

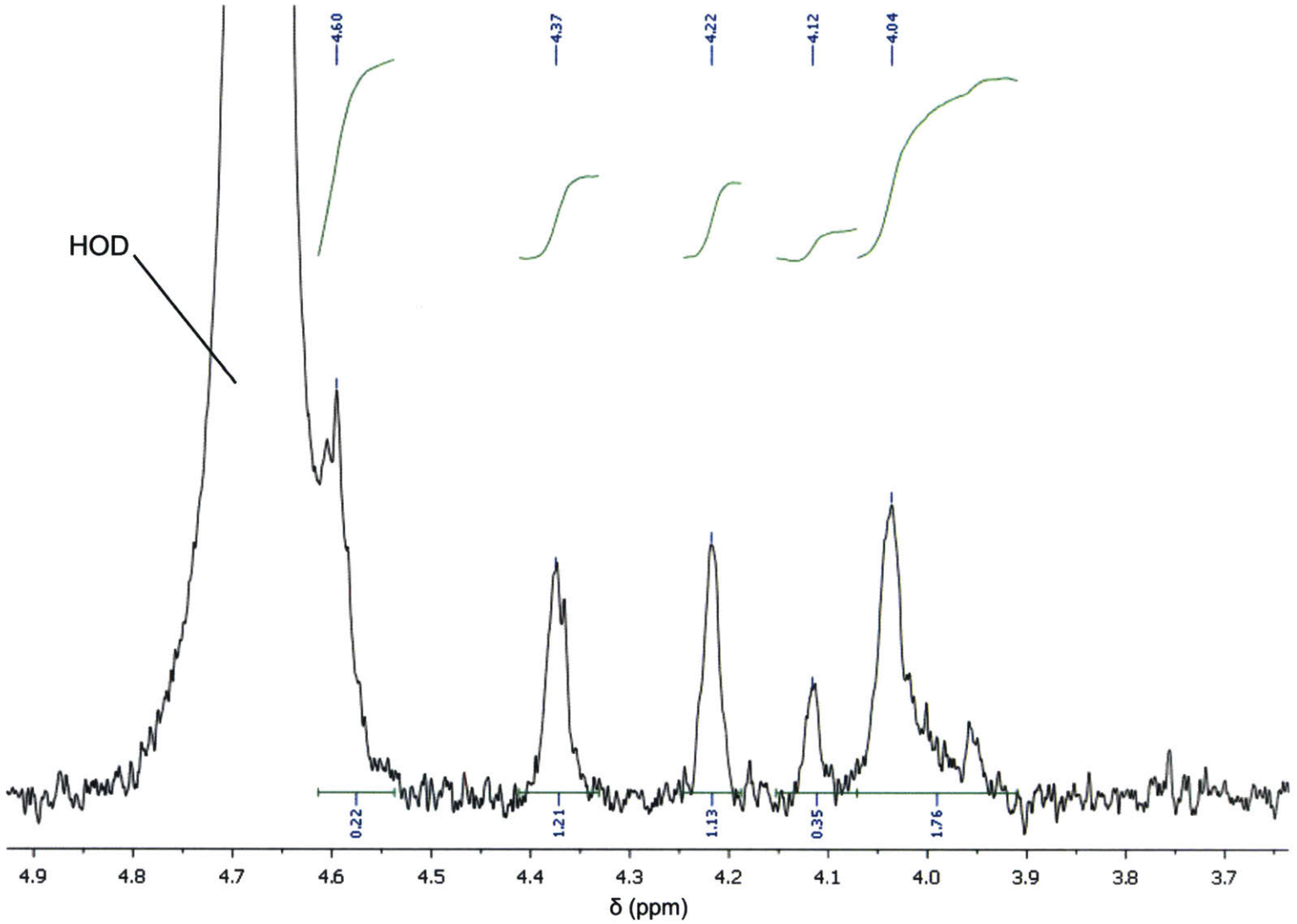


Figure A4.9. Blow up of the 1H-NMR spectrum in the region of the 3′, 4′, and 5′ hydrogen atom signals for the compound isolated from NrdE that eluted at ~390 mM TEABC.

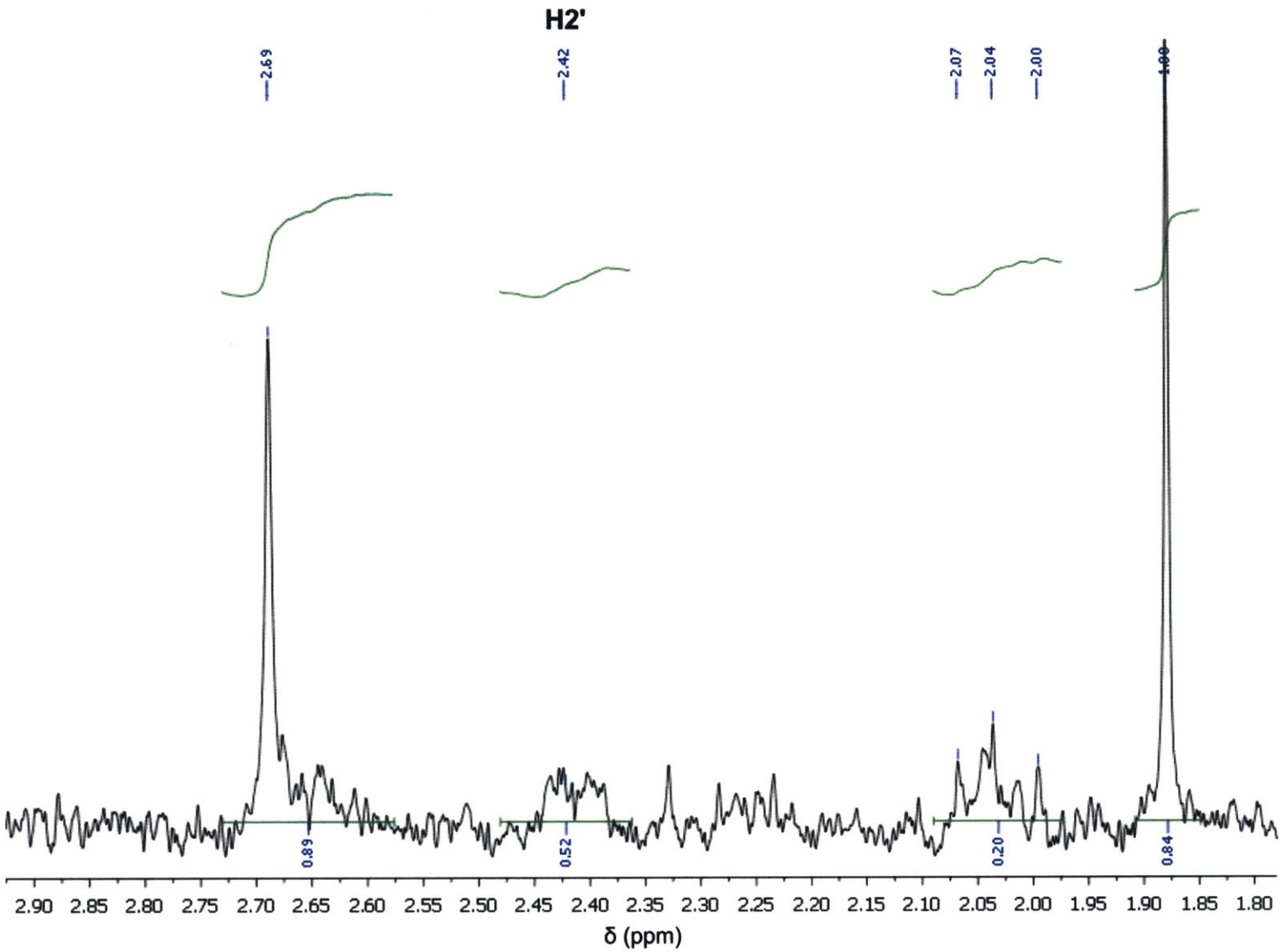


Figure A4.10. Blow up of the ^{1}H-NMR spectrum in the region of the 2′ hydrogens for the compound isolated from NrdE that eluted at ~390 mM TEABC.

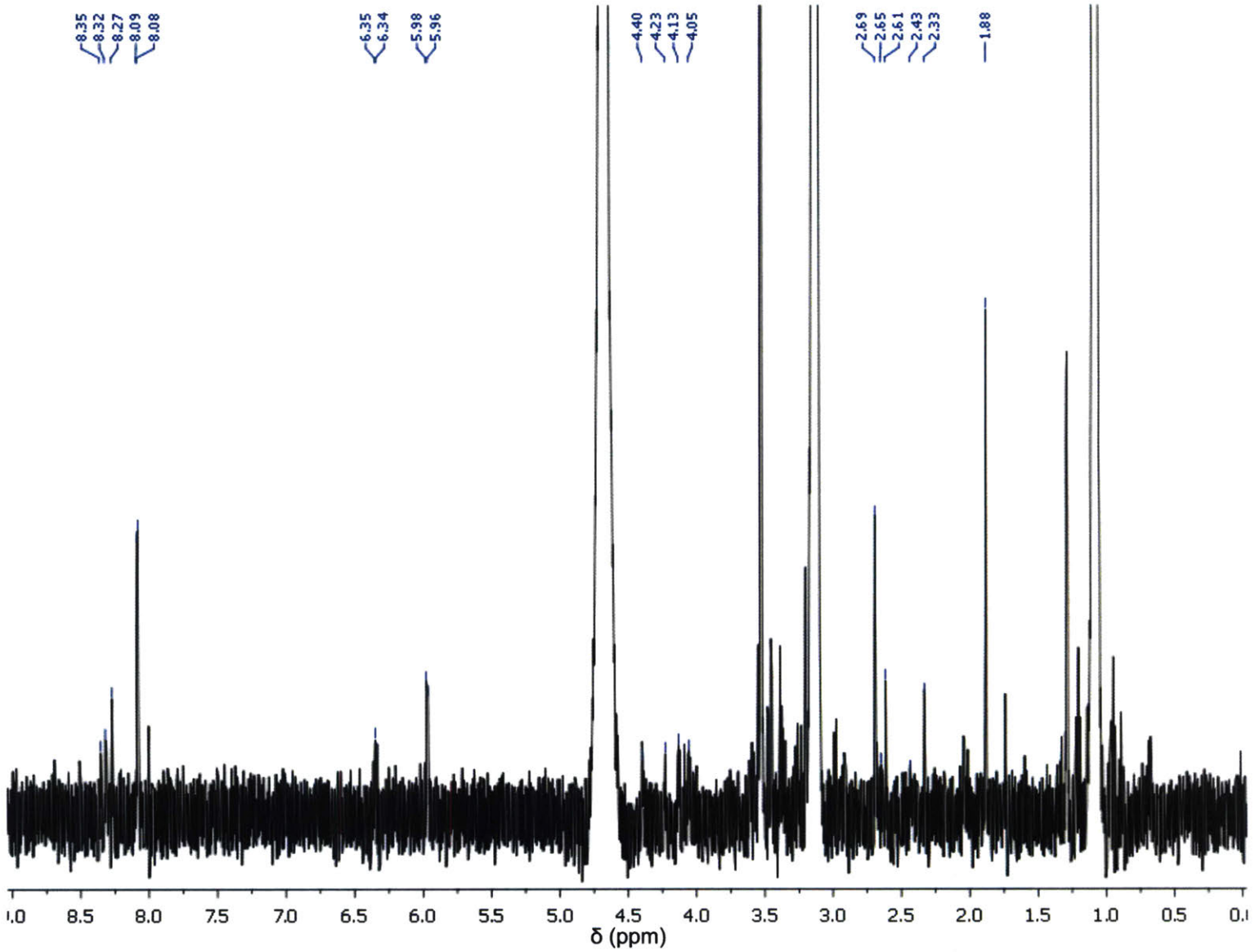


Figure A4.11. Full ^{1}H-NMR spectrum of the compound isolated from NrdE that eluted from the DEAE A25 column at ~550 mM TEABC. Peak chemical shifts are shown at the top of the spectrum.

Appendix 5

Miscellaneous Studies with the *B. subtilis* Class Ib RNR

A5.1. INTRODUCTION

Reported in this appendix are the results of two additional experiments carried out with the *B. subtilis* Ib RNR over the course of this thesis that were preliminary in nature and may be helpful in future studies of this and other class Ib RNRs. These include the cloning, expression, and purification of a ferredoxin/flavodoxin reductase called YumC and the α subunit, NrdEB, of the second, phage-associated class Ib RNR in *B. subtilis*. YumC was assessed for its ability to support the reconstitution of NrdF with the $Mn(III)_2$-Y• cluster by recycling NrdI via NADPH oxidation. NrdEB was tested for dATP inhibition and for the association of a tightly bound equivalent of dAMP. Each experiment will be presented individually with its own introduction, experimental, and results/discussion sections.

A5.2. Investigation of the role of the ferredoxin/flavodoxin reductase YumC in assembly of the $Mn(III)_2$-Y• cofactor in NrdF.

A5.2.1. Introduction.

As discussed in Chapter 2, one issue hampering studies of the class Ib RNRs is the inefficiency of the *in vitro* assembly of the $Mn(III)_2$-Y• cofactor in NrdF, resulting in enzyme preparations that are heterogeneous mixtures of apo- and holo-protein.[1] The ability to separate holo-*B. subtilis* NrdF from apo-protein using anion exchange chromatography is a rather serendipitous result as unpublished work from the Stubbe group on reconstituted NrdFs from *E. coli*, *S.* Typhimurium, and *S. sanguinis* have not had met with similar success (Cotruvo, Johansson, Makhlynets, Parker, Stubbe). Among other issues, the inability to recycle NrdI during reconstitution experiments has been attributed to the inefficiency of cluster assembly. If the endogenous NrdI reductant(s) can be identified and characterized, then it is possible that the

efficiency of cluster assembly in NrdFs could be improved and, therefore, facilitate further studies of the Ib RNRs in general.

In *E. coli* and *S.* Typhimurium, it is suspected that their generic flavodoxin reductases (Fpr) are the likely reductants for NrdI.[2] In contrast, a potential candidate for the NrdI reductase in many other bacteria, including *B. subtilis*, has remained unidentified in the years since the original discovery of the ability to reconstitute NrdF with $Mn(III)_2$-Y• *in vitro*.[3] Recently, however, the results of a genetic study of *B. subtilis* and *L. lactis* have tentatively established a link between ribonucleotide reduction and the ferredoxin(flavodoxin) reductase YumC.[4] YumC, also known as TrxB2 in some organisms, is a 36 kDa protein that adopts a thioredoxin reductase (TRR) – like tertiary structure and binds one equivalent of FAD, but does not have the CxxC motif that is characteristic of true TRRs.[5, 6] The protein is thought to be the equivalent of a generic ferredoxin/flavodoxin reductase in Gram-positive bacteria since in *B. subtilis*, the enzyme has been shown to be essential for cell viability,[7] to be able to reduce ferredoxin[6] and the generic flavodoxins,[8] and has been shown or inferred to support fatty acid desaturation,[9] NO biosynthesis,[8] and biotin biosynthesis.[10-13]

In the study by Chen *et al.*,[4] mutants of *L. lactis* in which *trxB2* had been inactivated grew fine under anaerobic conditions (the organism encodes a class III RNR[14]), but were unable to survive under aerobic conditions unless complemented with TrxB2 or *B. subtilis* YumC supplied *in trans*. A two-week adaptive evolution experiment using the *ΔtrxB2* mutants yielded two fast growing *L. lactis* suppressor strains, and genome sequencing revealed both had single nucleotide polymorphisms (SNPs) in the *nrdI* gene that resulted in the mutation of residues close to the FMN binding site.[15] Subsequent aerobic cultivation of the *L. lactis ΔtrxB2* mutant in the presence of exogenously supplied deoxynucleosides showed a restoration of growth to levels similar to

wildtype. The results of these experiments led to the conclusion that TrxB2 is likely acting as the NrdI reductase in *L. lactis* and is, therefore, critical to the production of dNTPs.

In contrast, conditional *yumC* knockout mutants of *B. subtilis* could be complemented by *L. lactis* TrxB2, but growth of the mutant was not restored by addition of deoxynucleosides to growth media despite the bacterium having the capability of taking up and converting the deoxynucleosides to the corresponding dNTPs.[4, 16] The result indicates that YumC has critical functions in *B. subtilis* in addition to or separate from dNTP production. To determine if YumC could support dNTP production in *B. subtilis* by acting as the reductant for NrdI, the gene was cloned and the protein overexpressed, purified, and tested in a preliminary experiment for the ability to initiate assembly of the $Mn(III)_2$-Y• cluster in NrdF in the presence of NADPH and NrdI. The results of this experiment showed that YumC catalyzed significant levels of O_2 reduction and, therefore, decoupled NADPH oxidation from NrdI reduction and cluster assembly.

A5.2.2. Experimental

A5.2.2A. General methodology. Chemicals, equipment, and routine methods were the same as described in previous chapters. Tagless NrdF and NrdI were prepared as described in Chapter 3. All YumC concentrations are expressed in terms of dimer.[6]

A5.2.2B. Cloning, expression, and purification of tagless YumC. *yumC* was amplified form JH624 genomic DNA with touchdown PCR[17] using the primers 5′-AGGAGGTCTCTAGGT **ATG**CGAGAGGATACAAAGGTTTATGATATTAC-3′ and 5′-GAGAACCTCGAG**TTA**TTTA TTTTCAAAAAGACTTGTTGAGTGAAGAGGC-3′ to introduce a 5′-*BsaI* and a 3′-*XhoI* restriction site (underlined, start and stop codons indicated in boldfaced font) into the amplicon. The amplicon was digested with BsaI and XhoI and ligated into similarly digested pE-SUMO. Sequencing (Quintara Biosciences) confirmed the successful cloning of *yumC*.

pE-SUMO-*yumC* was transformed into BL21 (DE3) cells and selected for on LB agar plates containing 100 μg mL^{-1} Kan. A one-liter culture of the strain was grown following the standard growth protocol described in Chapter 2. Briefly, 5 mL of a saturated overnight starter culture was diluted into 1 L of Kan-supplemented LB and grown at 37 °C with 150 rpm shaking to an OD_{600} of 0.6. The culture was then removed to a cold room while the shaker was cooled to 25 °C (~10 min). IPTG was added to the culture to a final concentration of 0.1 mM, and the culture was subsequently replaced in the shaker and incubated at 25 °C for 4 h prior to harvest (3500 x g, 15 min, 4 °C). A total of 4.6 g cell paste was recovered.

All steps of the purification procedure were completed at 4 °C. The cell paste (4.6 g) was resuspended in 23 mL of phosphate buffer (50 mM sodium phosphate, pH 7.6, 300 mM NaCl, 5% (w/v) glycerol) supplemented with 300 mM NaCl, one EDTA-free protease inhibitor cocktail tablet, 0.25 mM PMSF, and ~1 mg of lysozyme (from chicken egg whites, Sigma). The suspension was homogenized prior to lysing the cells with one passage through a French pressure cell at 14000 psi. The cell debris was pelleted (25000 x g, 15 min, 4 °C) and the supernatant frozen and stored at -80 °C. Later, the extract (13 mg mL^{-1}, 23 mL) was thawed and loaded onto a Ni-NTA column (3.5 x 1 cm, ~2.8 mL). The column was washed with 100 mL phosphate buffer supplemented with 10 mM imidazole and then eluted with the same buffer supplemented with 250 mM imidazole. Protein containing fractions were pooled, exchanged into storage buffer (50 mM sodium phosphate, pH 7.6, 150 mM NaCl), and concentrated with 10000 MWCO centrifugal filter devices (Amicon). A total of 2.8 mg His_6-Smt3-YumC was recovered (>95% pure as judged by SDS-PAGE). Protein concentrations were estimated with the BCA assay using BSA as a standard.[18]

The His_6-Smt3 tag was removed using a 250:1 molar ratio of His_6-Smt3-YumC:SUMO protease. Prior to protease addition, DTT was added to the sample of substrate to a final

concentration of 5 mM. The sample was incubated at 4 °C on a nutating platform for 3 h before being frozen and stored at -80 °C. Later, the sample was thawed and loaded onto the Ni-NTA column described above. Tagless YumC was eluted from the column using 30 mL of storage buffer supplemented with 20 mM imidazole, and fifteen 2 mL fractions were collected. Protein containing fractions were pooled, exchanged into storage buffer, and concentrated. A total of 2.2 mg tagless YumC was recovered (~70% pure as judged by SDS-PAGE). Protein concentration was estimated using $\varepsilon_{457} = 12300\ M^{-1}\ cm^{-1}$.[6]

A5.2.2C. Attempt to reconstitute $Mn(III)_2$-Y• in NrdF using YumC, NrdI, and NADPH. The reconstitution reaction (500 μL) contained 50 mM HEPES, pH 7.6, 5% (w/v) glycerol, 50 μM apo-β_2, 200 μM $MnCl_2$•$4H_2O$, 200 μM NADPH, 1 μM NrdI, and 1 μM YumC. Buffer, NrdF, and $MnCl_2$ were first combined in a 1.7 mL microcentrifuge tube and incubated at 25 °C for 20 min. NADPH and NrdI were then added prior to transferring the solution to a quartz cuvette (Starna) and incubating the reaction for 2 min at 25 °C. YumC was then added and the reaction was followed by A_{340} for 20 min before recovering assay solution from the cuvette and desalting it on a Sephadex G25 column (28 x 1.5 cm, 49 mL) equilibrated in 50 mM HEPES, pH 7.6, 5% (w/v) glycerol. Protein containing fractions were pooled, concentrated, and analyzed by EPR at 77 K using the following acquisition parameters: power = 1 mW, modulation amplitude = 1.5 G, gain = 2.52 x 10^4, time constant = 10.24 ms. A control reaction in which NrdI was omitted was run following the same procedure up to the desalting step.

A5.2.3. Results and Discussion.

His_6-Smt3-YumC was overexpressed and purified in high yield and purity following standard procedures. The purified protein was yellow in color and its UV-visible spectrum showed absorption features characteristic of flavins (**Figure A5.1**), indicating that the cofactor could still

be installed into protein recombinantly expressed in *E. coli*. Previous studies of YumC isolated from *B. subtlis*[5, 6] demonstrated that FAD was bound to the protein. Incubation of His_6-Smt3-YumC with SUMO protease at a 250:1 molar ratio for 3 h resulted in ~50% cleavage of the tag from the protein (**Figure A5.2**). Although the FAD loading of the protein was not established, the low extent of cleavage is reminiscent of that observed with His_6-Smt3-NrdI in Chapter 3 and, thus, suggests that flavinless YumC may sequester the tag and make it more difficult to remove.

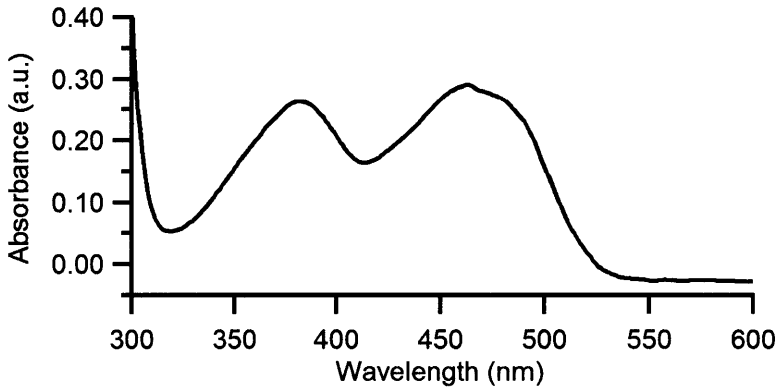


Figure A5.1. UV-visible spectrum of His_6-Smt3-YumC in the range of 300 – 600 nm, showing the absorption features characteristic of flavin cofactors (local λ_{max} = 382 nm, 463 nm). For comparison, local λ_{max} of 378 and 457 nm were reported previously for endogenously sourced YumC.[6]

Previously published data regarding *in vivo* protein concentrations in *B. subtilis* and the NrdI-NrdF interaction were taken into consideration in designing the experiment testing the ability of YumC to support $Mn(III)_2$-Y• formation in NrdF. Quantitative Western blotting experiments indicated NrdI was present in substoichiometric amounts (~10%) relative to NrdF, therefore indicating the flavodoxin-like protein acted catalytically during cluster assembly.[19] Furthermore, both NrdI and YumC react non-productively with O_2, thereby decoupling the electron flow during cluster assembly.[6, 20] As such, a relatively low concentration of 1 μM was chosen for these two proteins to minimize the decoupling reaction and meet the requirement of substoichiometric

amounts of NrdI relative to NrdF. A final concentration of 50 μM apo-β_2 was arbitrarily chosen, and the protein was incubated with 4 equiv Mn^{2+} for 20 min prior to adding NrdI, YumC, and NADPH to allow metal binding to NrdF to approach equilibrium. Given a K_d = 0.6 μM for the NrdI-NrdF interaction,[20] ~99% of the NrdI should be complexed under these conditions.

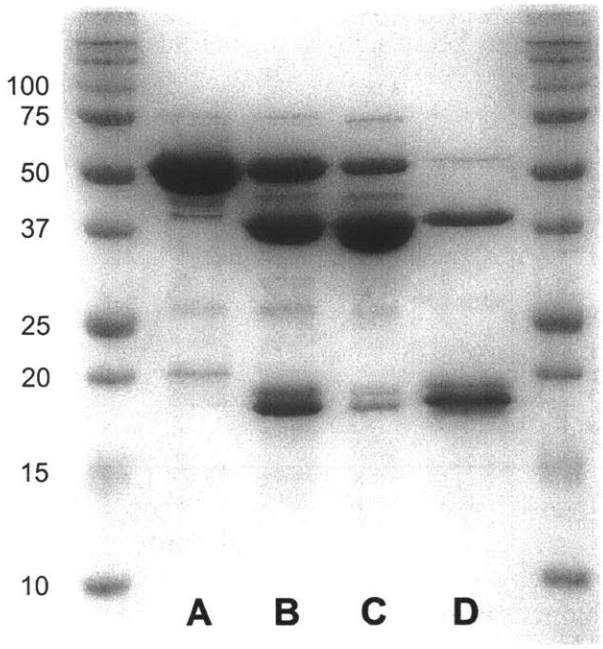


Figure A5.2. Preparation of tagless YumC (36.8 kDa). A SDS-PAGE gel (12.5% (w/v)) showing the different stages of tag removal: *lane A* = His_6-Smt3-YumC (49.2 kDa) prior to digestion, *lane B* = digest product after 3 h incubation at 4 °C with 250:1 ratio of substrate:SUMO protease, *lane C* = tagless YumC after removal of His_6-tagged components by subtractive IMAC, *lane D* = His_6-tagged components removed by subtractive IMAC.

Figure A5.3 shows a comparison of the NADPH consumption, as monitored by A_{340}, observed during the control (NrdI omitted) and full reconstitution experiments. NADPH consumption was observed only when YumC had been added to the assay, and, as shown in **Figure A5.3C**, the rate of its consumption was the same regardless of the presence of NrdI. This result indicates YumC was mostly catalyzing non-productive O_2 reduction under the conditions of the experiment and was, therefore, decoupled from NrdI reduction. From the linear portion of the

curve, turnover numbers of ~0.33 s^{-1} were calculated for both assays, which are comparable to those reported previously for the NADPH oxidase activity of YumC (0.13 s^{-1}).[6]

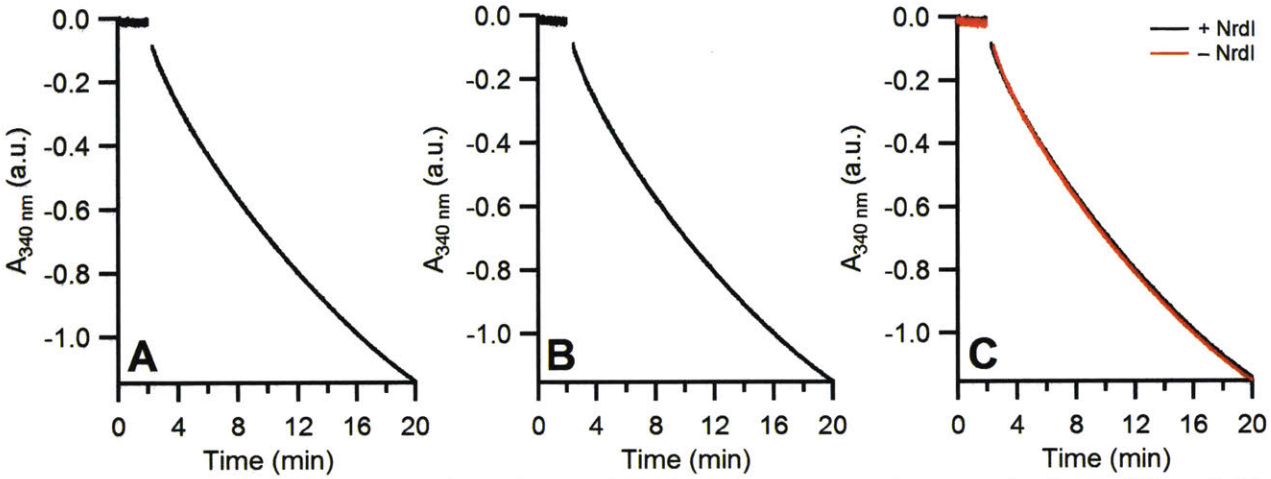


Figure A5.3. NADPH consumption observed during the attempted reconstitution of *B. subtilis* NrdF with $Mn(III)_2$-Y• using YumC and NrdI. (**A**) Apo-β_2 (50 μM) in 50 mM HEPES, pH 7.6, 5% (w/v) glycerol was incubated at 25 °C for 20 min with 4 equiv $MnCl_2$ prior to sequentially adding NADPH and NrdI (1 μM), incubating for 2 min, and initiating reconstitution by addition of YumC (1 μM). (**B**) Same conditions as in (**A**) except that NrdI was omitted from the reaction. (**C**) Overlay of the kinetic traces of the complete and control reactions.

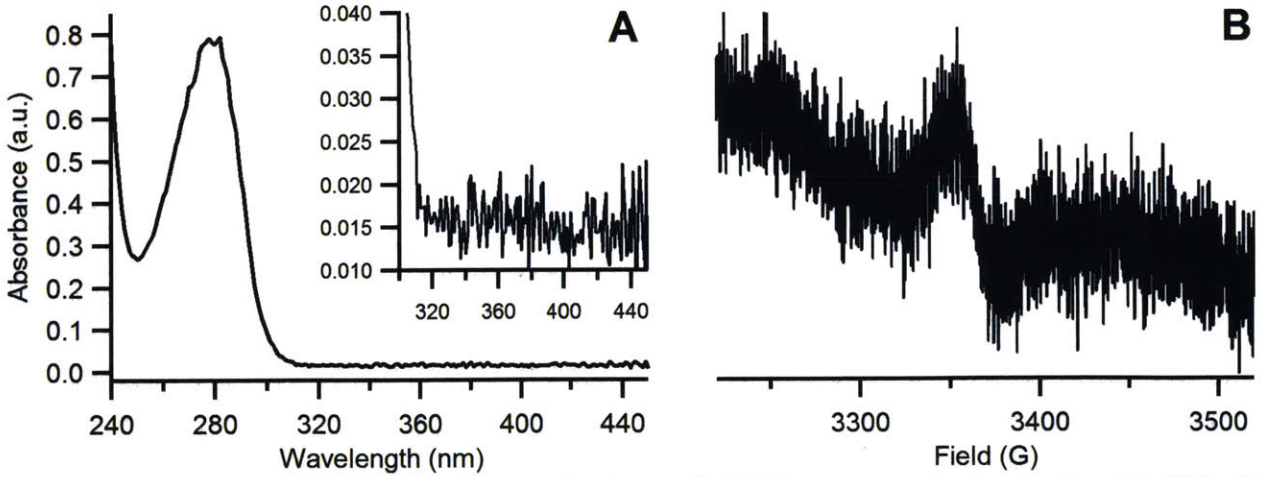


Figure A5.4. Spectroscopic characterization of NrdF recovered from the $Mn(III)_2$-Y• reconstitution experiment with NrdI, YumC, and NADPH. (**A**) UV-visible spectrum of the product in the range of 240 – 450 nm. *Inset:* blow up of the 300 – 450 nm region where Y• has characteristic absorption features. They are clearly not present in this spectrum, indicating low cluster reconstitution efficiency. (**B**) EPR spectrum (77 K) of the product. The radical signal was too low to accurately quantitate the amount of Y• generated.

To see if any $Mn(III)_2$-Y• had been reconstituted in NrdF, the protein from the complete reaction was desalted and analyzed by UV-vis and EPR spectroscopy (**Figure A5.4**). The UV-visible spectrum of the recovered protein only exhibited significant protein absorption at 280 nm (**Figure A5.4A**). Blowing up the spectrum in the range of 300 – 450 nm revealed no absorption features above background, indicating that if any Y• had formed, its concentration was too low to be observed by UV-vis. Additionally, no absorption characteristics of flavins were observed in the region as well. The acquired EPR spectrum of the product at 77 K (**Figure A5.4B**) revealed a tiny amount of signal that looked somewhat similar to that observed in fully reconstituted protein (see **Figure 2.5** for comparison). However, due to the low intensity of the signal, it could not be definitively established that the source of the signal was from reconstituted $Mn(III)_2$-Y• NrdF.

The results of the attempted reconstitution of NrdF with $Mn(III)_2$-Y• using NrdI and YumC are encouraging, but indicate that further studies are required in order to optimize the endogenous reconstitution system. As shown in **Figure A5.2**, the YumC preparation used in this experiment was not optimal as a significant amount of His_6-Smt3-tagged protein was recovered from the subtractive IMAC step. This could be a problem if the His_6-Smt3-tag prevents NrdI and YumC from interacting and, thus, it will be necessary to optimize the conditions for its removal. The interactions required for electron transfer between proteins are often transient in nature, therefore another issue with this experiment may have been using concentrations of NrdF and NrdI at which the two were almost completely complexed with one another. This may have prevented the necessary electron transfer step between YumC and NrdI from occurring efficiently, thus giving the former more time to react with O_2 non-productively. Successful reconstitution conditions will need to balance the concentrations of NrdF, NrdI, and YumC to maximize the efficiency of the electron transfer through the chain, therefore a K_d for the YumC-NrdI interaction would also be

beneficial. If successful conditions can be found, the ability to reconstitute NrdF with $Mn(III)_2$-Y• should aid in further studies of the interesting class Ib RNR from *B. subtilis*.

During the preparation of this dissertation, a report was published in which the YumC homolog from *B. cereus* was used to successfully reconstitute $Mn(III)_2$-Y• in NrdF.[21] A total of 0.6 Y•/β_2 was recovered under optimal conditions, which corresponds to the highest yield of cluster reconstitution ever reported for the *B. cereus* system.[22, 23] In general, the reported experimental design validated the improvements discussed in the previous paragraph for attempting to reconstitute *B. subtilis* NrdF using NrdI, YumC, and NADPH. The largest difference between the experiment reported in this appendix and the study of the *B. cereus* system was that a 1:1 ratio of NrdI:NrdF (80 μM of each protein) was used, whereas the concentration of the YumC homolog was present in a much smaller amount (0.5 μM). The reported K_ds for the interaction between *B. cereus* NrdF and NrdI and between YumC and NrdI are 30 μM,[22] and 20 μM, respectively.[21] Therefore, under optimal conditions, 55% and 0.5% of the total NrdI was complexed with NrdF and YumC, respectively, whereas 80% of the total YumC was complexed with NrdI.[21] The successful reconstitution of *B. cereus* NrdF under these conditions suggests that keeping YumC largely complexed with NrdI may minimize uncoupled O_2 reduction by the former, a result that should be confirmed with the *B. subtilis* system. The results of the reconstitution of *B. cereus* NrdF with $Mn(III)_2$-Y• using YumC highlight the promise of identifying and using the endogenous metallocofactor biosynthetic systems to generate high quality preparations of not only RNRs, but of any enzyme that incorporations such cofactors. Further studies aimed at improving the efficiency and yield of holo-NrdF should greatly aid in the study of the class Ib RNRs.

A5.3. Characterization of NrdEB, the α subunit of the SPβ bacteriophage Ib RNR encoded in the *B. subtilis* genome.

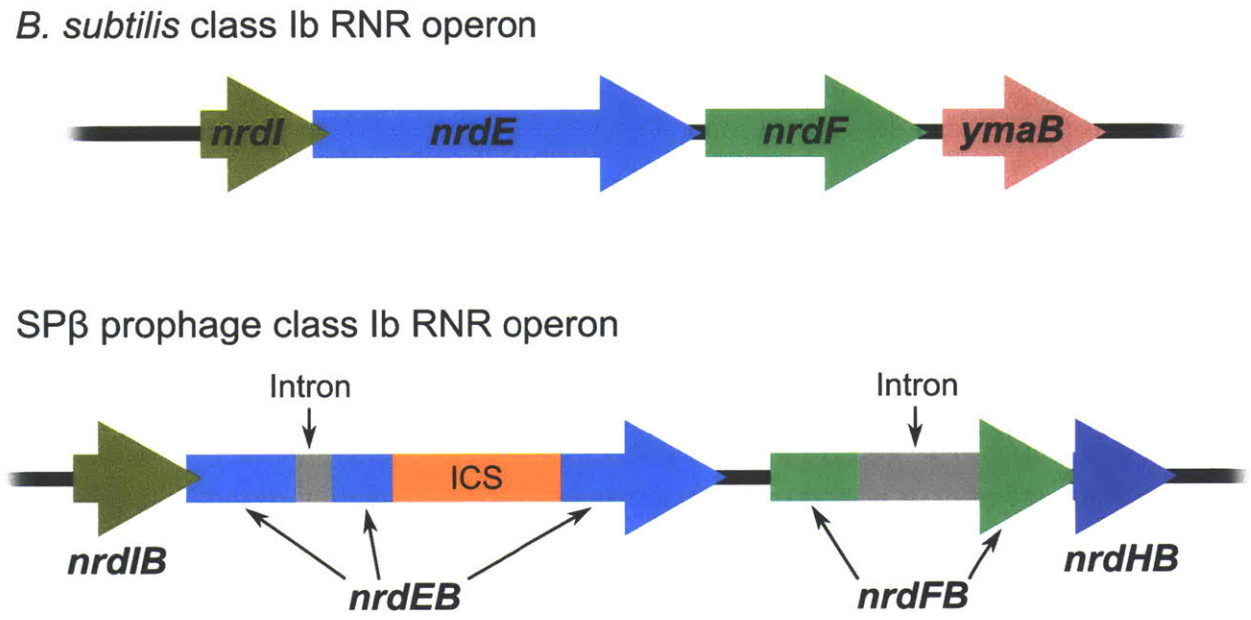


Figure A5.5. The class Ib RNR operons encoded in the genome of *B. subtilis*. ICS – intein coding sequence.

A5.3.1. Introduction

As described in Chapter 1, many strains of *B. subtilis* are highly unusual in that they encode two class Ib RNR operons, one of which is viral in origin and not essential for cell viability (**Figure A5.5**). An interesting observation, however, is the fact that the subunits and NrdI from both systems are extremely homologous to one another. After removal of the intron and intein coding sequences from the SPβ genes, a pairwise BLAST alignment reveals the following homology information: NrdIB vs NrdI – 43% identical (63% with positive substitution), NrdEB vs NrdE – 80% identical (90% with positive substitution), and NrdFB vs NrdF – 87% identical (94% with positive substitution). It is very peculiar that such high homology exists between a viral RNR and its host's system. Since it was expected that the two RNRs would have very similar structural and functional attributes, the SPβ RNR was initially not pursued. However, the subsequent discovery

that the *B. subtilis* RNR was inhibited by dATP and contained a dAMP molecule tightly associated with its α subunit prompted an investigation of the allosteric properties of the SPβ system.

The motivation for studying the allosteric regulation of the SPβ RNR was based on the precedent from previously characterized viral enzymes. As mentioned in Chapter 1, viral RNRs have been reported to show allosteric regulation that is distinct from the host's RNR, often as a result of a lack of overall activity regulation by dATP.[24-29] If the SPβ enzyme behaves similarly to other viral RNRs in this regard, high sequence homology between the *B. subtilis* and SPβ RNRs could provide an opportunity to identify plausible sites for dATP and/or dAMP binding on *B. subtilis* NrdE. A Clustal alignment, using default parameters, of the two α subunits is shown in **Figure A5.6**. NrdEB is shorter than NrdE by one residue and an additional 42 positions are not conserved between the two proteins. Mapping the location of these non-conserved positions plus those with weak conservation on the PHYRE model of *B. subtilis* NrdE from Chapter 2 revealed that none of the critical residues for RNR function and/or structure were affected (**Figure A5.7**). Instead, the vast majority of the differences between NrdE and NrdEB mapped to pockets on the surface of the protein that could be plausible binding sites for small molecules.

The clustering of the differences between NrdE and NrdEB to plausible pockets on the surface of the protein suggests that NrdEB, if it behaves differently, may provide critical insight into the allosteric regulation of the host's RNR. Therefore, *nrdEB* was cloned and the protein overexpressed and purified following standard procedures. It was found that NrdEB was active in NDP reduction using holo-$Mn(III)_2$-Y• NrdF and, therefore, the allosteric regulation of the subunit could be assessed. Contrary to expectations, NrdEB exhibited inhibition by dATP and was found to have a tightly associated nucleotide bound. This is an interesting result given the precedent in

the literature for viral RNRs being different from their host's system, but showed that NrdEB was not going to be useful for structural mapping of additional allosteric sites on NrdE.

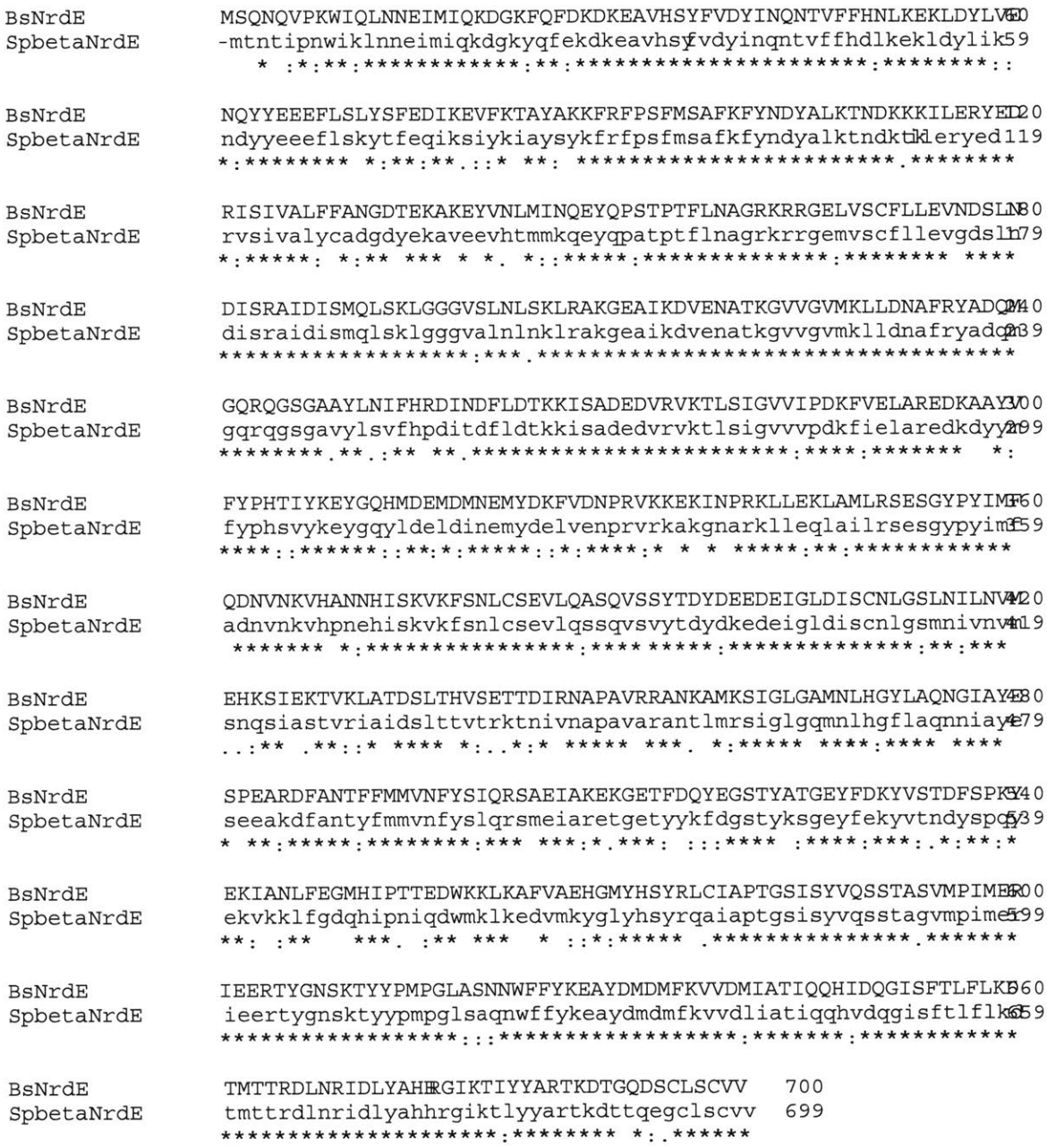
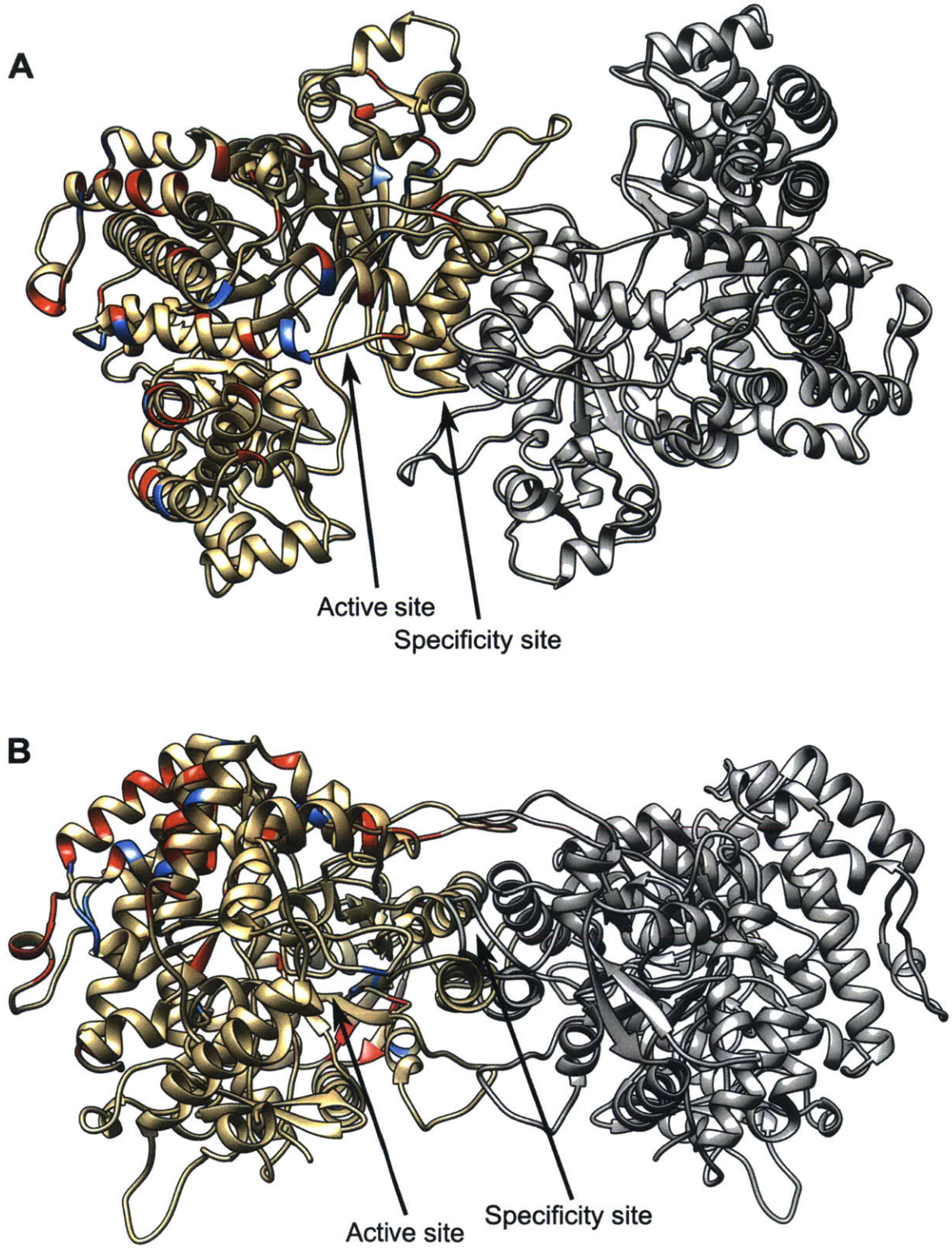


Figure A5.6. Clustal Omega sequence alignment of *B. subtilis* and the SPβ prophage NrdE subunits. Symbols mean the following: * = conserved, : = conserved with positive substitution, . = conserved with weakly similar properties.

A5.3.2. Experimental

A5.3.2A. General methodology. Chemicals, equipment, and routine methods were generally the same as described in the previous chapters. Tagless NrdF and NrdI were prepared as described in Chapter 3. NrdF was reconstituted with $Mn(III)_2$-Y• and holo-protein (0.87 Y•/β_2, 1300 nmol min^{-1} mg^{-1}) purified following the procedures described in Chapter 2. Unless noted otherwise, all concentrations of His_6-NrdEB are reported relative to monomer.

A5.3.2B. Cloning, overexpression, and purification of NrdEB. *nrdEB* was cloned from genomic DNA in three steps to remove the intron and intein sequences from the coding region of the protein. The touchdown PCR method[17] was used with the primers listed in **Table A5.1**. The 5′ end of *nrdEB* was amplified with Primer 1 and Primer 2, digested with NdeI and SacI, and ligated into similarly digested pET24a using T4 DNA ligase to create pET24a-$nrdEB_{Nterm}$. Next, the middle segment was amplified using Primer 3 and Primer 4, digested with BsaI and XhoI, and ligated into similarly digested pET24a-$nrdEB_{Nterm}$ to create pET24a-$nrdEB_{NM}$. In the final step, two amplicons were simultaneously ligated into pET28a digested with NdeI and XhoI using T4 DNA ligase. The first amplicon was produced by amplifying the *nrdEB* coding sequence from pET24a-$nrdEB_{NM}$ using Primer 1 and Primer 4 and subsequently digesting it with NdeI at 37 °C for 1 h followed by BsmFI at 65 °C for 1 h. The second amplicon was produced by amplifying the 3′ end of *nrdEB* from genomic DNA using Primer 5 and Primer 6 and subsequently digesting it with XhoI at 37 ° for 1 h followed by BsmFI at 65 °C for 1 h. Sequencing (Quintara Biosciences) confirmed the successful cloning of an intact copy of *nrdEB* into pET28a.


A
Active site
Specificity site
B
Active site
Specificity site

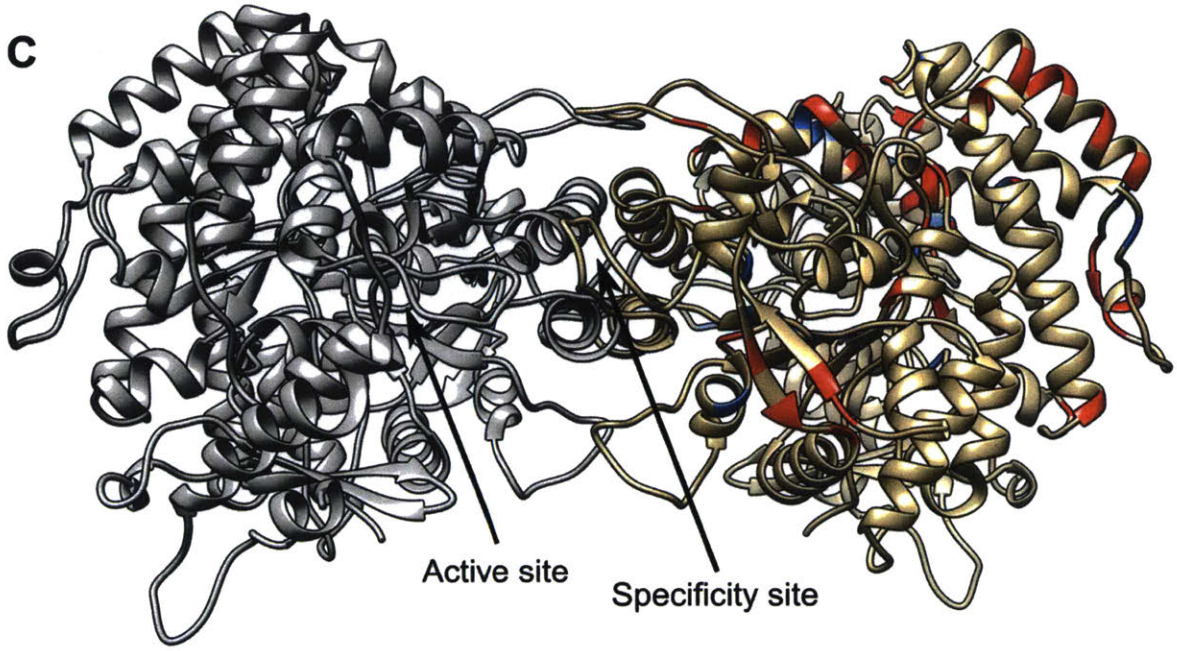


Figure A5.7. Mapping differences between the sequences of *B. subtilis* and SPβ NrdE. The PHYRE structure of NrdE from Chapter 2 was modeled as a dimer. One protomer is colored gray and the other tan except for residues that are weakly (blue) or not conserved (red) between the *B. subtilis* and SPβ NrdE. Three different views of the NrdE dimer are shown: (**A**, *opposite page*) top, (**B**, *opposite page*) front with active site of tan protomer facing the reader, and (**C**) back with active site of gray protomer facing the reader.

Table A5.1. Primers used in the cloning of *nrdEB*.

Segment	Primer name	Sequence[a]
5′ end	Primer 1	5′-AAAGGGCATATGACAAACACAATACCAAATTGGATCAAGC-3′
	Primer 2	5′-TTTAAAGAGCTCGGTCTCAACCCATTTGGTCGGCATATCTGAA-3′
Middle	Primer 3	5′-AAATTTGGTCTCAGGGTCAAAGACAAGGATCAGGAGC-3′
	Primer 4	5′-TTAAAACTCGAGGGGACCAGTAACGCACAAATTTGAAAACTTCAC-3′
3′ end	Primer 5	5′-AAAGCATGCGGGACTACTGGTAACTTTGTGTTCAGAGGTACTTC-3′
	Primer 6	5′-AATAATGAGCTCGAGTTAAACTACACAGGACAAGCATCCCTC-3′

[a] Start and stop codons are indicated in boldface font. The recognition sequences for restriction enzymes are color-coded as follows: red = *NdeI*, blue = *SacI*, green = *BsaI*, orange = *XhoI*, and purple = *BsmFI*.

pET28a-*nrdEB* was transformed into BL21 (DE3) CodonPlus RIL cells and selected for on LB agar plates supplemented with 100 μg mL^{-1} Kan and 50 μg mL^{-1} Cm. Overexpression and purification of His_6-tagged NrdEB followed the same procedures used for His_6-NrdE as described in Chapter 2. Four liters of the strain were cultivated with a total recovery of 10.3 g cell paste (~2.6 g paste L^{-1}). A total of ~285 mg His_6-NrdEB (>97% pure as judged by SDS-PAGE) was recovered from the purification procedure (~28 mg protein g^{-1} cell paste). Protein concentration was estimated using $\varepsilon_{280} = 92500\ M^{-1}\ cm^{-1}$.[30]

A5.3.2C. Analysis of NrdEB for a tightly bound nucleotide. Samples of NrdEB (100 μM, 90 μL) in 50 mM Tris, pH 7.6, 150 mM NaCl were heat denatured and the amount of compound and protein quantified following the procedure described in **Section 4.2.2A**.

A5.3.2D. Examining NrdEB for dATP inhibition with the spectrophotometric assay. CDP (1 mM) reduction activity catalyzed by NrdEB (0.5 μM) in the presence of $Mn(III)_2$-Y• NrdF (0.5 μM), 3 mM ATP, and the endogenous reducing system of TrxA/TrxB/NADPH was assessed at 37 °C using the spectrophotometric assay as described in Chapter 3. The ability to inhibit the enzyme was assessed by substituting 3 mM ATP with 1 mM dATP.

A5.3.3. Results and Discussion

The ability to clone *nrdEB* was facilitated by the use of two class IIS restriction endonucleases, BsaI and BsmFI, that cut double stranded DNA downstream of their recognition sequences and leave behind sticky ends. These enzymes are extremely useful as any sequence can be encoded downstream of the *BsaI* or *BsmFI* recognition sites, which allowed the segments of *nrdEB* to be joined seamlessly by designing primers with the correct junction sequences so that the intron and intein coding regions could be bypassed. The overexpression and purification of His_6-NrdEB following the procedures used for NrdE yielded large amounts of highly pure protein

(**Figure A5.8**). His_6-NrdEB migrated slightly lower than 75 kDa on SDS-PAGE gels, similar to what was observed with the *B. subtilis* enzyme in Chapter 2.

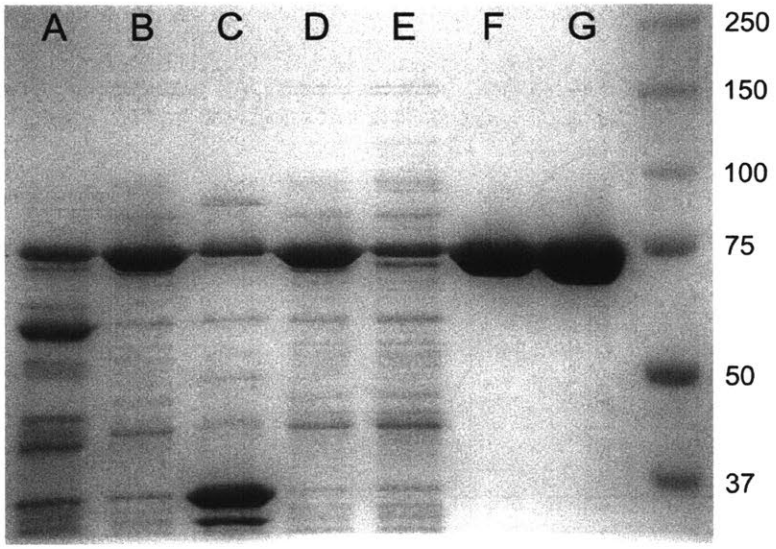
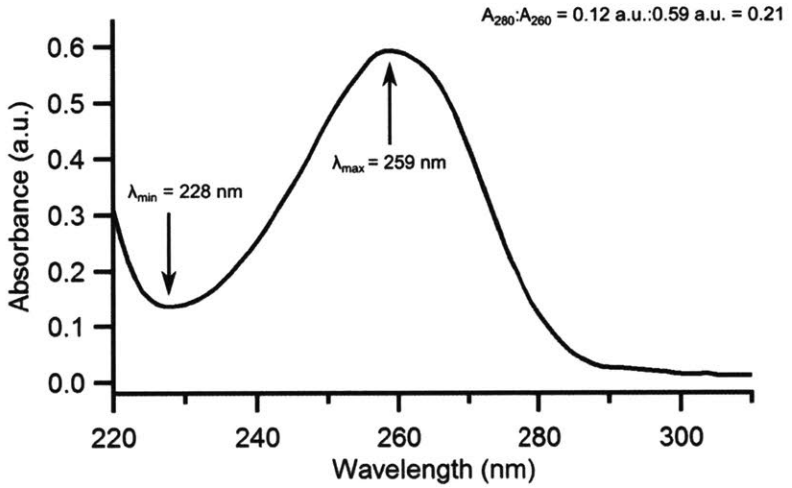


Figure A5.8. Purification of His_6-tagged NrdEB (80.7 kDa) as observed on a 7% (w/v) SDS-PAGE gel with 5 μg protein loaded per lane: *lane A*: insoluble cell debris, *lane B*: soluble extract, *lane C*: streptomycin sulfate precipitate, *lane D*: streptomycin treated extract, *lane E*: Ni-NTA loading flow through, *lane F*: Ni-NTA eluate, *lane G*: fully purified NrdEB.

His_6-NrdEB was first assessed to see if it, like NrdE, had (a) tightly associated nucleotide(s). This was found to be the case as examination of the supernatants of heat denatured samples of the protein by UV spectrophotometry revealed a spectrum with features similar to that observed for adenine-containing compounds: the spectrum had a λ_{max} = 259 nm, a λ_{min} = 228 nm, and an A_{280}:A_{260} = 0.21 (**Figure A5.9**). Given the high degree of homology between NrdEB and NrdE, the compound(s) was/were likely a mixture of deoxyadenosine analogs. A total of 3.5 nmol nucleotide(s) and 5.4 nmol protein was quantified, corresponding to a basal loading of ~0.6 equivalents of nucleotide bound per peptide. While it is interesting that both as-isolated NrdE and NrdEB have a tightly associated compound, and that both proteins are isolated with similar basal nucleotide loading, the purpose of the experiments reported in this section were to determine if

NrdEB exhibited differences from NrdE that could possibly allow for the mapping of nucleotide binding sites on the latter protein. Therefore, further characterization of the compound(s) was not pursued.



Literature values for pure compounds

	λ_{max} (nm)	λ_{min} (nm)	A_{280}:A_{260}
ATP	259	227	0.15
ADP	259	227	0.16
AMP	259	227	0.16
A	260	227	0.14
dATP	259	226	0.15
dAMP	259	226	0.15
dA	260	225	0.15

Figure A5.9. UV spectrum of compound(s) released from as-isolated His_6-NrdEB by heat denaturation into 50 mM Tris, pH 7.6, 150 mM NaCl. Tabulated to the right are values for λ_{max}, λ_{min}, and the A_{280}:A_{260} reported in the literature[31] for pure adenosine/deoxyadenosine compounds dissolved in water at pH 7.0.

The subunits of viral RNRs have been reported to be unable to cross-react with the subunits from the host's enzyme despite having fairly high sequence homology with one another (e.g. the α_2 subunit of the *E. coli* and T4 phage Ia RNRs are 54% identical (70% with positive substitutions)).[32] This was found not to be the case with NrdEB as the enzyme was able to catalyze CDP reduction using $Mn(III)_2$-Y• NrdF and the endogenous reducing system of the *B. subtilis*. Furthermore, the specific activity observed under routine conditions (600 ± 40 nmol min mg^{-1} α_2) was identical to that observed with NrdE (590 ±20 nmol min^{-1} mg^{-1} α_2). The ability of dATP to inhibit CDP reduction was quickly assessed by using a final dATP concentration of 1 mM, which was chosen since Ib RNRs other than the *B. subtilis* enzyme showed no loss of CDP reduction activity at this concentration.[22, 33-37] The specific activity measured under these conditions for

NrdEB was 60 ± 9 nmol min^{-1} mg^{-1} α_2, which corresponds to 10% of the activity observed with 3 mM ATP. In comparison, the relative difference in specific activity for NrdE in the presence of 1 mM dATP versus 3 mM ATP under the same conditions was 9%. The results of these experiments thus demonstrate that NrdEB shows dATP inhibition that is similar to that observed for the host-derived NrdE.

Given the results reported above, it is puzzling why two nearly identical RNRs are maintained in the *B. subtilis* genome. There remains a possibility that the SPβ Ib RNR might exhibit allosteric differences if reacted with its natural partner NrdFB. Efforts were made to test this possibility as the *nrdFB* and *nrdIB* genes were cloned concomitantly with *nrdEB*. However, both proteins were insoluble when overexpressed in *E. coli* under a few different growth conditions. Optimization of the expression of NrdFB and NrdIB is, therefore, required in order to examine the possibility of altered allosteric regulation of the viral Ib RNR.

On an evolutionary timescale, the high homology of the SPβ and *B. subtilis* Ib RNRs suggests that the phage recently acquired the genes from the host via horizontal gene transfer and, therefore, perhaps has not had sufficient time to alter them for its benefit.[38] A hypothesis also exists that host-like genes acquired by viruses can be expressed as an auxiliary metabolic system that supports the metabolism of the host during the infection process.[38] Therefore, it may be possible that the SPβ RNR acts to support *B. subtilis* viability until the virus has completed its reproduction and is ready for release into the environment. However, arguably the strongest evidence for why the SPβ Ib operon is maintained in the *B. subtilis* genome is that it appears to confer immunity to the host against other phages that could compete with SPβ.[39] Further studies of the enzymology of the SPβ RNR and the biology of *B. subtilis* should shed more light on this interesting system.

A5.4. REFERENCES

1. Cotruvo, J. A., and Stubbe, J. (2011) Class I Ribonucleotide Reductases: Metallocofactor Assembly and Repair *In Vitro* and *In Vivo*, In *Annual Review of Biochemistry* (Kornberg, R. D., Raetz, C. R. H., Rothman, J. E., and Thorner, J. W., Eds.), pp 733-767, Annual Reviews, Palo Alto.

2. Cotruvo, J. A., Jr., (2012) *Mechanism of the Biosynthesis of the Dimanganese-Tyrosyl Radical Cofactor of Class Ib Ribonucleotide Reductases*. PhD Thesis, Massachusetts Institute of Technology.

3. Cotruvo, J. A., and Stubbe, J. (2010) An Active Dimanganese(III)-Tyrosyl Radical Cofactor in *Escherichia coli* Class Ib Ribonucleotide Reductase, *Biochemistry 49*, 1297-1309.

4. Chen, J., Shen, J., Solem, C., and Jensen, P. R. (2015) A New Type of YumC-Like Ferredoxin (Flavodoxin) Reductase Is Involved in Ribonucleotide Reduction, *mBio 6*, 8.

5. Komori, H., Seo, D., Sakurai, T., and Higuchi, Y. (2010) Crystal Structure Analysis of *Bacillus subtilis* Ferredoxin-$NADP^+$ Oxidoreductase and the Structural Basis for its Substrate Selectivity, *Protein Sci. 19*, 2279-2290.

6. Seo, D., Kamino, K., Inoue, K., and Sakurai, H. (2004) Purification and Characterization of Ferredoxin-$NADP^+$ Reductase Encoded by *Bacillus subtilis yumC*, *Arch. Microbiol. 182*, 80-89.

7. Kobayashi, K., Ehrlich, S. D., Albertini, A., Amati, G., Andersen, K. K., Arnaud, M., Asai, K., Ashikaga, S., Aymerich, S., Bessieres, P., Boland, F., Brignell, S. C., Bron, S., Bunai, K., Chapuis, J., Christiansen, L. C., Danchin, A., Débarbouillé, M., Dervyn, E., Deuerling, E., Devine, K., Devine, S. K., Dreesen, O., Errington, J., Fillinger, S., Foster, S. J., Fujita, Y., Galizzi, A., Gardan, R., Eschevins, C., Fukushima, T., Haga, K., Harwood, C. R., Hecker, M., Hosoya, D., Hullo, M. F., Kakeshita, H., Karamata, D., Kasahara, Y., Kawamura, F., Koga, K., Koski, P., Kuwana, R., Imamura, D., Ishimaru, M., Ishikawa, S., Ishio, I., Le Coq, D., Masson, A., Mauël, C., Meima, R., Mellado, R. P., Moir, A., Moriya, S., Nagakawa, E., Nanamiya, H., Nakai, S., Nygaard, P., Ogura, M., Ohanan, T., O'Reilly, M., O'Rourke, M., Pragai, Z., Pooley, H. M., Rapoport, G., Rawlins, J. P., Rivas, L. A., Rivolta, C., Sadaie, A., Sadaie, Y., Sarvas, M., Sato, T., Saxild, H. H., Scanlan, E., Schumann, W., Seegers, J., Sekiguchi, J., Sekowska, A., Séror, S. J., Simon, M., Stragier, P., Studer, R., Takamatsu, H., Tanaka, T., Takeuchi, M., Thomaides, H. B., Vagner, V., van Dijl, J. M., Watabe, K., Wipat, A., Yamamoto, H., Yamamoto, M., Yamamoto, Y., Yamane, K., Yata, K., Yoshida, K., Yoshikawa, H., Zuber, U., and Ogasawara, N. (2003) Essential *Bacillus subtilis* Genes, *Proc. Natl. Acad. Sci. U. S. A. 100*, 4678-4683.

8. Holden, J. K., Lim, N., and Poulos, T. L. (2014) Identification of Redox Partners and Development of a Novel Chimeric Bacterial Nitric Oxide Synthase for Structure Activity Analyses, *J. Biol. Chem. 289*, 29437-29445.

9. Chazarreta-Cifre, L., Martiarena, L., de Mendoza, D., and Altabe, S. G. (2011) Role of Ferredoxin and Flavodoxins in *Bacillus subtilis* Fatty Acid Desaturation, *J. Bacteriol. 193*, 4043-4048.

10. Green, A. J., Munro, A. W., Cheesman, M. R., Reid, G. A., von Wachenfeldt, C., and Chapman, S. K. (2003) Expression, Purification and Characterisation of a *Bacillus subtilis* Ferredoxin: A Potential Electron Transfer Donor to Cytochrome P450 BioI, *J. Inorg. Biochem. 93*, 92-99.

11. Green, A. J., Rivers, S. L., Cheesman, M., Reid, G. A., Quaroni, L. G., Macdonald, I. D. G., Chapman, S. K., and Munro, A. W. (2001) Expression, Purification and Characterization of Cytochrome P450 BioI: A Novel P450 Involved in Biotin Synthesis in *Bacillus subtilis*, *J. Biol. Inorg. Chem. 6*, 523-533.

12. Lawson, R. J., von Wachenfeldt, C., Haq, I., Perkins, J., and Munro, A. W. (2004) Expression and Characterization of the Two Flavodoxin Proteins of *Bacillus subtilis*, YkuN and YkuP: Biophysical Properties and Interactions with Cytochrome P450 BioI, *Biochemistry 43*, 12390-12409.

13. Stok, J. E., and De Voss, J. J. (2000) Expression, Purification, and Characterization of BioI: A Carbon-Carbon Bond Cleaving Cytochrome P450 Involved in Biotin Biosynthesis in *Bacillus subtilis*, *Arch. Biochem. Biophys. 384*, 351-360.

14. Torrents, E., Buist, G., Liu, A., Eliasson, R., Kok, J., Gibert, I., Gräslund, A., and Reichard, P. (2000) The Anaerobic (Class III) Ribonucleotide Reductase from *Lactococcus lactis* - Catalytic Properties and Allosteric Regulation of the Pure Enzyme System, *J. Biol. Chem. 275*, 2463-2471.

15. One of the suppressor mutants had an additional SNP in the intergenic region between two genes involved in trehalose metabolism and sucrose transport as well as a large deletion of ~12 kbp encoding 12 genes putatively involved in tellurium resistance. These genomic modifications seem unrelated to nucleotide metabolism and probably did not affect the results of the study.

16. Schuch, R., Garibian, A., Saxild, H. H., Piggot, P. J., and Nygaard, P. (1999) Nucleosides as a Carbon Source in *Bacillus subtilis*: Characterization of the *drm-pupG* Operon, *Microbiology-(UK) 145*, 2957-2966.

17. Korbie, D. J., and Mattick, J. S. (2008) Touchdown PCR for Increased Specificity and Sensitivity in PCR Amplification, *Nat. Protoc. 3*, 1452-1456.

18. Smith, P. K., Krohn, R. I., Hermanson, G. T., Mallia, A. K., Gartner, F. H., Provenzano, M. D., Fujimoto, E. K., Goeke, N. M., Olson, B. J., and Klenk, D. C. (1985) Measurement of Protein Using Bicinchoninic Acid, *Anal. Biochem. 150*, 76-85.

19. Zhang, Y., and Stubbe, J. (2011) *Bacillus subtilis* Class Ib Ribonucleotide Reductase is a Dimanganese(III)-Tyrosyl Radical Enzyme, *Biochemistry 50*, 5615-5623.

20. Cotruvo, J. A., Stich, T. A., Britt, R. D., and Stubbe, J. (2013) Mechanism of Assembly of the Dimanganese-Tyrosyl Radical Cofactor of Class Ib Ribonucleotide Reductase: Enzymatic Generation of Superoxide Is Required for Tyrosine Oxidation via a Mn(III)Mn(IV) Intermediate, *J. Am. Chem. Soc. 135*, 4027-4039.

21. Lofstad, M., Gudim, I., Hammerstad, M., Røhr, A. K., and Hersleth, H. P. (2016) Activation of the Class Ib Ribonucleotide Reductase by a Flavodoxin Reductase in *Bacillus cereus*, *Biochemistry 55*, 4998-5001.

22. Crona, M., Torrents, E., Røhr, A. K., Hofer, A., Furrer, E., Tomter, A. B., Andersson, K. K., Sahlin, M., and Sjöberg, B. M. (2011) NrdH-Redoxin Protein Mediates High Enzyme Activity in Manganese-Reconstituted Ribonucleotide Reductase from *Bacillus anthracis*, *J. Biol. Chem. 286*, 33053-33060.

23. Gustafsson, T. N., Sahlin, M., Lu, J., Sjöberg, B. M., and Holmgren, A. (2012) *Bacillus anthracis* Thioredoxin Systems, Characterization and Role as Electron Donors for Ribonucleotide Reductase, *J. Biol. Chem. 287*, 39686-39697.

24. Averett, D. R., Lubbers, C., Elion, G. B., and Spector, T. (1983) Ribonucleotide Reductase Induced by Herpes Simplex Type 1 Virus - Characterization of a Distinct Enzyme, *J. Biol. Chem. 258*, 9831-9838.

25. Berglund, O. (1972) Ribonucleoside Diphosphate Reductase Induced by Bacteriophage T4. II. Allosteric Regulation of Substrate Specificity and Catalytic Activity, *J. Biol. Chem. 247*, 7276-7281.

26. Friedrich, N. C., Torrents, E., Gibb, E. A., Sahlin, M., Sjöberg, B. M., and Edgell, D. R. (2007) Insertion of a Homing Endonuclease Creates a Genes-In-Pieces Ribonudeotide Reductase that Retains Function, *Proc. Natl. Acad. Sci. U. S. A. 104*, 6176-6181.

27. Hendricks, S. P., and Mathews, C. K. (1997) Regulation of T4 Phage Aerobic Ribonucleotide Reductase - Simultaneous Assay of the Four Activities, *J. Biol. Chem. 272*, 2861-2865.

28. Huszar, D., and Bacchetti, S. (1981) Partial Purification and Characterization of the Ribonucleotide Reductase Induced by Herpes Simplex Virus Infection of Mammalian Cells, *J. Virol. 37*, 580-588.

29. Poncedeleon, M., Eisenberg, R. J., and Cohen, G. H. (1977) Ribonucleotide Reductase from Herpes Simplex Virus (Type 1 and Type 2) Infected and Uninfected KB Cells: Properties of the Partially Purified Enzymes, *J. Gen. Virol. 36*, 163-173.

30. Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D., and Bairoch, A. (2005) Protein Identification and Analysis Tools on the ExPASy Server., In *The Proteomics Protocols Handbook* (Walker, J. M., Ed.), pp 571-607, Humana Press, Totowa, NJ.

31. (1996) Characteristics of Nucleic Acids, In *Current Protocols in Molecular Biology*, pp A.1D.1-A.1D.11, John Wiley & Sons, Inc.

32. Berglund, O., Karlström, O and Reichard, P. (1969) A New Ribonucleotide Reductase System After Infection with Phage T4, *Proc. Natl. Acad. Sci. U., S. A. 62*, 829-835.

33. Eliasson, R., Pontis, E., Jordan, A., and Reichard, P. (1996) Allosteric Regulation of the Third Ribonucleotide Reductase (NrdEF Enzyme) from Enterobacteriaceae, *J. Biol. Chem. 271*, 26582-26587.

34. Jordan, A., Pontis, E., Åslund, F., Hellman, U., Gibert, I., and Reichard, P. (1996) The Ribonucleotide Reductase System of *Lactococcus lactis* - Characterization of an NrdEF Enzyme and a New Electron Transport Protein, *J. Biol. Chem. 271*, 8779-8785.

35. Jordan, A., Pontis, E., Atta, M., Krook, M., Gibert, I., Barbé, J., and Reichard, P. (1994) A Second Class I Ribonucleotide Reductase in *Enterobacteriaceae* - Characterization of the *Salmonella typhimurium* Enzyme, *Proc. Natl. Acad. Sci. U. S. A. 91*, 12892-12896.

36. Makhlynets, O., Boal, A. K., Rhodes, D. V., Kitten, T., Rosenzweig, A. C., and Stubbe, J. (2014) *Streptococcus sanguinis* Class Ib Ribonucleotide Reductase: High Activity with Both Iron and Manganese Cofactors and Structural Insights, *J. Biol. Chem. 289*, 6259-6272.

37. Yang, F. D., Curran, S. C., Li, L. S., Avarbock, D., Graf, J. D., Chua, M. M., Lui, G. Z., Salem, J., and Rubin, H. (1997) Characterization of Two Genes Encoding the *Mycobacterium tuberculosis* Ribonucleotide Reductase Small Subunit, *J. Bacteriol. 179*, 6408-6415.

38. Dwivedi, B., Xue, B. J., Lundin, D., Edwards, R. A., and Breitbart, M. (2013) A Bioinformatic Analysis of Ribonucleotide Reductase Genes in Phage Genomes and Metagenomes, *BMC Evol. Biol. 13*, 17.

39. Yee, L. M., Matsuoka, S., Yano, K., Sadaie, Y., and Asai, K. (2011) Inhibitory Effect of Prophage SPβ Fragments on Phage SP10 Ribonucleotide Reductase Function and its Multiplication in *Bacillus subtilis*, *Genes Genet. Syst. 86*, 7-18.