Mechanistic Investigations of Class III Anaerobic Ribonucleotide Reductases

by

Yifeng Wei

BSc, Chemistry, Imperial College London, 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

September 2015

Signature redacted

Signature of Author: _____
Department of Chemistry
August 31, 2015

Signature redacted

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

Signature redacted

Accepted by: _____
Robert W. Field
Haslam and Dewey Professor of Chemistry
Chairman, Departmental Committee on Graduate Students

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

Signature redacted

Alice Y. Ting
Ellen Swallow Richards Professor
Thesis Chair

Signature redacted

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

Signature redacted

Catherine L. Drennan
Professor of Chemistry and Biology
Investigator and Professor, Howard Hughes Medical Institute

To my parents

ACKNOWLEDGEMENTS

First I would like to thank my thesis advisor, Prof. JoAnne Stubbe, for giving me the intellectual and physical resources and freedom to pursue these scientific questions. Thank you for always being available to talk about science and for sharing your appreciation for the beauty of chemical reactions occurring in nature. It has been a privilege to work in the Stubbe lab and on ribonucleotide reductase. Thanks also to my thesis committee Prof. Alice Y. Ting and Prof. Catherine L. Drennan for guidance over the past several years.

This thesis was made possible in a large part due to the experimental and intellectual contributions of Dr. Guinevere Mathies and Prof. Robert G. Griffin for EPR spectroscopy; Michael A. Funk and Prof. Catherine L. Drennan for crystallography; Dr. Leonardo A. Rosado for phylogenetic studies; and Dr. Jiahao Chen for computational chemistry. Thanks also to Judy Baek for all the hours spent finding conditions to get soluble protein that was critical for the experiments to work. Thanks to Dr. Silvan Scheller for many helpful discussions about methanogens. Thank you for your help and it was a pleasure working together!

Thanks to my agency ASTAR Singapore for generous funding during the first 5 years of my study. I would like to thank Dr. Jun Yang, Dr. Patricia Hunt, Prof. Alan Armstrong and Prof. Jackie Y. Ying, your support has been critical in giving me the opportunity to pursue further studies, and I am very grateful.

To past lab members Yimon, Olga, Xuling, Ping, Crystal, Mimi, Joey, Ellen, Rachael and Leo, it was a privilege to work in the same lab as you. Special thanks to Ken for teaching me all about biochemistry when I first joined the lab, and for stimulating discussions about radical mechanisms. Thanks also to Henry and my good friend Yan for your company and sharing your research. To current lab members Lisa, Mac, Kanchana, Wankyu, Denyce, Julia, the last couple of years have been fun. Thanks in particular to Kanchana and Mac for keeping the lab running. To Hongik, Jeff, Jingnan, Ken, Marco, Michael, Tengfei, Vinita, I am very fortunate to have you as classmates. I will remember the fun we had on ski trips and birthday dinners at Mulan.

Thanks to my mother, father and three brothers, for being behind me throughout the process from applications to graduation. Special thanks to Aunt Cheryl and Uncle Lip Cher, my studies would not have been possible without your support.

Mechanistic investigations of Class III Anaerobic Ribonucleotide Reductases

by
Yifeng Wei

Submitted to the Department of Chemistry
on September 30, 2015 in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy in
Biological Chemistry

Abstract

Ribonucleotide reductases (RNRs) catalyze nucleotide reduction via complex radical chemistry, providing deoxynucleotides for DNA synthesis in all domains of life. The focus of this thesis is the class III RNR, found in anaerobic bacteria and archaea, which uses an O_2-sensitive glycyl radical cofactor (G•). The class III RNRs studied to date couple nucleotide reduction to the oxidation of formate to CO_2. We started by studying the *Escherichia coli* class III RNR (NrdD1), and found that reaction with CTP (substrate) and ATP (effector) in the absence of formate leads to loss of G• concomitant with stoichiometric formation of a new radical species and a "trapped" cytidine derivative, proposed to be 3′-keto-deoxycytidine. Addition of formate results in the recovery of G• and reduction of the cytidine derivative to dCTP. The new radical was identified by 9.5 and 140 GHz EPR spectroscopy to be a cysteine-methionine thiosulfuranyl radical $[RSSR_2]$•. Analogies with the disulfide anion radical proposed in the class I and II RNRs provide further evidence for the involvement of thiyl radicals in the reductive half reaction of all RNRs. A subsequent bioinformatics investigation led to the identification of a second subtype of class III RNR (NrdD2) with distinct active-site residues suggesting that, like the class I and II RNRs, reducing equivalents for nucleotide reduction are provided by a redoxin, which are ubiquitous proteins found in all organisms, instead of formate, a metabolite produced by some but not all organisms. The *Neisseria bacilliformis* NrdD2 was cloned and expressed, and found to catalyze nucleotide reduction using the thioredoxin / thioredoxin reductase / NADPH system. An active-site model based on a crystal structure of the homologous *Thermotoga maritima* enzyme showed conserved residues appropriately positioned to carry out chemistry. Phylogenetic studies suggest that NrdD2 is present in bacteria and archaea that carry out diverse types of anaerobic metabolism. The bioinformatics study also uncovered a third subtype of class III RNR (NrdD3) present in certain methanogenic archaea. The presence of a redoxin (NrdH) in the operon suggested redoxin-dependent chemistry like NrdD2. However, its distribution among the different types of methanogens suggested that reducing equivalents might come from reduced ferredoxin (Fdx) generated in methanogenesis, rather than from NADPH. The *Methanosarcina barkeri* class III RNR was cloned and expressed, and found to catalyze nucleotide reduction using a system involving Fdx, NrdH and a [4Fe4S] protein ferredoxin:thioredoxin reductase. The diversity of reducing equivalents used for anaerobic ribonucleotide reduction reflects the diversity of electron carriers used in anaerobic energy metabolism.

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

Table of Contents

List of Figures

List of Tables

List of Abbreviations

A_{280}	Absorbance at 280 nm
Amp	Ampicillin
ATP	Adenosine-5′-triphosphate
BME	β-mercaptoethanol
Cm	Chloramphenicol
CV	Column volume
CW	Continuous wave
Cyt	Cytosine
dATP	2′-deoxydenosine-5′-triphosphate
dCTP	2′-deoxycytidine-5′-triphosphate
DTT	Dithiothreitol
DTNB	5,5′-dithiobis-(2-nitrobenzoic acid)
Ec	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid
EPR	Electron paramagnetic resonance
Eq	Equivalent(s)
FAD	Flavin adenine dinucleotide
Fpr	Ferredoxin(flavodoxin)-NADPH reductase
G•	Glycyl radical
IPTG	Isopropyl-1-thio-β-galactopyranoside
Km	Kanamycin
LB	Luria-Bertani broth
Mb	*Methanosarcina barkeri*
NADPH	β-nicotinamide adenine dinucleotide phosphate, reduced form
Nb	*Neisseria bacilliformis*
Ni-NTA	Nickel nitriloacetate
NrdD	Class III RNR
NrdG	Class III RNR activating enzyme
NrdH	Thioredoxin-like protein that transfers electrons to some RNRs
NrdJ	Class II RNR
PMSF	Phenylmethylsulfonyl fluoride
RNR	Ribonucleotide reductase
RT	Room temperature
SDS-PAGE	Sodium dodecyl sulfate - polyacrylamide gel electrophoresis
Tm	*Thermotoga maritima*
TrxA	Thioredoxin
TrxB	Thioredoxin reductase
Y•	Tyrosyl radical

Chapter 1:

Introduction to the class III anaerobic ribonucleotide reductase

1. Introduction to the class III anaerobic ribonucleotide reductase

1.1. General introduction to ribonucleotide reductases

Ribonucleotide reductases (RNRs) catalyze the conversion of all four ribonucleotides (A,G,C,U) into deoxynucleotides, providing the monomeric precursors required for DNA synthesis and repair.[1,2] They provide the only known pathway for *de novo* biosynthesis of deoxynucleoside triphosphates (dNTPs) and thus perform an essential function in nearly all cellular organisms and many viruses. Their highly conserved allosteric regulation by dNTPs and ATP ensures the production of a balanced pool of dNTPs required for fidelity in DNA replication. All RNRs studied thus far share a structurally homologous active site architecture in the α subunit, and a partially conserved mechanism involving complex radical chemistry. The reaction is initiated by generating an organic radical on the 3′ position of the nucleotide substrate via a transient Cys thiyl radical on the "top" face of the ribose in the active-site.[3]

Although they carry out an essential function, the primary sequence of RNRs are not highly conserved and they are divided into three classes (I, II and III) according to the cofactor used to generate the initiating thiyl radical (**Figure 1.1**).[4]

Class I RNRs use cofactors that require O_2 and different combinations of metals (Fe and Mn) for their biogenesis.[5] These RNRs are only present in aerobic organisms. The class Ia RNR contains a $Fe^{III}Fe^{III}$-tyrosyl radical (Y•) cofactor and is present in all aerobic eukaryotes and in many aerobic bacteria. The class Ib RNR contains a $Mn^{III}Mn^{III}$-tyrosyl radical (Y•) generated in a reaction involving a flavoprotein NrdI, and is present only in bacteria, including many pathogens. The class Ic RNR is proposed to contain a $Fe^{III}Mn^{IV}$ cofactor and is present in certain bacteria and halophilic archaea. In all three subclasses, nucleotide reduction is believed to occur in a dimeric α_2 subunit (NrdA or NrdE), while the metallocofactor is situated in a dimeric β_2

subunit (NrdB or NrdF respectively). Based on structural models, initiation of the reaction is thought to involve radical translocation across ~35 Å by several proton-coupled electron transfer (PCET) steps involving aromatic amino acid radical intermediates.[6,7]

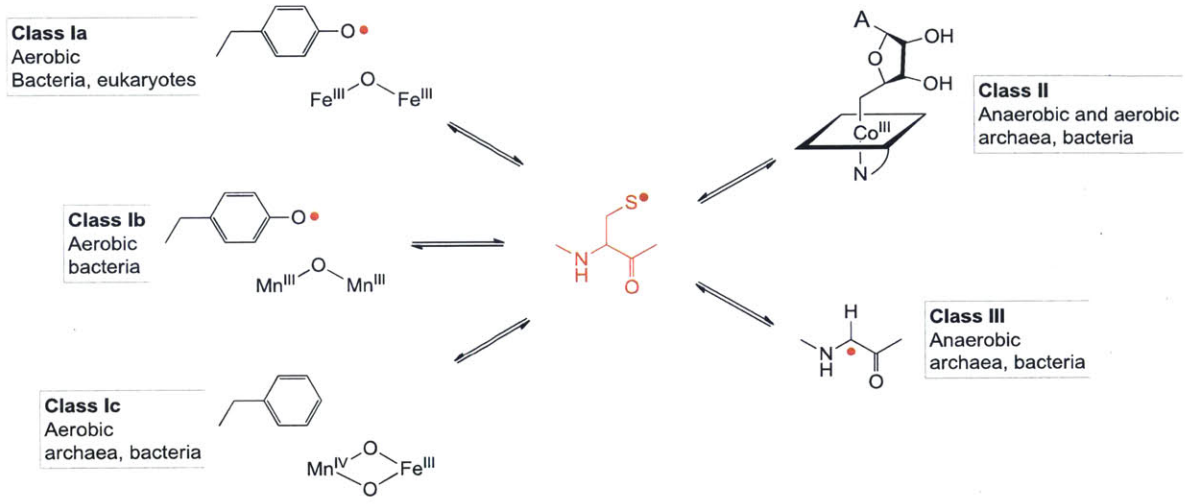


Figure 1.1 Classes of RNRs. RNRs are classified according to the cofactor used to generate the active site thiyl radical used to initiate the radical-dependent nucleotide reduction chemistry.

The class II RNRs use an adenosylcobalamin (AdoCbl) cofactor, bound in the α protein (NrdJ), that reversibly generates the active-site thiyl radical by direct H-atom abstraction.[3] The reaction is O_2-independent, and these RNRs are present in both aerobic and anaerobic bacteria and archaea.

Class III RNRs use a glycyl radical (G•),[8] situated in the α protein (NrdD), that reversibly generates the active-site thiyl radical by direct H-atom abstraction. The G• is O_2-sensitive, and these RNRs are found in facultative and obligate anaerobic bacteria and archaea. Generation G• requires a separate activating enzyme (NrdG),[9] and proceeds through chemistry involving S-adenosylmethionine (SAM) and a reduced $[4Fe4S]^{1+}$ cluster (**Figure 1.2**).[10,11]

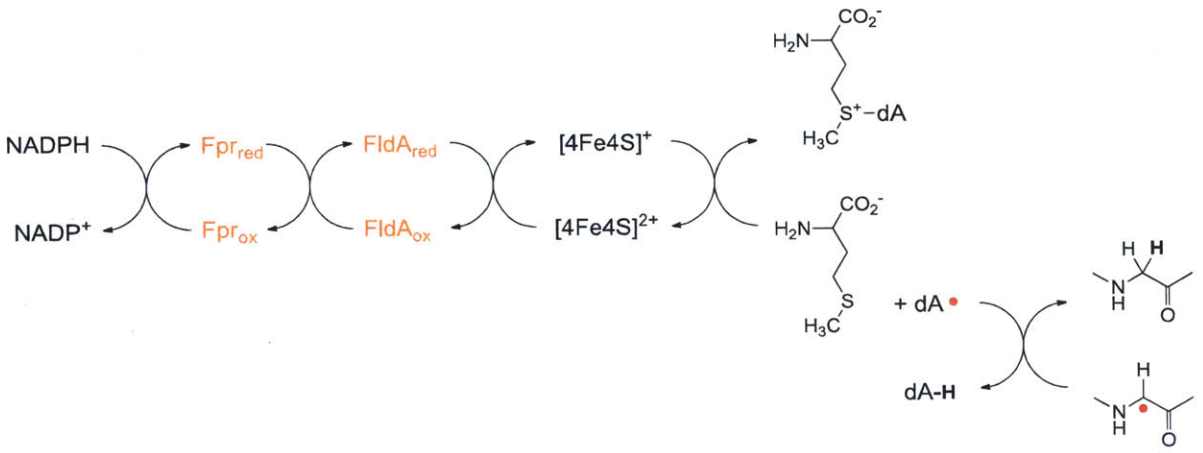


Figure 1.2 Proposed mechanism for generation of the G• cofactor on G• enzymes (GREs). The reaction involves a reduced $[4Fe4S]^{+}$ cluster on the activating enzyme (AE), which reductively cleaves a molecule of S-adenosylmethionine to form methionine and a 5′-deoxyadenosyl radical, which stereospecifically abstracts the α-H from a conserved Gly on the GRE. The reductant for some of the AEs, including PFLAE and NrdG of the *E. coli* class III RNR, is believed to be the flavodoxin (FldA) / flavodoxin reductase (Fpr) / NADPH system. The steps involving reduction of $[4Fe4S]^{2+}$ by FldA and cleavage of SAM are endergonic, and driven forward by the exergonic formation of G• (dA = 5′-deoxyadenosyl; orange = flavoproteins).

A further distinction between the three classes has been the source of the reducing equivalents for nucleotide reduction. In the class I and II RNRs, nucleotide reduction is coupled to the oxidation of a pair of conserved cysteines on the "bottom" face of the ribose in the active-site (**Figure 1.3A**).[12] The resulting disulfide is then reduced by a redoxin (thioredoxin, glutaredoxin, or NrdH). In contrast, in the class III RNRs from *Escherichia coli*, *Lactococcus lactis* and bacteriophage T4, characterized prior to the start of this thesis project, nucleotide reduction is coupled to the oxidation of formate to CO_2 (**Figure 1.3B**).[13] The occurrence of two varieties of redoxin-dependent class III RNRs from *Neisseria bacilliformis* and *Methanosarcina barkeri* are described in **Chapters 3** and **5** of this thesis respectively.

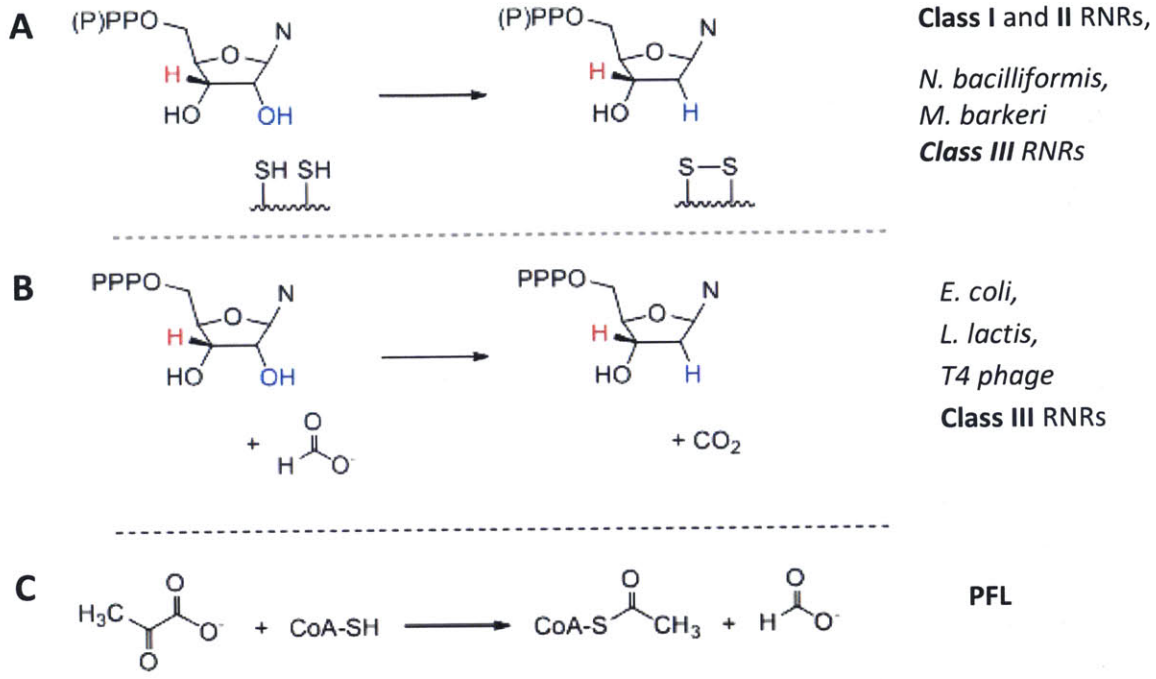


Figure 1.3 Reactions catalyzed by RNRs and pyruvate formate lyase (PFL). A) Nucleotide reduction in the class I and II RNRs, and in the *N. bacilliformis* and *M. barkeri* class III RNRs described in this thesis, is coupled to the formation of an active site disulfide, which is then reduced by a redoxin (thioredoxin, glutaredoxin or NrdH). B) Nucleotide reduction in the *E. coli*, *L. lactis* and bacteriophage T4 class III RNRs, studied prior to this thesis, is coupled to the oxidation of a molecule of formate to CO_2. C) PFL catalyzes the non-oxidative cleavage of a molecule of pyruvate to formate and acetyl-coA. The three enzymes are structurally related and the reactions proceed through a conserved Cys thiyl radical in the active site.

1.2. Review of studies of the class III RNR

The "long and winding" path to the current understanding of the class III RNR[14] was hampered by its requirement of stringent anaerobic conditions. In this section, we review the studies conducted on class III RNR systems to date, focusing on its discovery, relationship to the prototypical G• enzyme pyruvate formate lyase (PFL), generation of the G• cofactor, nature of the reductant, Zn binding site and allosteric regulation, and highlighting the characteristics and idiosyncrasies of the enzymes involved. Studies relevant to the mechanism of nucleotide

reduction by the class III RNR are reviewed in detail in the introduction of **Chapter 2**, and studies relevant to the distribution and evolution of RNRs are reviewed in **Chapter 4**.

1.2.1. A class of O_2-sensitive RNRs present in anaerobic organisms

While the class Ia and II RNRs have been extensively studied, less is known about the class III RNR. Prior to its discovery, the biosynthetic route for dNTPs in anaerobic bacteria and in archaea was unknown, and it was unclear whether dNTPs were synthesized from NTPs via an RNR, or if they were made by an alternative pathway, such as one involving deoxyribose-phosphate aldolase (DERA).[15]

The activity of an O_2-sensitive RNR distinct from the class I and II RNRs was first detected in extracts of the strictly anaerobic methanogenic archaeon, *Methanobacterium marburgensis* (formerly *Methanobacterium thermoautotrophicum* str. Marburg).[16] Radiolabeled cytidine, but not cytosine, in the growth medium was incorporated into DNA, suggesting that dNTPs in this archaeon are provided by an RNR. Cell extracts did not catalyze tritium exchange from 5′-[^{3}H]-AdoCbl, an activity characteristic of the class II RNR, ruling out its involvement in the pathway. Subsequently, adhering to strictly anaerobic conditions, RNR activity was detected in cell extracts.[17] Activity was independent of AdoCbl and only weakly inhibited by hydroxyurea, which quenches Y•, but dramatically decreased upon exposure to air.

Similar activity was later detected in the facultative anaerobic gram negative bacterium *E. coli*.[15,18] The class III RNR from this model bacterium has served as the prototype for this class of enzymes, and its discovery and early experiments are reviewed by Reichard.[19] During aerobic growth, *E. coli* uses a class Ia RNR, but under anaerobic conditions biosynthesis of the Y• cofactor is not possible and mutants of NrdA and NrdB showed that they were not essential for

anaerobic growth.[18] *E. coli* lacks a biosynthetic pathway for AdoCbl and does not require it under anaerobic conditions, therefore ruling out a class II RNR. Using radiolabeled nucleotides, RNR activity was detected under strict anaerobic conditions.[15] Activity was O_2-sensitive, not inhibited by antibodies raised against NrdB and only weakly inhibited by hydroxyurea. Also, while the class I RNRs use NDPs as substrates, this and all other class III RNRs studied in detail to date reduce NTPs.

Subsequent purification of the protein facilitated mapping and sequencing of the *nrdD* gene.[20] Apart from the N-terminal 95 amino acids (encoding the "ATP cone" allosteric domain), the sequence of *E. coli* NrdD displayed no significant similarity to *E. coli* NrdA or any other RNR that had been studied. It did, however, display 58% identity to a bacteriophage T4 intron-containing gene in *sunY* (later renamed *nrdD*). Although this gene had served as a model for splicing of group I introns, the function of its product was previously unknown, and the high sequence identity suggested that the phage codes for its own anaerobic RNR, in addition to its previously characterized class Ia RNR. Lysates of anaerobic *E. coli* infected with bacteriophage T4 exhibited 10-30-fold higher anaerobic RNR activity, and mutation of T4 *nrdD* severely reduced phage infectivity only under anaerobic conditions.[21]

A putative class III RNR sequence was also identified in a stretch of DNA that had been cloned from the gram positive bacterium *L. lactis*, with an amino acid sequence displaying 53% identity to *E. coli* NrdD.[22] Mutagenesis experiments showed that this gene is essential for anaerobic growth of the bacterium.

The *E. coli*, bacteriophage T4 and *L. lactis* enzymes have been cloned and characterized, and are the most extensively studied enzymes of this class. Subsequently, in the post-genomics

era, sequences encoding for putative class III RNRs have been found in diverse anaerobic and facultative anaerobic archaea, bacteria, large DNA phages and a small number of eukaryotes.[23,24]

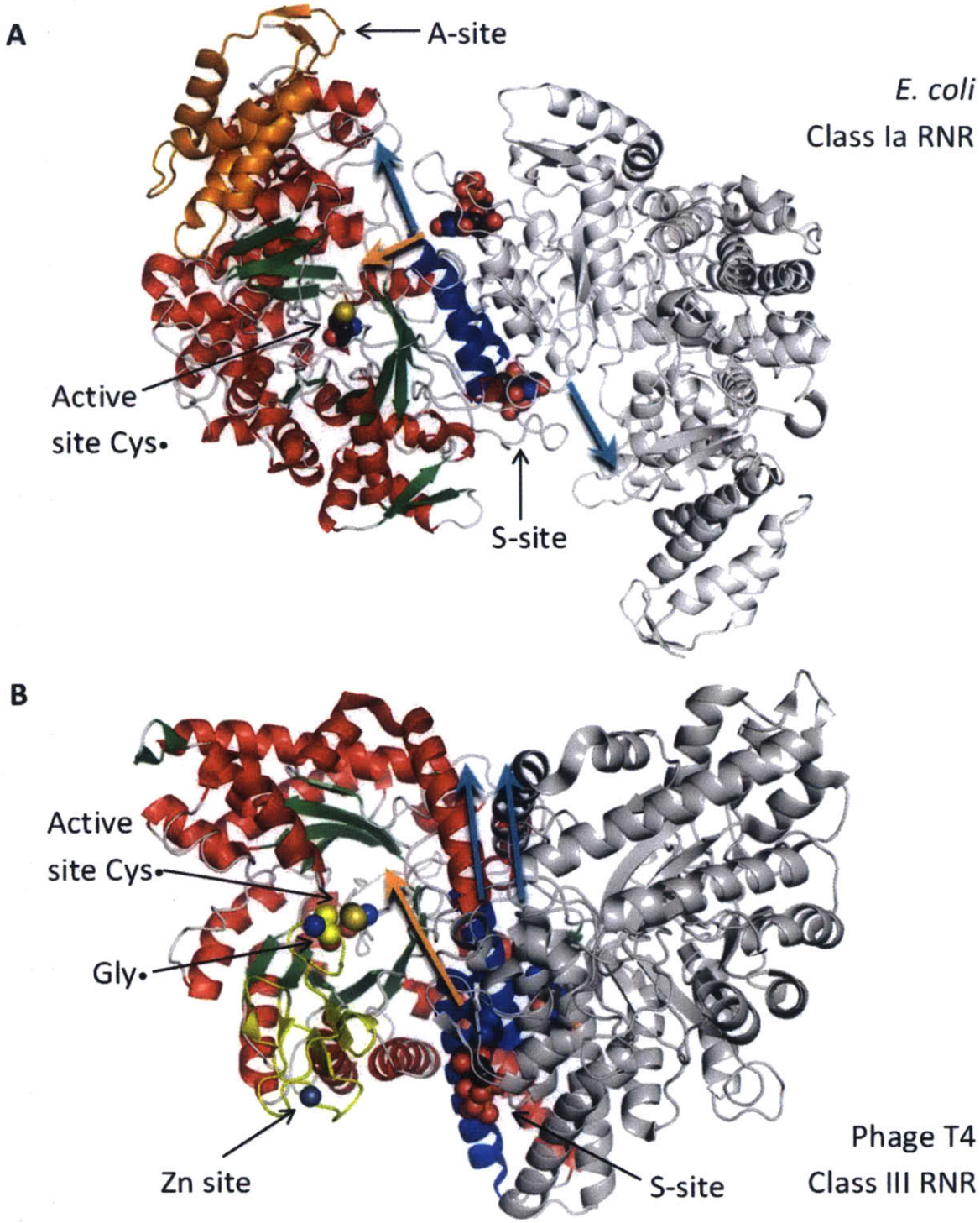


Figure 1.4 Crystal structures of A) *E. coli* class Ia RNR (PDB code 2R1R)[25] and B) bacteriophage T4 class III RNR (PDB code 1HK8)[26], showing the location of the allosteric and active sites, the orientation of the helices forming the dimer interface (blue arrows) and the separation between the allosteric and active sites (orange arrow). One monomer of the dimer structure is colored white. (Red = helices; green = sheets; blue = dimer interface; orange = ATP cone of class Ia RNR; yellow = Zn site of class III RNR; spheres = active site Cys• residue, Gly• residue, dNTP allosteric effectors, Zn in class III RNR)

1.2.2. The class III RNR is a G• enzyme with similarities to PFL

Initial studies of *E. coli* class III RNR involved fractionation of cell extracts to isolate components required to reconstitute RNR activity. The active component in one of the fractions was found to be SAM.[27] At the time of this discovery, apart from the role of SAM as a biological CH_3 donor, it was also known to participate in radical chemistry according to the mode of reactivity proposed for two enzymes: lysine 2,3-aminomutase (LAM)[28] involved in the lysine metabolism; and PFL activating enzyme (PFLAE)[29], involved in the generation of the G• cofactor in PFL, through pioneering work carried out in the laboratories of Perry A. Frey and Joachim Knappe respectively. These enzymes contain a conserved CXXXCXXC motif ligating a reduced $[4Fe4S]^+$ cluster, and initiate the radical-dependent reaction by reductive cleavage of a molecule of SAM to methionine and a 5′-deoxyadenosyl (5′-dA) radical (**Figure 1.2**). Later biochemical and bioinformatics studies expanded the scope of this "radical SAM superfamily" of enzymes to encompass a vast array of reactions in anaerobic biosynthesis of cofactors and secondary metabolites.[10,11,30]

In the pathway for anaerobic fermentation of sugars by *E. coli*, PFL catalyzes the non-oxidative cleavage of pyruvate to a molecule of acetyl-coA and a molecule of formate (**Figure 1.3C**).[31] The reaction requires an O_2-sensitive G• cofactor, which is generated on a conserved C-terminal peptide sequence by PFLAE through direct H-atom abstraction by the 5′-dA radical

(**Figure 1.2**). Once generated, the G• is stable under anaerobic conditions and acts as a radical "storage unit", used to reversibly generate a Cys thiyl radical required for initiation of the radical-dependent reaction in both PFL and RNR.

PFL has served as the paradigm for the "glycyl radical family" of enzymes, which includes the class III RNR, and is now known to include diverse enzymes involved in fermentation and anaerobic activation of hydrocarbons.[32,33] Similarities between PFL and RNR were noted soon after the discovery of the class III RNR. Investigations of the mechanism of G• generation in PFL and RNR have proceeded in parallel, and have established further links between the two enzymes, as reviewed by Reichard.[19]

Apart from the requirement for SAM, several other factors required for class III RNR were identified in the initial fractionation studies.[34] Activity for CTP reduction required Mg^{2+}, and was enhanced 5-fold by ATP, suggesting that, like the class I and II RNRs, the class III RNR is an allosteric enzyme. The active component of a metal-containing fraction was identified as K^+. The reaction also required NADPH, and two protein components essential for activity were found to be identical to NADPH:flavodoxin/ferredoxin reductase (cloned and named Fpr),[35] and flavodoxin (FldA),[36] previously known to provide electrons for activation of PFL (**Figure 1.2**). As with PFL, the FldA / Fpr / NADPH system could be substituted with photoreduced deazaflavin.[37] The reaction was inhibited by SAH,[27] a known inhibitor of radical SAM enzymes, and both methionine and 5′-dA were detected as products.[38] Based on the similarities with PFL, it was proposed that the class III RNR assays actually involved two separate steps: first an activation step involving SAM cleavage and generation of G•; and second the actual nucleotide reduction reaction.

Some conflicting observations were reported for the *M. marburgensis* RNR,[39] and the discrepancies are likely partly due to the *M. marburgensis* enzyme being isolated under strict anaerobic conditions with the preservation of G•, while the *E. coli* enzyme was isolated in the inactive form and thus required a separate activation step. Activity for CTP reduction by partially purified *M. marburgensis* enzyme was not affected by NADPH or NADH and is inhibited by SAM for unknown reasons. An air-sensitive organic radical at g = 2.0048 (width = 1.7 mT) was observed in the preparation.

Upon cloning of *nrdD*, it was apparent that the C-terminal peptide contained a sequence highly homologous to the peptide bearing G• in PFL.[20] The G• in PFL is known to react stoichiometrically with O_2, leading to cleavage of the protein backbone at that site, and the cleavage products and mechanism have been investigated in detail.[40] Like *E. coli* NrdA, *E. coli* NrdD binds tightly to dATP-sepharose and could thus be purified by affinity chromatography. The purified protein showed two bands of 77 and 74 kDa.[34] The relative intensity of the second band increased after activation and exposure to O_2, reflecting a loss of ~30 amino acids presumably due to cleavage at the G• site.[20] Similar observations were made for the bacteriophage T4 NrdD, which was known to undergo aerobic self-cleavage even before its function as an RNR was assigned.[21]

The organic radicals on activated *E. coli* and bacteriophage T4 NrdD were later detected by EPR (**Figure 1.5**).[8,37,41] Like PFL, the spectrum of the NrdD G• is dominated by a hyperfine from the α-H of Gly. Labeling with [^{2}H]- and [^{13}C]-Gly resulted in changes to the hyperfine coupling, confirming the assignment of the radical as G•. Mutation of the predicted radical-bearing G681 in *E. coli* NrdD or G580 in T4 NrdD to Ala resulted in no radical and no cleavage. The cleavage site in *E. coli* NrdD, mapped by mass spectrometry, was consistent with G681.[42]

The EPR spectrum of the PFL G• is more complex, with partially resolved hyperfine splittings likely arising from the nonexchangeable α protons of the two adjacent amino acids.[40] This is absent in the NrdD spectrum, suggesting a different conformation of the radical-bearing peptide. Also unlike the PFL G•, where exchange with D_2O is catalyzed by an active-site Cys, exchange does not occur in NrdDs. The precise g-values of several G• enzymes have since been measured using very high field EPR spectroscopy.[43]

Migration of NrdD during size exclusion chromatography suggested that it exists as a dimer, like PFL.[34] Also like PFL, a maximum of one G• is generated per NrdD dimer, proposed to reflect "half sites" reactivity.[44] The dimer structure and G• site was revealed in the crystal structure of the G580A mutant of T4 NrdD (the mutation was designed to prevent G• generation during protein expression).[26] The CH_3 group of A580 replaces the Pro-S α-H of Gly580 in the native enzyme, which faces toward the surface and is stereospecifically removed in PFL. A Gly α-C to Cys S distance of 5.2 Å suggested direct H-atom abstraction as the means for generation of the Cys thiyl radical. The structure of PFL showed essentially the same architecture despite almost no overall sequence similarity.[45,46]

A metabolic link between PFL and RNR was later established when formate, the product in the PFL reaction, was found to be a substrate for the class III RNR, serving as the reductant for nucleotide reduction.[13] Although RNR and PFL catalyze different reactions, commonalities in the structure, G• generation, method of substrate activation via a thiyl radical and reaction intermediates involved in the formation or reaction of formate were taken as evidence to suggest that PFL shares a common evolutionary root with the class III RNR and perhaps all RNRs.[32]

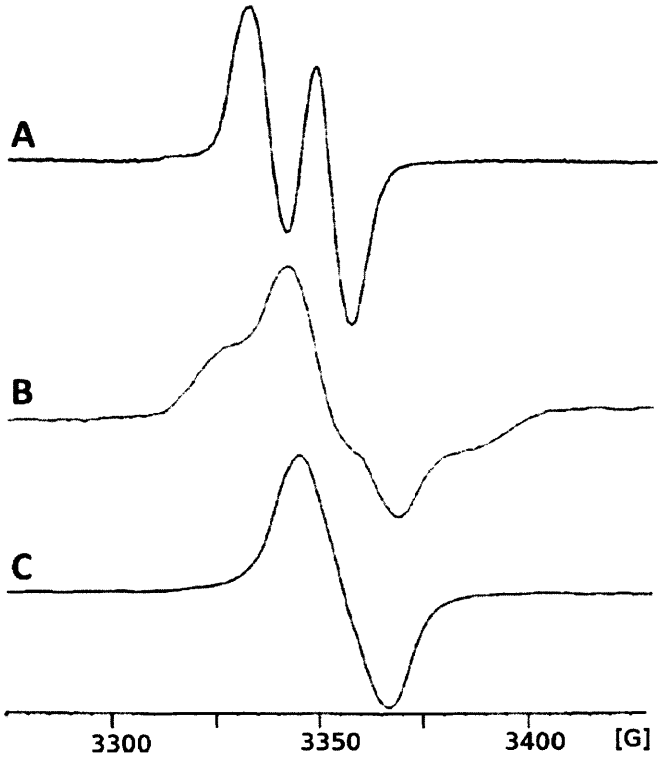


Figure 1.5 X-band EPR spectra of activated wild type and isotopically labeled NrdD, demonstrating that the organic radical is G•. A) Wild type NrdD. B) 2-[^{13}C]-Gly NrdD. C) [^{2}H]-Gly NrdD. (Temperature = 100 K) Figure slightly modified from reference[8].

1.2.3. Generation of the G• cofactor

1.2.3.1. A separate activase NrdG is needed for activity

Generation of G• in PFL was known to require a separate activating enzyme PFLAE bearing a [4Fe4S] cluster ligated by a conserved CXXXCXXC motif, encoded downstream of PFL in the same operon.[40] Initial preparations of purified *E. coli* NrdD contained an iron sulfur cluster,[37] and sequencing of *nrdD* revealed a conserved 4-Cys motif in the C-terminal domain of the protein, which was initially suggested to bind the cluster.[20] However, inspection of a sequenced fragment of the bacteriophage T4 genome showed that there was a small protein downstream of *nrdD* containing a CXXXCXXC motif like in PFLAE, even though the overall

sequence homology with PFLAE was weak.[47] This suggested that like PFL, a separate activating enzyme (named NrdG) might be required for G• generation in NrdD. By separate overexpression of the two T4 proteins, NrdG was shown to be required for NrdD activity in crude extracts.

Later, inspection of a stretch of DNA from *L. lactis* containing the *nrdD* gene showed that *nrdG* was present downstream in the same operon.[48] This motivated the further sequencing of a segment of the *E. coli* genome, where *nrdG* was also found downstream of *nrdD* in the same operon. *E. coli* NrdD and NrdG copurify as a tight complex, explaining the presence of the cluster in early NrdD purifications,[48] where NrdG eluded detection due to its small size (17.5 kDa). Preparations of NrdD from *E. coli* lacking *nrdG* did not contain an iron sulfur cluster. NrdG was found to be essential for anaerobic growth of *E. coli*, while deletion of PFLAE did not have any effect on anaerobic nucleotide reduction, showing that the activating enzymes are specific for their respective G• enzymes. *E. coli* NrdG was also unable to activate T4 NrdD.[47]

1.2.3.2. The active form of NrdG contains a 4Fe4S cluster ligated by three Cys

Unlike the stable iron sulfur clusters in electron transfer proteins, the [4Fe4S] cluster of PFLAE is labile, particularly in the presence of O_2.[49] It was proposed that this instability is due to ligation by only three protein-based Cys, as previously observed for the [4Fe4S] of aconitase. Determination of the active form of the iron sulfur cluster of PFLAE and NrdG occurred in parallel, and was complicated by this instability.

Initial attempts to reconstitute NrdG with iron and sulfide resulted in a [2Fe2S] cluster, with a loading of 0.5 per peptide.[50] Apo NrdG was monomeric, but dimerized upon cluster loading, leading to the suggestion of an interfacial cluster. The cluster form depended on purification conditions, and binding to NrdD followed by dATP affinity purification led to a

[3Fe-4S]$^{+}$ cluster.[51] However, upon reduction with dithionite, both these forms were converted into an EPR active [4Fe4S]$^{+}$ cluster (**Figure 1.6**).

Reaction of the *L. lactis* NrdG with various reductants and air was examined in detail.[52] Exposure of the [4Fe4S]$^{+}$ form to air resulted in the formation of a mixture of [3Fe4S] and [2Fe2S] clusters, which could be converted back to [4Fe4S]$^{+}$ by reduction with dithionite without addition of exogenous iron or sulfide, showing that the clusters were capable of rearrangements under activation conditions. Addition of excess dithionite resulted in the breakdown of the cluster to an inactive, presumably fully reduced form. Exposure of this mixture to air again forms a mixture of [3Fe4S] and [2Fe2S].

The formation of different cluster forms was later found to be caused by insufficiently anaerobic conditions during reconstitution, and stricter anaerobic conditions led to formation of [4Fe4S]$^{2+}$ clusters, with one cluster per NrdG peptide.[53] This ruled out binding of the cluster at the NrdG dimer interface, previously proposed. Exposure to air led to reversible conversion to [3Fe4S], and then to [2Fe2S], which could be converted back to [4Fe4S]$^{+}$ by reduction with dithionite. Thus it was concluded that like PFLAE, the active form of NrdG contains a [4Fe4S] cluster, bound to just three Cys in the CXXXCXXC motif conserved in radical SAM enzymes. Mutation of individual Cys residue to Ala residues resulted in NrdG that was still able to bind [4Fe4S], but was catalytically inactive, although binding of the [4Fe4S]-loaded mutants to NrdD was only slightly diminished.[54]

The unique Fe site of the [4Fe4S] cluster in radical SAM enzymes is occupied by SAM.[11] Filter binding experiments with NrdG showed that SAM only binds under reducing conditions, with a stoichiometry of one SAM per NrdG peptide and a K_d of 10 μM.[55] As with other radical SAM enzymes, binding of SAM resulted in a large change to the EPR signal of the [4Fe4S]$^{+}$

cluster (**Figure 1.6**).[55] The spectrum and SAM affinity were not altered in the presence of NrdD.[56] The mode of SAM binding was investigated using ENDOR and HYSCORE spectroscopy, which suggested coordination of the SAM NH_2 group to the open Fe of the [4Fe4S] cluster.[9]

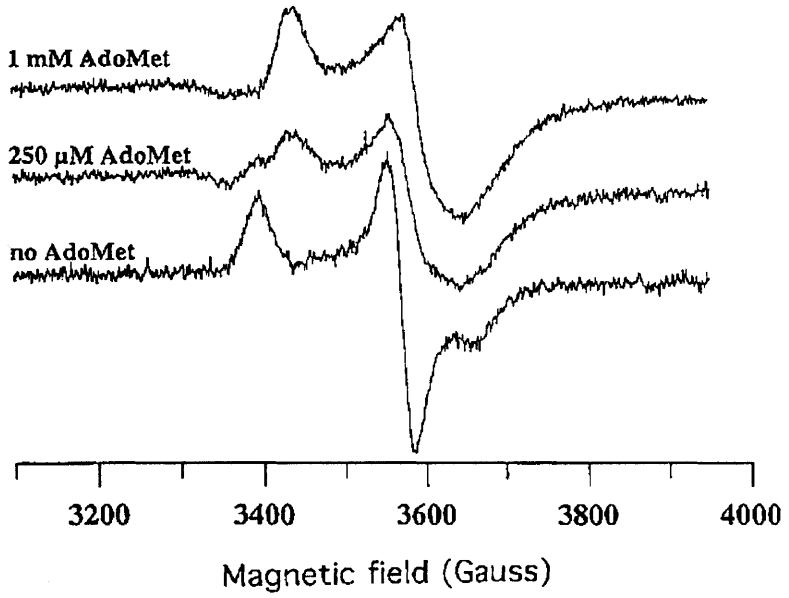


Figure 1.6 Effect of SAM on the EPR signal of the dithionite-reduced $[4Fe4S]^+$ center of NrdG. Binding of SAM to the unique Fe site results in changes to the axial signal of $[4Fe4S]^+$ (NrdG concentration = 140 µM; temperature = 10 K) Figure reproduced from reference[55].

1.2.3.3. NrdG is an "activase" required in catalytic amounts

Coexpressed *E. coli* NrdD and NrdG purify as a 1:1 complex by dATP affinity chromatography,[48] and SEC and sucrose gradient centrifugation under aerobic conditions with a slight excess of NrdG in the buffer, suggested an $\alpha_2\beta_2$ subunit structure.[57] NrdG alone exists as a monomer / dimer equilibrium mixture. Similar sucrose gradient centrifugation experiments with *L. lactis* NrdD and NrdG under aerobic conditions initially suggested that the two proteins did

not bind tightly to each other.[22] However, under these conditions, the ^{59}Fe label was lost on the top of the column, indicating that the cluster was labile. When the experiment was repeated under anaerobic conditions, an $\alpha_2\beta_2$ complex was observed with dATP as an effector, which was found to be necessary for NrdD dimerization and stability of the complex.[14]

In early assays, NrdG was used in stoichiometric complex with NrdD for the activation reaction. Subsequently, after optimization of cluster loading and thus NrdG activity under strict anaerobic conditions, NrdG was found to act catalytically, generating up to 5 G• per NrdG peptide.[53] This suggested that, like PFLAE, NrdG should be considered as an "activase" rather than a subunit of RNR, and that the binding of NrdG to NrdD, though tight, is reversible. The low microwave power saturation of G• suggested that it was not close to the NrdG cluster.

Unlike the *E. coli* enzyme, T4 NrdD and NrdG do not form a robust complex, but showed some copurification on a butyl-Sepharose column.[41] For the *L. lactis* enzyme, interaction between NrdD and NrdG is weak and, after activation, NrdD could be separated from NrdG by dATP affinity chromatography under anaerobic conditions without loss of G• and activity. This demonstrated that NrdG is only required for G• generation and not required for RNR activity.[14] The isolated NrdD G• was extremely stable on storage at 4°C in an anaerobic box, with only a 10% decrease in activity after 4 days.

Although the structure of the active complex of NrdG is not known, a high resolution structure has been determined for PFLAE bound with SAM and a 7-mer C-terminal peptide of PFL.[58] The [4Fe4S] cluster is coordinated by the CXXXCXXC motif, with the unique Fe coordinated by the NH_2 and CO_2^- of SAM, with a 3.2 Å distance between Fe and the sulfonium S of SAM, consistent with reductive cleavage by inner sphere electron transfer. The G• residue is positioned for direct H-atom abstraction, with a 4.1 Å distance between the 5′-position of the dA

moiety and the α-C of Gly. Compared to PFLAE (28 kDa), *E. coli* NrdG (17.5 kDa) is much smaller, missing a C-terminal portion of the enzyme. Also unlike PFLAE, NrdG forms a homodimer through an unknown interface.

The mode of interaction of NrdG with NrdD is not completely understood. The crystal structure of T4 NrdD shows that, like PFL, the G• residue is buried deep within the protein, and thus a significant conformational change is required to allow direct binding of NrdG to the G• peptide, possibly in an "open" conformation as proposed for PFL.[59] It is also not known how NrdG discriminates between activated and nonactivated NrdD to avoid interacting with NrdD that already contains G•. Compared to the G580A mutant, the crystal structure of wild-type T4 NrdD showed greater flexibility around G580, and it was proposed that formation of the planar delocalized G• might result in conformational changes lowering the affinity for NrdG.[60] Such a conformational change might also account for the formation of a maximum of one G• per dimer in both NrdD and PFL, by preventing the formation of the second G• or by destabilizing it.

1.2.3.4. Mechanism of activation of NrdD

The generation of G• by direct stereospecific abstraction of a Gly α-H by the 5′-dA radical (**Figure 1.2**), established for PFLAE, represents the simplest reaction catalyzed by a radical SAM enzyme.[11,61] As with other radical SAM enzymes, the generation of G• by NrdG is accompanied by an "uncoupling" reaction resulting in non-productive SAM cleavage. Using a complex of NrdD and NrdG photoreduced with deazaflavin, formation of 2 additional equivalents of methionine proceeded even after maximal G• and activity was achieved.[55]

To isolate NrdD and NrdG from the effects of excess reductants, NrdG was first pre-reduced to the $[4Fe4S]^{+}$ state with dithionite, followed by careful removal of excess reductant on

an anaerobic G-25 column.[62] Cleavage of SAM by pre-reduced NrdG is slow, but accelerated by at least 10 times in the presence of NrdD, concomitant with the generation of G• and oxidation of the $[4Fe4S]^{+}$ cluster to an EPR silent form. Under these conditions, a 1:1:1 stoichiometry of electron, Met and G• was observed. Support for a mechanism involving direct H atom abstraction was obtained using NrdD labeled with [^{2}H]-Gly.[9] Carrying out the reaction with pre-reduced NrdG resulted in generation of one [^{2}H]-5′-dA per G•.

In the absence of NrdD, SAM cleavage by pre-reduced NrdG is accelerated in the presence of DTT. It was suggested that DTT interacts with the cluster to trigger cleavage.[55] An alternative explanation is that a small population of 5′-dA radical forms in the absence of DTT, and the reaction is driven forward by its quenching by DTT. Curiously, up to 3 molecules of SAM were observed to be cleaved per reduced $[4Fe4S]^{+}$ cluster, which is oxidized to an EPR silent form.[62] The source of the extra electrons is unknown and has not been studied in detail, but a possible source is the disulfide anion radical that forms from one electron oxidation of DTT.[63] No reaction is observed with SAM and the oxidized $[4Fe4S]^{2+}$ cluster.

The FldA / Fpr / NADPH system serves as the electron source for NrdG in assays (**Figure 1.2**), and FldA is also believed to be the electron source *in vivo*.[56] The potential of the FldA oxidized / semiquinone (OX / SQ) couple is -256 mV, and the hydroquinone / semiquinone (HQ / SQ) couple is -440 mV (vs NHE at pH 7.0). Reduction of FldA by Fpr / NADPH gives only SQ, any HQ produced would disproportionate into SQ and OX. Both SQ and HQ can be produced by photoreduction with deazaflavin, and both support NrdD activation at same rate. During the anaerobic growth of *E. coli*, the major reduced form of FldA is believed to be HQ, produced by pyruvate-ferredoxin oxidoreductase.

The potential of the NrdG $[4Fe4S]^{2+/1+}$ couple was measured by direct electrochemistry on a carbon electrode and found to be very low (-550 mV vs NHE). Addition of SAM shifted is to an even lower value (-620 mV vs NHE). When $[4Fe4S]^{2+}$ NrdG is incubated with FldA SQ or HQ, no reduction to $[4Fe4S]^{+}$ is observed. Conversely, incubation of photoreduced NrdG with FldA results in formation of SQ. Using the FldA / Fpr / NADPH system for activation, acceleration of SAM cleavage occurred only in the presence of wild type NrdD and not with the G681A mutant, even though the mutation does not affect binding of NrdG or SAM. With FldA as the reductant, no SAM cleavage occurs in the absence of NrdD.[55]

The observations summarized in this section suggested that both the reduction of NrdG $[4Fe4S]^{2+}$ by FldA and cleavage of SAM to generate the 5′-dA are thermodynamically unfavorable, and only driven forward by the exergonic formation of G•, due to the large difference between the C-H bond dissociation energies of 5′-dA (100 kcal/mol) and Gly (79 kcal/mol).[56]

1.2.4. Formate is the reductant for the class III RNR

Reducing equivalents for nucleotide reduction in assays of the class Ia and class II RNRs can be provided by the thioredoxin (TrxA) / thioredoxin reductase (TrxB) / NADPH system, or by small molecule thiols like DTT.[12] However, early experiments suggested that this was not true for the class III RNR. Nucleotide reduction by anaerobically partially purified *M. marburgensis* class III RNR proceeded linearly only for 10 min and was not enhanced by a range of potential reductants, including DTT, dihydrolipoate, and *M. marburgensis* thioredoxin, suggesting that the reductant for this enzyme was different from the class I and II RNRs.[39]

Identification of the reductant in the *E. coli* system was hampered by the requirement of an electron source for the activation system, which was difficult to subsequently remove due to the sensitivity of the enzyme. Purification of the activated *E. coli* NrdD was eventually achieved through anaerobic dATP affinity chromatography, allowing experiments to determine the source of electrons for nucleotide reduction.[13] In the process of isolating the active component present in a heat stable, low molecular weight fraction of boiled *E. coli* extract, a large amount of activity was recovered and traced to the formate-containing buffer used in chromatography, suggesting that the reductant might be formate. Indeed, the active component of the extract displayed properties of an organic acid during purification, and formate was later detected in the extract. Activity assays with [^{14}C]-formate showed that CTP reduction is accompanied by oxidation of formate to [^{14}C]-CO_2, with a stoichiometry of 1:1. The reaction was also inhibited by the formate analog hypophosphite, a known inhibitor of PFL.

Separation of *L. lactis* NrdD from NrdG after activation by dATP affinity chromatography allowed the requirements of the activation reaction and the RNR reaction to be separately determined.[14] Reduction of CTP required K^+, Mg^{2+} and formate but not DTT, and was stimulated 5-fold by the effector ATP.

1.2.5. A Zn site in the C-terminal domain of NrdD

The C-terminal domain of *E. coli* NrdD contains a 4-Cys motif prior to the G•-containing peptide. Three of the Cys are conserved and one is partially conserved (**Figure 1.7**), and various theories have been proposed regarding the function of this motif. Initially, it was proposed to bind an iron sulfur cluster thought to be present in purified *E. coli* NrdD.[20] However the cluster

was later shown to be associated with copurifying NrdG, and is absent in NrdD preparations from bacteria lacking NrdG.[48]

Mutagenesis studies in T4 NrdD showed that the C-terminal Cys residues were individually essential for NrdD activation.[64] The Cys to Ala mutants do not abrogate the binding of NrdG, but decreased the rate of SAM cleavage, and prevented G• formation. Thus it was hypothesized that these residues may form a radical translocation pathway between the 5′-dA radical and Gly. Since SAM cleavage in the absence of NrdD was known to require DTT, another proposal was that the Cys residues bind and activate the 4Fe4S cluster of NrdG, facilitating SAM cleavage.

Subsequently, purified NrdD was shown to contain a disulfide formed between two of these Cys residues.[65] Formation of G• required reduction of the disulfide by either DTT or by thioredoxin, elucidating the role of DTT, previously shown to be essential for the NrdD activation. The 4-Cys motif was then proposed to function as a redox switch, preventing G• formation under oxidizing conditions.

The aforementioned theories were subsequently disputed based on an updated crystal structure of T4 NrdD.[60] In the initial structure, the 4-Cys motif was disordered, but a later structure showed that the site chelated a metal ion. Fe and Zn K-edge fluorescence showed the presence of Zn and a small amount of Fe. Anomalous dispersion was consistent with the metal being Zn. The motif was structurally similar to the Zn-ribbon motif present in transcription factors, and has some similarities to Fe-binding rubredoxins. Since heterologous expression of rubredoxin is known to lead to mixtures of Zn and Fe bound forms, it was unclear what the physiologically relevant metal was.

Later studies elucidated the identity of the bound metal as Zn. Expression of the protein in media containing 50 μM Zn acetate led to protein stoichiometrically loaded with Zn.[44] The metal bound with a high affinity constant ($K_a = 2.9 \times 10^{17}\ M^{-1}$), and was not completely removed even with the strong chelator Tetrakis(2-pyridylmethyl)ethane-1,2-diamine (TPEN). Exposure to air had no effect on the Zn content of NrdD, suggesting that the Zn site is not a redox switch, unlike in some other proteins with a similar motif. Activity obtained with this preparation was 1000-1500 nmol / min.mg for CTP reduction, which was 2-5 times larger than previous reports and in the range of that of the two other classes of RNR. Activity also increased with increasing Zn loading, suggesting that this was the physiologically relevant metal.

Binding of Zn protected the protein from proteolysis and prevented the formation of disulfides between the Cys residues. With the Zn-loaded protein, thioredoxin or DTT were dispensable for the formation of G•, although for unknown reasons they were still required for full CTP reduction activity. Curiously, the amount of G• formed was not fully accounted for by the Zn content of the protein, suggesting that radical can also be generated on Zn-free NrdD. A possible explanation is metalation by labile Fe from NrdG during the activation reaction.

Interestingly, despite their high redox potential of -100 to +50 mV, a rubredoxin has recently been identified as the reductant for the non-canonical radical SAM enzyme Dph2.[66] For NrdD, however, not all the four Cys residues in the motif are conserved (**Figure 1.7**), arguing against a conserved redox-active rubredoxin domain. The binding of a non-redox active Zn to the 4-Cys site suggested a structural role for the motif, in facilitating specific interactions with NrdG or promoting the stability of G•.[44] It was noted that a similar Cys motif of unknown function is also present at the C terminus of several class II RNRs.[60]

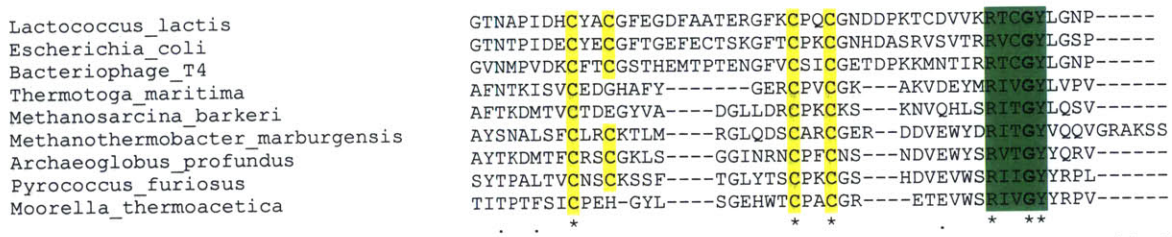


Figure 1.7 Alignment of a segment of the C-terminal domain of NrdD, showing the Zn-binding Cys residues (highlighted in yellow) and the consensus sequence for the G• peptide (highlighted in green).

1.2.6. Allosteric regulation in class III RNRs

Assays in cell extracts of *M. marburgensis* and later in *E. coli* suggested that the class III RNR reduces all four ribonucleotides (A,G,C,U), and that it is an allosteric enzyme.[15,17] Similarities with the allosteric scheme of the class I and II RNRs were taken as evidence suggesting a common ancestor for all RNRs.

Class Ia RNRs, like those in *E. coli* and in eukaryotes, were known to have two allosteric sites: a specificity site (S-site) at the dimer interface of α_2, which dictates which substrate (A,G,C,U) is reduced; and an activity site (A-site) consisting of an N-terminal ATP cone domain, which stimulates or inhibits activity when binding ATP or dATP respectively (**Figure 1.8A**).[67] The S-site is common to all RNRs, but the A-site cone is only present in some enzymes.[68] Class Ib RNRs, like that of *Corynebacterium ammoniagenes*, lack the ATP cone domain. The ATP cone is present in some class II RNRs (e.g. *Pyrococcus furiosus*) but is absent in others (e.g. *Lactobacillus leichmanii*).

The sequence of *E. coli* NrdD showed that the N-terminal 95 amino acids has a high degree of homology with *E. coli* NrdA,[20] suggesting an activity-regulating ATP cone domain is present in NrdD. This domain is also present in *L. lactis* and *M. marburgensis* NrdD, and the

activity of all three were shown to be stimulated by ATP and inhibited by dATP. The sequence of the T4 NrdD showed that it lacks the 95 N-terminal amino acids, and assays showed that it was not inhibited by dATP and was thus constitutively active,[69] similar to the T4 class Ia RNR previously characterized.

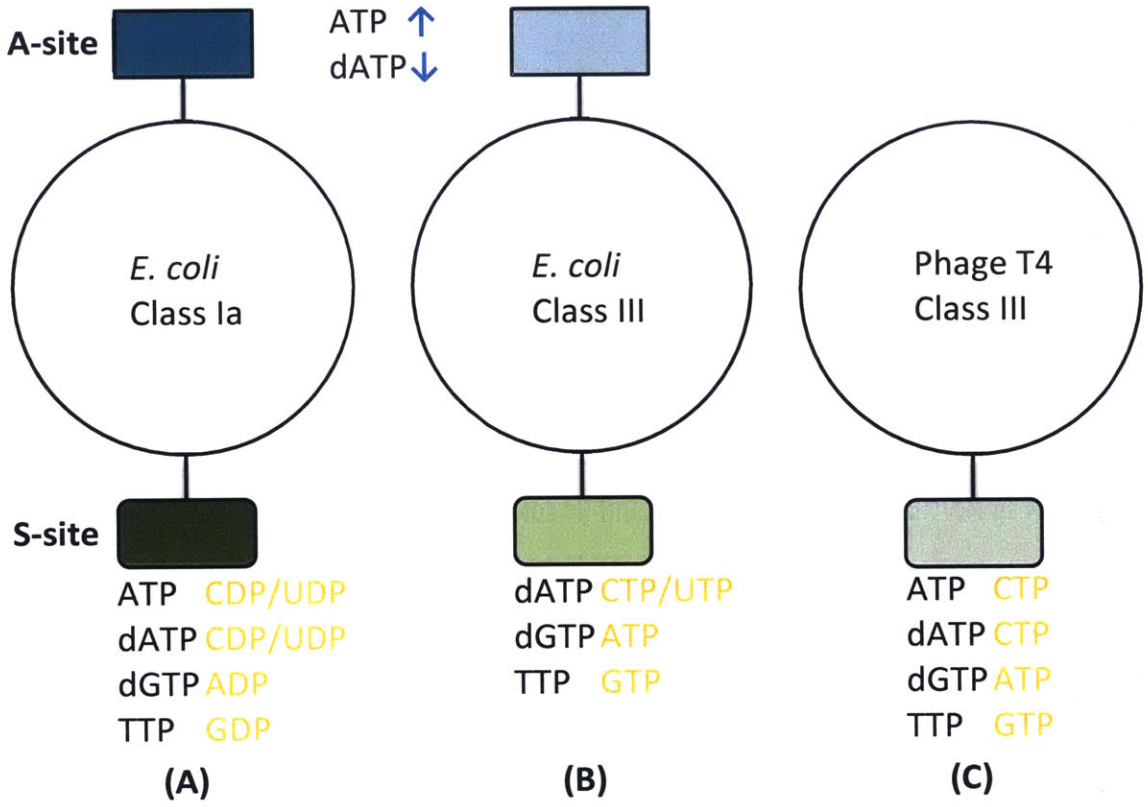


Figure 1.8 Allosteric regulation of the *E. coli* class Ia, *E. coli* class III and bacteriophage T4 class III RNRs. A) A schematic representation of the pattern of regulation of the *E. coli* class Ia RNR. The activity site (A-site) is in blue and the specificity site (S-site) in green. For simplicity the protein is represented as a single unit instead of a dimer. The specificity effectors are listed in black below each RNR, next to the nucleotides whose reduction they induce, in orange. The blue arrows next to the activity site represent overall up- and downregulation by ATP and dATP respectively. B) Allosteric regulation of the *E. coli* class III RNR. C) Allosteric regulation of the bacteriophage T4 class III RNR, where the activity site is absent, and the enzyme does not reduce UTP. Figure adapted from reference[70].

In the class Ia RNRs of *E. coli* and of eukaryotes, the active complex requires the association of NrdA and NrdB along a specific interface, and inhibition by dATP is thought to occur through the formation of inactive quaternary structures.[71,72] The active form of the class III RNR is thought to be a NrdD α_2 dimer, and the mechanisms of activity regulation in class III and class II RNRs have not been investigated in detail.

The nature of the nucleotide-binding allosteric sites in *E. coli* NrdD was investigated by activity assays with different nucleotide effectors, and by binding assays with radiolabeled nucleotides using a rapid filtration setup.[73] Activity was stimulated by up to 10-fold by the "correct" effector (dGTP for ATP reduction; ATP for CTP and UTP reduction; TTP for GTP reduction, **Figure 1.8B**). dGTP and TTP inhibited the reduction of the "incorrect" substrate, and dATP inhibited reduction of all four. The allosteric effectors alter the binding and K_M of the substrate, but not the k_{cat} of the reaction. No cooperativity was observed in the binding of substrate and effector molecules.

Assays with competing effectors suggested that there were two allosteric sites, which we refer to as the "S-site" and "A-site" because of shared properties with the analogous sites in *E. coli* class Ia RNR (**Figure 1.8**, early papers refer to these as the "purine site" and "pyrimidine site" respectively).[73] The S-site binds TTP, dGTP and dATP, and the A-site binds ATP and dATP. Unlike the *E. coli* NrdA, NrdD binds 1 rather than 2 ATP, which was taken to suggest that the S-site does not bind ATP. *E. coli* NrdD was also found to bind 2-3 dATP per monomer. For *L. lactis* NrdD, dATP was found to induce dimerization during sucrose gradient centrifugation, while ATP does not.[22]

The allosteric regulation of *L. lactis* is nearly identical to *E. coli* NrdD,[22] but that of T4 NrdD differs in some ways (**Figure 1.8C**).[69] T4 NrdD lacks activity regulation and only has one

allosteric site. It is also inactive in UTP reduction. This was rationalized by the presence in the phage of a highly active dCTP phosphatase, generating dCMP, which is converted to TMP by T4 dCMP deaminase and thymidylate synthase, and to the phage-specific DNA precursor hydroxymethyl-dCTP (hm-dCTP) by T4 dCMP hydroxymethyltransferase. In T4 NrdD, dCTP (and presumably hm-dCTP) is a very weak positive effector, like the T4 class Ia RNR.

Early attempts to map the allosteric sites employed photoaffinity labeling with radiolabeled 8-azido-ATP.[74] Subsequently, the allosteric S-site was revealed in the crystal structure of T4 NrdD.[26] The structure showed that despite the similarities in the allosteric scheme, this allosteric site at the dimer interface of NrdD was dramatically different from the class I RNR.

The S-site in Class I RNRs is located at the α_2 dimer interface, and consists of a short four-helix bundle, two helices from each α oriented in an antiparallel fashion (**Figure 1.4A**).[25] The structure of the interface is also shared with dimeric class II RNRs like in *Thermotoga maritima*.[75] The structure of the monomeric *Lactobacillus leichmanii* class II RNR revealed that, despite the different quaternary structure, it contains a two-helix insertion that structurally mimics this interface.[76]

The corresponding two helices that form the dimer interface in the class I and II RNRs also form the dimer interface of NrdD (**Figure 1.4B**).[26] However, in NrdD the helices in the two α units are oriented in a parallel fashion, and are lengthened and modified, creating an effector binding site that is completely different from the class I and II RNRs.

Allosteric regulation in the Class I and II RNRs is intermonomeric, with the effector binding adjacent to and regulating the active site of the other subunit. In contrast, in the class III RNR it is intramonomeric, with a larger separation between the corresponding allosteric and active sites. The crystal structure of T4 NrdD bound to all four allosteric effectors has been

determined, providing the structural basis for allosteric control.[70] Binding of different effectors results in remodelling of a loop terminating in the active site. Discrimination between effectors is largely achieved through side chain interactions, as opposed to backbone interactions in the class I RNRs. Despite the dramatic differences in the dimer interface, the allosteric information is transmitted to the same active site loop. Therefore it was proposed that, while effector recognition sites are highly divergent, interactions near the active site are conserved in all RNRs.

Unfortunately, structures could not be obtained with substrate bound in the active-site, either by soaking in or co-crystallization substrate (CTP / dATP), both with wild type NrdD and with the G580A mutant. However, soaking with high concentration of nucleotides resulted in binding in a "flipped" conformation, with the triphosphate moiety bound to the protein and the nucleoside moiety flipped out of the active site (this appears to be a general behavior as the substrate-bound structure of *T. maritima* class II RNR shows the NDP bound in a mixture of normal and "flipped" conformations)[75]. Nevertheless the position of the triphosphate provided a constraint that could be used to model the nucleoside in the active site.

1.3. Chapter preview

The objective of this thesis project is to investigate the mechanism of the class III anaerobic RNR. We started by characterizing the prototypical enzyme from *E. coli*. Subsequent bioinformatics investigations revealed the mechanistic diversity of class III RNRs, which led to the cloning and characterization of a further two subtypes of the enzyme from *N. bacilliformis* and *M. barkeri*.

Unlike the class I and II RNRs, where reducing equivalents for the reaction are delivered by a redoxin (thioredoxin, glutaredoxin or NrdH) via a pair of conserved active-site cysteines,

the class III RNRs examined prior to the start of this thesis project use formate as the reductant. In **Chapter 2**, we report that reaction of the *E. coli* class III RNR with CTP (substrate) and ATP (allosteric effector) in the absence of formate (reductant) leads to loss of the G• concomitant with stoichiometric formation of a new radical species and a "trapped" cytidine derivative, proposed to be 3′-keto-deoxycytidine. Addition of formate to the new species results in recovery of G• and reduction of the cytidine derivative to dCTP. The structure of the new radical was identified by 9.5 and 140 GHz EPR spectroscopy to be a thiosulfuranyl radical [RSSR2]•, composed of a cysteine thiyl radical stabilized by an interaction with a methionine residue. The presence of a stable radical species on the reaction pathway rationalizes the previously reported [^{3}H]-(k_{cat}/K_M) isotope effect of 2.3 with [^{3}H]-formate, requiring formate to exchange between the active site and solution during nucleotide reduction. Analogies with the disulfide anion radical proposed to provide the reducing equivalent to the 3′-keto-deoxycytidine intermediate by the class I and II RNRs provide further evidence for the involvement of thiyl radicals in the reductive half reaction catalyzed by all RNRs.

Redoxins, which serve as the reductant for the class I and II RNRs, are ubiquitous proteins present in all organisms. In contrast, formate, the reductant for the class III RNRs studied to date (from *E. coli*, *L. lactis* and bacteriophage T4), is not known to be produced by all anaerobic organisms containing this RNR. In **Chapter 3**, we report the cloning and heterologous expression of the *N. bacilliformis* class III RNR, and show that it can catalyze nucleotide reduction using the ubiquitous thioredoxin / thioredoxin reductase / NADPH system. We present a structural model based on a crystal structure of the homologous *Thermotoga maritima* class III RNR, showing its architecture and the position of conserved residues in the active site.

Phylogenetic studies suggest that this form of class III RNR is present in bacteria and archaea that carry out diverse types of anaerobic metabolism.

RNR has been proposed to provide the link between the RNA and DNA worlds. In **Chapter 4**, we review the distribution of different RNR subtypes, and possible scenarios for the evolution of RNR.

Certain methanogenic archaea contain a phylogenetically distinct third subtype of class III RNR, with distinct active-site residues. In the appendix **Chapter 5**, we report the cloning and heterologous expression of the *M. barkeri* class III RNR, and show that it couples ribonucleotide reduction to the oxidation of ferredoxin via a [4Fe4S] protein, ferredoxin : thioredoxin reductase, and a conserved thioredoxin-like protein, NrdH, present in the RNR operon. Ferredoxin is used as an electron carrier in methanogenesis, and the correlation between the electron donors used for ribonucleotide reduction and those used in energy metabolism is discussed.

1.4. References

(1) Nordlund, P.; Reichard, P. *Annu. Rev. Biochem.* **2006**, *75*, 681.

(2) Hofer, A.; Crona, M.; Logan, D. T.; Sjöberg, B.-M. *Crit. Rev. Biochem. Mol. Biol.* **2012**, *47*, 50.

(3) Licht, S.; Stubbe, J. In *Comprehensive Natural Products Chemistry*; Barton, S., Nakanishi, K., Meth-Cohn, O., Poulter, C., Eds.; Elsevier Science: New York, 1999; Vol. 5, p 163.

(4) Stubbe, J. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 2723.

(5) Cotruvo, J. A.; Stubbe, J. *Annu. Rev. Biochem.* **2011**, *80*, 733.

(6) Lenz, R.; Giese, B. *J. Am. Chem. Soc.* **1997**, *119*, 2784.

(7) Minnihan, E. C.; Nocera, D. G.; Stubbe, J. *Acc. Chem. Res.* **2013**.

(8) Sun, X.; Ollagnier, S.; Schmidt, P. P.; Atta, M.; Mulliez, E.; Lepape, L.; Eliasson, R.; Gräslund, A.; Fontecave, M.; Reichard, P.; Sjöberg, B. M. *J. Biol. Chem.* **1996**, *271*, 6827.

(9) Gambarelli, S.; Luttringer, F.; Padovani, D.; Mulliez, E.; Fontecave, M. *ChemBioChem* **2005**, *6*, 1960.

(10) Sofia, H. J.; Chen, G.; Hetzler, B. G.; Reyes-Spindola, J. F.; Miller, N. E. *Nucleic Acids Res.* **2001**, *29*, 1097.

(11) Wang, S. C.; Frey, P. A. *Trends Biochem. Sci.* **2007**, *32*, 101.

(12) Mao, S. S.; Holler, T. P.; Yu, G. X.; Bollinger, J. M., Jr.; Booker, S.; Johnston, M. I.; Stubbe, J. *Biochemistry* **1992**, *31*, 9733.

(13) Mulliez, E.; Ollagnier, S.; Fontecave, M.; Eliasson, R.; Reichard, P. *Proc. Natl. Acad. Sci. U. S. A.* **1995**, *92*, 8759.

(14) Torrents, E.; Eliasson, R.; Wolpher, H.; Gräslund, A.; Reichard, P. *J. Biol. Chem.* **2001**, *276*, 33488.

(15) Fontecave, M.; Eliasson, R.; Reichard, P. *Proc. Natl. Acad. Sci. U. S. A.* **1989**, *86*, 2147.

(16) Sprengel, G.; Follmann, H. *FEBS Lett.* **1981**, *132*, 207.

(17) Hogenkamp, H.; Follmann, H.; Thauer, R. *FEBS Lett.* **1987**, *219*, 197.

(18) Barlow, T. *Biochem. Biophys. Res. Commun.* **1988**, *155*, 747.

(19) Reichard, P. *Science* **1993**, *260*, 1773.

(20) Sun, X.; Harder, J.; Krook, M.; Jörnvall, H.; Sjöberg, B. M.; Reichard, P. *Proceedings of the National Academy of Sciences* **1993**, *90*, 577.

(21) Young, P.; Ohman, M.; Xu, M. Q.; Shub, D. A.; Sjöberg, B. M. *J. Biol. Chem.* **1994**, *269*, 20229.

(22) Torrents, E.; Buist, G.; Liu, A.; Eliasson, R.; Kok, J.; Gibert, I.; Gräslund, A.; Reichard, P. *J. Biol. Chem.* **2000**, *275*, 2463.

(23) Lundin, D.; Torrents, E.; Poole, A.; Sjöberg, B.-M. *BMC Genomics* **2009**, *10*, 589.

(24) Lundin, D.; Gribaldo, S.; Torrents, E.; Sjöberg, B.-M.; Poole, A. M. *BMC evolutionary biology* **2010**, *10*, 383.

(25) Eriksson, M.; Uhlin, U.; Ramaswamy, S.; Ekberg, M.; Regnström, K.; Sjöberg, B.-M.; Eklund, H. *Structure* **1997**, *5*, 1077.

(26) Logan, D. T.; Andersson, J.; Sjöberg, B.-M.; Nordlund, P. *Science* **1999**, *283*, 1499.

(27) Eliasson, R.; Fontecave, M.; Jornvall, H.; Krook, M.; Pontis, E.; Reichard, P. *Proc. Natl. Acad. Sci. U. S. A.* **1990**, *87*, 3314.

(28) Wang, S. C.; Frey, P. A. *Biochemistry* **2007**, *46*, 12889.

(29) Knappe, J.; Neugebauer, F. A.; Blaschkowski, H. P.; Gänzler, M. *Proceedings of the National Academy of Sciences* **1984**, *81*, 1332.

(30) Shisler, K. A.; Broderick, J. B. *Curr. Opin. Struct. Biol.* **2012**, *22*, 701.

(31) Knappe, J.; Sawers, G. *FEMS Microbiol. Lett.* **1990**, *75*, 383.

(32) Eklund, H.; Fontecave, M. *Structure* **1999**, *7*, R257.

(33) Lehtiö, L.; Goldman, A. *Protein Engineering Design and Selection* **2004**, *17*, 545.

(34) Eliasson, R.; Pontis, E.; Fontecave, M.; Gerez, C.; Harder, J.; Jornvall, H.; Krook, M.; Reichard, P. *J. Biol. Chem.* **1992**, *267*, 25541.

(35) Bianchi, V.; Reichard, P.; Eliasson, R.; Pontis, E.; Krook, M.; Jörnvall, H.; Haggård-Ljungquist, E. *J. Bacteriol.* **1993**, *175*, 1590.

(36) Bianchi, V.; Eliasson, R.; Fontecave, M.; Mulliez, E.; Hoover, D. M.; Matthews, R. G.; Reichard, P. *Biochem. Biophys. Res. Commun.* **1993**, *197*, 792.

(37) Mulliez, E.; Fontecave, M.; Gaillard, J.; Reichard, P. *J. Biol. Chem.* **1993**, *268*, 2296.

(38) Harder, J.; Eliasson, R.; Pontis, E.; Ballinger, M.; Reichard, P. *J. Biol. Chem.* **1992**, *267*, 25548.

(39) Sze, I. S. Y.; McFarlan, S. C.; Spormann, A.; Hogenkamp, H. P. C.; Follmann, H. *Biochem. Biophys. Res. Commun.* **1992**, *184*, 1101.

(40) Wagner, A. F.; Frey, M.; Neugebauer, F. A.; Schäfer, W.; Knappe, J. *Proc. Natl. Acad. Sci. U. S. A.* **1992**, *89*, 996.

(41) Young, P.; Andersson, J.; Sahlin, M.; Sjöberg, B.-M. *J. Biol. Chem.* **1996**, *271*, 20770.

(42) King, D. S.; Reichard, P. *Biochem. Biophys. Res. Commun.* **1995**, *206*, 731.

(43) Duboc-Toia, C.; Hassan, A. K.; Mulliez, E.; Ollagnier-de Choudens, S.; Fontecave, M.; Leutwein, C.; Heider, J. *J. Am. Chem. Soc.* **2003**, *125*, 38.

(44) Luttringer, F.; Mulliez, E.; Dublet, B.; Lemaire, D.; Fontecave, M. *J. Biol. Inorg. Chem.* **2009**, *14*, 923.

(45) Becker, A.; Fritz-Wolf, K.; Kabsch, W.; Knappe, J.; Schultz, S.; Wagner, A. V. *Nat. Struct. Mol. Biol.* **1999**, *6*, 969.

(46) Leppänen, V.-M.; Merckel, M. C.; Ollis, D. L.; Wong, K. K.; Kozarich, J. W.; Goldman, A. *Structure* **1999**, *7*, 733.

(47) Young, P.; Ohman, M.; Sjöberg, B. *J. Biol. Chem.* **1994**, *269*, 27815.

(48) Sun, X.; Eliasson, R.; Pontis, E.; Andersson, J.; Buist, G.; Sjöberg, B.-M.; Reichard, P. *J. Biol. Chem.* **1995**, *270*, 2443.

(49) Buis, J. M.; Broderick, J. B. *Arch. Biochem. Biophys.* **2005**, *433*, 288.

(50) Ollagnier, S.; Meier, C.; Mulliez, E.; Gaillard, J.; Schuenemann, V.; Trautwein, A.; Mattioli, T.; Lutz, M.; Fontecave, M. *J. Am. Chem. Soc.* **1999**, *121*, 6344.

(51) Mulliez, E.; Ollagnier-de Choudens, S.; Meier, C.; Cremonini, M.; Luchinat, C.; Trautwein, A. X.; Fontecave, M. *J. Biol. Inorg. Chem.* **1999**, *4*, 614.

(52) Liu, A.; Gräslund, A. *J. Biol. Chem.* **2000**, *275*, 12367.

(53) Tamarit, J.; Mulliez, E.; Meier, C.; Trautwein, A.; Fontecave, M. *J. Biol. Chem.* **1999**, *274*, 31291.

(54) Tamarit, J.; Gerez, C.; Meier, C.; Mulliez, E.; Trautwein, A.; Fontecave, M. *J. Biol. Chem.* **2000**, *275*, 15669.

(55) Ollagnier, S.; Mulliez, E.; Schmidt, P. P.; Eliasson, R.; Gaillard, J.; Deronzier, C.; Bergman, T.; Gräslund, A.; Reichard, P.; Fontecave, M. *J. Biol. Chem.* **1997**, *272*, 24216.

(56) Mulliez, E.; Padovani, D.; Atta, M.; Alcouffe, C.; Fontecave, M. *Biochemistry* **2001**, *40*, 3730.

(57) Ollagnier, S.; Mulliez, E.; Gaillard, J.; Eliasson, R.; Fontecave, M.; Reichard, P. *J. Biol. Chem.* **1996**, *271*, 9410.

(58) Vey, J. L.; Yang, J.; Li, M.; Broderick, W. E.; Broderick, J. B.; Drennan, C. L. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 16137.

(59) Peng, Y.; Veneziano, S. E.; Gillispie, G. D.; Broderick, J. B. *J. Biol. Chem.* **2010**, *285*, 27224.

(60) Logan, D. T.; Mulliez, E.; Larsson, K. M.; Bodevin, S.; Atta, M.; Garnaud, P. E.; Sjöberg, B. M.; Fontecave, M. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 3826.

(61) Frey, M.; Rothe, M.; Wagner, A. F.; Knappe, J. *J. Biol. Chem.* **1994**, *269*, 12432.

(62) Padovani, D.; Thomas, F.; Trautwein, A. X.; Mulliez, E.; Fontecave, M. *Biochemistry* **2001**, *40*, 6713.

(63) Elliot, A. J.; Simsons, A. S.; Sopchyshyn, F. C. *Radiation Physics and Chemistry (1977)* **1984**, *23*, 377.

(64) Andersson, J.; Westman, M.; Sahlin, M.; Sjöberg, B.-M. *J. Biol. Chem.* **2000**, *275*, 19449.

(65) Padovani, D.; Mulliez, E.; Fontecave, M. *J. Biol. Chem.* **2001**, *276*, 9587.

(66) Dong, M.; Su, X.; Dzikovski, B.; Dando, E. E.; Zhu, X.; Du, J.; Freed, J. H.; Lin, H. *J. Am. Chem. Soc.* **2014**, *136*, 1754.

(67) Reichard, P. *Arch. Biochem. Biophys.* **2002**, *397*, 149.

(68) Jonna, V. R.; Crona, M.; Rofougaran, R.; Lundin, D.; Johansson, S.; Brännström, K.; Sjöberg, B.-M.; Hofer, A. *J. Biol. Chem.* **2015**.

(69) Andersson, J.; Westman, M.; Hofer, A.; Sjöberg, B.-M. *J. Biol. Chem.* **2000**, *275*, 19443.

(70) Larsson, K.-M.; Andersson, J.; Sjöberg, B.-M.; Nordlund, P.; Logan, D. T. *Structure* **2001**, *9*, 739.

(71) Fairman, J. W.; Wijerathna, S. R.; Ahmad, M. F.; Xu, H.; Nakano, R.; Jha, S.; Prendergast, J.; Welin, R. M.; Flodin, S.; Roos, A. *Nat. Struct. Mol. Biol.* **2011**, *18*, 316.

(72) Zimanyi, C. M.; Ando, N.; Brignole, E. J.; Asturias, F. J.; Stubbe, J.; Drennan, C. L. *Structure* **2012**, *20*, 1374.

(73) Eliasson, R.; Pontis, E.; Sun, X.; Reichard, P. *J. Biol. Chem.* **1994**, *269*, 26052.

(74) Olcott, M. C.; Andersson, J.; Sjöberg, B.-M. *J. Biol. Chem.* **1998**, *273*, 24853.

(75) Larsson, K.-M.; Jordan, A.; Eliasson, R.; Reichard, P.; Logan, D. T.; Nordlund, P. *Nat. Struct. Mol. Biol.* **2004**, *11*, 1142.

(76) Sintchak, M. D.; Arjara, G.; Kellogg, B. A.; Stubbe, J.; Drennan, C. L. *Nat Struct Biol* **2002**, *9*, 293.

Chapter 2:

A Chemically Competent Thiosulfuranyl Radical on the *Escherichia coli* Class III Ribonucleotide Reductase

Adapted from Wei, Y., Mathies, G., Yokoyama, K., Chen, J., Griffin, R. G., and Stubbe, J. (2014) *J. Am. Chem. Soc.*, **136**(25), 9001-9013.

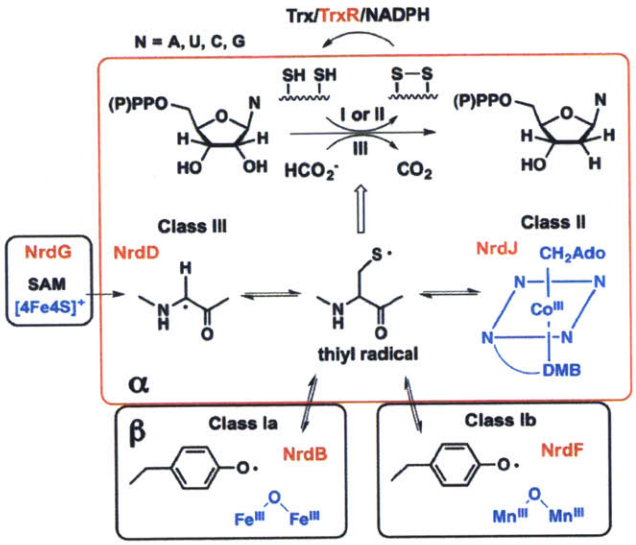
2. A Chemically Competent Thiosulfuranyl Radical on the *Escherichia coli* Class III Ribonucleotide Reductase

2.1. Introduction

Ribonucleotide reductases (RNRs) perform an essential function in all organisms, catalyzing the conversion of ribonucleotides into deoxynucleotides and providing the monomeric precursors required for DNA synthesis and repair (Eq 1).[1,2] All RNRs initiate nucleotide reduction *via* a transient protein-based thiyl radical[3] that abstracts a hydrogen atom from the 3′-position of the nucleotide[4]. RNRs have been divided into three classes (I, II and III) based on the metallo-cofactors required to generate this initiating thiyl radical.[5] This paper focuses on the *E. coli* class III RNR and the identification of a new radical species involving a three-electron bond between a cysteine and a methionine residue playing a role in the reductive half reaction. Despite the sequence differences between the three RNR classes, the first evidence for the involvement of thiyl radicals in the reductive half reaction of all classes is presented.

Many organisms possess multiple RNRs: *E. coli* possess three. Two of them are class I RNRs, one with a diferric-tyrosyl radical (Y•) cofactor (class Ia) and one with a dimanganese-Y• cofacor (class Ib). Both the Ia and Ib cofactors require O_2 for their biogenesis. The Ia enzyme is the workhorse in DNA replication, while the Ib RNR is expressed under iron-limitation and oxidative stress. The third RNR, which is the focus of this paper, is a class III enzyme and is expressed only under anaerobic conditions. It is encoded by the *nrdDG* operon.[6] Nucleotide reduction is catalyzed by the 80 kDa NrdD, which houses the O_2-sensitive glycyl radical[7] (G•) that is generated by the 17.5 kDa activating enzyme NrdG *via* chemistry involving S-adenosylmethionine (SAM) and a $[4Fe4S]^{1+}$ cluster (Eq 1).[8,9] The G•, as with the dimetallo-Y•s

in the class Ia and Ib RNRs, and adenosylcobalamin in the class II RNR, is reversibly involved in generation of the thiyl radical (Eq 1).



Equation 1 The class I, II and III RNRs catalyze the conversion of nucleoside di or triphosphates to deoxynucleotides. They differ in the metallo-cofactor that initiates the reduction process and in the reductant itself. Reproduced from Rhodes *et al.*[10]

Recent studies on the class Ia RNR[11] and earlier studies on the class II enzyme[12] established that the chemistry of nucleotide reduction is fast (~100 s^{-1}) and in both cases it is masked in the steady state by conformational gating[13,14] (k_{cat} of ~5 s^{-1}). The mechanism of nucleotide reduction by the class I and II RNRs has been investigated using many methods resulting in the model shown Figure 2.1A and B.[15] The overall reaction is divided into two half reactions. In the first half reaction (Figure 2.1A), a 3′-nucleotide radical[4,16,17] (2) is generated by 3′-hydrogen atom abstraction by the transient thiyl radical located on the top face of the sugar. Water loss is then facilitated by protonation by a bottom face cysteine, in an irreversible step, to generate intermediate 3. These first two steps are likely facilitated by removal of the proton of

the 3′-OH group.[18] In the second half reaction (Figure 2.1B), intermediate 3 is reduced to the 3′-keto-deoxynucleotide (4) concomitant with formation of the disulfide anion radical[19] which then, in a proton coupled electron transfer step[15], generates the 3′-deoxynucleotide radical 5. In the final step, the hydrogen atom that was initially removed from the 3′-position of the nucleotide is returned to the same position in the product.

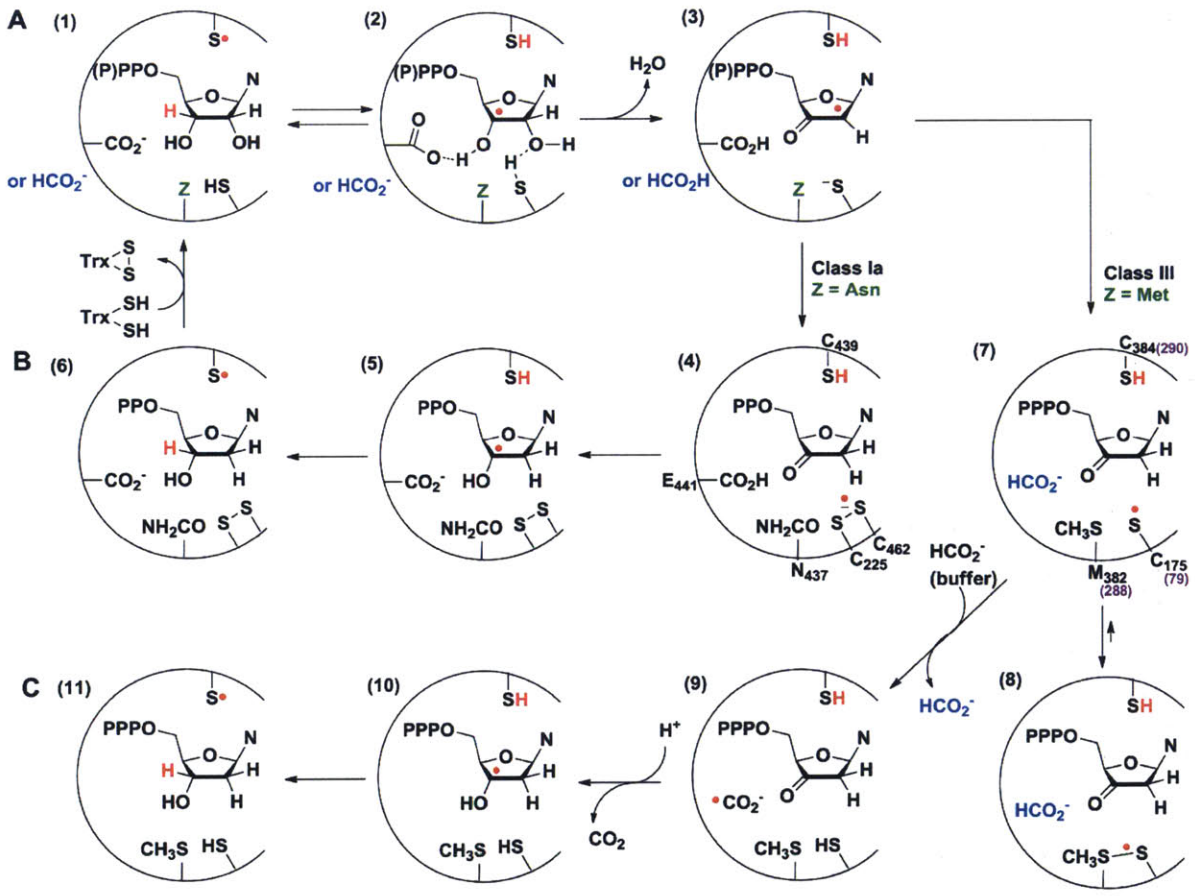


Figure 2.1. Mechanistic model for nucleotide reduction by the *E. coli* class Ia RNR and *E. coli* (or bacteriophage T4) class III RNR. A) First half reaction common to all RNRs; B) second half reaction of class Ia RNR; C) second half reaction of class III RNR.

The mechanism of nucleotide reduction by the class III RNRs has been less well characterized, but studies support a similar sequence of events for the first half reaction (Figure 2.1A).[20,21] A distinction between the RNR classes is apparent in the proposed mechanism of the second half reaction (compare Figures 2.1B and 2.1C), as a pair of conserved, bottom face, cysteine residues are essential for nucleotide reduction by the class I and II RNRs,[22] while formate is the reductant in the *E. coli* class III RNR.[23] Studies on the class III RNR in D_2O demonstrated that solvent replaces the 2′-OH with retention of configuration and that a small amount of exchange of deuterium into the 3′-position of the product takes place.[24] These results are similar to previous studies with the class I and II RNRs.[4,16,17] Studies with [^{3}H]-formate also provided insight about the reduction process.[23] First, ^{3}H was found in the solvent, consistent with the oxidation of formate to a $CO_2^{\bullet-}$ *via* the conserved bottom face cysteine. Second, an unusual [^{3}H]-(k_{cat}/K_m) isotope effect of 2.3 was observed, requiring exchange of formate between the active site and solvent prior to cleavage of the formate C-H bond and after the first irreversible step of nucleotide reduction, thought to be H_2O loss. Thus formate appears to be able to enter and leave the active site in the middle of nucleotide turnover.

A comparison of the active site structures of the *E. coli* class Ia α_2 (NrdA) and the bacteriophage T4 class III α_2 (NrdD, Figure 2.2A), reveals similar 10-stranded α/β barrels containing a finger loop with the essential cysteine C290 (equivalent to C439 in *E. coli* NrdA, C384 in *E. coli* NrdD) at its tip, that generates the substrate radical.[25] Of the two cysteines that donate the reducing equivalents in the class I and II RNRs (C225 and C462 in *E. coli* NrdA), only C79 (C225 in *E. coli* NrdA, C175 in *E. coli* NrdD) is conserved in the class III enzyme. Mutagenesis studies showed that both C79 and C290 are essential for catalysis in the class III enzyme.[26]

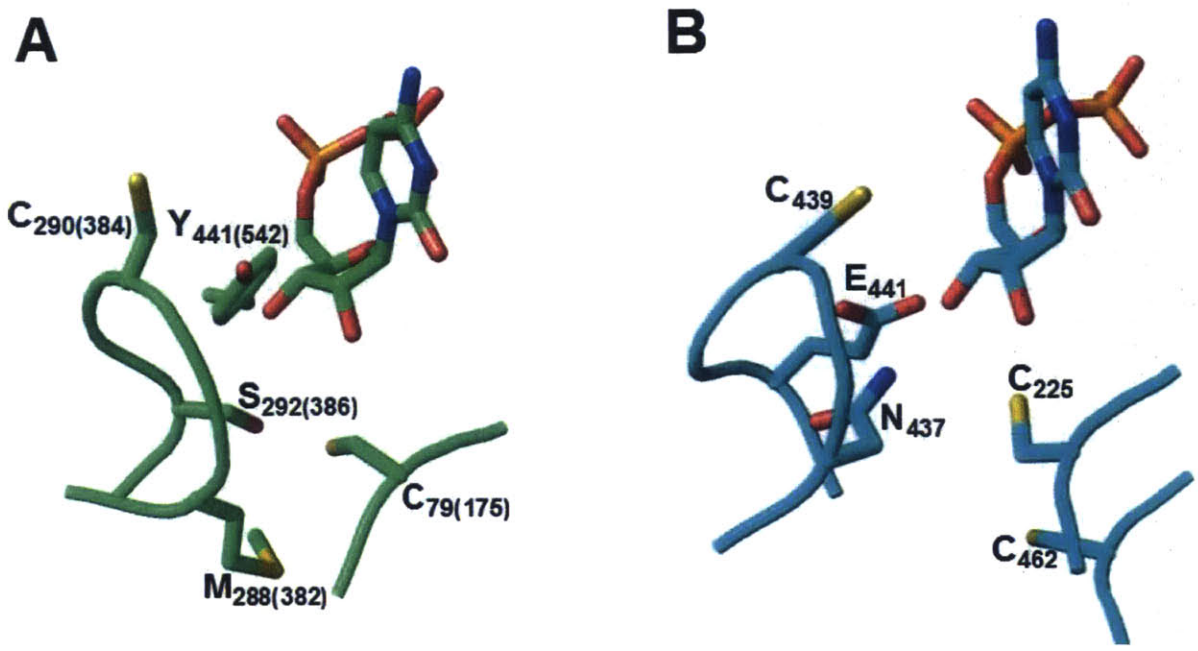


Figure 2.2. A) Active site from the crystal structure of T4 bacteriophage NrdD G580A mutant (PDB accession ID 1HK8)[25] (*E. coli* NrdD shares a 58% sequence identity with this protein, and its residues are numbered in the parentheses). CDP was docked by aligning the structures of NrdD and CDP-bound NrdA using pymol; B) Active site from the crystal structure of *E. coli* NrdA with substrate CDP. (Zimanyi, Drennan 2013 in preparation).

Additional similarities between the class I, II[27,28,29] and III[20] RNRs have been identified using mechanism-based inhibitors, 2′-chloro and fluoro 2′-deoxynucleotides, that can also function as alternative substrates (Figure 2.3). All RNRs react with these analogs to form 13, which can partition between the normal reduction reaction to generate the deoxynucleotide product (Figure 2.1B), or reduction from the top face or the bottom face of the sugar to generate a 3′-keto-deoxynucleotide (14), which dissociates from the active site and decomposes non-enzymatically to produce the nucleobase and pyrophosphate (tripolyphosphate). In the case of the class I and II RNRs, the furanone (15) has also been identified and its reaction with the enzyme has been shown to result in their time-dependent inactivation.[30] Similarities in the

reactivity of all three classes of RNRs towards these substrate analogs suggests a common first half reaction (Figure 2.1A) and provides a mechanism by which nucleobase (N) can be formed (Figure 2.3).

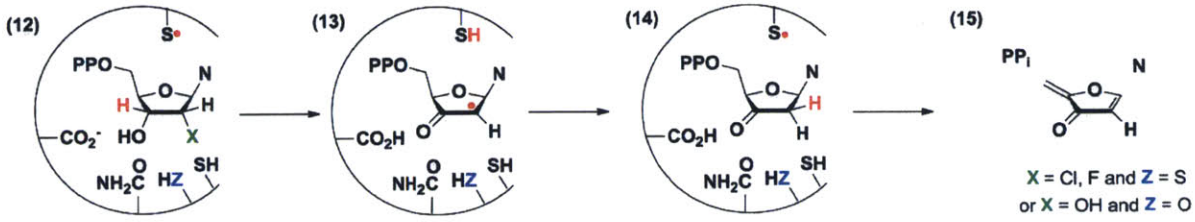


Figure 2.3. Mechanistic model for reaction of the *E. coli* class Ia RNR with a mechanism-based inhibitor or an active site-mutant (ZH) with CDP.

Reported here are our studies of the reaction of *E. coli* class III RNR that have provided insight about the reductive half reaction of this and the class I and II RNRs. Incubation of the class III RNR with CTP (substrate) and ATP (allosteric effector) in the absence of formate leads to the disappearance of the G•, concomitant with formation of a new radical and release of cytosine (Cyt), a known breakdown product of 3′-keto-deoxycytidine (Figure 2.1C, 7 and Figure 2.3, 14)[27]. Subsequent addition of formate leads to the disappearance of the new radical, recovery of G•, and formation of dCTP and a small amount of Cyt. Isotopic labeling of NrdD with [β-^{2}H]-cysteine, [ε-^{2}H]-methionine and [β,γ-^{2}H]-methionine, in combination with EPR spectroscopy of the new radical, identify this species as a thiosulfuranyl radical (Figure 2.1C, 8), generated from the C175 thiyl radical stabilized by an interaction with the conserved M382. The role of the thiosulfuranyl radical in the class III RNR second half reaction (Figure 2.1C) and its relationship to the disulfide radical anion in the second half reaction of the class Ia RNR (Figure 2.1B) is discussed.

2.2. Materials and Methods

2.2.1. Materials and General Methods

All chemical reagents were purchased from Sigma–Aldrich, unless otherwise indicated. Primers were purchased from Integrated DNA Technologies. UV-vis absorption spectroscopy was performed on an Agilent 8453 Diode Array spectrophotometer. Anaerobic procedures were carried out in a custom-designed MBraun glovebox in a cold room at 4 °C or at room temperature (RT). All solutions and proteins were made anaerobic on a Schlenk line by 3 cycles of evacuation (5 min) followed by flushing with Ar gas (10 min) before being brought into the glovebox. Nucleotides and SAM were brought into the glovebox as lyophilized solids. The plasmids pRSS and PN9 containing the genes for *E. coli* NrdD and NrdG respectively[31] were gifts from Professor Marc Fontecave, Collège de France (Paris, France).

2.2.2. Preparation of 5-[^{3}H]-CTP by phosphorylation of 5-[^{3}H]-CDP

The synthesis was carried out by minor modifications of the procedure of Lohman *et al.*[32] A reaction mixture (10 mL) containing 5-[^{3}H]-CDP (2 mM, Vitrax, specific activity of 26880/nmol), PEP (4 mM), Tris-HCl (50 mM, pH 7.5), KCl (80 mM), $MgCl_2$ (20 mM) and pyruvate kinase (rabbit muscle, 120 U/mL) was incubated at 37 °C for 1 h. The resulting 5-[^{3}H]-CTP was purified on a DEAE column (60 mL, 10 × 2 cm) washed with water (50 mL), and eluted with a 400 mL × 400 mL linear gradient from 0 to 750 mM triethylammonium bicarbonate (TEAB). The triphosphate containing fractions eluting at 550 mM TEAB were combined and the solvent removed *in vacuo*. The product 5-[^{3}H]-CTP was isolated in 87% yield and the structure confirmed by ^{1}H- and ^{31}P-NMR. The 5-[^{3}H]-CTP used in all experiments had a specific activity of 2860 cpm/nmol.

2.2.3. Construction of pET24a-nrdD to increase *nrdD* overexpression

The plasmid pRSS contains a 1800 bp insert between the promoter and *nrdD* start codon. An NdeI restriction site was introduced at the 5′ position of the *nrdD* ORF by site-directed mutagenesis using the primer GTTCTTAAAAATATGGAGCGCATATGACACCGC (mutation underlined). PCR was carried out using PfuUltraII polymerase (Stratagene) according to the manufacturer's protocol, followed by DpnI digestion of the methylated template plasmid. The resulting plasmid was digested with NdeI and EcoRI (NEB) and the *nrdD* fragment was ligated into pET24a (Novagen), which was linearized with the same restriction enzymes, to give pET24a-*nrdD*. The *nrdD* sequence and all cloned or mutant sequences were confirmed by DNA sequencing at the Massachusetts Institute of Technology Biopolymers Laboratory.

2.2.4. Construction of pTrc-*nrdD* for expression in a cysteine auxotroph

To facilitate insertion of *nrdD* into pTrcHisA (Invitrogen) without an N-terminal His_6-tag, an NdeI restriction site was introduced into the plasmid by site-directed mutagenesis using the primer CGATTAAATAAGGAGGAATAACATATGTATCGATTAAATAAGG (mutation underlined) and an internal NdeI restriction site present in the plasmid was mutated using the primer GGTATTTCACACCGCACATGGTGCACTC. The resulting plasmid was linearized using the restriction enzymes NdeI and EcoRI. The *nrdD* fragment was excised from pET24a-*nrdD* using the same restriction enzymes and ligated into pTrc to give pTrc-*nrdD*.

2.2.5. Construction of pET24a-*nrdD*-(S386E) and pTrc-*nrdD*-(S386E)

The S386E mutation was introduced by site-directed mutagenesis of the plasmids pET24a-*nrdD* and pTrc-*nrdD* using the primer CCGATGGGCTGCCGCGAGTTCCTCGGCGTGTGGG, (mutation underlined).

2.2.6. Expression of *E. coli nrdD* and protein purification

pET24a-*nrdD* was transformed into BL21 (DE3) cells (Invitrogen), grown on LB-agar plates with 50 μg/mL kanamycin (Kan) at 37 °C. A single colony was inoculated into 5 mL of LB (50 μg/mL Kan in all growths), grown at 37 °C until saturated (12 h), and transferred into 1.5 L of LB supplemented with 50 μM zinc acetate in a 6 L Erlenmeyer flask. The culture was grown at 37 °C with shaking at 220 rpm. At OD_{600} ~0.6, the temperature was lowered to 25 °C and isopropyl β-D-1-thiogalactopyranoside (IPTG, Promega) was added to a final concentration of 1 mM. After 14 h, the cells were pelleted by centrifugation (5,000 × g, 10 min, 4 °C) and frozen at -80 °C. Typically 4.7 g of cell paste per L of culture was obtained.

Cell paste (~7.1 g) was resuspended in 35 mL of buffer A (50 mM Tris, 5% glycerol (BDH Chemicals), 5 mM dithiothreitol (DTT, Promega), pH 7.5) containing 1 mM PMSF. The cells were lysed by a single passage through a French pressure cell (14,000 psi). DNA was precipitated by dropwise addition of 0.2 vol of buffer A containing 6% (w/v) streptomycin sulfate. The mixture was stirred for an additional 10 min, and the precipitated DNA was removed by centrifugation (20,000 × g, 10 min, 4 °C). Solid $(NH_4)_2SO_4$ was then added to 60% saturation. The solution was stirred for an additional 20 min, and the precipitated protein was isolated by centrifugation (20,000 × g, 10 min, 4 °C).

The pellet was dissolved in 20 mL of buffer A containing 0.8 M $(NH_4)_2SO_4$ and loaded onto a butyl-Sepharose column (2 × 16 cm, 50 mL). The column was eluted with 2 column volumes (CV) each of buffer A containing 0.8 M, 0.6 M, 0.4 M, 0.2 M and 0.0 M $(NH_4)_2SO_4$ and the fractions were analyzed by SDS-polyacrylamide gel electrophoresis (SDS-PAGE, 10% gel). This analysis led to re-chromatography of the protein on the butyl-Sepharose column with a gentler step-gradient. The fractions eluting at 0.6-0.4 M $(NH_4)_2SO_4$ were pooled and precipitated with $(NH_4)_2SO_4$ (60% saturation). The pellet was redissolved in buffer A containing 0.8 M $(NH_4)_2SO_4$ and loaded onto the butyl-Sepharose column (2 × 16 cm, 50 mL). The column was then eluted with 2 CV each of buffer A containing 0.8 M, 0.7 M, 0.6 M, 0.5 M and 0.4 M $(NH_4)_2SO_4$. The fractions eluting at 0.7-0.5 M $(NH_4)_2SO_4$ were pooled and precipitated with $(NH_4)_2SO_4$ (60% saturation).

The pellet was dissolved in 100 mL of buffer A and was loaded onto a DEAE-Sepharose column (2 × 16 cm, 50 mL) and washed with 100 mL of buffer A. The flow-through and wash fractions were pooled and precipitated with $(NH_4)_2SO_4$ (60% saturation). The pellet was dissolved in a minimal volume of buffer A (10 mL) and desalted using a Sephadex G-25 column (2 × 50 cm, 150 mL). The protein-containing fractions were pooled and concentrated by ultrafiltration (Amicon YM-30, Millipore). This procedure yielded 55 mg NrdD per g of cells (ε_{280} = 77020 $M^{-1}cm^{-1}$), judged pure by SDS-PAGE.

2.2.7. Expression of *E. coli nrdG* and protein purification

PN9 was transformed into BL21 (DE3) cells (Invitrogen), grown on LB-agar plates with 100 μg/mL ampicillin (Amp) at 37 °C. A single colony was inoculated into 5 mL of LB (100 μg/mL Amp in all growths), grown at 37 °C until saturated (12 h), and transferred into 6 × 1.5 L

of LB in 6 L Erlenmeyer flasks. The cultures were grown at 37 °C with shaking at 220 rpm. At OD_{600} ~0.7, the temperature was lowered to 28 °C and IPTG was added to a final concentration of 1 mM. After 12 h, the cells were pelleted by centrifugation (4,000 × g, 10 min, 4 °C) and frozen at -80 °C. Typically 2.1 g of cell paste per L of culture was obtained.

Cell paste (~12.9 g) was resuspended in 50 mL of buffer B (30 mM Tris, 50 mM KCl, 10 mM DTT, pH 7.5) containing 0.1 mM PMSF. The cells were lysed by a single passage through a French pressure cell (14,000 psi). DNA was precipitated by dropwise addition of 0.2 vol of buffer A containing 6% (w/v) streptomycin sulfate. The mixture was stirred for an additional 10 min, and the precipitated DNA was removed by centrifugation (15,000 × g, 10 min, 4 °C). Solid $(NH_4)_2SO_4$ was then added to 40% saturation. The solution was stirred for an additional 20 min, and the precipitated protein was isolated by centrifugation (15,000 × g, 40 min, 4 °C).

The pellet was dissolved in 10 mL of buffer B and desalted using a Sephadex G-25 column (5 × 50 cm, 1 L) equilibrated with buffer B. The colored fractions were combined and loaded onto a DEAE-Sepharose column (5 × 10 cm, 200 mL) and washed with 300 mL of buffer B. The flow-through and wash fractions were pooled and precipitated with $(NH_4)_2SO_4$ (40% saturation). The pellet was dissolved in 4 mL of buffer B and loaded onto a Sephadex G-75 column (3 × 100 cm, 700 mL) equilibrated with buffer B. Fractions (15 mL) were collected and analyzed by SDS-PAGE. The fractions containing NrdG were pooled and concentrated by ultrafiltration (Amicon YM-10), yielding 4 mg NrdG per g of cells (ε_{280} = 28480 $M^{-1}cm^{-1}$), pure by SDS-PAGE.

2.2.8. Reconstitution of the [4Fe4S] cluster of NrdG

The procedure was carried out in a glovebox in a 4 °C cold room. Solutions of Na_2S and of $Fe(NH_4)_2(SO_4)_2$ in water (100 mM) were freshly prepared in the glovebox. A solution of NrdG (250 μM, 1 mL) was made anaerobic on a Schlenk line and brought into the glovebox. A solution of DTT (100 mM) was added to 10 mM, followed by ordered addition of a solution of Na_2S (5 eq.) and $Fe(NH_4)_2(SO_4)_2$ (5 eq.), and incubation for 12 h. The solution was concentrated to 200 μL by ultrafiltration (Amicon YM-30). EDTA (5 eq.) was then added and the solution was desalted using a Sephadex G-25 column (1 × 9 cm, 7 mL) equilibrated with triethanolamine (TEA) buffer (30 mM, pH 7.5). The final material typically contained ~2 atoms of Fe per peptide determined by the ferrozine assay[33].

2.2.9. Minimization of formate levels in reaction buffers

In order to carry out reactions in the absence of formate, substitutes were used for components of the NrdD storage buffer found to contain a formate contaminant (Tris base, glycerol and DTT, see Results section). Inositol and TCEP, which contain undetectable levels of formate, effectively replaced glycerol and DTT in preserving NrdD activity. Tris buffer was replaced with Bicine, which also served as a coreductant for the photoreduction system. Bicine was later found to contain small amounts of formate, and was thus replaced with TEA for experiments where complete removal of formate is desired.

2.2.10. Generation of the NrdD G• using a catalytic amount of NrdG

To generate NrdD-G•, a solution of NrdD (200 μM, 1.5 mL) was concentrated to 300 μL by ultrafiltration (Amicon YM-30). The protein was exchanged into Bicine buffer (30 mM, pH 7.5), 3% inositol, 5 mM tris(2-carboxyethyl)phosphine (TCEP), using a Sephadex G-25 column (1 × 20 cm, 15 mL). The protein-containing fractions were pooled and made anaerobic on a Schlenk line and brought into a glovebox in a 4 °C cold room.

A 50 μL mixture of NrdD (100 μM), NrdG (10 μM), SAM (1 mM), Bicine buffer (30 mM, pH 7.5) and Rose Bengal (50 μM) was exposed to a fluorescent lamp in the glovebox for 2 h. For inspection by X-band EPR spectroscopy, the solution was diluted to 250 μL with Bicine buffer (30 mM, pH 7.5), 3% inositol to give a final concentration of 20 μM NrdD, and sealed in an EPR tube with a rubber stopper. The solution was quenched in liquid N_2 immediately after removal from the glovebox. The amount of G• in the solution was determined by comparing the EPR signal intensity to that of a $CuSO_4$ standard[34]. A typical yield of 0.45-0.50 radicals per NrdD polypeptide was reproducibly obtained.

2.2.11. Activity assay for dCTP formation

The assay mixture in 100 μL contained NrdD (0.2 μM), ATP (1 mM), 5-[^{3}H]-CTP (1 mM, 2860 cpm/nmol) in Tris (30 mM, pH 7.5), KCl (30 mM), $MgSO_4$ (10 mM), HCO_2Na (10 mM) and was incubated at 25 °C. Four, 20 μL aliquots were removed at one min intervals and quenched with 2% perchloric acid (20 μL). Subsequent to removal of the phosphates using calf intestine alkaline phosphatase (Roche), dCTP formation was analyzed by the method of Steeper and Steuart[35]. One unit of activity is equivalent to 1 nmol of dCTP/min. The specific activity of NrdD is ~1500 U/mg, consistent with literature values[36].

2.2.12. Assay for Cyt formation

The method of Steeper and Steuart involves chromatography on a Dowex-1-borate column where the cytidine is retained as a complex with borate, while dC is eluted. To assay for Cyt released from 3′-keto-dCTP, the dephosphrylation step was omitted. Controls showed that CTP and dCTP are retained on the Dowex anion-exchange column, while the Cyt passes directly through the column. 5-[^{3}H]-CTP was lyophilized before the experiments, eliminating any ^3H_2O that would also pass directly through the column. Further characterization by reverse phase HPLC is described below.

2.2.13. Reaction of NrdD with CTP, ATP and analysis by X-band EPR spectroscopy

The final reaction mixture at 4 °C contained in a volume of 250 μL: NrdD (40 μM, activated as described above), CTP (1.5 mM), ATP (1.5 mM), Bicine (30 mM, pH 7.5), KCl (30 mM), $MgSO_4$ (10 mM), 3% inositol. The reaction was initiated by addition of NrdD and the sample was rapidly transferred and sealed in an EPR tube with a rubber stopper, removed from the glovebox and frozen in liquid N_2. The total reaction time was 1 min for NrdD-wild-type (WT) and 60 min for NrdD-(S386E).

For reaction with formate, NrdD was reacted first with CTP and ATP for 1 min at 4 °C as described above, followed by addition of sodium formate to 10 mM (2.5 μL, 1 M). The total reaction time with formate was 1 min.

2.2.14. X-band EPR spectroscopy

Continuous wave (cw) X-band EPR spectra were recorded at 77 K in the MIT Department of Chemistry Instrumentation Facility on a Bruker ESP-300 X-band spectrometer equipped with a quartz finger Dewar filled with liquid N_2. Experimental conditions were as follows: microwave frequency, 9.45 GHz; modulation amplitude, 0.15 mT; modulation frequency, 100 kHz; time constant, 5.12 ms; scan time, 41.9 s. A microwave power of 10 μW and an average of 10 scans was used for G•, while a power of 160 μW and 100 scans was used for the thiosulfurnanyl radical.

2.2.15. Sample preparation for 140 GHz EPR analysis

The 140 GHz EPR spectrometer requires a more concentrated protein sample. A solution of NrdD and NrdG was concentrated by ultrafiltration (Amicon YM-30) in a glovebox in a 4 °C cold room. The final mixture for G• generation (5 μL) contained NrdD (800 μM), NrdG (80 μM), SAM (1.5 mM), Bicine (pH 7.5, 30 mM), KCl (30 mM), $MgSO_4$ (10 mM), Rose Bengal (50 μM), 3% inositol. The mixture was exposed to the fluorescent lighting in the glovebox at 4 °C for 2 h to generate G•. The reaction with the nucleotide was initiated by adding a solution (1 μL) containing CTP and ATP in reaction buffer, to give a final reaction concentration of 2 mM each. Loading and freezing the small 140 GHz EPR tubes (ID 0.5 mm, effective sample volume 200 nL) could not be done as swiftly as with the X-band tubes and the total reaction time was extended to 2.5 min at 4 °C.

2.2.16. 140 GHz EPR spectroscopy (spectra were acquired by Dr. Guinevere Mathies, MIT)

Echo-detected 140 GHz EPR spectra were obtained on a spectrometer developed by Smith *et al.*[37] Pulsed EPR spectra were acquired at a temperature of 85 K using a Hahn echo pulse sequence ($\pi/2$ pulse = 35 ns and τ = 250 ns). The applied microwave power was 100 mW. At each field position, 400 shots were acquired with a repetition time of 1 ms. The ^{1}H resonance frequency of a small water sample that resides just below the sample space in the magnet was used to set the magnetic field[38].

2.2.17. Preparation of [β-^{2}H]-cys-NrdD

This procedure was adapted from existing protocols.[39] pTrc-*nrdD* or pTrc-*nrdD*-(S386E) was transformed into the *E. coli* cysteine auxotroph JW3582-2 (Yale *E. coli* Genetic Stock Center), containing the mutation ΔcysE720::kan[40]. A single colony was inoculated into 5 mL of LB (100 μg/mL Amp in all growths), grown at 37 °C until saturated (12 h), harvested by centrifugation (3,000 × g, 10 min, 4 °C) and transferred into M9 medium supplemented with L-amino acids and cofactors. The growth medium contained in 1 L: Na_2HPO_4 (6 g), KH_2PO_4 (3 g), NH_4Cl (1 g), NaCl (1 g), glycerol (4 g), $MgCl_2.6H_2O$ (210 mg), and $CaCl_2.2H_2O$ (14 mg), alanine (0.50 g), arginine (0.40 g), aspartic acid (0.40 g), asparagine (0.40 g), glutamine (0.4 g), sodium glutamate (0.74 g), glycine (0.55 g), histidine hydrochloride (0.13 g), isoleucine (0.24 g), leucine (0.23 g), lysine hydrochloride (0.43 g), methionine (0.25 g), phenylalanine (0.15 g), proline (0.10 g), serine (2.10 g), threonine (0.23 g), tryptophan (0.06 g), valine (0.23 g), thiamine (50 mg) and racemic [β,β′-^{2}H]-cystine (120 mg, 98%, Cambridge Isotope Laboratories). The cultures were grown at 37 °C with shaking at 220 rpm. At OD_{600} ~0.6, the temperature was lowered to 25 °C and IPTG was added to a final concentration of 1 mM. After 14 h, the cells

were pelleted by centrifugation (5,000 × g, 10 min, 4 °C) and frozen at -80 °C, yielding ~2 g of cell paste per L of culture.

2.2.18. Preparation of [ε-^{2}H]-met- and [β,γ-^{2}H]-met-labeled NrdD

The procedure was adapted from existing protocols[41]. pET24a-*nrdD* or pET24a-*nrdD*-(S386E) was transformed into BL21 (DE3) cells, grown on LB-agar plates with 50 μg/mL Kan. A single colony was inoculated into 5 mL of LB (50 μg/mL Kan in all growths), grown at 37 °C until saturated (12 h), harvested by centrifugation (3,000 × g, 10 min, 4°C) and transferred into M9 medium containing in 1 L : Na_2HPO_4 (6 g), KH_2PO_4 (3 g), NH_4Cl (1 g), NaCl (1 g), glycerol (4 g), $MgCl_2.6H_2O$ (210 mg), $CaCl_2.2H_2O$ (14 mg) and thiamine (50 mg). The cultures were grown at 37 °C with shaking at 220 rpm. At OD_{600} ~0.3, the L-amino acids lysine (0.1 g), phenylalanine (0.1 g), threonine (0.1 g), isoleucine (0.05 g), leucine (0.05 g) and valine (0.05 g) were added. Additionally, 50 mg of [ε-^{2}H]-methionine (98%, Cambridge Isotope Laboratories) or [β,γ-^{2}H]-methionine (98%, CDN isotopes) was added per L of medium, followed by shaking for 20 min. The temperature was lowered to 25 °C and IPTG was added to a final concentration of 1 mM. After 14 h, cells were pelleted by centrifugation (5,000 × g, 10 min, 4 °C) and frozen at -80 °C. Typically 2 g of cell paste was obtained per L of culture.

2.2.19. Purification of [β-^{2}H]-cys-, [ε-^{2}H]-met- and [β,γ-^{2}H]-met-NrdD and NrdD-(S386E)

Cell paste (~2 g) was suspended in 20 mL of buffer A containing 1 mM PMSF. The cells were lysed by a single passage through a French pressure cell (14,000 psi). DNA was precipitated by dropwise addition of 0.2 vol of buffer A containing 6% (w/v) streptomycin sulfate. The mixture was stirred for an additional 10 min, and the precipitated DNA was removed

by centrifugation (20,000 × g, 10 min, 4 °C). Solid $(NH_4)_2SO_4$ was then added to 60% saturation. The solution was stirred for an additional 20 min, and the precipitated protein was isolated by centrifugation (20,000 × g, 10 min, 4 °C). The pellet was dissolved in a minimal volume (0.3 mL) of buffer A and desalted using a Sephadex G-25 column (1 × 20 cm, 15 mL) equilibrated with buffer containing Bicine (30 mM, pH 7.5), TCEP (5 mM) and 3% inositol. This procedure yielded NrdD that was > 90% pure by SDS-PAGE.

2.2.20. Exchange of proteins into D_2O buffer

Reaction buffer containing Bicine (30 mM, pH 7.5), KCl (30 mM), $MgSO_4$ (10 mM) and 3% inositol was prepared in D_2O and lyophilized to remove exchangeable protons, brought into the glovebox and redissolved in D_2O (99.9%, Cambridge Isotope Laboratories). NrdD and NrdG were exchanged into this buffer by repeated dilution and concentration by ultrafiltration (Amicon YM-30), such that < 1% H_2O remained.

2.2.21. Single turnover of NrdD with 5-[^{3}H]-CTP and ATP in the absence (A) or presence (B) of formate

(A) To minimize levels of formate in the assay, components of the protein storage and reaction buffers found to contain a formate contaminant (Tris base, glycerol and DTT) were substituted with compounds containing undetectable levels of formate (TEA, inositol, and TCEP). The reaction was initiated in a glovebox in a 4 °C cold room by mixing a solution of activated NrdD (20 μL) with a solution containing 5-[^{3}H]-CTP, ATP and all other components of the reaction buffer (5 μL). The final reaction mixture contained NrdD (100 μM), 5-[^{3}H]-CTP (1.5

mM), ATP (1.5 mM), TEA (30 mM, pH 7.5), KCl (30 mM) and $MgSO_4$ (10 mM). After 40 s, the reaction was quenched by addition of 2% perchloric acid (20 μL).

(B) NrdD was incubated with 5-[^{3}H]-CTP and ATP as described above, followed by addition of a mixture (1 mL) of HCO_2Na (10 mM), $MgSO_4$ (3 mM), TEA (10 mM, pH 7.5), and unlabeled CTP (1 mM). Subsequent to mixing, the sample was immediately quenched with 17 μL of 60% perchloric acid. The total reaction time with formate prior to quenching with acid was < 5 s.

In both cases, the samples were neutralized with KOH while cooling on ice and the precipitated $KClO_4$ was removed by centrifugation. Quantification of 5-[^{3}H]-Cyt and 5-[^{3}H]-dC was carried out as described above. To a 7 mL portion of the eluate of the Dowex-1-borate column, containing a mixture of 5-[^{3}H]-Cyt and 5-[^{3}H]-dC, was added carrier Cyt and dC (10 nmol each). The mixture was concentrated by lyophilization, redissolved in water, cooled on ice and the precipitated borate salts removed by centrifugation. The supernatant was analyzed by HPLC using an Alltech Econosil column (C18, 10 μm, 250 × 4.6 mm) on a Waters 515 HPLC system equipped with a 2996 photodiode array detector. The compounds were eluted with KPi (20 mM, pH 6.8) at a flow rate of 1.0 mL/min. Fractions were collected (0.5 mL) and analyzed by scintillation counting. Cyt and dC were identified by coelution with a standard.

2.2.22. Electronic structure calculations on a model thiosulfuranyl radical (calculations were carried out by Dr. Jiahao Chen, MIT)

A*b initio* calculations were performed using the Q-Chem software package, version 4.1.[42] Unless otherwise specified, calculations were performed in the gas phase, using spin-unrestricted Kohn-Sham density functional theory (KS-DFT) with the B3LYP approximate functional[43] and

the 6-31++G** basis set.[44] All spin-unrestricted calculations were verified to have spin contamination of less than 0.5% based on the calculated expectation of the spin operator $<S^2>$ and its deviation from the ideal value of 0.75 (for a doublet system with spin quantum number *s* = 0.5). Visualizations were generated using the Visual Molecular Dynamics (VMD)[45] program using orthographic projection, with isosurface values of 0.05 for the orbitals and 0.0025 (0.05^2) for the spin densities respectively.

The thiosulfuranyl radical formed from the methanethiyl radical and ethylmethylsulfide was used as a minimal model system to investigate the electronic structure of the NrdD thiosulfuranyl radical. The model was constructed using the atomic coordinates of the side chains of C79 and M288 in the crystal structure of bacteriophage T4 NrdD (PDB accession ID 1HK8, Figure 2.2A).[25] Hydrogen atoms were added and their positions optimized while fixing the heavy atoms in place. A plausible reaction coordinate for the formation of a thiosulfuranyl radical was then calculated. At each point along the reaction profile, the (methanethiyl) S_1-(ethylmethylsulfide) S_2 distance was fixed and all the other coordinates were allowed to relax, to obtain the potential energy of the structure as a function of S_1-S_2 separation. The stationary points on the potential energy surfaces were verified to be true minima by ensuring that they contain no imaginary frequencies. The resulting reaction coordinate represents the enthalpic contribution to creating a structure of a given S_1-S_2 separation in the gas phase at 0K.

2.3. Results

Two experimental challenges had to be overcome in order to study the reaction of NrdD, CTP and ATP in the absence of the reductant formate. The first was to optimize G• formation,

and the second was to remove formate contaminants in the reaction mixture components so that the first half reaction of RNR could be studied.

2.3.1. Generation of NrdD G•

E. coli NrdD and NrdG were expressed and purified by minimal modification of reported procedures[36,46], with the inclusion of 50 μM zinc sulfate in the expression media for NrdD. The generation of G• requires the activating enzyme NrdG, SAM and a reductant for the NrdG-[4Fe4S] cluster. Previously studies have shown that generation of G• by NrdG can be carried out catalytically with the *E. coli* flavodoxin (FldA)/ flavodoxin reductase (Fpr) system[36,47,48], or with a stoichiometric amount of NrdG and photoreduced deazaflavin[49]. We wanted to activate NrdD catalytically to minimize possible interference from the activating components, which are difficult to routinely remove due to the O_2-sensitivity of G•. We also wanted to avoid the FldA / Fpr system because an EPR signal from the FldA semiquinone might complicate analysis of the EPR spectra. We initially focused on using a catalytic amount of NrdG activated with photoreduced deazaflavin. However, this method proved unsuccessful as this photoreducant resulted in the reversible quenching of G•.

Several photoreductant / coreductant pairs were screened, including all combinations of photoreductants acriflavin, fluorescein and Rose Bengal and coreductants EDTA, tetramethylethylenediamine, TEA, Bicine and cysteine. The Rose Bengal / Bicine pair was chosen because it did not reduce the NrdD-G•. This pair was used for all EPR experiments. However, Bicine contains a small amount of formate (see next section), and thus it was replaced with Rose Bengal / TEA for all experiments involving product analysis, where the complete removal of formate is desired. NrdD purified and activated in either way had ~0.5 G• per peptide

(Figure 2.4i) and a specific activity of ~1500 U/mg for CTP reduction, consistent with literature values[36].

2.3.2. Removal of formate from reaction and protein storage buffers

When NrdD, [^{3}H]-CTP and allosteric effector ATP were incubated in the absence of formate and analyzed for product, we were surprised to detect dCTP. Since the formation of dCTP requires a reductant, we suspected that formate contaminated one or more of our reagents. Several methods were used to assess its presence. Using a coupled assay with formate dehydrogenase, formate was detected in the Tris base and glycerol used in the protein storage and reaction buffers. These components were thus replaced with TEA and inositol. A second more sensitive assay monitored [^{3}H]-dCTP formation, which suggested formate presence in DTT. DTT had been shown to be important for high activity of NrdD.[36] We found that NrdD activity could be maintained by replacing it with 5 mM TCEP in the NrdD storage buffer. Using this formate-free buffer, dCTP production from the reaction of NrdD with CTP was largely eliminated.

2.3.3. A new radical generated by the reaction of NrdD with CTP and ATP in the absence of formate

Figure 2.4i shows the cw X-band EPR spectrum of the G• that is generated by NrdG. Incubation of activated NrdD with CTP and allosteric effector ATP for 1 min at 4 °C led to the loss of G• and the appearance of a new radical shown in Figure 2.4ii, accompanied by a spin loss of ~5%. Subsequent addition of saturating amounts of formate (10 mM, K_M = 0.2 mM)[23] led to the disappearance of the new radical and return of the G• spectrum with a yield of ~80%.

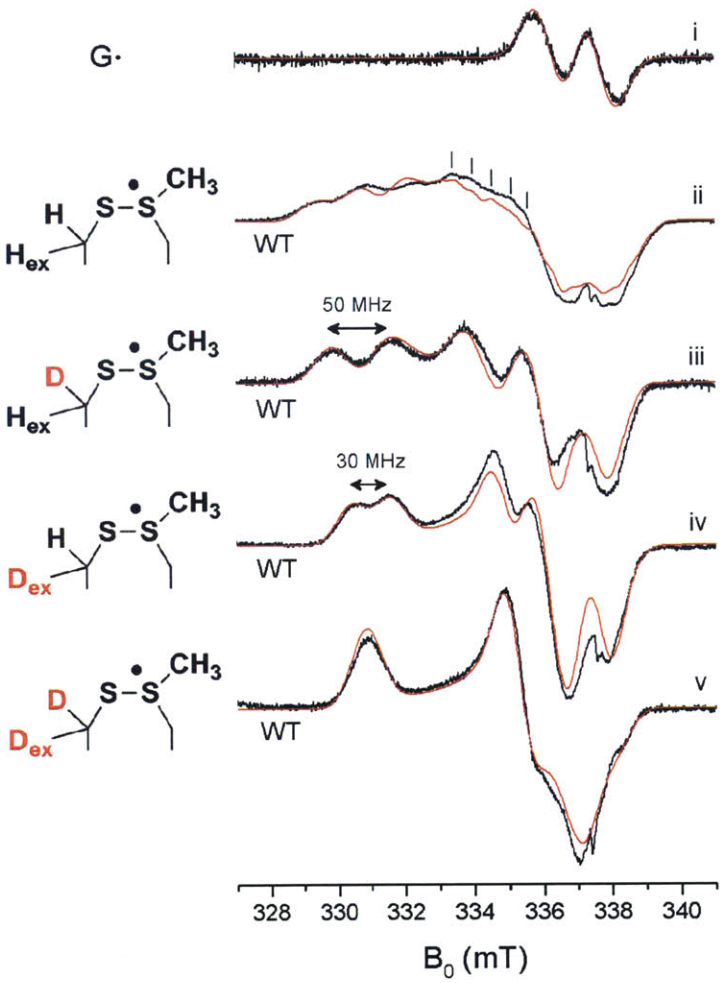


Figure 2.4. X-band cw EPR spectra (black) and simulations (red) of NrdD glycyl and thiosulfuranyl radicals. See the text and Table 2.1 for the simulation parameters. All experimental spectra were scaled to the simulation. The narrow signal of the G• was scaled by a factor of $^{1}/_{10}$ for comparison with the broader signals of the new radical. The spectral artifact at 337.5 mT (g_e) is due to an impurity in the cryostat. i) The G• in activated NrdD. ii-v) The thiosulfuranyl radical generated by reaction of CTP and ATP in the absence of formate with: ii) NrdD in H_2O, iii) [β-^{2}H]-cys-NrdD in H_2O, iv) NrdD in D_2O, and v) [β-^{2}H]-cys-NrdD in D_2O.

The X-band spectrum of the new radical shows significant g-anisotropy. In addition, three peaks at 329, 331, and 332 mT and structure in the region of 333-340 mT suggest hyperfine interactions with several protons.

In analogy with the mechanism of the class Ia RNRs, we hypothesize that in the absence of formate, NrdD catalyzes cleavage of the 3′ C-H bond and loss of H_2O from CTP to form 3′-keto-dCTP, generating a thiyl radical on C175 (Figure 2.1C,7). Because the new radical is long-lived ($t_{1/2}$ ~10 min at 4 °C), we propose that it might be the C175 thiyl radical stabilized by a bonding interaction with the conserved M382 residue on the thiyl radical loop, in the form of a thiosulfuranyl radical (Figure 2.1C, 8),[50,51,52] resembling the disulfide anion radical (Figure 2.1B, 4) of the class Ia RNR. In the crystal structure of T4 NrdD,[25] the distance between the sulfur atoms of C79 and M288 is 4.4 Å (Figure 2.2A).

To test this hypothesis, we pursued further EPR spectroscopy of the new radical at high microwave frequency to separate g-anisotropy from hyperfine interactions, in combination with isotope labeling of NrdD at X-band frequency to identify the proton hyperfine interactions contributing to the spectrum.

Characterization of the new radical by EPR spectroscopy

2.3.4. g-values support a thiosulfuranyl radical

Figure 2.5 shows the echo-detected 140 GHz EPR spectrum of activated NrdD reacted with CTP and ATP. The spectrum is dominated by a broad signal stretching from 4.902 T to 4.998 T arising from a radical with a rhombic g-tensor, but also shows a narrow signal around 4.993 T. The narrow signal is readily assigned to residual G• present in the sample.[53] The broad signal must therefore be due to a new radical(s). The principal values of its g-tensor are g_x =

2.040, $g_y = 2.013$, $g_z = 2.0016$. These g-values are inconsistent with a thiyl radical, where the g_x value is expected to be shifted further down-field,[54,55] but are similar to g-values previously reported for putative thiosulfuranyl radicals.[55,56]

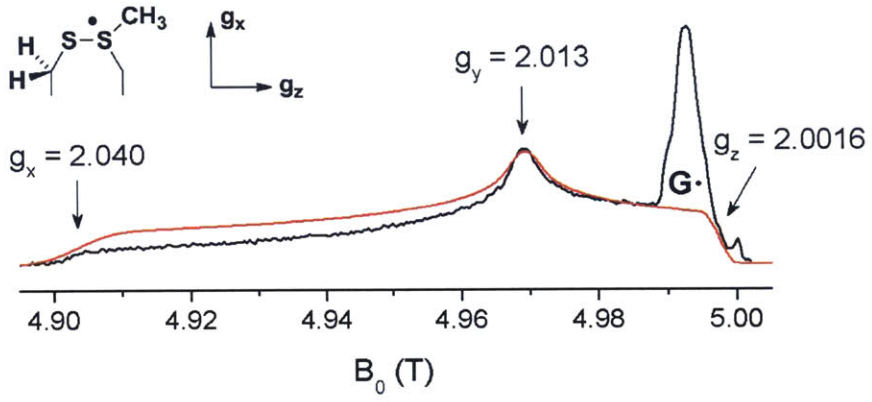


Figure 2.5. 140 GHz echo-detected EPR spectrum (black) and simulation (red) of the NrdD thiosulfuranyl radical in H_2O. See the text for the simulation parameters. At 5.00 T, an impurity from the sample tube is apparent.

2.3.5. [β-^{2}H]-cys-NrdD reveals a hyperfine interaction with a cysteine β-proton

If the new radical is partially localized on C175 it will likely exhibit hyperfine interactions with its β-protons. [β-^{2}H]-cys-NrdD was prepared and reacted with CTP and ATP. The resulting X-band EPR spectrum is shown in Figure 2.4iii. It reveals, in comparison with the unlabeled NrdD (Figure 2.4ii), loss of a strong hyperfine interaction, providing evidence for localization of the radical on cysteine. Another, strong hyperfine interaction remains and now dominates the spectrum.

2.3.6. Reaction in D_2O suggests a hyperfine interaction with a solvent-exchangeable proton

In search of the origin of this other proton hyperfine interaction (Figure 2.4iii), we obtained the X-band spectrum of the reaction of NrdD with CTP and ATP carried out in D_2O (Figure 2.4iv). Again, the spectrum has clearly lost a strong hyperfine interaction, which suggests that the unpaired electron interacts with a solvent exchangeable proton. At 4 °C this exchange is complete within the 1 min required for sample preparation, and longer incubation does not lead to further changes in the signal profile.

Finally, carrying out the reaction of [β-^{2}H]-cys-NrdD in D_2O leads to a spectrum that no longer shows any resolved hyperfine interactions (Figure 2.4v). Comparison of this spectrum with unlabeled NrdD in D_2O reveals an isotropic hyperfine interaction of ~30 MHz with the non-exchangeable β-proton of cysteine. Comparison of this spectrum with [β-^{2}H]-cys-NrdD in H_2O, on the other hand, reveals a hyperfine interaction of ~50 MHz associated with the exchangeable proton, possibly the second β-proton of cysteine.

2.3.7. NrdD-(S386E) mutation prevents detectable proton exchange on the radical

The exchange of a β-proton of cysteine is unexpected and thus requires further verification. We hypothesized that if this exchange involved deprotonation to form a thiyl radical anion (Figure 2.6A), then introduction of a negative charge near the radical could disfavor this process. The structure of T4 NrdD suggests that C175 of *E. coli* NrdD is positioned close to S386, which resides on the thiyl radical loop and aligns with E441 in *E. coli* NrdA, a residue thought to act as a general base/acid catalyst in the class Ia RNR (compare Figures 2.2A and 2.2B). We constructed the NrdD-(S386E) mutant and reacted it with CTP and ATP in D_2O. NrdD-(S386E) has activity < 1% that of NrdD-(WT), and the new radical also formed at a much

slower rate than with NrdD-(WT), requiring a reaction time of ~1 h. The X-band spectrum of NrdD-(S386E) in D_2O resembles the spectrum of NrdD-(WT) in H_2O (compare Figure 2.6Bi with Figure 2.4ii), which suggests that the proton exchange is blocked and that the structure of the radical remains largely unaffected.

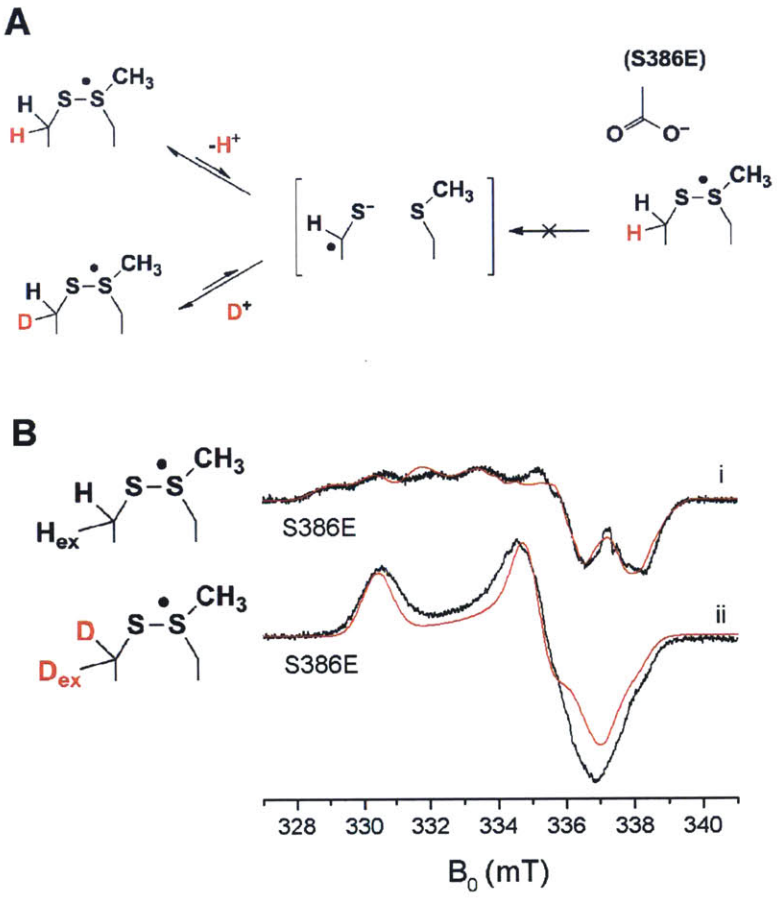


Figure 2.6. A) Proposed mechanism for proton exchange on C175, which is prevented in NrdD-(S386E). B) X-band cw EPR spectra (black) and simulations (red) of i) NrdD-(S386E) in D_2O and ii) [β-^{2}H]-cys-NrdD-(S386E) in H_2O. See the text and Table 2.1 for the simulation parameters. The spectral artifact at 337.5 mT (~g_e) is due to an impurity in the cryostat.

2.3.8. [β-^{2}H]-cys-NrdD-(S386E) establishes that the exchangeable proton is a cysteine β-proton

The suppression of the proton exchange by NrdD-(S386E) allows the site of the exchangeable proton to be unambiguously established. A reaction of [β-^{2}H]-cys-NrdD-(S386E) in H_2O with CTP and ATP revealed the spectrum in Figure 2.6Bii, which is very similar to that of [β-^{2}H]-cys-NrdD-(WT) in D_2O (Figure 2.4v). Thus the exchangeable proton is a β-proton of C175, and the exchange occurs stereospecifically.

2.3.9. [ε-^{2}H]-met- and [β,γ-^{2}H]-met-NrdD reveal hyperfine interaction with methionine ε- and γ-protons

We hypothesize that the new radical involves M382 (Figure 2.1C 7 and 8). To test this model, [ε-^{2}H]-met-NrdD was prepared and the resulting radical examined for the effects of isotopic substitution. Comparison of the spectrum of [ε-^{2}H]-met-NrdD (Figure 2.7i) and unlabeled NrdD (Figure 2.4ii) in H_2O, revealed sharpening of the three peaks between 329 and 332 mT. Moreover, in the region 333-340 mT, a fine structure consisting of lines separated by 0.5 mT (marked with vertical lines in Figure 2.7i) has become clear. This fine structure is also present in the spectrum of unlabeled NrdD in H_2O (Figure 2.4ii, see vertical lines), but is not fully resolved. This shows that in unlabeled NrdD, the fine structure is masked by a broadening of the peaks due to unresolved hyperfine interactions with the three ε-protons of methionine.

To further investigate the fine structure, we obtained the spectrum of [ε-^{2}H]-met-NrdD-(S386E) in D_2O (Figure 2.7ii). D_2O exchanges the envelope protons around the active site, sharpening the features associated with the thiosulfuranyl radical, while the S386E mutation prevents exchange of the C175 β proton. Comparison of this spectrum (6ii) with that of the

corresponding NrdD-(WT) (6i) reveals an almost identical radical signal with sharpened features. The fine structure could be reproduced in the simulation by including two additional hyperfine interactions of ~14 MHz, which we propose to be associated with the two γ protons of M382.

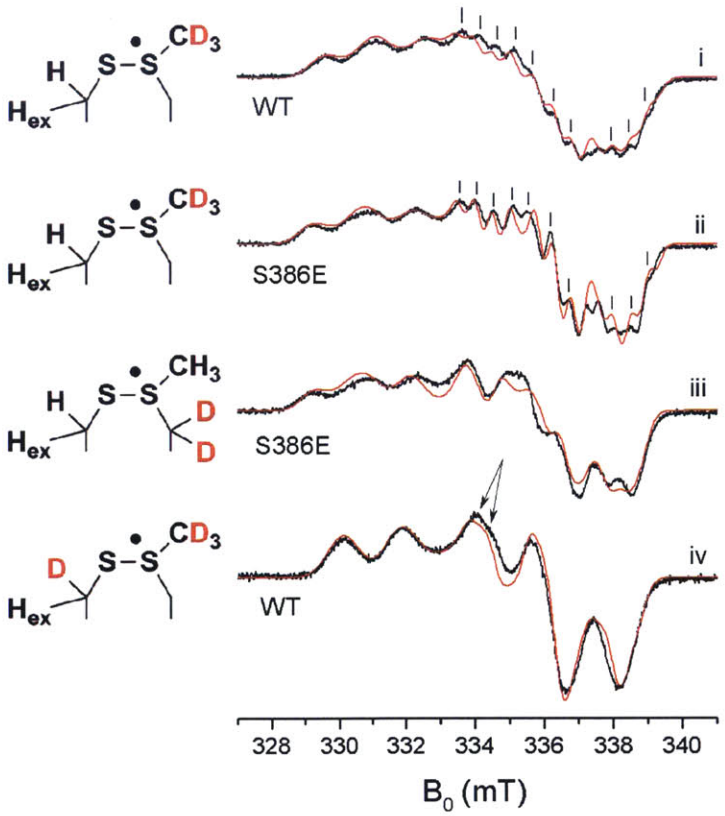


Figure 2.7. X-band cw EPR spectra (black) and simulations (red) of the NrdD thiosulfuranyl radical in i) [ε-^{2}H]-met-NrdD in H_2O; ii) [ε-^{2}H]-met-NrdD-(S386E) in D_2O; iii) [β,γ-^{2}H]-met labeled NrdD-(S386E) in H_2O, iv) [β-^{2}H]-cys,[ε-^{2}H]-met-NrdD in H_2O. See the text and Table 2.1 for the simulation parameters.

Evidence for this hypothesis was obtained from the spectrum of [β,γ-^{2}H]-met-NrdD-(S386E) in H_2O (only the doubly labeled methionine is commercially available). The results shown in Figure 2.7iii reveal the loss of the two ~14 MHz hyperfines, consistent with our model. Thus our studies using a combination of [ε-^{2}H]-met- and [β,γ-^{2}H]-met-labeled proteins support the role of met in the thiosulfuranyl radical.

2.3.10. Simulation of the EPR spectra (spectra were simulated by Dr. Guinevere Mathies, MIT)

As shown in Figures 2.4-2.7, simulations of the EPR spectra using the EPR simulation package EasySpin[57] were successfully performed, providing the underlying g-values and hyperfine interactions that support our electronic and structural assignments. The simulation of the spectrum of NrdD-(WT) in H_2O (Figure 2.4ii) was built step-by-step. First, g-values obtained at 140 GHz were used. Then, the hyperfine interactions were estimated for each of the four different types of protons from the appropriately labeled NrdDs in H_2O and D_2O (Figure 2.4ii-v and Figure 2.7). Finally all hyperfine interactions were optimized in a manual, global fit of all acquired spectra. These hyperfine interactions are summarized in Table 2.1.

The quality of all of the 9 and 140 GHz simulations improved when we included an isotropic line-broadening of 0.4 mT peak-peak and a g-strain 10 % of $|g_i-g_e|$, with i = x,y,z and g_e the g-value for the free electron. Furthermore, all samples were found to contain a residual 5 to 10% G•. To simulate its contribution to the spectra, its g-values (g_z = 2.0044, g_y = 2.0035, g_x = 2.0023),[53] an isotropic hyperfine interaction with a proton of a_{iso} = 39 MHz, and an isotropic line-broadening 1 mT peak-peak was included. For simulations of the thiosulfuranyl radical on NrdD-(S386E), the parameters derived from the spectra of NrdD-(WT) were used, except that g_x was adjusted to 2.042.

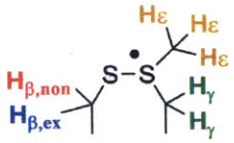
Additional observations concerning the simulations require comment. First, the match between our spectra and the simulations improved with the introduction of a small anisotropy to three of the four proton hyperfine interactions listed in Table 2.1. This anisotropic contribution arises from a dipole-dipole interaction and is expected to be small between a "π-electron" and a "π-proton" that lies above or below the plane of the electron spin distribution[58] (see Figure 2.9A and the quantum chemical calculations in the next section). Because these dipolar contributions are small, we did not attempt to quantify them by using more advanced fitting methods as they depend on three unknown Euler angles for every proton, which cannot be accurately determined from our current data set.

Second, on the low-field side of the 140 GHz spectrum (Figure 2.5) the intensity is lower than predicted by the simulation. This is caused by anisotropy of the relaxation properties of the thiosulfuranyl radical, also observed for thiyl radicals.[59,60]

Third, although the spectrum of [β-^{2}H]-cys-NrdD in H_2O (Figure 2.4iii) does not show any resolved fine structure due to the two γ-protons of methionine, the quality of the simulation improves when these hyperfine interactions are included. The same is true for the spectrum of [β-^{2}H]-cys and [ε-^{2}H]-met NrdD in H_2O (Figure 2.7iv), where including the γ-proton hyperfine interactions in the simulation reproduces the unresolved splitting of the 334 mT peak (marked with arrows).

Fourth, and in contrast with the above observation, in all samples where the exchangeable cysteine β-proton is replaced by deuterium (Figure 2.4iv,v and Figure 2.6ii), inclusion of the hyperfine interactions with the methionine γ-protons in the simulation gives rise to features that are not observed in the experimental spectra. More advanced EPR techniques such as [^{2}H]-ENDOR spectroscopy are required to investigate the presence of any interaction of the radical

with the methionine γ-protons in these and other samples, and to more accurately measure the hyperfine interactions that we have assigned.



	A_z	A_y	A_x
[ex β-^{1}H]-cys	50	49	48
[non-ex β-^{1}H]-cys	33	28	28
[γ-^{1}H]-met	11	14	14
[ε-^{1}H]-met	10	10	10

Table 2.1. The proton hyperfine interactions in MHz used to simulate the EPR spectra of the NrdD thiosulfuranyl radical shown in Figures 2.4-2.7.

2.3.11. Molecular orbitals of the model thiosulfuranyl radical (calculations were carried out by Dr. Jiahao Chen, MIT)

To better understand the electronic structure and bonding in thiosulfuranyl radicals, we undertook a computational study using a minimal model system consisting of methanethiyl radical and ethylmethylsulfide. DFT calculations were performed to obtain a plausible reaction coordinate for the formation of a thiosulfuranyl radical. This analysis yielded the reaction profile shown in Figure 2.8, as a function of the separation of the thiyl (S_1) and thioether (S_2) sulfur

atoms. An energetically minimal structure was found at a S_1-S_2 distance of 3.2 Å, with a binding energy of -1.8 kcal/mol relative to dissociated methanethiyl radical and ethylmethylsulfide.

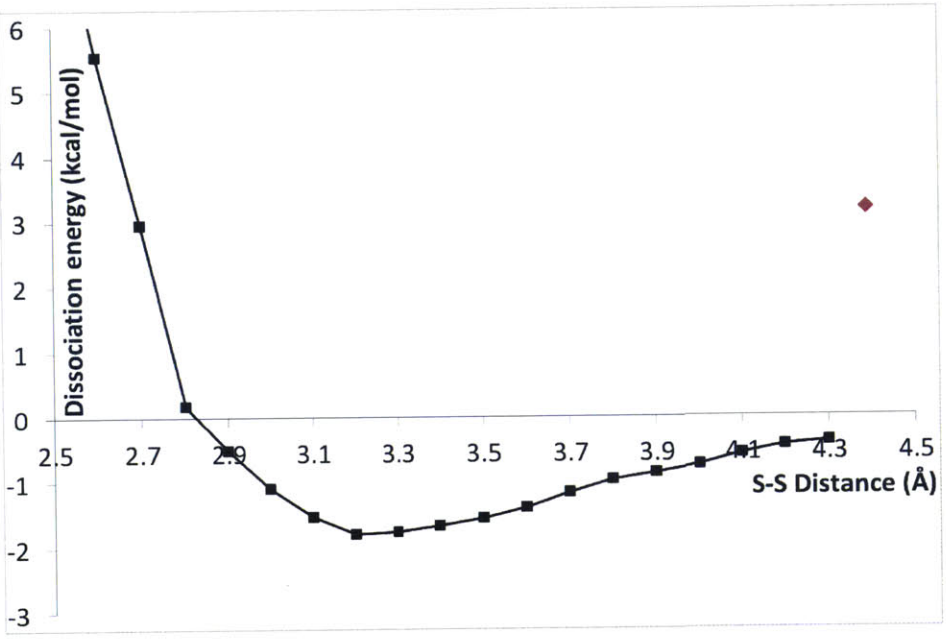


Figure 2.8. Minimal energy reaction profile along the S_1-S_2 dissociation coordinate (black line with squares). Shown for comparison is the native structure (red diamond), which does not lie on the minimal energy reaction coordinate.

The molecular orbitals involving the interacting sulfur 3p orbitals of the energetically minimal structure are shown in Figure 2.9A. The unrestricted DFT calculation yields separate spin orbitals for the α and β (up and down) spin electrons. The S_1 and S_2 $3p_z$ orbitals engage in a bonding interaction, with a σ* antibonding singly occupied molecular orbital (SOMO), and a σ bonding SOMO-1. Together, they account for a S_1-S_2 three-electron σ-bond of order ½. The frontier orbitals show significant electron amplitude on two of the hydrogen atoms on the methanethiyl fragment. Localization of a hydrogen atom within the lobe of the $\beta\sigma^*$ orbital could account for its lability as observed in the EPR experiments.

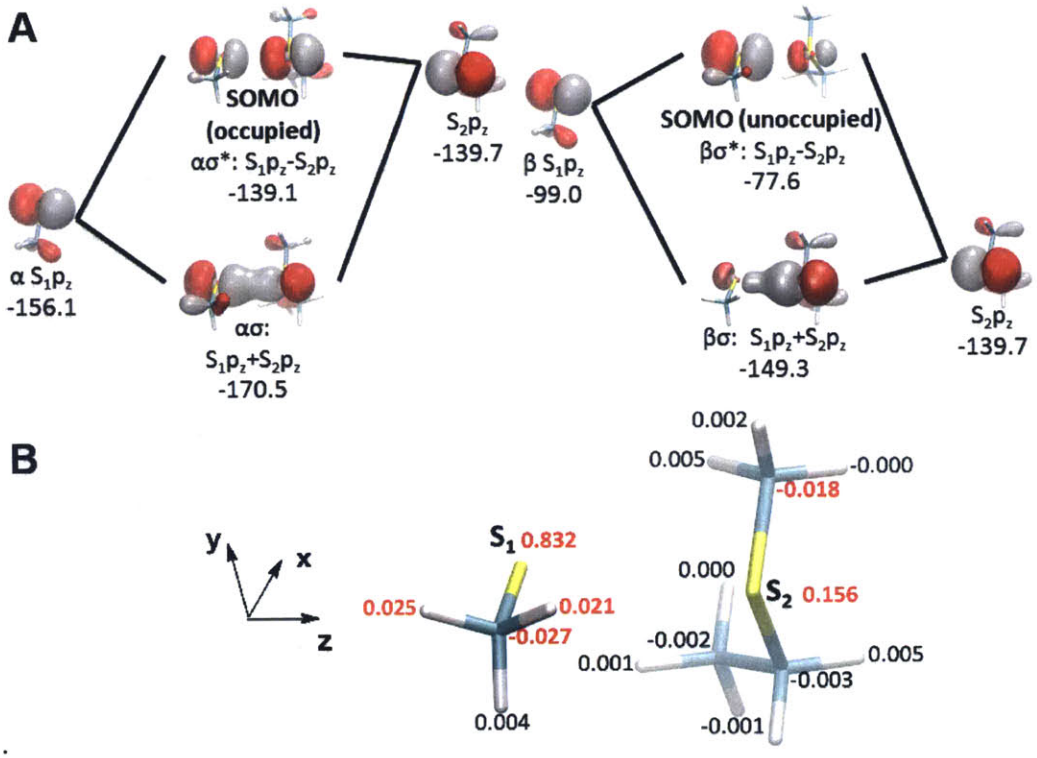


Figure 2.9. A) Frontier molecular orbitals for the model thiosulfuranyl radical formed from methanethiyl radical and ethylmethylsulfide. Diagrams for the α and β spin orbitals are shown separately on the left and right respectively, together with orbital energies in kcal/mol, which represent the ionization enthalpy of the electron in that orbital. For each diagram, the leftmost and rightmost orbitals represent the fragment molecular spin orbital of the methanethiyl and methylethylsulfide fragments respectively, which mix to produce the σ bonding and σ* antibonding orbitals. S_np_z refers to the $3p_z$ orbital on the sulfur atom of fragment n (1=methanethiyl, 2=ethylmethylsulfide). B) Atomic spin populations obtained from Löwdin population analysis on the model thiosulfuranyl radical, showing excess spin density localized primarily on the sulfur atoms, and secondarily on two of the hydrogen atoms of the methanethiyl fragment. Values in red indicate significant spin populations exceeding 0.01 in magnitude.

To quantify the spin excesses on an atomic level, Löwdin population analysis[61] was performed on the spin density matrix, and the resulting spin charges are shown in Figure 2.9B. Most of the spin (83%) is localized on the methanethiyl sulfur, with a smaller population (15%) on the thioether sulfur. In addition, a small amount of excess spin is localized on two methanethiyl hydrogens (2.1% and 2.5%), and to a lesser extent on several protons adjacent to

the thioether group (0.1 – 0.5%), which is consistent with the observation of hyperfine interactions with the cysteine β-protons and methionine γ- and ε-protons in the NrdD radical. In NrdD, the relative magnitude of the hyperfine interactions likely depends on the angle of rotation about the cysteine and methionine C-S bonds, which are dictated by the constraints of the protein scaffold.

2.3.12. Reaction of NrdD with CTP and ATP in the absence of formate generates cytosine

As shown in Figure 2.1C, we propose that formation of the thiosulfuranyl radical involves NrdD catalyzing the conversion of CTP to 3′-keto-dCTP (Figure 2.1C, 7 and 8). To investigate the fate of CTP during the reaction, the thiosulfuranyl radical was generated by incubating NrdD with 5-[^{3}H]-CTP and ATP in the formate-free buffer for 40 s at 4 °C followed by quenching in 1% perchloric acid. Workup of the reaction using a Dowex-1-borate column as shown in Figure 2.10, gave 1.0 ± 0.14 eq. of 5-[^{3}H]-Cyt and 0.38 ± 0.25 eq. of dC (identified by co-migration with a standard using HPLC). The former is a breakdown product of 3′-keto-dCTP (Figure 2.3 and Figure 2.10), consistent with our proposed model, and the latter is attributed to our inability to completely remove formate from the reactions components.

2.3.13. Addition of formate to the thiosulfuranyl radical species results in dCTP formation

As noted above, addition of 10 mM formate to the thiosulfuranyl radical results in the recovery of 80% of the initial G•. We propose that during this reaction, NrdD-bound 3′-keto-dCTP is reduced to dCTP, completing the catalytic cycle. The experiment was repeated as described above, incubating NrdD, 5-[^{3}H]-CTP and ATP in formate-free buffer for 40 s at 4 °C. A 1 mL solution containing 10 mM formate and 1 mM CTP was then added, followed by

immediate quenching by hand with 1% perchloric acid. The unlabeled CTP was added to dilute 5-[^{3}H]-CTP 27-fold to limit the contribution of [^{3}H]-dCTP to the first turnover.

The workup is as described (Figure 2.10) and the results reveal 0.12 ± 0.03 eq. of 5-[^{3}H]-Cyt and 1.57 ± 0.21 eq. of 5-[^{3}H]-dC (see Table 2.2). Comparison with the products formed by acid quench without addition of formate reflects a decrease in the amount of Cyt and increase in the amount of dC detected. This suggests that 0.88 ± 0.14 eq. of enzyme-bound 3′-keto-dCTP is consumed and converted to dCTP, and that NrdD carries out only 0.3 additional turnovers between addition of formate and acid quench.

To account for the 0.12 eq. of Cyt remaining after reaction with formate, we suggest that during the reaction time of 40 s prior to addition of formate, some amount of the 3′-keto-dCTP dissociates from the enzyme and decomposes to Cyt before or during the acid quench. However, it is remarkable that the results suggest that the 3′-keto-dCTP remains largely sequestered in the active site during this period.

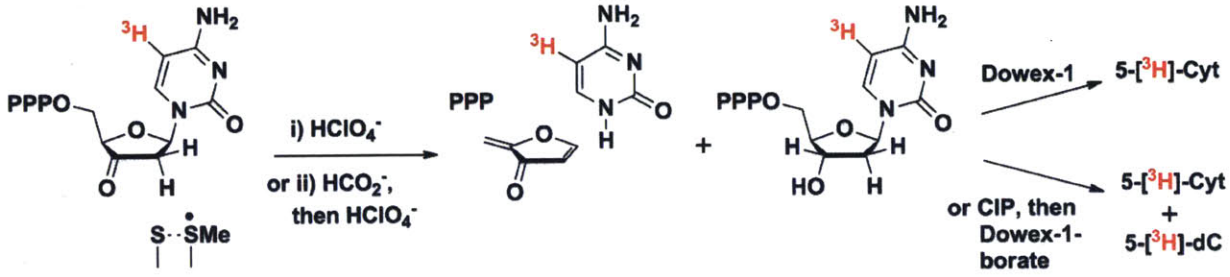


Figure 2.10. Products formed after quenching of the thiosulfuranyl radical generated by reaction of NrdD with 5-[^{3}H]-CTP and ATP. i) Quenching with perchloric acid leads to release of 5-[^{3}H]-Cyt from breakdown of 5-[^{3}H]-3′-keto-dCTP. ii) Addition of formate prior to quench leads to conversion of 3′-keto-dCTP to 5-[^{3}H]-dCTP. HPLC analysis distinguishes 5-[^{3}H]-Cyt from 5-[^{3}H]-dC. CIP is calf intestinal alkaline phosphatase.

Workup	Cyt (eq.)	dC (eq.)
i)	1.00 ± 0.14	0.38 ± 0.25
ii)	0.12 ± 0.03	1.57 ± 0.21

Table 2.2. Amount of 5-[^{3}H]-Cyt and 5-[^{3}H]-dC formed per G• after quenching the thiosulfuranyl radical *via* method i) or ii) in Figure 2.10, average of 5 experiments.

2.4. Discussion

Thiyl radicals have long been proposed to be involved in a diverse range of enzymatic reactions from the pyruvate-ferredoxin oxidoreductase[62] to pyruvate formate lyase[63] and all classes of RNRs[5]. More recent examples include mechanistically diverse glycyl radical enzymes involved in fermentation[64-66], and in the environmentally important processes of anaerobic hydrocarbon activation[67-69], and the non-canonical glycyl radical enzyme involved in the formation of methane from methyl phosphonate[70]. However, the spectroscopic observation of thiyl radicals and the demonstration of their kinetic competence has been challenging due to their short lifetimes and the large spin-orbit coupling with the sulfur atom, which results in short relaxation times and large g-anisotropy and as a consequence broad EPR lines.[71]

In the class I and II RNRs, three cysteines are involved in nucleotide reduction (Figure 2.1A,B): the transient thiyl radical located on the top face of the nucleotide and two cysteines located on its bottom face that are converted to a disulfide in the reduction process. Only in the case of the *Lactobacillus leichmannii* class II RNR has evidence for the involvement of a kinetically competent top face thiyl radical in the first half-reaction (Figure 2.1A) been demonstrated.[3,54] The evidence in support of this mechanism for the class I and III RNRs is

inferred based on structural homology of all the classes of RNRs[5] and extensive biochemical studies.

In the second half-reaction (Figure 2.1B), the involvement of a thiyl radical in the nucleotide reduction by the bottom face cysteines has been postulated.[15] In the *E. coli* class Ia RNR, E441 is believed to provide the proton required for the reduction of the 3′-keto-deoxynucleotide by the disulfide anion radical *via* proton coupled electron transfer (Figure 2.1B, 4 and 5). Support for this proposal has been provided by studies with the E441Q mutant that causes accumulation of the disulfide anion radical on the 10 s timescale, allowing its detection by pulsed, high-field EPR spectroscopy.[19] However, a limitation from the mutant studies is that the trapped radical is not chemically competent for deoxynucleotide formation because the essential proton for the PCET step is missing. Thus the detailed mechanism of the bottom face reduction chemistry, and specifically the involvement of thiyl radicals, has remained elusive.

Unlike class I and II RNRs, the class III RNRs studied to date use formate as a reductant.[23] We hoped that examination of this RNR might shed light on the second half reaction of all RNRs. Here we have shown that omission of formate from our assays leads to accumulation of a new radical upon reaction of NrdD with CTP and ATP, accompanied by formation of a bound cytidine species proposed to be 3′-keto-dCTP (Figure 2.1C, 7 and 8). Furthermore, addition of formate converts the bound nucleotide into dCTP with recovery of the G•. Our studies suggest that the new species is a thiosulfuranyl radical, and demonstrate its chemical competence. These results provide further evidence for the involvement of a thiyl radical in the reductive half-reaction of RNRs in general.

Evidence for the thiosulfuranyl radical: Assignment of the structure of the new radical is based on results from isotopic labeling studies with various NrdDs accompanied by analysis of their X band EPR spectra, g-values obtained by 140 GHz EPR studies, structural insight from the bacteriophage T4 NrdD, and computational studies. Important insight is also provided by recent density functional theory and correlated *ab initio* calculations by Van Gastel *et al.*[72] on the electronic structure of the cysteine thiyl radical and its unusual EPR parameters. Thiyl radicals exhibit broad g anisotropy with g_x-values(x axis parallel to the C_β-S bond) that range from 2.10 to 2.49.[60] In the cysteine thiyl radical, the EPR g-values are sensitive to the energy difference between the nearly-degenerate singly occupied orbital (SOMO with a predominant p_y character in Van Gastel's system) and one of the lone-pair orbitals (p_z).[72] This unusual property makes g_x highly sensitive to radical conformation and H bonds and requires an expression for the g-tensor in which third order corrections must be taken into account.

The g-tensor for our new radical from the 140 GHz EPR spectrum is also anisotropic (Figure 2.5), with a g_x value of 2.040. This value is distinct from the thiyl radicals and is consistent with the range of values reported for candidate thiosulfuranyl radicals generated by pulse radiolysis and laser flash photolysis of a range of thiols (2.027- 2.058).[55,56,73] The S-S bonds in these structures are weak and have only been observed in small molecules containing electronegative alkyl substituents (CF_3SSR_2• or $R'COSSR_2$•)[74], in glasses/solid matrices, or in intramolecular cases.[52] Arguments in favor of the assignment of these species as thiosulfuranyl radicals are described by Symons and coworkers.[56] In NrdD, the formation of the thiosulfuranyl radical is likely favored by the juxtaposition of the C175 and M382 sulfur atoms within the active site by the protein.

Our calculations on a model thiosulfuranyl radical provide an explanation for our observed g_x value relative to that of a thiyl radical. They suggest that the σ^* antibonding SOMO and the σ bonding SOMO-1 together account for a weak, S1-S2 three-electron σ-bond of order ½ due to the interaction between the sulfur $3p_z$ orbital of the thiyl radical and a nonbonding orbital of the thioether. This interaction is expected to perturb both the symmetry and degeneracy of the thiyl radical $3p_y$ and $3p_z$ orbitals, shifting the g_x value closer to that of the free electron.

Our isotopic labeling studies with NrdD / CTP / ATP provide strong evidence for localization of the radical on cysteine and methionine as hyperfine interactions associated with both cysteine β-protons and the methionine ε- and γ-protons are observed. The spin densities from our DFT calculations also suggest that excess spin of the three-electron bond includes small packets of spin localized on the cysteine β-protons, and to a lesser extent on the adjacent ε and γ protons of methionine (Figure 2.9B).

One initially puzzling observation from our EPR analyses was that one of the cysteine β–protons on the thiosulfuranyl radical was exchangeable with the solvent. However, perusal of the literature focused on examining the fate of thiyl radials generated by a variety of different methods in small molecules (cysteine, penicillamine, glutathione) in D_2O revealed reversible deuterium incorporation into both the β and / or α positions.[73,75,76] In cysteamine radicals generated by pulse radiolysis in acidic solution, hydrogen transfer reactions lead to the equilibration of $^{+}H_3NCH_2CH_2$-S• and $^{+}H_3NCH_2$•CH–SH species with very fast forward and reverse rates of $k_{12} \approx 10^5\ s^{-1}$ and $k_{21} \approx 1.5 \times 10^5\ s^{-1}$ respectively.[75] In the case of NrdD C175, we propose that its β-CH bond might be weakened by a hyperconjugative interaction with the singly-filled sulfur p_z orbital, facilitating its deprotonation and allowing exchange with deuterium from the solvent, although the mechanism remains to be determined. This result is

also consistent with our DFT calculations showing the localization of a proton within a lobe of the S-S σ^* orbital, and the observation that of the two diasteriotopic β-protons on cysteine (we do not know which one), the one with the largest hyperfine interaction is exchanged. The exchange is prevented in NrdD-(S386E), where the negative charge adjacent to the radical was designed to destabilize the conjugate anion and thus disfavor deprotonation (see Figure 2.2). Finally, the rapid exchange of hydrogens adjacent to the thiyl radical might provide a diagnostic for thiyl radicals in general which have been, for reasons outlined above, challenging to observe.

Mechanistic insight into the reductive half reactions of RNRs in general: Based on earlier studies[20,23,24,26,77] and our observations reported herein, we propose the mechanistic model for nucleotide reduction shown in Figure 2.1C for the class III RNR. The reaction is initiated by the generation of the CTP 3′-radical (2) by the C384 thiyl radical. In the absence of basic residues in the active site, formate has been proposed to act as a base to catalyze water loss to form a ketyl radical (3).[77] The ketyl radical gains an electron from C175 and a proton from formic acid, forming 3′-keto-dCTP and generating the thiosulfuranyl radical (8). For the experiments in formate-free buffer, 3′-keto-dCTP can still be generated without base catalysis. This observation is analogous to the E441Q mutant in the class I RNR where the 3′-keto-dCDP is still formed, but at a much slower rate.

Exchange of formate between the active site and buffer after formation of this stable radical rationalizes the [^{3}H]-formate isotope effect of 2.3 measured on the reduction process.[23] Oxidation of formate by the thiosulfuranyl radical generates $CO_2^{\bullet-}$ (9). A similar reaction between formate and DTT radicals generated by pulse radiolysis has previously been reported.[78] The relative reduction potentials of $CO_2^{\bullet-}$ (E° = -1.90 V) and $(CH_3)_2CO^{\bullet-}$ (E° = -2.10 V),[79]

suggest that reduction of 3′-keto-dCTP to the dCTP 3′-radical (10) may require the concerted delivery of a proton from an unknown source. Reduction of the product radical generates dCTP and regenerates the C384 thiyl radical (11).

In our mechanistic model, the C175 thiyl radical serves as the oxidant for formate. Although the function of the thiosulfuranyl radical, a stabilized form of the C175 thiyl radical, is not understood, its stoichiometric formation and ability to catalyze the formation of deoxynucleotide upon formate addition suggest its mechanistic importance. We hypothesize that since formate can enter and leave the active site even after generation of radical intermediates, formation of this species might protect the C175 thiyl radical *in vivo* under conditions when formate levels are low and the enzyme is saturated with substrate and effector nucleotides.

Bioinformatics and the nature of the class III RNR reductant: The unanticipated identification of a residue (M382) that plays a part in the reaction with formate, provides us with a handle to search for NrdDs that use alternative reductants among the many metabolically diverse organisms with sequenced genomes. An examination of the RNRdb[80] shows that M382 is not strictly conserved (Figure 2.11). However, all annotated archaeal and bacterial NrdD sequences lacking this M382 residue contain a cysteine residue in place of G383 on the thiyl radical loop adjacent to C384 (Figure 2.12), in a position that may allow formation of a disulfide with the bottom face thiol. In addition, the bacterial proteins contain a conserved glutamate residue aligning with Y542 in the active site (compare Figures 2.2A and 2.12), placing it in a position to act as a general base/acid catalyst analogous to the role of E441 in *E. coli* NrdA. These residues are shown in the homology model of the *Thermotoga maritima* NrdD given in Figure 2.12. Finally, archaeal NrdDs, like those found in *Archaeoglobus veneficus* and

Methanosarcina barkerii, contain a thioredoxin-like protein in the *nrdDG* operon. All these observations suggest that there exist NrdDs that use disulfide chemistry instead of formate for nucleotide reduction. The recombinant production and characterization of some of these proteins are the focus of ongoing studies.

Figure 2.11. NrdD alignments created using ClustalW[81]. Blue text denotes formate-utilizing NrdDs. Black and red text denote putative disulfide-utilizing NrdDs. Red text denotes NrdDs with a thioredoxin in the operon. Putative residues involved in chemistry highlighted in green.

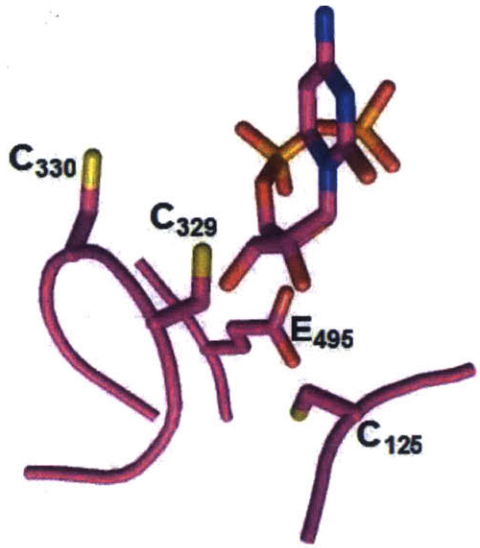


Figure 2.12. Homology model of *Thermotoga maritima* NrdD created using Phyre2.[82] Residues C330 and C125 are conserved in all RNRs. In addition, residues C329 and E495 are conserved in NrdDs related to *T. maritima* NrdD. We hypothesize that nucleotide reduction by *T. maritima* NrdD is accompanied by formation of a disulfide bond between C125 and C329.

Summary: Despite the many enzymatic reactions proposed to involve thiyl radicals,[5,62,63] there have been few systems that have allowed their spectroscopic and chemical characterization. In this study we found that reaction of the *E. coli* class III RNR with CTP in the absence of formate resulted in stoichiometric accumulation of a thiosulfuranyl radical, comprising a cysteine thiyl radical stabilized by a three-electron bond to a methionine residue. This new sulfur-based radical joins other sulfur-based radicals observed in the other two classes of RNRs: the class Ia disulfide anion radical[19] proposed to be involved directly in nucleotide reduction and the class II exchange coupled thiyl radical-cob(II)alamin[54] involved in 3′-hydrogen atom abstraction, that have been experimentally detected. Our results suggest that detecting thiyl radicals and controlling their reactivity may require an orchestrated constellation of residues adjacent to the

thiyl radical, that protects it from alternative chemistry while allowing the reaction to proceed rapidly when the substrate (effectors) are in the appropriate conformation.

2.5. Acknowledgements

We thank Dr. Guinevere Mathies who acquired the 140 GHz spectra and carried out all the EPR simulations, and Prof. Robert G. Griffin for guidance regarding EPR experiments and analysis. We thank Dr. Jiahao Chen who carried out all DFT calculations. We thank Prof. Kenichi Yokoyama who initiated the class III RNR project in this lab. We are grateful to Prof. Marc Fontecave for the gift of plasmids pRSS and pN9, and Profs. Willem Koppenol, Christian Schöneich and Bernhard Jaun for advice regarding the electronic structure and reactivity of thiyl and thiosulfuranyl radicals.

2.6. References

(1) Nordlund, P.; Reichard, P. *Annu. Rev. Biochem.* **2006**, *75*, 681.

(2) Hofer, A.; Crona, M.; Logan, D. T.; Sjöberg, B.-M. *Crit. Rev. Biochem. Mol. Biol.* **2012**, *47*, 50.

(3) Licht, S.; Gerfen, G. J.; Stubbe, J. *Science* **1996**, *271*, 477.

(4) Stubbe, J.; Ackles, D. *J. Biol. Chem.* **1980**, *255*, 8027.

(5) Stubbe, J. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 2723.

(6) Fontecave, M.; Eliasson, R.; Reichard, P. *Proc. Natl. Acad. Sci. U. S. A.* **1989**, *86*, 2147.

(7) Sun, X.; Ollagnier, S.; Schmidt, P. P.; Atta, M.; Mulliez, E.; Lepape, L.; Eliasson, R.; Graslund, A.; Fontecave, M.; Reichard, P.; Sjöberg, B. M. *J. Biol. Chem.* **1996**, *271*, 6827.

(8) Sofia, H. J.; Chen, G.; Hetzler, B. G.; Reyes-Spindola, J. F.; Miller, N. E. *Nucleic Acids Res.* **2001**, *29*, 1097.

(9) Gambarelli, S.; Luttringer, F.; Padovani, D.; Mulliez, E.; Fontecave, M. *ChemBioChem* **2005**, *6*, 1960.

(10) Rhodes, D. V.; Crump, K. E.; Makhlynets, O.; Snyder, M.; Ge, X.; Xu, P.; Stubbe, J.; Kitten, T. *J. Biol. Chem.* **2013**, *289*, 6273.

(11) Yokoyama, K.; Uhlin, U.; Stubbe, J. *J. Am. Chem. Soc.* **2010**, *132*, 15368.

(12) Licht, S. S.; Lawrence, C. C.; Stubbe, J. *J. Am. Chem. Soc.* **1999**, *121*, 7463.

(13) Ge, J.; Yu, G.; Ator, M. A.; Stubbe, J. *Biochemistry* **2003**, *42*, 10071.

(14) Seyedsayamdost, M. R.; Stubbe, J. *J. Am. Chem. Soc.* **2006**, *128*, 2522.

(15) Licht, S.; Stubbe, J. In *Comprehensive Natural Products Chemistry*; Barton, S., Nakanishi, K., Meth-Cohn, O., Poulter, C., Eds.; Elsevier Science: New York, 1999; Vol. 5, p 163.

(16) Stubbe, J.; Ator, M.; Krenitsky, T. *J. Biol. Chem.* **1983**, *258*, 1625.

(17) Stubbe, J.; Ackles, D.; Segal, R.; Blakley, R. L. *J. Biol. Chem.* **1981**, *256*, 4843.

(18) Lenz, R.; Giese, B. *J. Am. Chem. Soc.* **1997**, *119*, 2784.

(19) Lawrence, C. C.; Bennati, M.; Obias, H. V.; Bar, G.; Griffin, R. G.; Stubbe, J. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 8979.

(20) Eliasson, R.; Pontis, E.; Eckstein, F.; Reichard, P. *J. Biol. Chem.* **1994**, *269*, 26116.

(21) Eklund, H.; Fontecave, M. *Structure* **1999**, *7*, R257.

(22) Mao, S. S.; Holler, T. P.; Yu, G. X.; Bollinger, J. M., Jr.; Booker, S.; Johnston, M. I.; Stubbe, J. *Biochemistry* **1992**, *31*, 9733.

(23) Mulliez, E.; Ollagnier, S.; Fontecave, M.; Eliasson, R.; Reichard, P. *Proc. Natl. Acad. Sci. U. S. A.* **1995**, *92*, 8759.

(24) Eliasson, R.; Reichard, P.; Mulliez, E.; Ollagnier, S.; Fontecave, M.; Liepinsh, E.; Otting, G. *Biochem. Biophys. Res. Commun.* **1995**, *214*, 28.

(25) Logan, D. T.; Andersson, J.; Sjöberg, B.-M.; Nordlund, P. *Science* **1999**, *283*, 1499.

(26) Andersson, J.; Westman, M.; Sahlin, M.; Sjöberg, B.-M. *J. Biol. Chem.* **2000**, *275*, 19449.

(27) Stubbe, J. A.; Kozarich, J. W. *J. Biol. Chem.* **1980**, *255*, 5511.

(28) Ator, M. A.; Stubbe, J. *Biochemistry* **1985**, *24*, 7214.

(29) Harris, G.; Ashley, G. W.; Robins, M. J.; Tolman, R. L.; Stubbe, J. *Biochemistry* **1987**, *26*, 1895.

(30) Harris, G.; Ator, M.; Stubbe, J. *Biochemistry* **1984**, *23*, 5214.

(31) Sun, X.; Eliasson, R.; Pontis, E.; Andersson, J.; Buist, G.; Sjöberg, B.-M.; Reichard, P. *J. Biol. Chem.* **1995**, *270*, 2443.

(32) Lohman, G. J.; Gerfen, G. J.; Stubbe, J. *Biochemistry* **2010**, *49*, 1396.

(33) Fish, W. *Methods Enzymol.* **1988**, *158*, 357.

(34) Malmström, B. G.; Reinhammar, B.; Vänngård, T. *Biochim. Biophys. Acta (BBA)-Bioenergetics* **1970**, *205*, 48.

(35) Steeper, J.; Steuart, C. *Anal. Biochem.* **1970**, *34*, 123.

(36) Luttringer, F.; Mulliez, E.; Dublet, B.; Lemaire, D.; Fontecave, M. *J. Biol. Inorg. Chem.* **2009**, *14*, 923.

(37) Smith, A. A.; Corzilius, B.; Bryant, J. A.; DeRocher, R.; Woskov, P. P.; Temkin, R. J.; Griffin, R. G. *J. Magn. Reson.* **2012**, *223*, 170.

(38) Maly, T.; Bryant, J.; Ruben, D.; Griffin, R. G. *J. Magn. Reson.* **2006**, *183*, 303.

(39) van der Donk, W. A.; Stubbe, J.; Gerfen, G. J.; Bellew, B. F.; Griffin, R. G. *J. Am. Chem. Soc.* **1995**, *117*, 8908.

(40) Baba, T.; Ara, T.; Hasegawa, M.; Takai, Y.; Okumura, Y.; Baba, M.; Datsenko, K. A.; Tomita, M.; Wanner, B. L.; Mori, H. *Mol. Syst. Biol.* **2006**, *2*.

(41) Van Duyne, G. D.; Standaert, R. F.; Karplus, P. A.; Schreiber, S. L.; Clardy, J. *J. Mol. Biol.* **1993**, *229*, 105.

(42) Shao, Y.; Molnar, L. F.; Jung, Y.; Kussmann, J.; Ochsenfeld, C.; Brown, S. T.; Gilbert, A. T.; Slipchenko, L. V.; Levchenko, S. V.; O'Neill, D. P. *Phys. Chem. Chem. Phys.* **2006**, *8*, 3172.

(43) Clark, T.; Chandrasekhar, J.; Spitznagel, G. W.; Schleyer, P. V. R. *J. Comput. Chem.* **1983**, *4*, 294.

(44) Becke, A. D. *The Journal of Chemical Physics* **1993**, *98*, 5648.

(45) Humphrey, W.; Dalke, A.; Schulten, K. *Journal of molecular graphics* **1996**, *14*, 33.

(46) Tamarit, J.; Gerez, C.; Meier, C.; Mulliez, E.; Trautwein, A.; Fontecave, M. *J. Biol. Chem.* **2000**, *275*, 15669.

(47) Bianchi, V.; Eliasson, R.; Fontecave, M.; Mulliez, E.; Hoover, D. M.; Matthews, R. G.; Reichard, P. *Biochem. Biophys. Res. Commun.* **1993**, *197*, 792.

(48) Bianchi, V.; Reichard, P.; Eliasson, R.; Pontis, E.; Krook, M.; Jörnvall, H.; Haggård-Ljungquist, E. *J. Bacteriol.* **1993**, *175*, 1590.

(49) Ollagnier, S.; Mulliez, E.; Schmidt, P. P.; Eliasson, R.; Gaillard, J.; Deronzier, C.; Bergman, T.; Graslund, A.; Reichard, P.; Fontecave, M. *J. Biol. Chem.* **1997**, *272*, 24216.

(50) Giles, J. R. M.; Roberts, B. P. *J. Chem. Soc., Chem. Commun.* **1978**, 623.

(51) Anklam, E.; Steenken, S. *Journal of Photochemistry and Photobiology A: Chemistry* **1988**, *43*, 233.

(52) Anklam, E.; Margaretha, P. *Res. Chem. Intermed.* **1989**, *11*, 127.

(53) Duboc-Toia, C.; Hassan, A. K.; Mulliez, E.; Ollagnier-de Choudens, S.; Fontecave, M.; Leutwein, C.; Heider, J. *J. Am. Chem. Soc.* **2003**, *125*, 38.

(54) Gerfen, G. J.; Licht, S.; Willems, J. P.; Hoffman, B. M.; Stubbe, J. *J. Am. Chem. Soc.* **1996**, *118*, 8192.

(55) Engström, M.; Vahtras, O.; Ågren, H. *Chem. Phys. Lett.* **2000**, *328*, 483.

(56) Nelson, D. J.; Petersen, R. L.; Symons, M. C. R. *Journal of the Chemical Society, Perkin Transactions 2* **1977**, 2005.

(57) Stoll, S.; Schweiger, A. *J. Magn. Reson.* **2006**, *178*, 42.

(58) Heller, C.; McConnell, H. M. *The Journal of Chemical Physics* **1960**, *32*, 1535.

(59) Kolberg, M.; Bleifuss, G.; Sjöberg, B.-M.; Gräslund, A.; Lubitz, W.; Lendzian, F.; Lassmann, G. *Arch. Biochem. Biophys.* **2002**, *397*, 57.

(60) Lassmann, G.; Kolberg, M.; Bleifuss, G.; Graslund, A.; Sjoberg, B.-M.; Lubitz, W. *Phys. Chem. Chem. Phys.* **2003**, *5*, 2442.

(61) Lowdin, P.-O. *Adv. Quantum Chem.* **1970**, *5*, 185.

(62) Ragsdale, S. W. *Chem. Rev. (Washington, DC, U. S.)* **2003**, *103*, 2333.

(63) Reddy, S. G.; Wong, K. K.; Parast, C. V.; Peisach, J.; Magliozzo, R. S.; Kozarich, J. W. *Biochemistry* **1998**, *37*, 558.

(64) O'Brien, J. R.; Raynaud, C.; Croux, C.; Girbal, L.; Soucaille, P.; Lanzilotta, W. N. *Biochemistry* **2004**, *43*, 4635.

(65) Yu, L.; Blaser, M.; Andrei, P. I.; Pierik, A. J.; Selmer, T. *Biochemistry* **2006**, *45*, 9584.

(66) Craciun, S.; Balskus, E. P. *Proceedings of the National Academy of Sciences* **2012**, *109*, 21307.

(67) Leuthner, B.; Leutwein, C.; Schulz, H.; Hörth, P.; Haehnel, W.; Schiltz, E.; Schägger, H.; Heider, J. *Mol. Microbiol.* **1998**, *28*, 615.

(68) Jarling, R.; Sadeghi, M.; Drozdowska, M.; Lahme, S.; Buckel, W.; Rabus, R.; Widdel, F.; Golding, B. T.; Wilkes, H. *Angewandte Chemie International Edition* **2012**, *51*, 1334.

(69) Heider, J. *Curr. Opin. Chem. Biol.* **2007**, *11*, 188.

(70) Kamat, S. S.; Williams, H. J.; Dangott, L. J.; Chakrabarti, M.; Raushel, F. M. *Nature* **2013**, *497*, 132.

(71) Symons, M. C. *Journal of the Chemical Society, Perkin Transactions 2* **1974**, 1618.

(72) van Gastel, M.; Lubitz, W.; Lassmann, G.; Neese, F. *J. Am. Chem. Soc.* **2004**, *126*, 2237.

(73) Hofstetter, D.; Thalmann, B.; Nauser, T.; Koppenol, W. H. *Chem. Res. Toxicol.* **2012**, *25*, 1862.

(74) Giles, J. R. M.; Roberts, B. P. *Journal of the Chemical Society, Perkin Transactions 2* **1980**, 1497.

(75) Nauser, T.; Koppenol, W. H.; Schöneich, C. *The Journal of Physical Chemistry B* **2012**, *116*, 5329.

(76) Mozziconacci, O.; Williams, T. D.; Schöneich, C. *Chem. Res. Toxicol.* **2012**, *25*, 1842.

(77) Andersson, J.; Bodevin, S.; Westman, M.; Sahlin, M.; Sjöberg, B.-M. *J. Biol. Chem.* **2001**, *276*, 40457.

(78) Elliot, A. J.; Simsons, A. S.; Sopchyshyn, F. C. *Radiation Physics and Chemistry (1977)* **1984**, *23*, 377.

(79) Schwarz, H.; Dodson, R. *The Journal of Physical Chemistry* **1989**, *93*, 409.

(80) Lundin, D.; Torrents, E.; Poole, A.; Sjöberg, B.-M. *BMC Genomics* **2009**, *10*, 589.

(81) Larkin, M. A.; Blackshields, G.; Brown, N. P.; Chenna, R.; McGettigan, P. A.; McWilliam, H.; Valentin, F.; Wallace, I. M.; Wilm, A.; Lopez, R.; Thompson, J. D.; Gibson, T. J.; Higgins, D. G. *Bioinformatics* **2007**, *23*, 2947.

(82) Kelley, L. A.; Sternberg, M. J. *Nat. Protoc.* **2009**, *4*, 363.

Chapter 3:

The class III ribonucleotide reductase from *Neisseria bacilliformis* can utilize thioredoxin as a reductant

Adapted from Wei, Y., Funk, M. A., Rosado, L. A., Baek, J., Drennan, C. L., and Stubbe, J. (2014) *Proc. Natl. Acad. Sci. U.S.A.*, **111**(36), E3756-E3765.

3. The class III ribonucleotide reductase from *Neisseria bacilliformis* can utilize thioredoxin as a reductant

3.1. Introduction

The class III ribonucleotide reductases (RNRs) are glycyl radical enzymes present in many strict and facultative anaerobes that catalyze the conversion of nucleotides to deoxynucleotides [1,2] *via* a mechanism involving complex free radical chemistry, and are largely responsible for providing the balanced pool of deoxynucleotides required for DNA synthesis and repair [3]. The class III RNRs that have been characterized thus far obtain the reducing equivalents required to make the deoxynucleoside triphosphates (dNTPs) from the oxidation of formate to CO_2 [4]. Here we report a second subtype of class III RNR from *Neisseria bacilliformis*, which can obtain its reducing equivalents from the thioredoxin (TrxA) / thioredoxin reductase (TrxB) / NADPH system.

RNRs provide the only pathway for *de novo* biosynthesis of dNTPs [5]. They share a structurally homologous active site architecture in the α subunit, and a partially-conserved, radical based reduction mechanism. RNRs have been isolated and characterized from all kingdoms of life and, based on the characterization of these proteins thus far, are divided into three classes (I, II and III) according to the metallo-cofactor used to generate a thiyl radical that initiates the radical-dependent reduction chemistry [6]. The class I RNRs use cofactors generated by the reaction of reduced metals (Fe, Mn and Fe/Mn) and O_2, and are present only in aerobic organisms. The class II RNRs use adenosylcobalamin in an O_2-independent reaction, and are present in both aerobes and anaerobes. The class III RNR uses an O_2-sensitive glycyl radical (G•) [2] situated in the α protein (NrdD), which is generated by a separate activating enzyme (NrdG) *via* radical S-adenosylmethionine (SAM)-$[4Fe4S]^{1+}$ chemistry [7,8]. The class III RNRs are

only found in facultative and obligate anaerobes. A second distinction between the three classes has been the source of the reducing equivalents for nucleotide reduction. In the class I and II RNRs, they are provided by a redoxin (thioredoxin, glutaredoxin, or NrdH) which is rereduced by thioredoxin reductase and NADPH [9-11]. In contrast, for the bacteriophage T4 [12], its gram negative host *Escherichia coli* [1], and the gram-positive *Lactococcus lactis* [13], the only class III RNRs characterized in detail to date, nucleotide reduction is coupled to the oxidation of formate to CO_2 [4].

Formate in *E. coli* and *L. lactis* is provided by carrying out the fermentation of sugars to acetate and formate *via* a pathway involving pyruvate-formate lyase (PFL) [14,15]. For *E. coli* growing in the absence of electron acceptors, formate induces the formate-hydrogenlyase pathway in which it is converted to the waste products H_2 and CO_2 by formate dehydrogenase (FDH) and hydrogenase [16]. However, there are many proteins annotated as class III RNRs present in diverse bacteria and archaea [17,18], many of which do not possess PFL or generate formate as an intermediate or end-product in their primary metabolism [19], suggesting that an alternative reducing system for class III RNRs might be involved. This variability in the presence of formate-producing pathways is in contrast to the ubiquitous distribution of thioredoxin-like proteins used by the class I and II RNRs. This observation prompted us to carry out a bioinformatics search for candidate class III RNRs that employ disulfide chemistry similar to the class I and II enzymes.

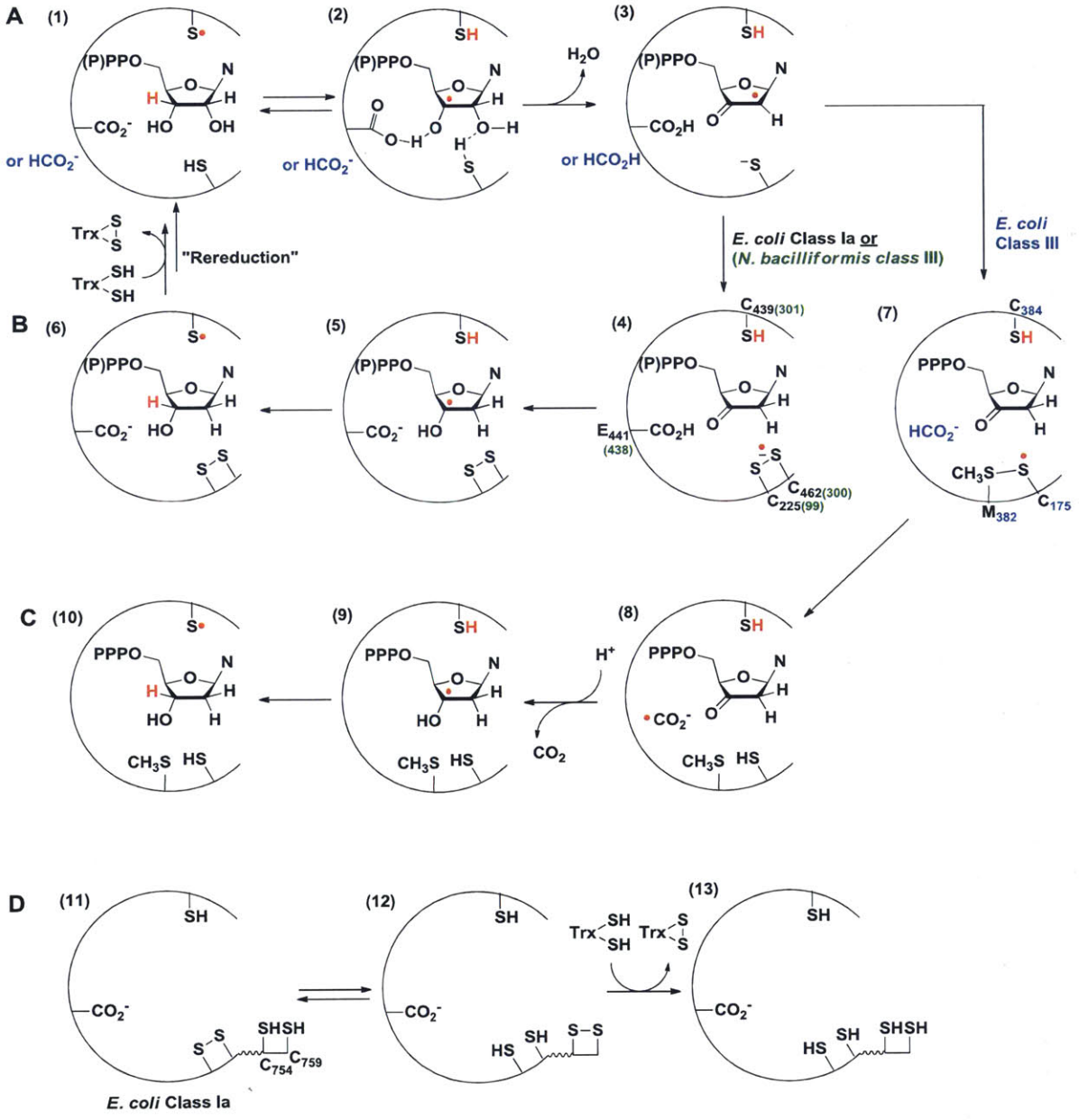


Figure 3.1. Mechanistic model for nucleotide reduction by RNRs. A) First half reaction common to all RNRs; B) second half reaction of *E. coli* class Ia and *N. bacilliformis* class III RNR; C) second half reaction of *E. coli* class III RNR. M382 in *E. coli* class III RNR is located two residues from the top face thiyl radical (C384), in a position similar to the conserved N437 in *E. coli* class Ia RNR which makes a hydrogen bond to the 2′-OH group of the substrate. D) Mechanistic model for rereduction of the active-site disulfide in class I and II RNRs *via* a pair of conserved cysteines on the C-terminal tail of α.

The generic mechanism of nucleotide reduction by all three classes of RNRs can be divided into two half reactions: the radical initiation process and the reduction process (Figure 3.1) [20,21]. In all RNR classes, nucleotide reduction is initiated by generating a 3′-nucleotide radical [22-24] (Figure 3.1A, **2**) *via* a transient, conserved, top face thiyl radical [25] (**1**) on the Cys loop in the active-site. This reaction likely involves general base catalysis by a conserved glutamate (class I and II RNRs), and perhaps a formate (class III RNRs) [26], which facilitates loss of water to form a ketyl radical (**3**). In the class I and II RNRs, reduction of the ketyl radical to a 3′-keto-deoxynucleotide intermediate is accompanied by the oxidation of the conserved cysteines [27] on the bottom face of the nucleotide to a disulfide anion radical [28] (Figure 3.1B, **4**), which serves as the reductant for the ketonucleotide, forming a 3′-deoxynucleotide radical and a disulfide (**5**). Product formation is accompanied by regeneration of the top face thiyl radical (**6**). Rereduction of the active-site disulfide by thioredoxin occurs *via* disulfide exchange with a pair of conserved cysteines on the C-terminal tail of α (Figure 3.1D) [27].

In the class III RNR, only one of the disulfide-forming cysteines in the active site is conserved [21,29,30], and we recently showed [31] that reduction of the ketyl radical to the 3′-keto-deoxynucleotide is accompanied by the formation of a thiosulfuranyl radical (Figure 3.1C, **7**) between the bottom face cysteine thiyl radical and a methionine residue. The thiosulfuranyl radical, in equilibrium with the thiyl radical, then oxidizes formate to a $\bullet CO_2^-$ radical [21] (**8**), proposed to serve as the reductant for the 3′-keto-deoxynucleotide. In all classes, the 3′-keto-deoxynucleotide intermediate (Figure 3.1B, **4** or Figure 3.1C, **7**) is proposed to be reduced by proton coupled electron transfer [32], with the source of the proton being the conserved glutamate in the class I and II RNRs, and unknown in the class III RNR.

Because of the role of this methionine residue in the reaction with formate in the *E. coli* class III RNR, we expected that it would be conserved in all formate-dependent NrdDs. However, just as the pathways for formate production are not conserved, our search in the RNRdb [33] showed that this methionine is missing in a set of NrdD sequences. In addition, all annotated NrdD sequences lacking this methionine residue invariably contain an additional cysteine residue immediately preceding the conserved thiyl radical on the Cys loop. This location may allow formation of a disulfide between the additional cysteine and the conserved bottom face thiol, thus allowing the reducing equivalents to be provided by chemistry similar to that in the class I and II RNRs.

To establish whether some class III RNRs use a formate-independent reduction strategy, a number of candidate class III RNRs were cloned and expressed. We now report the characterization of the NrdD and NrdG proteins from *Neisseria bacilliformis* (NbNrdD and NbNrdG) [34,35]. This organism lacks the fermentative pathway terminating in PFL, a major source of formate for the class III RNRs studied to date, and its NrdD lacks the active site methionine. In addition, we were able to clone, express, isolate and crystallize a related NrdD from the deep-branching thermophilic bacterium *Thermotoga maritima* (TmNrdD, 30% sequence identity with NbNrdD). The mesophilic NbNrdD proved more amenable to biochemical studies, and we demonstate that it is a G• enzyme and show that, like the previously characterized class III RNRs, NTPs are substrates. We also show that formate is unable to provide the reducing equivalents to make dNTPs. This organism possesses a TrxA that has 61% sequence identity with *E. coli* TrxA, and activity can be reconstituted *in vitro* using the *E. coli* TrxA / TrxB / NADPH system. The X-ray structure of TmNrdD reported here provides our framework for modeling the conserved residues in this newly discovered class III RNR subtype, and supports the hypothesis for the

NbNrdD that three cysteines and a glutamate are located in the active site region where they can play a role in catalysis. The distribution and significance of this new form of class III RNR are discussed.

3.2. Materials and methods

3.2.1. Materials and general methods

All chemical reagents were purchased from Sigma–Aldrich, unless otherwise indicated. Primers were purchased from Integrated DNA Technologies. UV-vis absorption spectroscopy was performed on an Agilent 8453 Diode Array spectrophotometer. Anaerobic procedures were carried out in a custom-designed MBraun glovebox equipped with a chiller at 15°C. All solutions and proteins were made anaerobic on a Schlenk line by 3 cycles of evacuation (5 min) followed by flushing with Ar gas (10 min) before being brought into the glovebox. Nucleotides, SAM and NADPH were brought into the glovebox as lyophilized solids. E. coli TrxA and TrxB were purified according to published procedures[36,37]. *N. bacilliformis* and *T. maritima* genomic DNA were gifts from Dr. Xiang-Yang Han, University of Texas M. D. Anderson Cancer Center, and Prof. Kenneth Noll, University of Connecticut respectively.

3.2.2. Cloning of NbNrdD, NbNrdG, and TmNrdD

N. bacilliformis nrdD and *nrdG* were obtained by PCR from genomic DNA (gift from Dr. Xiang-Yang Han, University of Texas M. D. Anderson Cancer Center) and *T. maritima nrdD* was obtained by PCR from genomic DNA (gift from Prof. Kenneth Noll, University of Connecticut), using Phusion polymerase (NEB) and the primers TmNrdDf/r, NbNrdDf/r and NbNrdGf/r (Table 3.1), and inserted into pET28a (Novagen), linearized with NdeI and HindIII, by Gibson

isothermal assembly[38], yielding the plasmids pET28a-NbNrdD, pET28a-NbNrdG and pET28a-TmNrdD. Because of the insolubility of the protein expressed from pET28a-NbNrdG, the *nrdG* gene was excised using the restriction enzymes NdeI and XhoI, and ligated into pSV272, which contains a maltose-binding protein (MBP) with an N-terminal His_6-tag, and linearized with the same enzymes to obtain pMBP-NbNrdG. The plasmids pET28a-NbNrdD(C300A), pET28a-NbNrdD(C301A), and pET28a-NbNrdD(E438Q), containing the C300A, C301A and E438Q mutation respectively, were constructed by site-directed mutagenesis from pET28a-NbNrdD using the primers NbNrdD(C300A)f/r, NbNrdD(C301A)f/r, and NbNrdD(E438Q)f/r (Table 3.1). All constructs were confirmed by DNA sequencing at the Massachusetts Institute of Technology Biopolymers Laboratory. The pET28a plasmid contains an N-terminal His_6 affinity purification tag followed by a thrombin cleavage site (MGSSH HHHHH SSGLV PRGSH-).

Name	Sequence
TmNrdDf	GCGGCCTGGTGCCGCGCGGCAGCCATATGAAGGTTCAGTATTCGTTCGAAAGAG
TmNrdDr	GGTGGTGCTCGAGTGCGGCCGCTCACCTTCTGATTGTCAGCG
NbNrdDf	GCGGCCTGGTGCCGCGCGGCAGCCATATGATTCGGCTGTATCCCGAACAG
NbNrdDr	GGTGGTGCTCGAGTGCGGCCGCCTAAGCCGCCTCGCGCTGCTTTTGC
NbNrdGf	GCGGCCTGGTGCCGCGCGGCAGCCATATGCCGAGCCTGAAATTCACCACCG
NbNrdGr	GGTGGTGCTCGAGTGCGGCCGCTCAGGCCGTCTGAACAAACGCCC
NbNrdD(C300A)f	CTCGCTCGCCTCCGCCTGCCGCCTGCGC
NbNrdD(C300A)r	GCGCAGGCGGCAGGCGGAGGCGAGCGAG
NbNrdD(C301A)f	CGCCTCCTGCGCCCGCCTGCGCAACGCC
NbNrdD(C301A)r	GGCGTTGCGCAGGCGGGCGCAGGAGGCG
NbNrdD(E438Q)f	CTCGCTCGCCTCCTCCTGCCGCCTGCGC
NbNrdD(E438Q)r	GCGCAGGCGGCAGGAGGAGGCGAGCGAG

Table 3.1 Primers used in cloning and mutagenesis (mutations underlined).

3.2.3. Expression and purification of NbNrdD and NbNrdG

Expression and purification of the NbNrdD and MBP-NbNrdG followed a similar protocol. The plasmids pET28a-NbNrdD and pMBP-NbNrdG were separately transformed into BL21 (DE3) codon plus (RIL) cells (Stratagene), grown on LB-agar plates with 50 μg/mL kanamycin (Kan). A single colony was inoculated into 5 mL starter culture of LB (50 μg/mL

Kan in all growths), grown at 37°C until saturated (12 h), and transferred into 200 mL of LB. Media for expression of NbNrdD was supplemented with 50 μM $ZnSO_4$. The cultures were grown at 37°C with shaking at 200 rpm. At OD600 ~0.6, the temperature was decreased to 20°C and IPTG (Promega) was added to a final concentration of 0.1 mM. After 12 h, cells were pelleted by centrifugation (4,000 × g, 10 min, 4°C) at and frozen at -80°C. Typical yield was ~5 g of cell paste per L.

Cell paste (~1 g) was resuspended in 25 mL of 20 mM Tris pH 7.5, 5 mM dithiothreitol (DTT), 1 mM PMSF (buffer A). The cells were lysed by a single passage through a French pressure cell (14,000 psi). DNA was precipitated by dropwise addition of 5 mL of buffer A containing 6% (w/v) streptomycin sulfate. The mixture was shaken for an additional 10 min, and the precipitated DNA was removed by centrifugation (20,000 × g, 10 min, 4°C). Solid $(NH_4)_2SO_4$ was then added to 60% saturation. The solution was shaken for an additional 20 min, and the precipitated protein was isolated by centrifugation (20,000 × g, 10 min, 4°C).

The pellet was dissolved in 30 mL of 20 mM Tris pH 7.5, 5 mM tris(2-carboxyethyl)phosphine (TCEP), 0.2 M KCl (buffer B) and incubated with 2 mL of TALON resin (Clontech) with shaking for 30 min. The column was then packed (0.8 × 4 cm) and washed with 10 CV of buffer B, followed by 10 CV of buffer B containing 5 mM imidazole. Protein was eluted with 5 CV of buffer B containing 150 mM imidazole. The eluted protein was precipitated with solid $(NH_4)_2SO_4$ to 60% saturation, and isolated by centrifugation (20,000 × g, 10 min, 4°C). The pellet was dissolved in 0.5 mL of buffer B and desalted using a Sephadex G-25 column (1.5 cm × 8.5 cm, 15 mL), pre-equilibrated with 20 mM Tris pH 7.5, 5% glycerol, 1 mM DTT (buffer C). The eluted protein was concentrated to ~200 μM by ultrafiltration (Amicon YM-30), frozen

in aliquots in liquid N2 and stored at -80°C. The final yield was ~10 mg per g cells for NbNrdD (ε280 = 73410 M-1cm) and ~1 mg per g cells for MBP-NbNrdG (ε280 = 95760 M-1cm).

3.2.4. Preparation of [SeMet]-labeled TmNrdD for crystallography

The procedure was adapted from existing protocols [39]. The plasmid pET28a-TmNrdD was transformed into BL21 (DE3) codon plus (RIL) cells (Stratagene), grown on LB-agar plates with 50 μg/mL kanamycin (Kan). A single colony was inoculated into 5 mL starter culture of LB (50 μg/mL Kan in all growths), grown at 37°C until saturated (12 h), harvested by centrifugation (3,000 × g, 10 min, 4°C) and transferred into 200 mL of M9 minimal medium supplemented with: glucose (0.4%), thiamine (50 mg/L), Kan (50 mg/L), $ZnSO_4$ (50 μM), $FeCl_3$ (10 μM), $MgCl_2$ (2 mM), $CaCl_2$ (0.1 mM), and the L-amino acids lysine (100 mg/L), phenylalanine (100 mg/L), threonine (100 mg/L), isoleucine (50 mg/L), leucine (50 mg/L) and valine (50 mg/L). The culture was grown at 37°C with shaking at 220 rpm. At OD_{600} ~0.3, SeMet (50 mg/L) was added, followed by shaking for 20 min. The temperature was lowered to 20°C and IPTG (Promega) was added to a final concentration of 0.2 mM. After 12 h, cells were pelleted by centrifugation (4,000 × g, 10 min, 4°C) and frozen at -80°C. The yield was ~1 g of cell paste and purification was carried out according to the procedure described for NbNrdD. The final yield was ~2 mg per g cells (ε_{280} = 106830 $M^{-1}cm$).

3.2.5. Reconstitution of the NbNrdG [4Fe4S] cluster

The procedure was carried out in a glovebox. Solutions of Na_2S and of $Fe(NH_4)_2(SO_4)_2$ in water (100 mM) were freshly prepared in the glovebox. A solution of MBP-NbNrdG (200 μM, 0.3 mL) was made anaerobic on a Schlenk line and brought into the glovebox. A solution of

DTT (dithiothreitol, 1 M) was added to 10 mM, followed by ordered addition of the solution of Na_2S (5 eq.) and $Fe(NH_4)_2(SO_4)_2$ (5 eq.). The mixture was incubated for 12 h at 4°C. EDTA (ethylenediaminetetraacetic acid, 5 eq.) was then added and the solution was desalted using a Sephadex G-25 column (1 × 9 cm, 7 mL) equilibrated with Tris buffer (20 mM, pH 7.5). The final material contained ~2.5 atoms of Fe per peptide determined by the ferrozine assay [40].

3.2.6. Generation of the NbNrdD G•

In a 1.5 mL polypropylene Eppendorf tube, a 50 μL mixture of NbNrdD (40 μM), NbNrdG (20 μM), SAM (0.5 mM), Bicine potassium salt pH 7.5 (30 mM) and acriflavin (10 μM) was placed 1 m away from a fluorescent lamp in the glovebox at 15°C for 3 h. For inspection by X-band EPR spectroscopy, the solution was diluted to 200 μL with Tris buffer (20 mM, pH 7.5), 5% glycerol to give a final concentration of 10 μM NbNrdD, and sealed in an EPR tube with a rubber stopper. The solution was quenched in liquid N_2 immediately after removal from the glovebox. The amount of G• in the solution was determined by comparing the EPR signal intensity to that of a $CuSO_4$ standard [41]. A typical yield of 0.25-0.30 radicals per NbNrdD polypeptide was reproducibly obtained.

3.2.7. X-band EPR spectroscopy

Continuous wave X-band EPR spectra were recorded at 77 K in the MIT Department of Chemistry Instrumentation Facility on a Bruker ESP-300 X-band spectrometer equipped with a quartz finger Dewar filled with liquid N_2. Experimental conditions were as follows: microwave frequency, 9.45 GHz; modulation amplitude, 0.15 mT; modulation frequency, 100 kHz; time constant, 5.12 ms; scan time, 41.9 s; microwave power, 20 μW.

3.2.8. Preparation of [^{2}H]-Gly-labeled NbNrdD to establish the location of the radical on glycine

The procedure was identical to the preparation of [SeMet]-TmNrdD, except that instead of SeMet, the culture contained L-methionine (50 mg/L) and [^{2}H]-glycine (6 mM, 98% isotopic enrichment, Cambridge Isotope Labratories). The yield was ~1 g of cell paste and purification was carried out according to the procedure described for the unlabeled protein. The EPR sample was then prepared as described above.

3.2.9. Determination if H_α of glycine in NbNrdD exchanges

Tris buffer (20 mM, pH 7.5) was prepared in D_2O (99.9%, Cambridge Isotope Laboratories) in the glovebox. NbNrdD and NbNrdG were exchanged into this buffer by repeated dilution and concentration by ultrafiltration (Amicon YM-30), such that < 1% H_2O remained, and incubated for 12 h at 4°C to allow for proton exchange. The activation reaction was carried out as described above, but with all components made up in D_2O in the glovebox, followed by preparation of the EPR sample as described above.

3.2.10. Activity assay for dCTP formation by NbNrdD

The assay mixture contained in 100 μL: NbNrdD (4 μM, ~1 μM G•), ATP (1 mM), 5-[^{3}H]-CTP (1 mM, 4170 cpm/nmol), *E. coli* TrxA (30 μM), *E. coli* TrxB (1 μM), NADPH (1 mM) in assay buffer (30 mM Tris pH 7.5, 30 mM KCl, 10 mM $MgSO_4$) and was incubated at 30°C. Aliquots (20 μL) were removed at 2 min intervals and quenched with 2% perchloric acid (20 μL). Subsequent to removal of the phosphates using calf intestine alkaline phosphatase (Roche), dCTP formation was analyzed by the method of Steeper and Steuart [42]. One unit of activity is

equivalent to 1 nmol of dCTP/min. The specific activity of NbNrdD is 49 U/mg NrdD protein (~0.24 s^{-1} per G•).

3.2.11. Stoichiometry of NADPH consumption and dCTP production

The assay mixture was divided into 20 μL aliquots containing: NbNrdD (4 μM, ~1 μM G•), dATP (0.1 mM), 5-[^{3}H]-CTP (1 mM, 4170 cpm/nmol), *E. coli* TrxA (5 μM), *E. coli* TrxB (1 μM), NADPH (0, 70, 140 or 210 μM) in assay buffer and was incubated at 30°C for 3 h to allow for complete consumption of NADPH. Workup of the samples was carried out as described above to quantify dCTP formed.

3.2.12. Activity assay for cytosine release by NbNrdD(C300A)

The assay mixture contained in 100 μL: NbNrdD(C300A) (33 μM, ~8 μM G•), 5-[^{3}H]-CTP in assay buffer and was incubated at 30°C. Aliquots (20 μL) were removed at 2, 4, 8 and 16 min and quenched with 2% perchloric acid (20 μL). dCTP formation was analyzed by the method of Steeper and Steuart [42]. Formation of Cyt was analyzed by passing the mixture through an anion exchange column to remove the nucleoside triphosphates as previously described [31].

To a 7 mL portion of the eluate of the Dowex-1-borate column was added carrier Cyt and dC (10 nmol each). The mixture was concentrated by lyophilization, redissolved in water, cooled on ice and the precipitated borate salts removed by centrifugation. The supernatant was analyzed by HPLC using an Alltech Econosil column (C18, 10 μm, 250 × 4.6 mm) on a Waters 515 HPLC system equipped with a 2996 photodiode array detector. The compounds were eluted with KPi (20 mM, pH 6.8) at a flow rate of 1.0 mL/min. Fractions were collected (0.5 mL) and analyzed by scintillation counting. Cyt was identified by coelution with a standard at 5 min (Figure 3.4C).

3.2.13. Crystallization and crystal structure of TmNrdD (crystallography was carried out by Michael A. Funk, MIT)

Crystals of SeMet-TmNrdD were grown aerobically by sitting drop vapor diffusion at 21 °C. Protein, with the His-tag intact, at 13 mg/mL in a buffer containing 25 mM HEPES pH 7.6, 15 mM $MgCl_2$, 20 mM KCl, 0.5 mM TCEP, 1 mM dGTP, and 5 mM ATP was screened against commercial screens (Hampton Research, Microlytic, and Qiagen) at a 1:1 ratio of protein to precipitant. A Phoenix pipetting robot (Art Robbins Instruments) was used for dispensing 150 nL drops in screening trays. Diffraction quality crystals grew over several weeks in wells containing 0.085 M tri-sodium citrate pH 5.6, 0.17 M ammonium acetate, 25.5% PEG 4000, and 15% glycerol. Crystals were flash frozen in liquid nitrogen without additional cryoprotection.

The structure of TmNrdD was solved by single-wavelength anomalous dispersion. A 1.64 Å resolution Se peak anomalous dataset (12664.1 eV) was collected at the Advanced Photon Source using a mini-kappa goniometer to collect Friedel pairs on a single image. Data were collected on a Pilatus 6M detector (Dectris). The data were indexed and scaled with HKL2000 [43] in space group $P2_1$ with unit cell constants a = 78.3, b = 98.8, c = 86.6, β = 111.7. The resulting unit cell volume of 622000 $Å^3$ is consistent with two molecules per asymmetric unit with a solvent content of ~40%. Forty-five initial Se sites were found with SHELXD in the package HKL2MAP [44] with a resolution cutoff of 2.1 Å (d''/sig = 0.8). The resulting sites were refined using data to 1.8 Å resolution and density modification was performed in Phenix AutoSol.[45] The resulting maps at 1.8 Å resolution were adequate for chain tracing and manual building of the entire protein model and surrounding water molecules in Coot [46]. The resolution was increased to 1.64 Å, and iterative refinement of the model was performed with phenix.refine. Composite omit maps were used to verify residue and ligand placement. All native residues are present at the N

and C termini of chain A; chain B is missing three C-terminal residues. Both chains have additional density at the N terminus corresponding to eight residues of the thrombin cleavage site and linker; no residues of the His-tag are observed although the His-tag was not cleaved prior to crystallization. Residues 54-62 of chain A are unstructured. Residues 330-349 of chain A and 328-350 of chain B are not observed in the density at all as a consequence of peptide bond cleavage near residue 330 (Figure 3.7). The final model contained 98.3% of residues in the favored region of the Ramachandran plot, with 1.6% in additionally allowed regions, and 0.08% (1 residue, G621) disallowed. G621 is converted in to the G• and thus is expected to adopt a slightly strained conformation in the unactivated protein. Figures were created in PyMOL Version 1.4.1 (Schrödinger, LLC). A structural model for TmNrdD Cys loop was constructed by superimposing the structures of T4 bacteriophage (1H79) [47] and *T. maritima* class III RNR using residues in the β-barrel surrounding the active site loops. Since there are no structures of a nucleoside triphosphate-bound RNR, CTP was modeled in the active site based on the position of CDP in the class II *T. maritima* NrdJ (1XNJ) [48]. The resulting hybrid model should be considered purely as a tool for guiding discussions of the possible chemistry occurring in TmNrdD and was not subjected to energy minimizations or dynamics.

3.2.14. Phylogenetic analysis of NrdDs (phylogenetic tree construction was carried out by Dr. Leonardo A. Rosado, MIT)

To determine the existence of different NrdD subtypes, 59 amino acid sequences were chosen to include phylogenetically and metabolically diverse bacteria and archaea and aligned with Muscle (MEGA5 software) [49]. Subsequently, a neighbor-joining phylogenetic tree [50] was generated, and the JTT matrix method [51] was employed to determine the evolutionary distances.

The unequal rate of variation among amino-acid sites was modeled with a γ distribution with shape parameter of 4 [52]. Additionally, a complete deletion of the sites containing gaps was employed in order to palliate the length and divergence variation, focusing on more conserved amino acids clusters. The level of confidence for the branches was determined based on 2,000 bootstrap replicates [53]. The resulting consensus tree was rendered using the web-based program iTOL [54] (Figure 3.9B).

3.3. Results

Our bioinformatics analysis, leading to the identification of a potential new class III RNR subtype that couples nucleotide reduction to disulfide bond formation, is described below. To test our hypothesis, we attempted to clone, express and purify six of the class III RNRs of this subtype, including the enzymes from *N. bacilliformis*, *T. maritima*, *Shewanella sediminis*, *Pyrococcus furiosus*, *Pseudomonas aeruginosa*, and *Schizosaccharomyces japonicas*. The NbNrdD and NbNrdG were soluble and could be obtained in reasonable amounts and thus became the focus of our attention.

3.3.1. NbNrdD is a G• enzyme

To generate active NbNrdD for biochemical studies, we incubated NbNrdD with NbNrdG and SAM in the presence of the acriflavin / bicine photoreduction system [31], resulting in the generation of a radical with an EPR signal shown in Figure 3.2A. The spectrum reveals a dominant hyperfine coupling constant of 40 MHz, proposed to be associated with the H_α of glycine, consistent with its assignment as G•. Uniform labeling of NrdD with [^{2}H]-glycine resulted in the collapse of the signal into a singlet (Figure 3.2C) establishing the assignment.

Additional spectral features (indicated with arrows in Figure 3.2A), which persist even when the activation reaction is carried out in D_2O (Figure 3.2B), are also visible for the PFL G•, though less well resolved [55]. These features are attributed to hyperfine coupling with additional non-exchangeable protons, likely the α-protons of the two adjacent amino acids in the sequence [55]. These interactions are predicted to be affected by the conformation of the peptide backbone, leading to variations in the G• EPR spectra between different G• enzymes. The lack of exchange of the H_α of glycine with D_2O, previously shown to occur with PFL [56], is similar to observations for the G• of *E. coli* class III RNR [57].

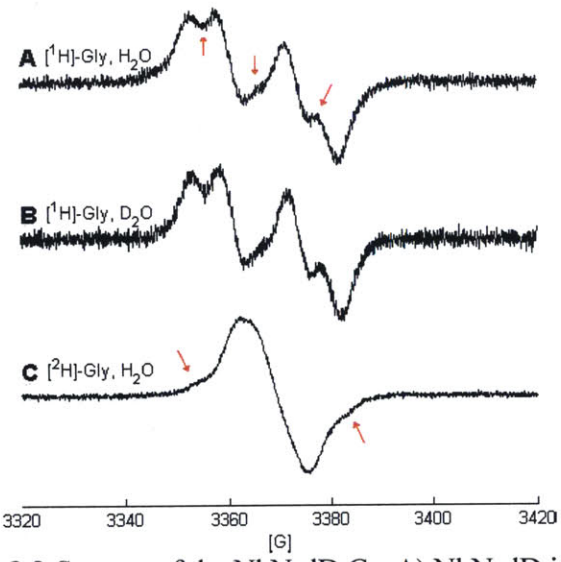


Figure 3.2 Spectra of the NbNrdD G•. A) NbNrdD in H_2O, arrows indicate features arising from hyperfine coupling to non-exchangeable protons; B) NbNrdD in D_2O; and C) [^{2}H]-Gly-NbNrdD in H_2O, arrows indicate features possibly due to contaminating unlabeled NbNrdD.

3.3.2. NbNrdD catalyzes CTP reduction using TrxA / TrxB / NADPH

Our hypothesis is that reducing equivalents for nucleotide reduction by NbNrdD are delivered by a redoxin, similar to the class I and II RNRs (Figure 3.1 A and B). To identify a candidate redoxin for NbNrdD, we first carried out a BLAST search using the NbNrdD sequence. This search yielded a set of related sequences with ~50% pairwise identity in diverse organisms, including *Shewanella sediminis* (gammaproteobacterium), *Bacteroides ovatus* (CFB group) and *Clostridium citroniae* (Firmicutes). Examination of the redoxins present in these organisms revealed that only TrxA / TrxB is conserved. The high sequence identity between *N. bacilliformis* TrxA and *E. coli* TrxA (61% identity), which has been used in assays for other class I and II RNRs [58-60], suggested that *E. coli* TrxA could be used in our activity assays.

The assays were thus carried out with NbNrdD (0.25 G•/α) and *E. coli* TrxA / TrxB / NADPH, and the results are summarized in Table 3.2. NbNrdD was active for reduction of CTP to dCTP with ATP as an effector, but nearly inactive for CDP reduction (~3% of the activity for CTP reduction, Table 3.2). Catalytic activity was dependent on the presence of G• and TrxA (Table 3.2). Formate failed to produce any dCTP. Unexpectedly the activity was the same in the absence or presence of allosteric effectors (ATP or dATP, Table 3.2). NbNrdD lacks the ATP cone domain that controls the activity of many RNRs by binding the activator (ATP) or the inactivator (dATP) [61]. Thus in NbNrdD, both ATP and dATP would be predicted to bind to the specificity site and activate nucleotide reduction. The activity that we have obtained with NbNrdD is ~0.24 s^{-1} per G•, which is 20 fold lower than that of *E. coli* NrdD, which is ~4 s^{-1} per G• [62]. We hypothesize that the low activity and insensitivity to allosteric effectors is a result of *E. coli* TrxA functioning as a sub-optimal reductant, making rereduction of the active-site disulfide, rather than nucleotide reduction, rate limiting [63]. Further study of the allosteric regulation of this

enzyme will likely be facilitated by cloning, expressing and utilizing the *N. bacilliformis* Trx system in our assays.

Reaction conditions	Activity (U/mg)
Complete (CTP, ATP)	49
- SAM (no G• formed)	N.D.
-TrxA	N.D.
-TrxA, + formate (10 mM)	N.D.
-ATP	47
-ATP, + dATP (0.1 mM)	51
-CTP, + CDP (1 mM)	~1.5

Table 3.2 Requirements for dCTP formation by NbNrdD. *N.D. = activity not detected, <3 turnovers per G• over 10 min.

The number of dCTPs formed per NADPH in the reaction mixture is 0.97 (Figure 3.3), suggesting a 1:1 stoichiometry, in agreement with the proposal that the reducing equivalents are provided by the TrxA / TrxB / NADPH system (Figure 3.1A,B). The ~26 turnovers per G• that occur without addition of NADPH are attributed to the reduction of TrxA by residual DTT carried over from the NbNrdD storage buffer (~27 μM), which is essential for maintaining enzymatic activity.

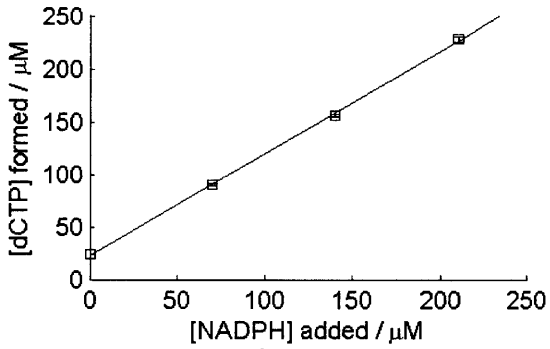


Figure 3.3. Amount of 5-[^{3}H]-dCTP formed after incubation of NbNrdD with 5-[^{3}H]-CTP, dATP, TrxA, TrxB and limiting amounts of NADPH at 30°C for 3 h. Stoichiometry of dCTP produced per NADPH added is 0.97. The concentration of NbNrdD G• in the reaction is ~1 μM, and the ~26 turnovers per G• that occur without addition of NADPH are attributed to the reduction of TrxA by residual DTT carried over from the NbNrdD storage buffer (~27 μM).

3.3.3. NbNrdD(C301A) is inactive and reaction of NbNrdD(C300A) with CTP generates cytosine (Cyt)

To test our hypothesis that C301 forms the top face thiyl radical that initiates nucleotide reduction (Figure 3.1A,B), the NbNrdD(C301A) mutant was generated. This mutant is inactive in dCTP and Cyt formation, despite having 0.25 G•/α2, consistent with our model. We propose that C300 in NbNrdD, which is adjacent to the top face thiyl radical residue C301 on the Cys loop (Figure 3.1B and Figure 3.8A), plays a role analogous to that of C462 in the *E. coli* class Ia α protein (NrdA) (Figure 3.1B), donating reducing equivalents for nucleotide reduction by generating a disulfide with C99 (Figure 3.1B). To test this hypothesis, we generated the NbNrdD(C300A) mutant and reacted it with CTP. The analogous C462A mutation in *E. coli* NrdA results in the generation of a 3′-keto-deoxycytidine intermediate (Figure 3.4B, **16**) that decomposes to release Cyt (**17**) [27].

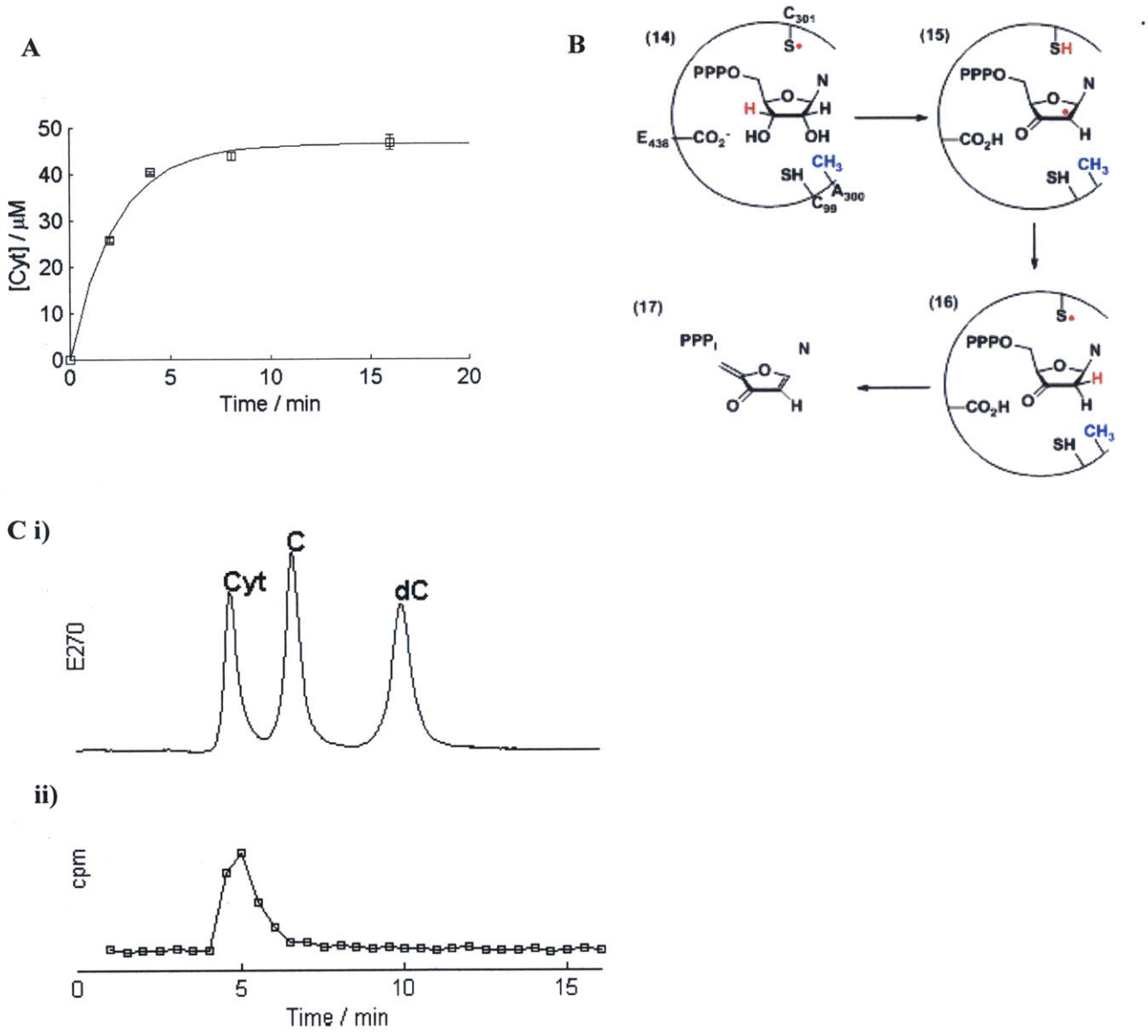


Figure 3.4. A) Time-dependent 5-[^{3}H]-Cyt release in the reaction of NbNrdD(C300A) with 5-[^{3}H]-CTP. The concentration of G• in the reaction is ~8 µM. B) Proposed mechanism for Cyt release by NbNrdD(C300A). HPLC detection of Cyt released in the reaction of NbNrdD(C300A) with 5-[3H]-CTP. C i) Standards, monitored by absorbance at 270 nm. C ii) 5-[3H]-Cyt detected by scintillation counting.

The G• of the NbNrdD(C300A) has an EPR spectrum identical to that of the wild-type (WT)-NbNrdD, and is generated with a similar efficiency: Reaction of NbNrdD(C300A) with 5-[^{3}H]-CTP leads to the time-dependent release of ~5.5 eq. of 5-[^{3}H]-Cyt per G• (Figure 3.4A), identified by HPLC (Figure 3.4C), with an initial rate of 9.3 U/mg (~2.5 min^{-1} per G•). No dCTP is detected and the same amount of Cyt is produced in the presence or absence of the Trx system. A control with WT-NbNrdD shows no Cyt release either in the presence or absence of TrxA / TrxB/ NADPH.

3.3.4. Crystal structure of TmNrdD allows modeling of active site residues in redoxin-dependent NrdDs (crystallography was carried out by Michael A. Funk, MIT)

Bioinformatics analysis (described subsequently) suggested that in addition to the three cysteines in the active site, a glutamate will also be present. To determine if these residues are located in the active site of this class of redoxin-dependent NrdDs, we wanted to obtain a structure of a representative of this NrdD subtype. SeMet-labeled NrdD from *T. maritima* (TmNrdD), which is related to NbNrdD (30% sequence identity), was successfully crystallized, and the structure was solved by single-wavelength anomalous dispersion to 1.64 Å resolution. We observe a $(\beta/\alpha)_{10}$ barrel architecture similar to the T4 bacteriophage class III RNR [29] (RMSD 2.5 Å), including the C-terminal G• domain and Zn-binding site (Figure 3.5A) but with four additional helices at the N terminus (Figure 3.6A and Figure 3.5B). The Cys loop containing essential cysteines C329 and C330 (equivalent to C300 and C301 in NbNrdD), however, is not present in its expected conformation within the barrel. In fact, residues 330-349 are not visible in the electron density map in either monomer, and SDS-PAGE analysis of the crystals indicates that protein cleavage occurs within this stretch of amino acids (Figure 3.7). The molecular

weights of the fragment bands (39 and 36 kDa) suggest that the missing density may be due to disorder around a single cleavage site rather than a missing stretch of residues. Regardless, cleavage within residues 330-349 has a dramatic affect; adjacent residues 320-329 and 350-365 move, and residues 350-365, which show no sequence or structural similarity to the Cys loop, now occupies the Cys loop position (Figure 3.6B). Despite the absence of the Cys loop in our structure, the barrel architecture is intact and the G• domain is ordered. The C terminus beyond the G• loop is also ordered but adopts two distinct conformations in the two different molecules in the asymmetric unit (Figure 3.6C).

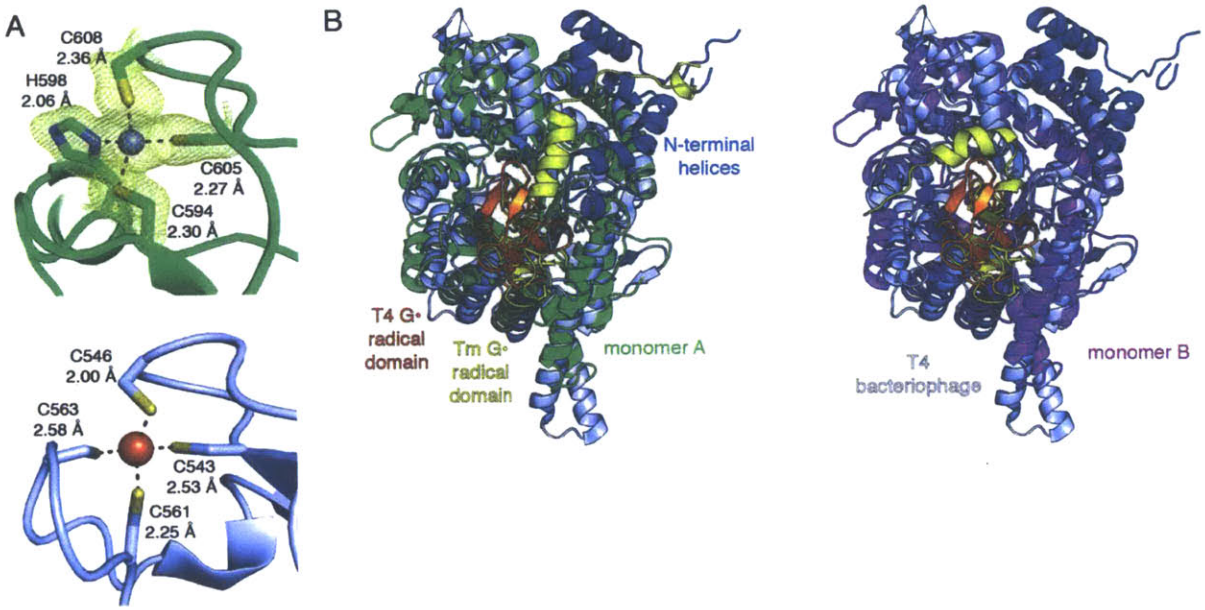


Figure 3.5. Comparison of *T. maritima* and T4 bacteriophage class III RNRs. A) The Zn-binding site of TmNrdD (green, monomer A) is similar to the T4 bacteriophage NrdD (light blue) (PDB code 1H79) [47] with the exception that H598 replaces C563. Composite omit maps around the Zn and ligating residues are contoured at 1σ (yellow). B) Comparison of TmNrdD monomer A (green) and TmNrdD monomer B (magenta) to the T4 bacteriophage NrdD (light blue). The N-terminal helices in TmNrdD (dark blue) have no counterpart in the T4 NrdD. The T4 G• domain (orange) and Tm G• domain (yellow) are shifted as seen in Figure 3.6D.

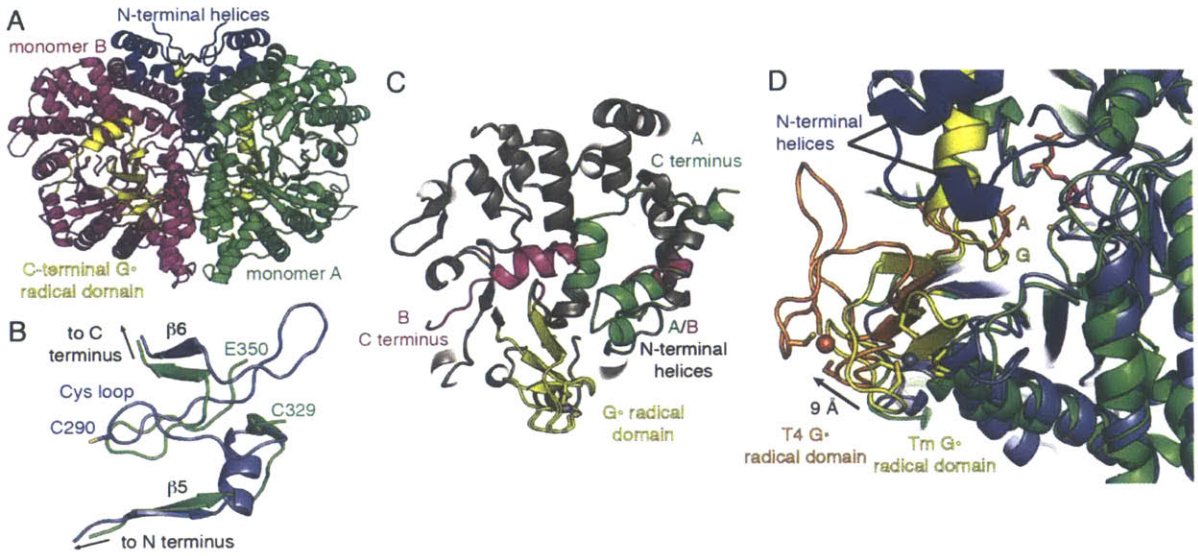


Figure 3.6. Overall structure of *T. maritima* class III RNR. A) The structure of the dimer is shown with monomer A (green) and monomer B (magenta) joined by an interface that contains four N-terminal helices specific to TmNrdD (blue). B) Superposition of residues 265-312 of the T4 bacteriophage structure (PDB code 1H79) [47], which includes the Cys loop (blue) with residues 314-327 and 351-369 of the TmNrdD structure (green). Residues 350-356 of TmNrdD occupy the position left vacant by the absence of the Cys loop (residues 324-334). C) Monomers of TmNrdD differ by more than 1 Å in two regions (green, monomer A; magenta, monomer B): the tip of the N-terminal helices that contact the G• domain and the C-terminal helix. D) G• domains of TmNrdD (yellow, Gly shown in sticks) in both chains are positioned further inside the barrel relative to that of T4 bacteriophage (orange, G580A in sticks). The C-terminal helix of TmNrdD is shown in its monomer A orientation (yellow). Zinc (TmNrdD, gray) and iron (T4 phage, orange) atoms, from Tm and T4 structures, respectively, are shown as spheres. CDP (magenta) is modeled based on the *T. maritima* class II RNR structure (PDB code 1XJN) [48].

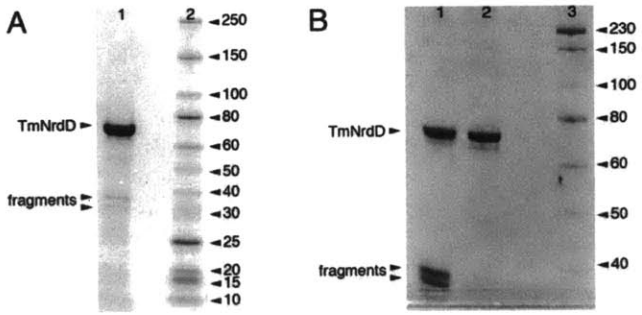


Figure 3.7. TmNrdD fragments form during purification and are enriched by crystallization (SDS-PAGE, 10% gels). A) Purified TmNrdD contains intact enzyme (75 kDa) and a small amounts of two fragments (~36 and ~39 kDa). Molecular weight markers in lane 2. B) Unwashed crystals of TmNrdD (lane 1) show enrichment of the fragment bands, whereas the surrounding protein-precipitant solution (lane 2) is depleted in the fragmented protein. This differential suggests selective crystallization of the two fragments over intact enzyme. Molecular weight markers in lane 3.

Using the TmNrdD structure and T4 bacteriophage NrdD structure (PDB code 1H79) [47], we constructed a hybrid structural model (Figure 3.8A). The Cys loop was modeled from T4 NrdD (T4 residues 287-293; equivalent to TmNrdD 327-333). Since there are no structures of a nucleoside triphosphate-bound RNR, CTP was modeled in the active site based on the position of CDP in the *T. maritima* class II RNR (NrdJ) (PDB code 1XJN) [48]. Structural superpositions based on aligning residues in the β strands surrounding the active site yield all-atom RMSD values between TmNrdD and T4 NrdD of 0.9 Å and between TmNrdD and *T. maritima* NrdJ of 2.0 Å.

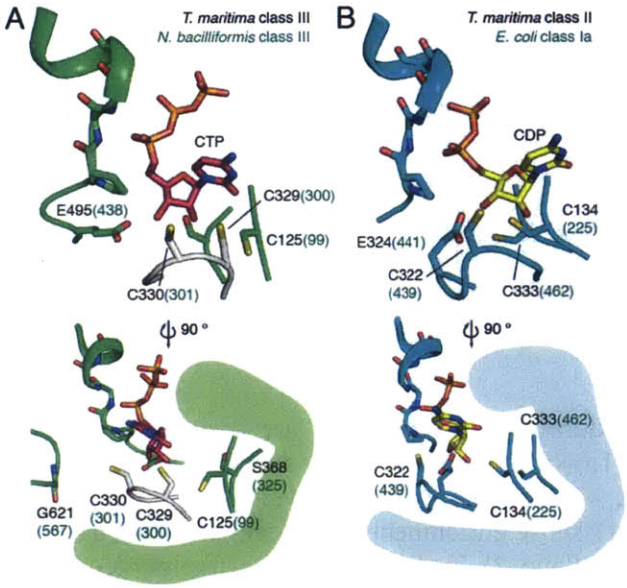


Figure 3.8. Model for disulfide formation in class III RNR. A) The crystal structure of TmNrdD is shown in green (the inserted loop, residues 350-365, is not shown for clarity). The Cys loop based on bacteriophage T4 NrdD (PDB code 1H79) [47] has been modeled in white. CTP has been modeled based on the *T maritima* class II RNR structure (magenta). The side view shows the position of the G• loop. B) The structure of CDP (yellow) bound *T. maritima* class II RNR (PDB code 1XJN) [48].

In the TmNrdD model, C330 is positioned at the tip of the Cys loop to initiate catalysis (Figure 3.8A). The modeled C329 is directly adjacent to C125 (Cβ-Cβ distance: 3.8 Å), sufficiently close to form the proposed disulfide (see Figure 3.1B), as occurs in the *E. coli* class Ia (Cβ-Cβ distance: 4.1 Å) and *T. maritima* class II enzymes (Cβ-Cβ distance: 4.2 Å) (Figure 3.8B). The precise orientation of the Cys loop and the placement of C329 in relation to the ribose of the CTP cannot be determined from this model, but the general locations of C329 and C125 are consistent with their proposed mechanistic role. In the class Ia RNR from *E. coli*, reduction of the ketyl radical intermediate (Figure 3.1A, **3**) is proposed to occur by hydrogen atom transfer from C225 from the bottom face of the nucleotide, followed by eventual disulfide formation with C462 (Figure 3.1B), which is deeply buried on the inner-most side of the active site. TmNrdD

C125 is observed in the crystal structure to be positioned equivalently to *E. coli* class Ia C225, consistent with it playing the same catalytic role. However, TmNrdD does not have a cysteine equivalent to *E. coli* class Ia C462 in the rear of the active site; S368 occupies this space in the structure. Instead, the location of C329 on the Cys loop requires formation of the disulfide at the front of the active site, in an orientation distinct from the class I/II enzymes, possibly in a position to facilitate its rereduction by redoxins (discussed below).

The reduction of the 3′-keto-deoxynucleotide (Figure 3.1B, **4**) requires a proton in addition to the electron from the disulfide radical anion. We proposed from sequence alignments that residue E495 could function as the proton source, taking the place of E441 in *E. coli* class Ia RNR. This glutamate is not conserved in the *E. coli*-type NrdD, but is conserved in the *T. maritima*-type NrdD (see Figure 3.9 and the discussion below). Here we find that the location of E495 in the TmNrdD structure overlaps quite well with the position of E441 (Figure 3.8A). Unfortunately, the mutant of the corresponding residue in NbNrdD, E438Q, displayed low G• content, and the enzyme was inactive in forming either dCTP or Cyt, which prevented us from carrying out experiments analogous to the *E. coli* NrdA(E441Q) mutant [28,64].

A major puzzle with respect to this newly discovered class III RNR subtype is how the active site disulfide, which must be formed on each round of catalysis, is rereduced by a redoxin. The class I and II RNRs require five cysteine residues for catalysis: three in the active site (Figure 3.1B) and two located in the C-terminal tail (Figure 3.1D) that are involved in rereduction of the disulfide generated during dNDP (dNTP) formation so that multiple turnovers can occur. All class III RNRs lack a C-terminal tail containing a pair of cysteines [27] that could function in this capacity. The TmNrdD structural model shows that C329 and C125 are found at the outer edge of the active site, in contrast to the deeply buried cysteine pair found in the class

I/II enzymes. Nonetheless, a large conformational change would seem be required to allow the direct access of TrxA to the oxidized cysteine pair. The active site in our structure is buried primarily by the presence of the G• domain and two N-terminal helices specific to the TmNrdD-type enzymes. Although the active site appears buried, the $(\beta/\alpha)_{10}$ barrel architecture conserved in all classes of RNRs and all G• enzymes is known to undergo conformational changes that can either close or expose the active site. In the class II RNR from *Lactobacillus leichmannii*, coenzyme B_{12} binding causes a shift in a small number of residues exterior to the barrel, which closes the barrel and shields the cofactor from solvent [65]. In PFL, the G• domain must exit the active site to become accessible to its partner activating enzyme for radical generation, and CD spectroscopy shows that binding of PFL to its activase is accompanied by enhanced enzyme conformational flexibility [66,67]. The recent structure of the G• enzyme benzylsuccinate synthase [68] reveals a clamshell-like opening of the barrel, allowing release of the G• domain from the interior of the protein, which would again allow for G• formation. In this work, support for the proposal that the G• domain is flexible comes from the observation that the domain position in TmNrdD is shifted by 9 Å (~15° rotation) relative to that of T4 bacteriophage (Figure 3.6D). Although the relevance, if any, of this shift is not established at this time, it illustrates the wide range of motion available to the G• domain in class III RNRs. Also notable for understanding active site access in TmNrdD is the observation that the G• domain C-terminal helix has two distinct conformations (Figure 3.6C): in monomer A, the helix is extended and completely blocks active site access; in monomer B, the helix kinks and wraps around the protein, similar to other G• enzymes, revealing an opening to the active site. Although this active site cavity is not large enough to accommodate thioredoxin, the flexibility observed in the C- and N-terminal

helices as well as in the G• domain indicates the likelihood that further opening of the active site that would permit rereduction by thioredoxin.

3.3.5. Bioinformatics analysis suggests three chemically distinct NrdD subtypes

(phylogenetic tree construction was carried out by Dr. Leonardo A. Rosado, MIT)

Further support for the existence of distinct subtypes of NrdDs was obtained by a bioinformatics analysis. A phylogenetic tree of NrdDs (Figure 3.9B) was constructed using 59 sequences from the RNRdb [33], chosen to include phylogenetically and metabolically diverse bacteria and archaea. Analysis of the conserved residues proposed to be important for chemistry (the three active site cysteines Cys(1-3), Met and Glu in Figure 3.9A) led us to propose that there are three chemically distinct NrdD subtypes which we label: NrdD1, NrdD2 and NrdD3. The top face Cys(1) on the Cys loop, and the bottom face Cys(2) are conserved in all NrdDs, and Cys(3), Met and Glu are variously conserved in the different NrdD subtypes. The phylogeny of these NrdDs suggests that horizontal gene transfer plays a large role in accounting for their distribution among the diverse bacteria and archaea, as postulated for all classes of RNR [18].

NrdD1 includes the previously studied bacteriophage T4, *E. coli* and *L. lactis* enzymes that use formate as a reductant [4]. Met, thought to play a role in the reaction with formate (Figure 3.1A,C) [31], is conserved in all NrdD1s. None of the NrdD1 sequences examined have the Cys(3) and the Glu. Nrd1s are present almost exclusively in fermentative bacteria which have a PFL, and in hydrogenotrophic methanogens, where the F420-dependent formate dehydrogenase provides a pathway for formate generation [69]. In certain fermentative bacteria, like *E. coli* and *L. lactis*, PFL plays a role in energy metabolism by converting pyruvate to acetyl-CoA, which leads

to ATP production *via* acetyl phosphate, and formate, which is a waste product [16]. In some bacteria, PFL also plays a role in providing C1 intermediates for biosynthesis [70,71].

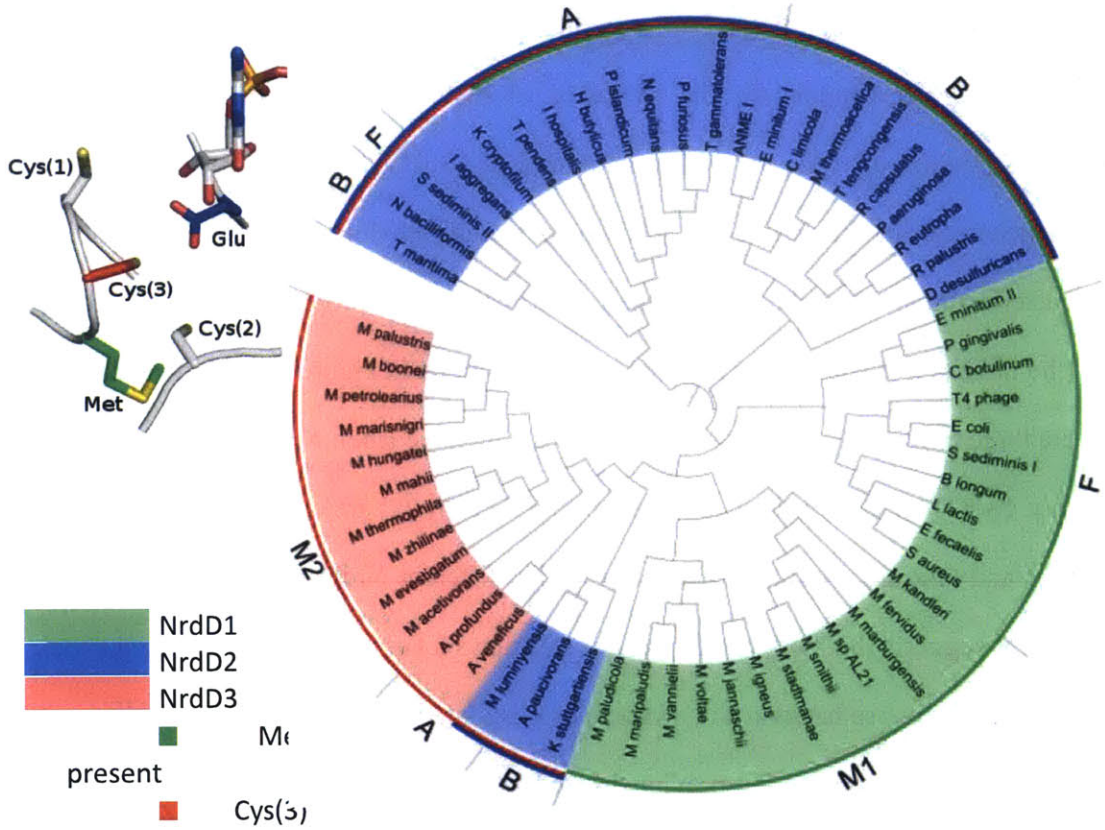


Figure 3.9. A) Model of a generic NrdD active site showing residues proposed to be important for chemistry. Cys(1) and Cys(2) are conserved in all NrdDs, and Met, Cys(3) and Glu are variously conserved in the different NrdD subtypes. B) Phylogenetic tree of NrdDs. NrdD subtypes are indicated by highlighting (green=NrdD1, blue=NrdD2, red=NrdD3). Presence of Met, Cys(3) and Glu are indicated by colored bars (green, red and blue respectively) along the circumference of the plot. F=bacteria with PFL, B=other bacteria, M1=Type I methanogens that use the cytosolic electron-bifurcating heterodisulfide reductase, M2=other methanogens, A=other archaea.

NrdD2 includes the enzymes from *N. bacilliformis* and *T. maritima*, which are proposed to obtain reducing equivalents from a redoxin. In these proteins, three Cys(1-3) and a Glu, which enable nucleotide reduction *via* disulfide chemistry (Figure 3.1A,B) are conserved. In these enzymes, Met is often present but is not conserved. In NrdD2s lacking Met, this position is most commonly occupied by Ser, as in NbNrdD and TmNrdD, but also Ala, Glu or Phe in some deeply rooted bacteria and archaea. The residue corresponding to the catalytic E441 in *E. coli* NrdA (Figure 3.1A,B), which is expected to be close to the formate binding site, is replaced with either Ser or Thr in NrdD1s. However in NrdD2s, it is replaced with Ile, Leu or Phe, which may hinder the access of formate. In contrast to the restricted distribution of NrdD1, NrdD2 is present in non-methanogenic archaea and in bacteria with diverse types of anaerobic metabolism. In support of the nature of the proposed reductant, several organisms have NrdD2s with associated redoxins, including *Acetobacterium woodii* [72], in the form a C-terminal fusion to NrdD2, and *Desulfarculus baarsii* [73] and *Methanomassiliicoccus luminyensis* [74], which have a thioredoxin-like protein in the operon.

NrdD3, present only in certain methanogens and the closely related *Archaeoglobi* [75], show important differences with respect to NrdD1 and NrdD2. Although they share up to 50% pairwise sequence identity with NrdD1s from other methanogens, they lack Met. Also, while they contain all three Cys(1-3) like the NrdD2s, they lack Glu. Additional observations support designation of a third NrdD subtype and that these proteins will possess RNR activity. NrdD3s all contain the G• consensus sequence, and some of them (*e.g. Archaeoglobus veneficus*) contain an N-terminal ATP cone domain. In some sequenced organisms, NrdD3 is the only annotated RNR (e.g. *Archaeoglobus profundus* , *Methanospirillum hungatei*), although other organisms

also contain a NrdJ. Interestingly, some NrdD3s (*Archaeoglobus veneficus, Methanoscarcina barkerii*) are found in the same operon as a thioredoxin like protein.

The designation of the three NrdD subtypes largely correspond to the phylogeny of the protein (Figure 3.9B), with the exception of a deeply-rooted branch containing *M. luminyensis*, *Aminomonas paucivorans* and *Kuenenia stuttgartiensis* NrdD2. An additional observation suggesting the existence of NrdDs with distinct types of chemistry is that although it is uncommon for a single organism to contain two copies of the same NrdD subtype, there many organisms that contain both NrdD1 and NrdD2 (*e.g. Shewanella sediminis, Bacteroides ovatus, Elusimicrobium minitum*). We hypothesize that these class III RNR variants are used under different growth conditions, similar to the case of the class Ia and Ib RNRs in *E. coli*.

3.4. Discussion

Because of their essential role in the *de novo* production of deoxynucleotides, RNRs are ubiquitous enzymes in nearly all cellular organisms and many viruses [18,33,76]. The complex chemistry involved in nucleotide reduction requires initial generation of the transient thiyl radical. The class of RNR used by an organism reflects the mechanism of this radical formation and a combination of factors including the presence or absence of oxygen [77] and availability of metals [78]. Here we report a new subtype of class III RNR in *N. bacilliformis* where, like the class I and II RNRs, nucleotide reduction is facilitated by a redoxin, which are ubiquitous proteins found in all organisms. This proposed reliance on a redoxin is unlike the class III RNRs studied to date that couple nucleotide reduction to the oxidation of formate, a metabolite produced by some but not all organisms as part of their primary metabolism.

3.4.1. Biochemistry structure, and bioinformatics support the existence of a second NrdD subtype

The cloning of *N. bacilliformis* class III RNR was motivated by a bioinformatics analysis, which led to the identification of a potential second NrdD subtype containing three cysteines necessary to carry out nucleotide reduction *via* a mechanism analogous to that of the class I and II RNRs (Figure 3.1A,B). Our biochemical investigations showed that NbNrdD catalyzes the reduction of NTPs, which are also the substrates of formate-dependent class III RNRs. DeoxyNTP formation required the presence of the TrxA / TrxB / NADPH system but not formate, and the 1 : 1 stoichiometry of NADPH consumption and dCTP production demonstrates that the reducing equivalents for dCTP generation are provided by the Trx system.

As predicted, the NbNrdD(C301A) mutant is inactive as no initiating thiyl radical can be produced. In support of the active-site disulfide between C99 and C300 (Figure 3.1B), our structural modeling using the related TmNrdD enzyme as the molecular scaffold is consistent with these cysteines being close enough to form a disulfide bond. Furthermore, the NbNrdD(C300A) mutant behaves like the *E. coli* class Ia RNR C462A mutant in that reaction with substrate (CTP in the case of NbNrdD) results in the formation of Cyt (Figure 3.4B), showing that the mutant is competent for formation of an 3′-keto-dCTP intermediate, but unable to carry out nucleotide reduction. Thus our assays establish the presence of a second NrdD subtype that is distinct from *E. coli* NrdD in being able to use thioredoxin rather than formate as a reductant for nucleotide reduction.

3.4.2. Distribution of NrdD subtypes correlates with metabolism

Our bioinformatics study revealed that there are in fact three NrdD subtypes (NrdD1, NrdD2 and NrdD3), the distribution of which shows a striking correlation with the organism's anaerobic metabolism. Among bacteria, NrdD1 is localized almost exclusively in fermentative bacteria that utilize PFL. An interesting exception is the deep branching *Elusimicrobium minitum* (termite group 1, Figure 3.9B) [79], which has two NrdD subtypes (NrdD1 and NrdD2) but no PFL: its NrdD1 operon contains a formyl-THF synthetase as a possible source of formate. In contrast, NrdD2 is present in bacteria carrying out a diverse range of anaerobic metabolism.

Although it is tempting to explain the distribution of NrdD subtypes by the availability of intracellular formate, a counterexample is provided by acetogens: although formate is an intermediate of acetogenesis [80], acetogens invariably contain NrdD2. Instead we hypothesize that the energy metabolism of the organism and the redox window that it inhabits determines whether it is more thermodynamically efficient to use formate or other reductants for nucleotide reduction. In fermentative bacteria, formate is produced by PFL as a waste product, thus NrdD1 may be preferred. In contrast, in bacteria that carry out anaerobic respiration, formate is not known to be produced as part of their energy metabolism, and oxidation of any available formate from the environment can be coupled to the generation of ATP, which rationalizes why NrdD2 may be preferred. In acetogens, coupling of ribonucleotide reduction by NrdD2 to the TrxA / TrxB system could provide oxidized pyridine nucleotides that are substrates for energy conservation [72].

All archaea examined contain exclusively NrdD2, except for methanogens and their relatives, which contain either NrdD1 or NrdD3. In the case of the methanogens, the NrdD subtype used correlates with its mechanism for energy conservation [81], which is in turn related to its formate metabolism. NrdD1 is present only in Type I hydrogenotrophic methanogens that

carry out energy conservation by means of the cytosolic electron-bifurcating heterodisulfide reductase [81]. The Type I methanogens examined contain the F420-dependent formate dehydrogenase [69] as a possible means of synthesizing formate, and many can carry out methanogenesis with formate as a substrate. Apart from ribonucleotide reduction, the other known role of formate in the primary metabolism of Type I methanogens is in purine biosynthesis, in an ATP-dependent reaction with formate to form formylphosphate catalyzed by PurT (glycineamide ribonucleotide synthetase) [82,83], suggesting the presence of intracellular formate. A case in point is *Methanocella paludicola* (Rice Cluster I) [84], a methanogen with cytochromes that nevertheless employs the Type I pathway for methanogenesis [81], and contains NrdD1.

Other methanogens, including methylotrophic and acetoclastic methanogens, that carry out energy conservation by other means, and their relatives the sulfate-reducing *Archaeoglobi*, use NrdD3 instead. Unlike type I methanogens, these organisms are unable to use formate as a substrate for methanogenesis and several sequenced members (e.g. *Methanosarcina acetivorans, Methanosarcina mazei*) do not contain formate dehydrogenase [85]. They carry out purine biosynthesis in a manner dependent not on formate but on ^{10}N-formyl-tetrahydrofolate, catalyzed by PurN (glycineamide ribonucleotide synthase) [82], suggesting that formate may not be present in the cell. Although many NrdD3 operons contain a redoxin, methanogens lack a conserved thioredoxin reductase, and the origin of the reducing equivalents for NrdD3 is unclear, but could be linked to the heterodisulfide reductase system common to these organisms.

Several interesting methanogens and relatives also possess NrdD2. The obligate methylotroph *Methanococcoides burtonii* [86], has both NrdD2 and NrdD3, suggesting that the source of reducing equivalents for these two RNRs may be different. The methanogen relative

ANME-1 [87,88], which carries out anaerobic methane oxidation, has NrdD2 but no NrdD3, in conjunction with its unique metabolism. *Methanomassiliicoccus luminyensis* [74], which carries out methanogenesis from methanol but is a relative of the non-methanogenic *Thermoplasmatales*, contains a deeply-rooted NrdD2 with an associated redoxin in the operon. Because of the central role of formate and thiols in many anaerobic processes, the distribution of class III RNRs among different organisms may shed light on aspects of anaerobic biochemistry, particularly in organisms that are unculturable.

3.4.3. A clue regarding the ancestral NrdD

RNR has been proposed to provide the link between the RNA and DNA worlds [89,90], with the class III RNR proposed to precede the class I and II enzymes. Given the central role of formate and thiols in many anaerobic processes, the identity of the original NrdD has implications regarding the metabolism of the first organism with a DNA genome. Unfortunately, the NrdD1 and NrdD2 sequences are too divergent to allow us to convincingly root the NrdD phylogenetic tree. However, we note that the residue Met (Figure 3.9A) proposed to be involved in formate chemistry in NrdD1 [31], is also present in many but not all NrdD2s, possibly as an evolutionary relic, suggesting that NrdD1 may precede NrdD2. If true, this suggests that formate is present in the ancestral organism, possibly as an intermediate in an ancient Wood-Ljungdahl pathway [72].

3.5. Acknowledgements

We thank Michael A. Funk who carried out all crystallographic experiments and analysis, and Prof. Catherine L. Drennan for insight and guidance in the structural analysis. We thank Dr. Leonardo A. Rosado who constructed the phylogenetic trees. We thank Jiyeon Baek for help in development of procedures for protein purification. We are grateful to Dr. Xiang-Yang Han for the gift of *N. bacilliformis* genomic DNA, to Prof. Kenneth Noll for the gift of *T. maritima* genomic DNA and to Dr. Greg Fournier, Dr. Silvan Scheller and Prof. Rudolf K. Thauer for helpful discussions. This work is based upon research conducted at the Advanced Photon Source on the Northeastern Collaborative Access Team beamlines.

3.6. References

(1) Fontecave, M.; Eliasson, R.; Reichard, P. *Proc. Natl. Acad. Sci. U. S. A.* **1989**, *86*, 2147.

(2) Sun, X.; Ollagnier, S.; Schmidt, P. P.; Atta, M.; Mulliez, E.; Lepape, L.; Eliasson, R.; Graslund, A.; Fontecave, M.; Reichard, P.; Sjöberg, B. M. *J. Biol. Chem.* **1996**, *271*, 6827.

(3) Nordlund, P.; Reichard, P. *Annu. Rev. Biochem.* **2006**, *75*, 681.

(4) Mulliez, E.; Ollagnier, S.; Fontecave, M.; Eliasson, R.; Reichard, P. *Proc. Natl. Acad. Sci. U. S. A.* **1995**, *92*, 8759.

(5) Hofer, A.; Crona, M.; Logan, D. T.; Sjöberg, B.-M. *Crit. Rev. Biochem. Mol. Biol.* **2012**, *47*, 50.

(6) Stubbe, J. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 2723.

(7) Sofia, H. J.; Chen, G.; Hetzler, B. G.; Reyes-Spindola, J. F.; Miller, N. E. *Nucleic Acids Res.* **2001**, *29*, 1097.

(8) Gambarelli, S.; Luttringer, F.; Padovani, D.; Mulliez, E.; Fontecave, M. *ChemBioChem* **2005**, *6*, 1960.

(9) Holmgren, A. *J. Biol. Chem.* **1979**, *254*, 3672.

(10) Jordan, A.; Åslund, F.; Pontis, E.; Reichard, P.; Holmgren, A. *J. Biol. Chem.* **1997**, *272*, 18044.

(11) Avval, F. Z.; Holmgren, A. *J. Biol. Chem.* **2009**, *284*, 8233.

(12) Young, P.; Ohman, M.; Xu, M. Q.; Shub, D. A.; Sjöberg, B. M. *J. Biol. Chem.* **1994**, *269*, 20229.

(13) Torrents, E.; Buist, G.; Liu, A.; Eliasson, R.; Kok, J.; Gibert, I.; Gräslund, A.; Reichard, P. *J. Biol. Chem.* **2000**, *275*, 2463.

(14) Sawers, G.; Böck, A. *J. Bacteriol.* **1988**, *170*, 5330.

(15) Knappe, J.; Sawers, G. *FEMS Microbiol. Lett.* **1990**, *75*, 383.

(16) Axley, M. J.; Grahame, D. A. *J. Biol. Chem.* **1991**, *266*, 13731.

(17) Torrents, E.; Aloy, P.; Gibert, I.; Rodríguez-Trelles, F. *J. Mol. Evol.* **2002**, *55*, 138.

(18) Lundin, D.; Gribaldo, S.; Torrents, E.; Sjöberg, B.-M.; Poole, A. M. *BMC evolutionary biology* **2010**, *10*, 383.

(19) Kim, B. H.; Gadd, G. M. *Bacterial physiology and metabolism*; Cambridge university press, 2008.

(20) Licht, S.; Stubbe, J. In *Comprehensive Natural Products Chemistry*; Barton, S., Nakanishi, K., Meth-Cohn, O., Poulter, C., Eds.; Elsevier Science: New York, 1999; Vol. 5, p 163.

(21) Eklund, H.; Fontecave, M. *Structure* **1999**, *7*, R257.

(22) Stubbe, J.; Ackles, D. *J. Biol. Chem.* **1980**, *255*, 8027.

(23) Stubbe, J.; Ackles, D.; Segal, R.; Blakley, R. L. *J. Biol. Chem.* **1981**, *256*, 4843.

(24) Stubbe, J.; Ator, M.; Krenitsky, T. *J. Biol. Chem.* **1983**, *258*, 1625.

(25) Licht, S.; Gerfen, G. J.; Stubbe, J. *Science* **1996**, *271*, 477.

(26) Andersson, J.; Bodevin, S.; Westman, M.; Sahlin, M.; Sjöberg, B.-M. *J. Biol. Chem.* **2001**, *276*, 40457.

(27) Mao, S. S.; Holler, T. P.; Yu, G. X.; Bollinger, J. M., Jr.; Booker, S.; Johnston, M. I.; Stubbe, J. *Biochemistry* **1992**, *31*, 9733.

(28) Lawrence, C. C.; Bennati, M.; Obias, H. V.; Bar, G.; Griffin, R. G.; Stubbe, J. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 8979.

(29) Logan, D. T.; Andersson, J.; Sjöberg, B.-M.; Nordlund, P. *Science* **1999**, *283*, 1499.

(30) Andersson, J.; Westman, M.; Sahlin, M.; Sjöberg, B.-M. *J. Biol. Chem.* **2000**, *275*, 19449.

(31) Wei, Y.; Mathies, G.; Yokoyama, K.; Chen, J.; Griffin, R. G.; Stubbe, J. *J. Am. Chem. Soc.* **2014**, *136*, 9001.

(32) Lenz, R.; Giese, B. *J. Am. Chem. Soc.* **1997**, *119*, 2784.

(33) Lundin, D.; Torrents, E.; Poole, A.; Sjöberg, B.-M. *BMC Genomics* **2009**, *10*, 589.

(34) Han, X. Y.; Hong, T.; Falsen, E. *J. Clin. Microbiol.* **2006**, *44*, 474.

(35) Masliah-Planchon, J.; Breton, G.; Jarlier, V.; Simon, A.; Benveniste, O.; Herson, S.; Drieux, L. *J. Clin. Microbiol.* **2009**, *47*, 1973.

(36) Russel, M.; Model, P. *J. Bacteriol.* **1985**, *163*, 238.

(37) Chivers, P. T.; Prehoda, K. E.; Volkman, B. F.; Kim, B.-M.; Markley, J. L.; Raines, R. T. *Biochemistry* **1997**, *36*, 14985.

(38) Gibson, D. G.; Young, L.; Chuang, R.-Y.; Venter, J. C.; Hutchison, C. A.; Smith, H. O. *Nat Meth* **2009**, *6*, 343.

(39) Van Duyne, G. D.; Standaert, R. F.; Karplus, P. A.; Schreiber, S. L.; Clardy, J. *J. Mol. Biol.* **1993**, *229*, 105.

(40) Fish, W. *Methods Enzymol.* **1988**, *158*, 357.

(41) Malmström, B. G.; Reinhammar, B.; Vänngård, T. *Biochim. Biophys. Acta (BBA)-Bioenergetics* **1970**, *205*, 48.

(42) Steeper, J.; Steuart, C. *Anal. Biochem.* **1970**, *34*, 123.

(43) Otwinowski, Z.; Minor, W. In *Methods Enzymol.*; Charles W. Carter, Jr., Ed.; Academic Press: 1997; Vol. Volume 276, p 307.

(44) Pape, T.; Schneider, T. R. *J Appl Crystallogr* **2004**, *37*, 843.

(45) Adams, P. D.; Afonine, P. V.; Bunkoczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L. W.; Kapral, G. J.; Grosse-Kunstleve, R. W.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; Terwilliger, T. C.; Zwart, P. H. *Acta Crystallogr D Biol Crystallogr* **2010**, *66*, 213.

(46) Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. *Acta Crystallogr D Biol Crystallogr* **2010**, *66*, 486.

(47) Larsson, K.-M.; Andersson, J.; Sjöberg, B.-M.; Nordlund, P.; Logan, D. T. *Structure* **2001**, *9*, 739.

(48) Larsson, K.-M.; Jordan, A.; Eliasson, R.; Reichard, P.; Logan, D. T.; Nordlund, P. *Nat. Struct. Mol. Biol.* **2004**, *11*, 1142.

(49) Tamura, K.; Peterson, D.; Peterson, N.; Stecher, G.; Nei, M.; Kumar, S. *Mol. Biol. Evol.* **2011**.

(50) Saitou, N.; Nei, M. *Mol. Biol. Evol.* **1987**, *4*, 406.

(51) Jones, D. T.; Taylor, W. R.; Thornton, J. M. *Comput. appl. biosci.* **1992**, *8*, 275.

(52) Eyre-Walker, A.; Keightley, P. D. *Nat. Rev. Genet.* **2007**, *8*, 610.

(53) Felsenstein, J. *Syst. Biol.* **1978**, *27*, 401.

(54) Letunic, I.; Bork, P. *Bioinformatics* **2007**, *23*, 127.

(55) Wagner, A. F.; Frey, M.; Neugebauer, F. A.; Schäfer, W.; Knappe, J. *Proc. Natl. Acad. Sci. U. S. A.* **1992**, *89*, 996.

(56) Parast, C. V.; Wong, K. K.; Lewisch, S. A.; Kozarich, J. W.; Peisach, J.; Magliozzo, R. S. *Biochemistry* **1995**, *34*, 2393.

(57) Mulliez, E.; Fontecave, M.; Gaillard, J.; Reichard, P. *J. Biol. Chem.* **1993**, *268*, 2296.

(58) Laurent, T. C.; Moore, E. C.; Reichard, P. *J. Biol. Chem.* **1964**, *239*, 3436.

(59) Moore, E. C.; Reichard, P.; Thelander, L. *J. Biol. Chem.* **1964**, *239*, 3445.

(60) Goulian, M.; Beck, W. S. *J. Biol. Chem.* **1966**, *241*, 4233.

(61) Reichard, P. *Arch. Biochem. Biophys.* **2002**, *397*, 149.

(62) Luttringer, F.; Mulliez, E.; Dublet, B.; Lemaire, D.; Fontecave, M. *J. Biol. Inorg. Chem.* **2009**, *14*, 923.

(63) Ge, J.; Yu, G.; Ator, M. A.; Stubbe, J. *Biochemistry* **2003**, *42*, 10071.

(64) Persson, A. L.; Sahlin, M.; Sjöberg, B.-M. *J. Biol. Chem.* **1998**, *273*, 31016.

(65) Sintchak, M. D.; Arjara, G.; Kellogg, B. A.; Stubbe, J.; Drennan, C. L. *Nat Struct Biol* **2002**, *9*, 293.

(66) Vey, J. L.; Yang, J.; Li, M.; Broderick, W. E.; Broderick, J. B.; Drennan, C. L. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 16137.

(67) Peng, Y.; Veneziano, S. E.; Gillispie, G. D.; Broderick, J. B. *J. Biol. Chem.* **2010**, *285*, 27224.

(68) Funk, M. A.; Judd, E. T.; Marsh, E. N. G.; Elliott, S. J.; Drennan, C. L. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 10161.

(69) Wood, G. E.; Haydock, A. K.; Leigh, J. A. *J. Bacteriol.* **2003**, *185*, 2548.

(70) Thauer, R.; Rupprecht, E.; Jungermann, K. *FEBS Lett.* **1970**, *8*, 304.

(71) Thauer, R. K.; Kirchniawy, F. H.; Jungermann, K. A. *Eur. J. Biochem.* **1972**, *27*, 282.

(72) Poehlein, A.; Schmidt, S.; Kaster, A.-K.; Goenrich, M.; Vollmers, J.; Thürmer, A.; Bertsch, J.; Schuchmann, K.; Voigt, B.; Hecker, M. *PLoS One* **2012**, *7*, e33439.

(73) Sun, H.; Spring, S.; Lapidus, A.; Davenport, K.; Del Rio, T. G.; Tice, H.; Nolan, M.; Copeland, A.; Cheng, J. F.; Lucas, S.; Tapia, R.; Goodwin, L.; Pitluck, S.; Ivanova, N.; Pagani, I.; Mavromatis, K.; Ovchinnikova, G.; Pati, A.; Chen, A.; Palaniappan, K.; Hauser, L.; Chang, Y. J.; Jeffries, C. D.; Detter, J. C.; Han, C.; Rohde, M.; Brambilla, E.; Göker, M.; Woyke, T.; Bristow, J.; Eisen, J. A.; Markowitz, V.; Hugenholtz, P.; Kyrpides, N. C.; Klenk, H. P.; Land, M. *Stand. Genomic Sci.* **2010**, *3*, 276.

(74) Gorlas, A.; Robert, C.; Gimenez, G.; Drancourt, M.; Raoult, D. *J. Bacteriol.* **2012**, *194*, 4745.

(75) Klenk, H.-P.; Clayton, R. A.; Tomb, J.-F.; White, O.; Nelson, K. E.; Ketchum, K. A.; Dodson, R. J.; Gwinn, M.; Hickey, E. K.; Peterson, J. D.; Richardson, D. L.; Kerlavage, A. R.; Graham, D. E.; Kyrpides, N. C.; Fleischmann, R. D.; Quackenbush, J.; Lee, N. H.; Sutton, G. G.; Gill, S.; Kirkness, E. F.; Dougherty, B. A.; McKenney, K.; Adams, M. D.; Loftus, B.; Peterson, S.; Reich, C. I.; McNeil, L. K.; Badger, J. H.; Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J. D.; Weidman, J. F.; McDonald, L.; Utterback, T.; Cotton, M. D.; Spriggs, T.; Artiach, P.; Kaine, B. P.; Sykes, S. M.; Sadow, P. W.; D'Andrea, K. P.;

Bowman, C.; Fujii, C.; Garland, S. A.; Mason, T. M.; Olsen, G. J.; Fraser, C. M.; Smith, H. O.; Woese, C. R.; Venter, J. C. *Nature* **1997**, *390*, 364.

(76) Dwivedi, B.; Xue, B.; Lundin, D.; Edwards, R. A.; Breitbart, M. *BMC evolutionary biology* **2013**, *13*, 33.

(77) Poole, A. M.; Logan, D. T.; Sjöberg, B.-M. *J. Mol. Evol.* **2002**, *55*, 180.

(78) Cotruvo Jr, J. A.; Stubbe, J. *Metallomics* **2012**, *4*, 1020.

(79) Herlemann, D. P. R.; Geissinger, O.; Ikeda-Ohtsubo, W.; Kunin, V.; Sun, H.; Lapidus, A.; Hugenholtz, P.; Brune, A. *Appl. Environ. Microbiol.* **2009**, *75*, 2841.

(80) Ragsdale, S. W.; Pierce, E. *Biochim. Biophys. Acta (BBA)-Proteins and Proteomics* **2008**, *1784*, 1873.

(81) Thauer, R. K.; Kaster, A.-K.; Seedorf, H.; Buckel, W.; Hedderich, R. *Nat. Rev. Microbiol.* **2008**, *6*, 579.

(82) White, R. H. *J. Bacteriol.* **1997**, *179*, 3374.

(83) Fricke, W. F.; Seedorf, H.; Henne, A.; Krüer, M.; Liesegang, H.; Hedderich, R.; Gottschalk, G.; Thauer, R. K. *J. Bacteriol.* **2006**, *188*, 642.

(84) Sakai, S.; Takaki, Y.; Shimamura, S.; Sekine, M.; Tajima, T.; Kosugi, H.; Ichikawa, N.; Tasumi, E.; Hiraki, A. T.; Shimizu, A. *PLoS One* **2011**, *6*, e22898.

(85) Maeder, D. L.; Anderson, I.; Brettin, T. S.; Bruce, D. C.; Gilna, P.; Han, C. S.; Lapidus, A.; Metcalf, W. W.; Saunders, E.; Tapia, R.; Sowers, K. R. *J. Bacteriol.* **2006**, *188*, 7922.

(86) Allen, M. A.; Lauro, F. M.; Williams, T. J.; Burg, D.; Siddiqui, K. S.; De Francisci, D.; Chong, K. W. Y.; Pilak, O.; Chew, H. H.; De Maere, M. Z.; Ting, L.; Katrib, M.; Ng, C.; Sowers, K. R.; Galperin, M. Y.; Anderson, I. J.; Ivanova, N.; Dalin, E.; Martinez, M.; Lapidus, A.; Hauser, L.; Land, M.; Thomas, T.; Cavicchioli, R. *ISME J* **2009**, *3*, 1012.

(87) Michaelis, W.; Seifert, R.; Nauhaus, K.; Treude, T.; Thiel, V.; Blumenberg, M.; Knittel, K.; Gieseke, A.; Peterknecht, K.; Pape, T.; Boetius, A.; Amann, R.; Jørgensen, B. B.; Widdel, F.; Peckmann, J.; Pimenov, N. V.; Gulin, M. B. *Science* **2002**, *297*, 1013.

(88) Milucka, J.; Ferdelman, T. G.; Polerecky, L.; Franzke, D.; Wegener, G.; Schmid, M.; Lieberwirth, I.; Wagner, M.; Widdel, F.; Kuypers, M. M. M. *Nature* **2012**, *491*, 541.

(89) Reichard, P. *Science* **1993**, *260*, 1773.

(90) Stubbe, J. *Curr. Opin. Struct. Biol.* **2000**, *10*, 731.

Chapter 4:

Diversity and evolution of ribonucleotide reductases

4. Diversity and evolution of ribonucleotide reductases

4.1. Interesting RNRs from sequences and literature

Apart from detailed studies of model enzymes like the *Escherichia coli* class Ia ribonucleotide reductase (RNR), mechanistic studies have capitalized on unique properties of certain RNR subtypes. For example, studies of the *Lactobacillus leichmannii* class II[1] and *E. coli* class III[2] RNRs allowed the observation of sulfur radicals, not visible in the wild type *E. coli* class Ia RNR prototype, providing evidence for the involvement of sulfur radicals in the first and second half reactions respectively. Studies of the *Chlamydia trachomatis* class Ic RNR, which lacks Y•, have complemented studies of radical propagation in the *E. coli* class Ia RNR.[3] The *Saccharomyces cerevisiae* heterodimeric ββ′ complex has allowed the intra- and inter-monomeric models of radical propagation to be distinguished.[4] The *Methanosarcina barkeri* class III RNR lacks the active-site Glu proposed to deprotonate the substrate 2′-OH in the class I and II RNRs, and further study of this protein will test the catalytic role of this residue.

In the post genomics era, sequence information suggests that the diversity of RNRs is greater than anticipated. This thesis project has involved looking through many RNR sequences, during which a number of RNRs with distinct sequence features have been noted. Several RNRs with interesting biochemical properties have also been described in literature, but not subjected to detailed study. These RNRs may have properties that facilitate structural or biochemical studies, or even serve to question our current models.

In this section, we describe the diversity of RNRs, their occurrence in different organisms and combine multiple sources of information including sequence, phylogeny and metabolic context to discuss the possible characteristics of unusual RNRs and their evolutionary history.

The class I RNRs have been extensively reviewed[5] and thus we will focus instead on the diversity class II and III RNRs.

4.1.1. Class II RNRs

The first class II RNR to be studied in detail is the monomeric enzyme from *L. leichmannii*, which reduces NTPs.[6-8] Subsequently, several dimeric class II RNRs have been cloned and biochemically characterized, and some reduce NTPs and others NDPs (**Figure 4.1**). These and additional sources of variability are described below. Detailed phylogenetic studies have been reported for the class II RNR,[9] which enable us to trace the evolution of certain characteristics of class II RNRs.

Organism	**4-Cys**	**ATP cone**	**Substrate**	**4° structure**	**Ref.**
Lactobacillus leichmannii	×	×	NTP	Monomer	7
Euglena gracilis	×	×	NTP	Monomer	10,11
Pseudomonas aeruginosa	✓	×	NTP	Dimer (split)	12
Alcaligenes eutrophus	✓	×	NTP	Dimer	*
Pyrococcus furiosus	✓	✓	NDP	Dimer	13
Thermoplasma acidophila	×	✓	NDP	Dimer	14
Thermotoga maritima	✓	×	NDP	Dimer	15
Deinococcus radiodurans	✓	×	NDP	Dimer	15
Chloroflexus aurantiacus	✓	×	?	Dimer	15

Figure 4.1 Characteristics of biochemically characterized class II RNRs. *The recombinant *Alcaligenes eutrophus* class II RNR was found to reduce NTPs (unpublished).

4.1.1.1. Monomeric vs. dimeric class II RNRs

The sequences of monomeric and dimeric class II RNRs are highly divergent, suggesting an early evolutionary split between of the two forms.[9,16] The compact dimer interface containing the allosteric specificity site in class II RNRs is formed by two helices from each monomer, aligned antiparallel to one another.[17] The crystal structure of the *L. leichmanii* class II RNR revealed the insertion of a domain mimicking the dimer interface,[18] suggesting that the monomeric RNR evolved from a dimeric ancestor. The monomeric class II RNR is less common than the dimeric form, but is present in diverse organisms, including *Lactobacillus*, cyanobacteria, bacteriophages and Eukaryotes.[9] The adaptive advantage of this change in oligomeric state is not known.

In view of the proposal that the advent of oxygenic photosynthesis drove the evolution of the O_2-tolerant class II RNR from the O_2-sensitive class III RNR,[19] the presence of the deep-branching monomeric class II RNR in cyanobacteria is particularly interesting since these organisms may have been the first to experience intracellular O_2 in the gradual transition towards light-dependent oxidation of H_2O.[20]

4.1.1.2. Reduction of NDPs vs NTPs

While all class III RNRs studied to date reduce NTPs and all class I RNRs reduce NDPs, both NTP- and NDP-reducing class II RNRs have been found. The reported activities (**Figure 4.1**), together with the crystal structures of the *L. leichmannii*[18] and substrate-bound *T. maritima*[17] class II RNRs enabled us to identify a conserved sequence motif close to where the β-phosphate of the NDP or NTP substrate binds (analysis carried out by Michael A. Funk). The motif was distinct in the reported NDP and NTP-reducing class II RNRs (**Figure 4.2**), enabling

the substrate of the class II RNRs to be predicted from their primary sequence (**Figure 4.3**). All monomeric class II RNRs are thought to reduce NTPs, and mapping of the motifs onto the phylogenetic tree of dimeric class II RNRs suggests that the transition between NTP and NDP reduction in dimeric class II RNRs occurred only once.

NTP-reducing class II RNRs are found in diverse mesophilic bacteria. NDP-reducing class II RNRs are found in archaea and several deep branching thermophilic bacterial groups (*Thermus*, *Thermotoga*). The archaeal sequences are interspersed with many bacteria sequences, some from mesophiles and others from thermophiles.

NTP-reducing class III and II RNRs often have promiscuous activity for NDP reduction (e.g. *N. bacilliformis* class III RNR,[21] *A. eutrophus* class III RNR unpublished). The converse is not true for NDP-reducing RNRs, possibly because the larger NTP is not accommodated in the active site. Although the identity of the ancestor of class II RNRs cannot be determined solely by using phylogenetic analysis, we suggest that it is an NTP-reducing enzyme, with subsequent NDP-reduction evolving from promiscuous activity. This is consistent with the greater sequence diversity and thus probable antiquity of the NTP reducing branches of class II RNR. Class I RNRs, which are all NDP reducers, have a greater sequence similarity to NDP-reducing class II RNRs, and have been proposed to have evolved from this subtype.[13,14]

It was suggested that the transition from reduction of NTPs to NDPs facilitates finer control of substrate concentration and allosteric regulation.[16] A possible advantage is that it avoids incorporating dUTP into DNA. We note that there are other examples of NTP-dependent enzymes replaced with NDP-dependent paralogs, for example ADP-dependent kinases in thermophilic archaea, although the advantage of using a diphosphate is unclear.[22,23] Because of

the prevalence of NDP-reducing class II RNRs in thermophiles and archaea, we propose that NDP reduction may be an adaptation to high temperature or to archaeal nucleotide metabolism.

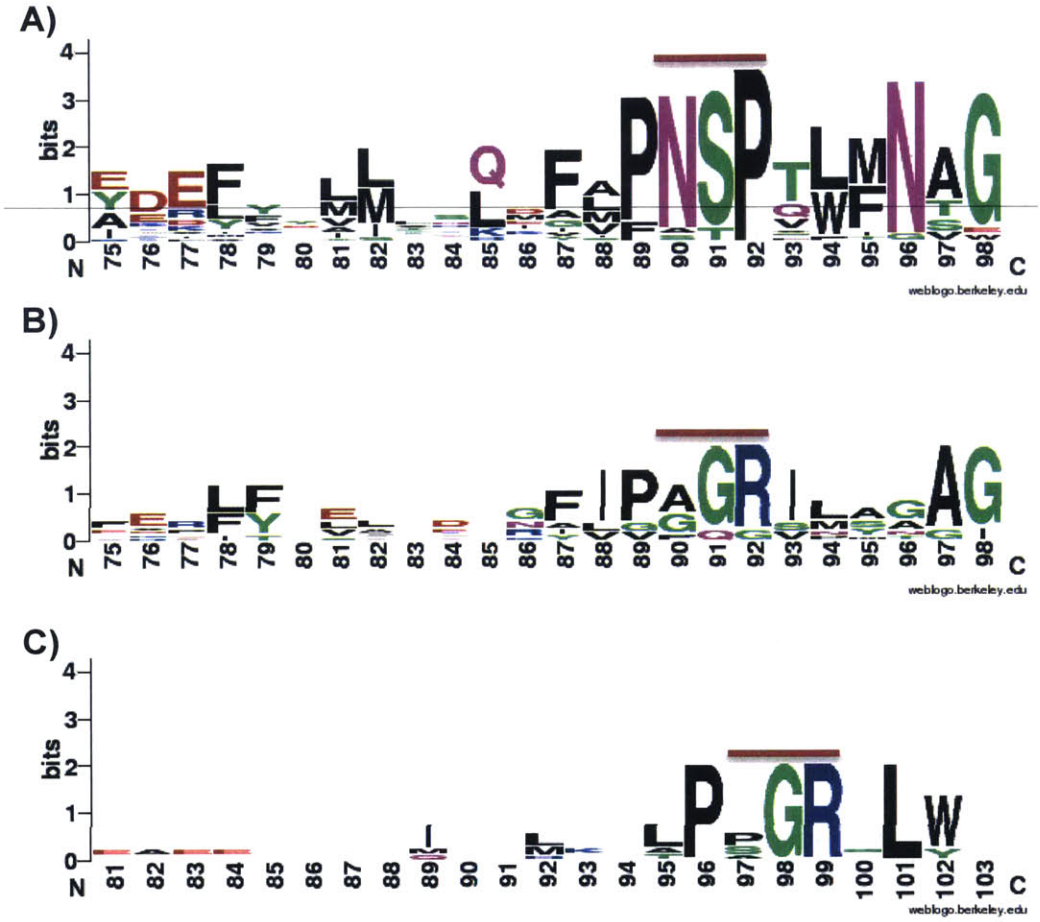


Figure 4.2 Sequence motifs for conserved residues adjacent to the β-phosphate of the NDP or NTP substrate in class II RNRs. The red bar indicates residues adjacent to the β-phosphate of the substrate. A) Dimeric, NDP-reducing class II RNRs. B) Dimeric NTP-reducing class II RNRs. C) Monomeric NTP-reducing class II RNRs. Plotted using WebLogo.[24]

The stability and decomposition pathways for AdoCbl differ depending on temperature, pH and presence of oxygen, and it is unclear whether class II RNRs have any adaptations to stabilize the AdoCbl cofactor. The high K_M for AdoCbl reported for the divergent *L. leichmanii* (1.3 μM)[7] and *P. furiosus* (1 μM)[13] class II RNRs suggest that the cofactor is not bound tightly,

but more investigations are required. In the *L. leichmanii* enzyme, one AdoCbl can service multiple RNRs, proposed to be an adaptation to economize on the biosynthetically costly AdoCbl.[25] In the crystal structure of the *T. maritima* enzyme, contacts were observed between AdoCbl and the substrate, suggesting cofactor release after every turnover.[26]

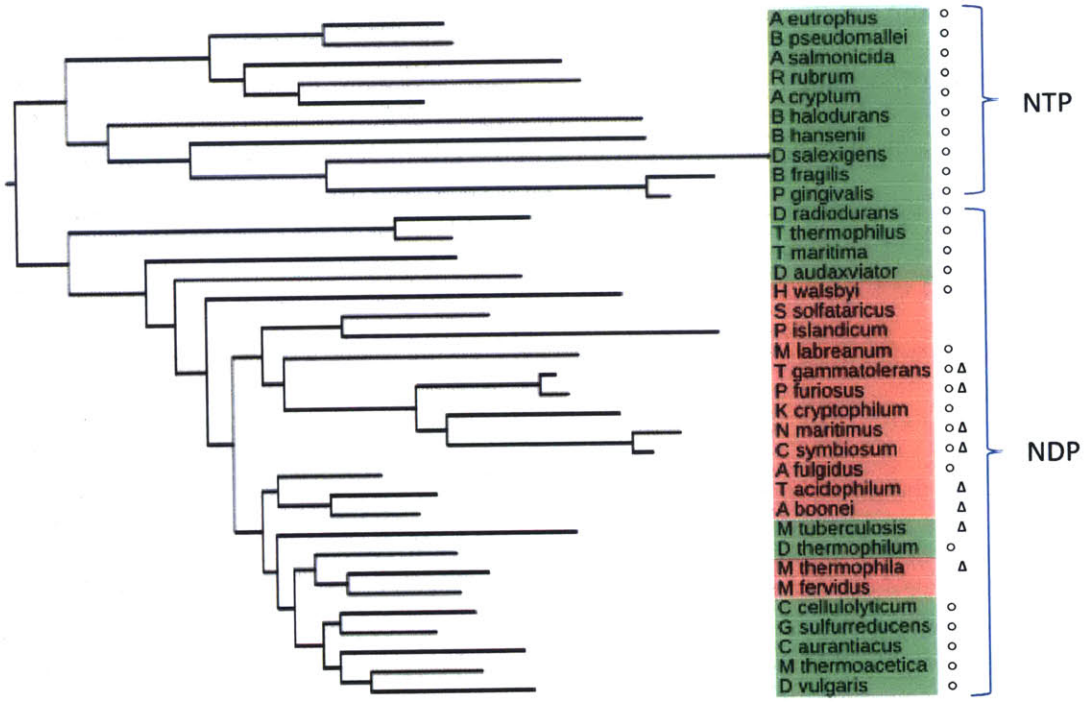


Figure 4.3 Phylogenetic tree of dimeric class II RNRs in bacteria (highlighted green) and archaea (highlighted red), showing the presence of ATP cones (Δ) and C-terminal 4-Cys motif (○). The position of the root was placed to separate inferred NTP- and NDP-reducing class II RNRs, based on the consensus sequences in **Figure 4.2**. Constructed using PhyML[27] and plotted using iTOL[28].

4.1.1.3. A potential metal-binding site and split class II RNRs

It was noted that most dimeric class II RNRs contain a C-terminal 4-Cys motif,[15] upstream of the conserved C-terminal Cys pair involved in the reduction of the active site disulfide in class I and II RNRs (**Figure 4.4**). Its widespread distribution suggests that it is likely

present in the ancestor of dimeric class II RNRs (**Figure 4.3**). Although the function of this motif has not been investigated, possible functional or evolutionary links have been suggested with the Zn-binding site in class III RNR[29] or perhaps with NrdR[30].

The 4-Cys motif is present in *Thermotoga maritima* class II RNR, which has been cloned and studied. Expression in *E. coli* gave a mixture of full length protein and a proteolyzed fragment,[31] where the C-terminal 4-Cys motif is truncated. However, in assays using DTT as a reductant, both forms displayed equal activity, possibly because DTT can directly reduce the active site disulfide. The crystal structure obtained for *T. maritima* class II RNR utilized a mutant lacking the C-terminal domain,[17] thus the structure of this domain is still unknown. Unlike the C-terminal Cys pair in class I RNRs, which is appended to the RNR by a variable and likely flexible linker, it is likely that this domain in class II RNRs is ordered and might provide structural insight into the mechanism of reduction of the active site disulfide.

The *Pseudomonas aeruginosa* class II RNR has been cloned and studied.[12] The enzyme is expressed in two separate ORFs, with the 4-Cys motif contained in the second smaller ORF. Both peptides were necessary for nucleotide reduction activity in vitro, using DTT as the reductant. Phylogenetic analysis indicated that the split form evolved only once, though the adaptive advantage of a split protein is not known.

```
P_aeruginosa        YE---AAQGTSKVEPGEGFPAGAQLCGKCNSMA-----VVQMDGCATCLNCGHS-KCG--
B_halodurans        ----------VRTTDVTIGSEVGDTCPVCREGT-----VEDLGGCNTCTNCGAQLKCGL-
B_fragilis          LK-------KYVQ---DGTEAKGKKCPNCGNET-----LVYQEGCLICTTCGAS-RCG--
D_radiodurans       ------APSAPASTGVSVDGLGQERCPVCEEKA-----VIREEGCLKCQACGYS-KCG--
P_furiosus          ------VRPGNIPEEKIRELLGVVYCPVCYEKEGKLVELKMESGCATCPVCGWS-KCVIS
T_maritima          ----IVAHNLRWQSGYYVDDEGNVYCPVCLSKNS----LIKQEGCVSCKNCGWS-KCE--
C_aurantiacus       TATTPVSHPH--SDHGHGTIEAAEICPECQN-AT----MLNEEGCRKCHSCGYS-EC---
                                             *  *  .       :    **  *  **  .  .*
```

Figure 4.4 Sequence alignment of C-terminal portion of diverse class II RNRs containing the 4-Cys motif (highlighted yellow) and disulfide-forming Cys pair (highlighted green). Generated using ClustalO.[32]

4.1.1.4. ATP cones in class II RNRs

Most class II RNRs, including all monomeric class II RNRs, lack the N-terminal ATP cone (**Figures 1** and **3**),[16] and in several of those that possess it, the sequences resemble ATP cones of class III RNRs. It was therefore proposed that the ancestor of modern class II RNRs did not have an ATP cone.

4.1.2. Class III RNRs

As discussed in **Chapter 3**, the subclassification of the class III RNR into the NrdD1, NrdD2 and NrdD3 subtypes is related to the source of electrons for ribonucleotide reduction, which we believe is related to the electron carriers used in energy metabolism. Additional sources of variability and interesting cases are described below.

4.1.2.1. ATP cones in class III RNRs

The bacterial and archaeal NrdD1s and NrdD2s form distinct branches on the phylogenetic tree,[21] and the broad coherence with the phylogeny of the organisms is evidence for the antiquity of the class III RNR and the early divergence of NrdD1 and NrdD2. Another characteristic that distinguish the bacterial and archaeal enzymes of each subtype is the presence of the N-terminal ATP cone. Bacterial NrdD1s and NrdD2s contain one ATP cone domain. Archaeal NrdD2s have no ATP cone domain. The archaeal NrdD1s, present only in methanogens, have two tandem ATP cone domains, the second of which is predicted to be inactive (an exception is the deep branching *Methanopyrus* NrdD1, which has no ATP cone). Several bacterial and archaeal NrdD1s have been demonstrated to have activity regulation.[33-35]

The branch containing archaeal NrdD1s also contain NrdD3s, present in methanogens and their relatives Archaeoglobi, and a version of NrdD2 that appears to be derived from NrdD3, present in diverse bacteria and archaea (**Figure 4.9**). These enzymes have lost the N-terminal ATP cone, but retained the second putatively inactive one, suggesting a lack of allosteric activity regulation.

An exception is the NrdDs of bacteriophages, which are closely related to the NrdDs of their host,[36] but often lack an ATP cone and thus activity regulation, for example the NrdD1 from bacteriophage T4.[37]

4.1.2.2. Bacteroides ovatus RNRs

Bacteroides species are obligate anaerobic intestinal bacteria, which carry out fermentation and possess a PFL. Most species contain the bacterial NrdD1, which has an ATP cone domain. However, several species, including *B. ovatus*, contain a deep branching NrdD2, related to *T. maritima* and *Neisseria bacilliformis* NrdD2.[21] These NrdD2s are more closely related to archaeal NrdD2s and generally lack an ATP cone. However, *B. ovatus* NrdD2 contains an N-terminal ATP cone. A BLAST search for the origin of this ATP cone shows a 48% identity with the ATP cone of *B. ovatus* NrdD1, suggesting that the ATP cone is derived from NrdD1 (**Figure 4.5**).

Whether or not this domain is active is not known. If it indeed confers allosteric activity regulation on NrdD2, it would serve as a model illustrating the modularity of this allosteric domain. The mechanism of allosteric activity regulation by the ATP cone of class II and class III RNRs is not known, but the modularity of the ATP cone domain would disfavor mechanisms

requiring subtle conformational changes, and instead favor mechanism involving changes in topology or oligomeric state like the class Ia RNR.[38,39,67]

Another curiosity that has been noted in these organisms is that even though they are obligate anaerobes, they possess a class Ia RNR, likely acquired from a eukaryote.[9] There is evidence that this enzyme functions in facilitating recovery from exposure to O_2.[40] Many obligate anaerobic clostridia also contain class Ia and Ib RNRs.

```
B_ovatus_NrdD2      MISSEIFIIKRDGKKEAFSLDKIKNAISKAFLSVGSFATQDVITNVLSRVS---ISDGTN
B_ovatus_NrdD1      MNYAEICIIKRDGKREDFSISKIKNAISKAFSATGIQDEQQLVADITMNVISQFTTPTIT
                    *  :** *******:* **:.********** :.*    *:::::: .*     :   .

B_ovatus_NrdD2      VEEIQNQVEVALMAEHYYSVAKAYMLYRQKHLEDREVRDKLKFLMDYCDASNPASGSKYD
B_ovatus_NrdD1      VEEIQDLVEKSLMKV-RPEVAKKYIIYREWRNTERDKKTQMKQVMDGIVAIDKND-----
                    *****: ** :**     .*** *::**: :  :*: : ::* :**   * :  .

B_ovatus_NrdD2      ANANVENKNIATLIGELPKSNFIRLNRR---------LLTDRLRDMYGKEASDRYLELLN
B_ovatus_NrdD1      --VNLSNANM---SSHTPAGQMMTFASEVTKDYTYKYLLPKRFAEAH------QLGDIHI
                      .*:.* *:    .. * .::: :  .         ** .*: : :      :  ::

B_ovatus_NrdD2      HHFIYKNDETNLANYCASITMYPWLIAGTTAVGGNSTAPTNLKSFCGGFINMVFIVSSML
B_ovatus_NrdD1      HDLDYYP---TKTTTCIQYDMDDLFERGFRTKNGSIRTPQSIQSYATLATIIFQTNQNEQ
                    *.: *     . :. * .  *   :  * :  *.  :* .::*:.     :.   ..
```

Figure 4.5 Sequence alignment of N-terminal portion of *B. ovatus* NrdD1 and NrdD2 showing a 48% identity of the ATP cone domains (highlighted in yellow). Generated using ClustalO.[32]

4.1.2.3. C-terminal fusions to NrdD

Some eukaryotes contain PFL, acquired from bacteria.[41] It was noted that the gene from the marine diatom *Thalassiosira pseudonana* showed that its PFL contains a C-terminal fusion to PFLAE. Similarly, the existence of eukaryotic NrdD, acquired from bacteria, containing a C-terminal NrdG fusion was also noted.[9] It is not clear why a fusion protein is used by eukaryotes, especially since the activating enzymes are known to act catalytically. The reason that NrdD and NrdG fusion occurs more often than the PFL-PFLAE fusion may be related to the much lower activity of NrdG compared to PFLAE.

It was noted that gene fusions have also been observed in other proteins acquired from bacteria, for example glutamate synthase.[41] A possible reason is that, unlike in bacteria and archaea, genes in eukaryotes are not expressed as an operon, and the fusion proteins may facilitate coordinated expression, or facilitate horizontal gene transfer.

A version of NrdD2 with a C-terminal fusion to NrdH was noted in bacteria including *Acetobacterium woodii*. Again the function of this fusion is unclear.

4.1.2.4. Split class III RNRs

A class III RNR was identified in the megaplasmid of *Alcaligenes eutrophus* (*Ralstonia eutropha*), and preliminary mutagenesis and biochemical studies were carried out.[42] Deletion of *nrdD* or *nrdG* abolished the ability of the bacterium to grow anaerobically. The G• site was identified as G650 from sequence alignments, and the G650A mutation mimicked the *nrdD* deletion phenotype, suggesting to the authors that G650 was indeed the G• site. However, we noticed that G650 is located between two putative Zn-binding ligands, and the amino acids surrounding G650 is atypical of a G• site. A BLAST search revealed that this pattern is displayed by several other NrdD2 sequences, including *Pseudomonas aeruginosa* NrdD2 (**Figure 4.6**). Closer examination of the surrounding DNA sequence revealed that, in all these cases, there was a separate ORF (NrdD2b) encoding the G• peptide downstream of the annotated NrdD (NrdD2a).

Failure to detect anaerobic RNR activity in extracts of *A. eutrophus* is likely due to the use of assay conditions designed for the formate-dependent *E. coli* class III RNR, as noted by the authors (we now believe that the *A. eutrophus* class III RNR is thioredoxin-dependent). The same reason may underlie failure to detect class III RNR activity in extracts of *P. aeruginosa*.[43]

The "split" NrdD appears to have arisen independently multiple times. In *A. eutrophus* and *P. aeruginosa* NrdD2, the split occurs after the Zn site. In the NrdD3 of *Methanomicrobiales*, the split occurs after the second Zn ligand. These organisms also contain a class II RNR, and the G• peptide is located in a separate genomic locus together with AdoCbl salvage proteins.

The reason for expressing the G• peptide in a separate ORF is unclear, but a similar case has been described for *E. coli* PFL. In the event of cleavage of the G• peptide on exposure to O_2, activity can be reconstituted by complexation with a separate protein YfiD, which contains the G• peptide.[44]

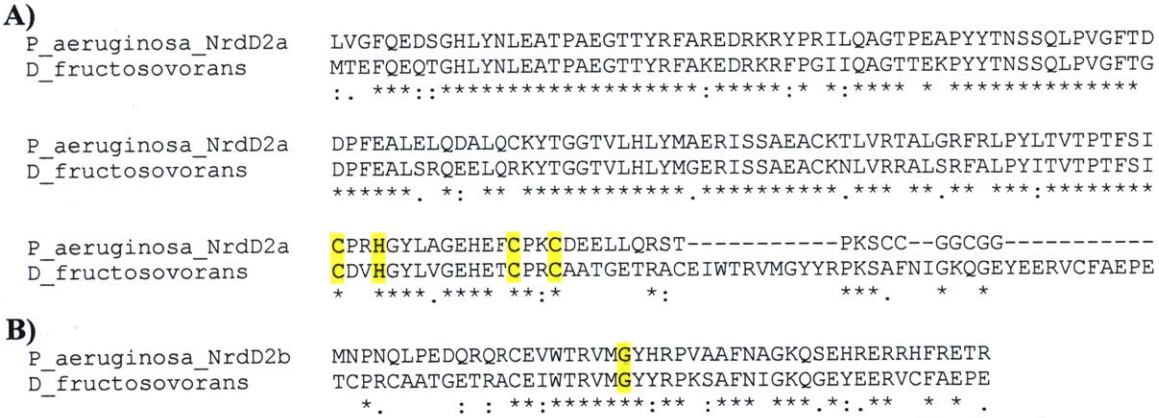


Figure 4.6 Sequence alignment of *P. aeruginosa* split NrdD2 with *Desulfovrbrio frutosovorans* full length NrdD2. A) C-terminal portion of NrdD2a (putative Zn ligands highlighted). B) NrdD2b encoding the separate G• peptide (G• residue highlighted). Generated using ClustalO.[32]

4.1.2.5. Class III RNRs of the termite symbiont Elusimicrobium minitum

An interesting case involves split NrdDs is the deep branching bacterium *E. minitum*, isolated from the termite gut.[45] This organism has NrdD1 and NrdD2 and no other RNRs. Both the NrdDs are split in the manner of *A. eutrophus* NrdD2, but are not closely related to the other split NrdDs.

The termite gut is an intensively studied model system for microbial ecology, contains a community of cellulolytic fermenters, acetogens and methanogens, and is rich in formate, which is thought to either accumulate, get oxidized to CO_2 or reduced to acetate by acetogens.[46] The acetogenic members are Spirochetes, including *Treponema primitia*. While most other acetogens contain NrdD2, *T. primitia* contains NrdD1, possibly because of availability of formate from the environment.

Unlike most other bacteria with NrdD1, *E. minitum* does not contain PFL. Instead, the operon for NrdD1 contains a formyl-THF synthetase. A possible role of this enzyme is to provide the source of formate for NrdD1. However, because of the presumed abundance of formate in its environment, we propose instead that the formyl-THF has an anabolic function and is co-induced with NrdD1 in the presence of formate. This would be consistent with our thesis that class III RNRs utilize electron sources that are abundant in the cell, rather than dedicated pathways.

4.1.3. Other interesting organisms and RNRs

A key realization in recent years is that the current distribution of RNRs is a result of widespread inter-domain horizontal gene transfer.[9] Thus the RNR present in the genome of an organism reflects its particular lifestyle, and not only its phylogeny. The presence of multiple RNRs of the same or different class within the same organism is particularly informative, often suggesting chemical differences between the RNR subtypes related to different growth conditions or cofactor availability (rather than the null hypothesis of functional redundancy as has been suggested[9]).

4.1.3.1. Preference for the class II RNR and regulation by cobalamin riboswitches

The main factor determining the use of class I or class III RNR by an organism is proposed to be presence or absence of O_2.[19,47] For the class II RNR, which is active under both aerobic and anaerobic conditions, one factor determining its use is the availability of AdoCbl, which is complex and costly to biosynthesize. *Streptomyces coelicolor* has a class Ia and a II RNR, either of which supports vegetative growth.[48] Expression of the class Ia RNR is repressed in the presence of cobalamin, by means of a cobalamin riboswitch. The advantage of using a class II RNR rather than a class I RNR in aerobic conditions is unclear.

The opportunistic human pathogen *P. aeruginosa* contains all three classes of RNR. Under aerobic conditions, the class Ia RNR is used during expotential growth and the class II during stationary phase.[12] During infection of *Drosophila*, class Ia RNR expression decreases, while class II and III RNR expression increases and contributed to virulence.[49] *P. aeruginosa* has an aerobic cobalamin biosynthesis pathway and under anaerobic conditions, despite the presence of NrdD2, lack of cobalamin in the medium results in defective growth, filamentation and biofilm formation.[50] This bacterium does not carry out fermentation, and anaerobic growth is strictly dependent on the presence NO_3^- as the electron acceptor. The failure of its split NrdD2 to fully support anaerobic growth is unclear, but we speculate that it may be due to reactive nitrogen species generated during NO_3^- reduction, or by immune cells during infection.

We noticed that in the strict anaerobe *Desulfuromonas acetoxidans* has a class II RNR and two class III RNRs (both NrdD2), one of which is preceded by cobalamin riboswitch. Both NrdD2s of this organism are split, the G• peptide is contained in a separate downstream ORF. We suggest that in addition to the possible salvage pathways for G•, the class II RNR may offer another strategy for O_2 tolerance in this organism. It is unclear whether the class II RNR confers

any advantage over the class III RNR under strict anaerobic conditions, and further investigations might reveal whether or not there was any driving force for evolution of the class II RNR before the advent of oxygenic photosynthesis.

4.1.3.2. Pseudomonas aeruginosa, an organism with RNRs of all three classes

P. aeruginosa is one of a minority of organisms that contain all three classes of RNR,[43,51] and the regulation of the expression of its RNRs has been studied in great detail.[49] Each of its RNRs contains unique features, and characteristics of its split class II[12] and III RNRs have been described above.

Its class Ia RNR is also unique and has been studied in detail.[43,50,52,53] The NrdB in this organism has a short-lived Y• that is not detectable during steady state turnover. Nucleotide reduction activity required O_2 in the assay mixture, where Y• was proposed to be efficiently regenerated.[52] This is unlike the atypical *E. coli* NrdB, which forms a stable "met" form, and requires a separate conserved ferredoxin YfaE for reactivation.[54,55] The *P. aeruginosa* NrdB is more closely related to the eukaryotic and viral NrdBs,[56] which have less stable diferric-Y•s. In these proteins, a single mechanism may be involved in activation and "reactivation". The reason for the contrasting characteristics of the class Ia RNRs from these two γ-proteobacterial species is unkown.

Unlike most other NrdAs with one ATP cone, this enzyme has two tandem ATP cones. The N-terminal one resembles the ATP cone of bacterial class III RNRs, and the second one has mutations suggesting that it is no longer functional. This pattern is also present in closely related but chemically distinct *C. trachomatis* class Ic RNR[57], and in other proteins.[58] Unlike most other class I RNRs, the *P. aeruginosa* NrdA and NrdB form a tight complex. A recent study on the

mechanism of activity regulation of this enzyme concluded that the active form is $\alpha_2\beta_2$, while dATP induces an α_4 structure mediated by the N-terminal ATP cone, that forms an inactive $\alpha_4\beta_2$ complex with NrdB. Furthermore, unlike other class I RNRs where the ATP cone binds one molecule of ATP, the N-terminal ATP cone of the *Pseudomonas* enzyme was found to bind 2 molecules of dATP. This study highlighted the diversity of mechanisms of activity regulation by the ATP cone module.

Formation of the active complex in class I RNRs is thought to be mediated in part by interaction of NrdA with a C-terminal peptide of NrdB. In *P. aeruginosa* NrdB, there is an additional N-terminal extension consisting of a variable length linker attached to a conserved 4-residue motif, resembling a motif present in the C-terminal of the same protein. This suggests that it may play a role in forming unusual quaternary structures, either in the active or inactive form of this RNR.

The reason for the presence of three highly unusual RNRs in this organism is unclear and may be related to its opportunistic lifestyle.

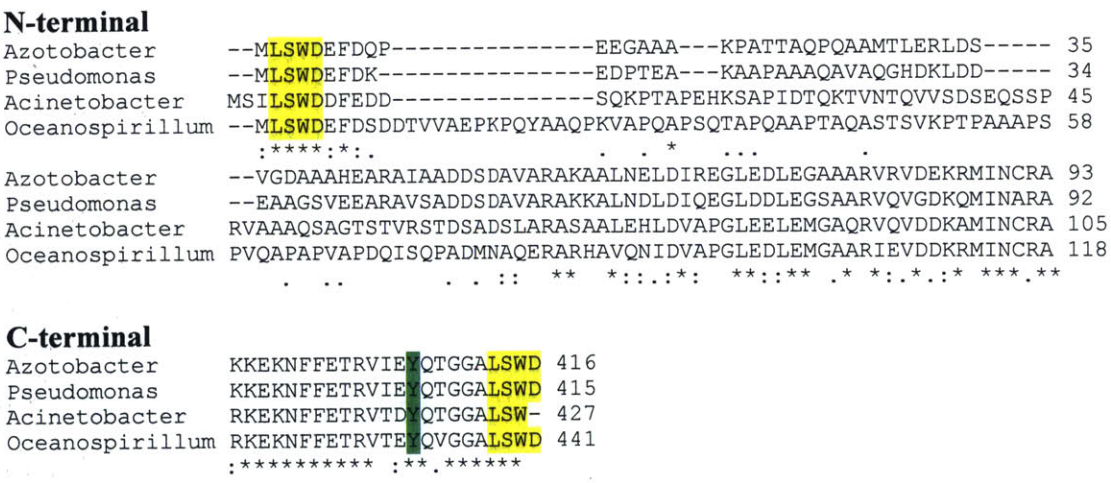


Figure 4.7 Sequence alignment of NrdBs related to *P. aeruginosa* NrdB showing the N-terminal repeat of the C-terminal motif likely involved in binding NrdA (highlighted yellow), and the Tyr residue involved in radical propagation (highlighted green). Generated using ClustalO.[32]

4.1.3.3. Bacteriophage RNRs

In addition to their nearly universal presence in cellular organisms, RNRs are encoded in a large number of viruses. The distribution of RNRs in bacteriophage metagenomes has been recently reviewed.[36] RNRs were found to be present in many large DNA phages. In many cases, RNRs of phages that infect a certain type of bacteria are more closely related to each other than to their host RNR. But they are more closely related to their host RNRs than to the RNRs of other types of bacteria that they do not infect, possibly reflecting an adaptation to the common cellular environment.

A model organism for bacteriophage RNRs is the lytic bacteriophage T4,[59,60] which infects *E. coli* and encodes a class Ia and a class III RNR. Both RNRs are highly similar to the host enzyme (58% identity for NrdD, 55% for NrdA, 53% for NrdB). Deletion of NrdD resulted in 5-fold decreased yield of the phage, showing that they play an important role in the phage lifecycle. Part of the difference in sequence is due to differences in allosteric regulation of the phage and host enzymes. An interesting difference between the phage and host class Ia RNRs is that while the *E. coli* class Ia RNR operon contains the ferredoxin YfaE,[54,55] it is absent in these short lifecycle lytic bacteriophages, consistent with its proposed role in cofactor maintenance rather than biosynthesis. This is true even for the T4-like lytic *Acinetobacter* phage ZZ1,[61] which infects a host with a related to *Pseudomonas* possessing a class Ia RNR lacking YfaE.

A reason why many large DNA phages encode their own RNRs, in addition to other genes for nucleotide metabolism and phosphate acquisition,[36] is suggested by recent biophysical modeling that quantified differences in the elemental stoichiometry of DNA phages compared with their hosts.[62] The authors suggest that while the protein content scales with the surface area of the capsid, the DNA content scales with its volume, with larger phages having larger DNA to

protein ratios. RNRs in these large DNA phages may thus serve to change the balance of protein and DNA synthesis in the cell, redirecting resources toward DNA synthesis. This is consistent with the observation that the RNRs of many phages, like bacteriophage T4, lack activity regulation. Both the T4 class Ia and class III RNRs are not inhibited by dATP. NrdD lacks an N-terminal ATP cone, while the N-terminal ATP cone in NrdA is inactive.

The "arms race" between host and phage has given rise to a variety of virus-specific DNA modifications, leading to the proposal that viruses may have played a role in the transition from the use of RNA (A,G,C,U) to the use of DNA (A,G,C,T) as the genetic material.[63] The bacteriophage T4 incorporates hydroxymethyl-dC (hm-dC) into its DNA in place of C, after which the hydroxymethyl group is glycosylated,[64] which rationalized differences in substrate specificity and allosteric regulation between the RNRs of the phage and host.[37,65] In T4 class Ia RNR, hm-dCTP and dCTP stimulate the reduction of pyrimidine nucleotides. dCTP has the same effect for T4 class III RNR, where hm-dCTP has not been tested. The T4 class III RNR does not reduce UTP, and UDP is a poor substrate for T4 class Ia RNR. This was rationalized by the presence of the efficient T4 dCTPase, which excludes dC from T4 DNA and produces dCMP, a precursor of both dUTP and the T4-specific hm-dCTP.[37]

While viruses are often considered as "particles", some features of their nucleotide metabolism may be understood by focusing on the intracellular stage of their life cycle, where they behave like "molecular" parasites.[66] Unlike "cellular" parasites, the phage is not bound by a membrane during infection, allowing it greater access to resources of the host, but also resulting in the coexistence of two organisms (phage and host) within the same cytoplasmic mix with no membrane barrier demarcating the phage. Because of this, a form of chemical orthogonality is required to prevent cross-talk of host and phage proteins.

The class Ia RNRs of *E. coli* and bacteriophage T4 form such an orthogonal system: T4 NrdA does not react with the *E. coli* NrdB, and vice versa.[65] The specificity of NrdB for its cognate NrdA despite high sequence identity of the host and phage enzymes is likely due to the divergent sequences of the NrdB C-terminal tails, which mediate interactions with NrdA. The T4 NrdB tail is two residues longer, and contains more hydrophobic residues (**Figure 4.8B**). Unlike the *E. coli* enzyme, the T4 NrdA and NrdB form a tight complex. The tight active complex of the T4 enzyme is also consistent with lack of activity regulation, which in the *E. coli* enzyme is mediated by structural interconversions.[67]

The bacteriophage T4 encodes its own thioredoxin, which is a substrate of *E. coli* thioredoxin reductase, but has a sequence and structure more similar to *E. coli* glutaredoxin and has residual glutaredoxin activity.[68] Like the orthogonality of the phage and host NrdA-NrdB combinations, the T4 thioredoxin does no reduce *E. coli* NrdA and vice versa. The origin of this specificity may be the NrdA C-terminal tail sequences (**Figure 4.8A**): *E. coli* NrdA has a CXXXXC sequence, while T4 NrdA has a CXXC sequence, similar to all class Ib RNRs which often require a dedicated redoxin NrdH.

Another interesting case that has been noted is in *Clostridium* phage c-st.[69] While the bacterial host encodes NrdD1, the phage genome encodes NrdD2, NrdG, NrdH and a ferredoxin:thioredoxin reductase, suggesting that it may hijack a different electron source for nucleotide reduction. A particularly striking arrangement occurs in some strains of *C. botulinum*, where the four genes occur contiguously in an operon adjacent to other genes involved in deoxynucleotide metabolism, and together with a homing endonuclease, suggesting that they are part of a mobile genetic element.

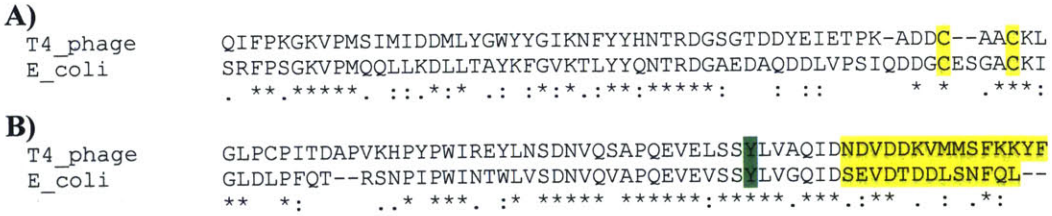


Figure 4.8 Sequence alignment of the C-terminal portion of *E. coli* and bacteriophage T4 NrdA and NrdB. A) NrdA, showing the divergent sequences of the C-terminal Cys pair (highlighted yellow) that serve as the substrate for their respective thioredoxins. B) NrdB, showing the divergent sequences of the C-terminal peptide that interacts with NrdA (highlighted yellow), and the Tyr residue involved in radical propagation (highlighted green). Generated using ClustalO.[32]

4.1.3.4. Introns, inteins and mobile genetic elements in RNRs

A puzzling observation is the prevalence of group I self-splicing introns in DNA biosynthesis genes of bacteriophages.[36,70] The bacteriophage T4, where these introns were first observed, contains introns in *nrdB*, *nrdD* and thymidylate synthase (prior to the discovery of the class III RNR, the *nrdD* gene was named *sunY*, and had served as a model for intron splicing)[59]. Subsequently, introns have been found in RNR genes of many other bacteriophages.[71,72] Because of this distribution, they had been proposed to play a regulatory role in regulation of DNA biosynthesis. However, it was oberved that the introns are not conserved in closely related phages, and deletion of the introns in bacteriophage T4 led to no observable deleterious phenotype.[70]

In contrast, bacterial introns are largely present in non-coding RNA, mostly tRNA.[70] Translation in bacteria like *E. coli* occurs co-transcriptionally, forming polyribosomes that interfere with intron splicing. This was proposed to account for deleterious effect of introns in bacterial protein-coding genes. Splicing of introns in bacteriophage protein-coding genes is facilitated by a stop codon close to the start of the intron sequence that releases the ribosome,[70]

and the absence of negative effects of this abortive translation may be related to the reduced requirements for protein vs. nucleic acid biosynthesis of the phage.

Some of these introns are associated with mobile genetic elements like homing endonucleases, which may facilitate its horizontal transfer and evolution of RNRs. An example of the ability of mobile genetic elements to drive RNR diversification is the NrdA of *Aeromonas hydrophila* phage Aeh1, where insertion of a sequence by a homing endonuclease resulted in the gene being split into two separate ORFs, each containing half of the 10-stranded barrel.[73] Soluble and enzymatically active NrdA could be obtained by coexpression of the two fragments in a recombinant system.

In contrast to phages, the mobile elements associated with bacterial and archaeal RNRs are often inteins. The location of the intein in all classes of RNR is often in between the two half-barrels on the conserved sequence encoding the thiyl radical loop, with the thiyl radical cysteine putatively involved in the intein excision chemistry.

4.2. Possible scenarios for the evolution of RNRs

The discovery that deoxynucleotides are enzymatically synthesized from ribonucleotides led to the proposal by Reichard that RNR may provide a direct evolutionary pathway for the transition from an RNA-protein world[74] to the modern world where genetic information is encoded by DNA, a scenario that has been the subject of numerous reviews.[16,47,75,76] The basis of this proposal is that, as far as we know, the conversion of NTPs to dNTPs is absolutely dependent on complex radical chemistry enabled by RNRs.

Alternative and less direct pathways have been postulated for synthesis of deoxynucleotides involving deoxyribose phosphate aldolase (DERA),[77] which catalyzes the

reversible synthesis of deoxyribose phosphate from glyceraldehyde and acetaldehyde. However, in modern organisms studied to date, DERA is a catabolic enzyme and its involvement in dNTP synthesis has not been demonstrated, despite advantages presented by a pathway that avoids the reliance radical chemistry requiring specific metals or AdoCbl, which are scarce for many environmental and pathogenic bacteria.

Based on the structural homology, common scheme for allosteric regulation and partially conserved radical-based reaction mechanism for all three classes of RNR, there is general consensus that the current diversity of RNRs arose from divergent evolution of a common ancestor.[19,51] The conserved pattern of allosteric regulation despite lack of structural homology between the allosteric sites of the class I and II vs. class III RNR is a mystery and discussed in several reviews.[16,78]

In this section, we add to the discussion of the evolutionary history of RNR in view of the recently described class III RNR subtypes and new G• enzymes. A distinction has been made between the first enzyme with RNR activity (proto-RNR) and the last common ancestor of modern RNRs (ur-RNR).[16] Here we discuss only the "ur-RNR", suggesting that it is a formate-dependent class III RNR.

4.2.1. Nature of the ancestor of class I and II RNRs

Because of the reliance of class I RNR on O_2, thought to be absent in the primordial atmosphere, there is general consensus that it was the last class of RNR to evolve.[19,47,76] Cloning and sequencing of the first class II RNR, the monomeric NTP-reducing enzyme from *L. leichmannii*,[7] showed that it exhibits negligible sequence identity with the class I RNRs. Subsequently, the cloning of several dimeric, NDP-reducing class II RNRs showed that they

exhibit significant sequence identity to the class Ia RNRs.[13,14] The conserved architecture of active and allosteric sites, use of NDP substrate, and mechanism of reaction with thioredoxin via a pair of C-terminal tail cysteines suggest that class I RNRs evolved from dimeric NDP-reducing class II RNRs.

We suggested above that the ancestor of class II RNRs is dimeric, reduces NTPs and contains a possible C-terminal metal-binding site. These features are shared with the class III RNR, and possibly reflect the early divergence of these two classes.

4.2.2. Nature of the ancestor of class III RNRs

The phylogenetic analysis and subclassification of class III RNRs into NrdD1, NrdD2 and NrdD3 according to active-site residues is discussed in **Chapter 3**. The archaeal and bacterial NrdD1s and NrdD2s are in separate branches of the tree, suggesting that the enzymes had likely already diverged in the last common ancestor of bacteria and archaea. This is consistent with the proposed antiquity of the class III RNR, but makes it impossible to infer whether NrdD1 or NrdD2 was the ancestral RNR using phylogenetic analysis.

However, some information may be obtained by examining the branch containing the class III RNRs of Euryarchaeota, which includes methanogens and their relatives and contains all three NrdD subtypes (**Figure 4.9**). The phylogeny of this group of organisms has been studied in detail, and broadly corresponds to the phylogeny of their class III RNRs. This suggests that the diverse class III RNRs in these organisms evolved from a protein present in their last common ancestor.

The basal methanogenic orders (*Methanopyrus*, *Methanococcales*, *Methanobacteriales*) contain NrdD1, suggesting that this is the ancestral class III RNRs in this group. More

evolutionarily recent orders of methanogens and their relatives (*Archaeoglobales*, *Methanomicrobiales*, *Methanosarcinales*) have NrdD3. The high sequence identity of up to 53% (between *Archaeoglobus profundus* NrdD3 and *Methanocella arvoryzae* NrdD1) is evidence for the recent divergence of these enzymes.

This branch also contains a NrdD2 that is distantly related to NrdD2s in the other bacterial and archaeal branches. This NrdD2 is present in some members of the newly identified seventh order of methanogens (*Methanomassiliicoccales*),[79] but also in other diverse bacteria, likely through horizontal gene transfer. The high sequence identity of up to 44% (between *A. profundus* NrdD3 and *Methanomassiliicoccus luminyensis* NrdD2), suggests convergent evolution of the NrdD2 active site from NrdD3 precursor by gaining an active site Glu.

These observations above establish a precedent for the evolution of a NrdD2-type active site from a NrdD1 precursor. We extend this to propose that NrdD2 in other bacteria and archaea evolved from NrdD1 as well. Evidence comes from the presence of an active site Met in most of the archaeal and bacterial NrdD2s.[21] In NrdD1, this Met is involved in forming a thiosulfuranyl radical that acts as the oxidant for formate during nucleotide reduction.[2] This role is thought not to be required for NrdD2, which is proposed to use the same catalytic mechanism as the class I and II RNRs.[21] The presence of this Met in most but not all NrdD2s suggests that it may be an evolutionary relic, consistent with the high degree of conservation of residues on the thiyl radical loop. We thus propose that the ancestor of class III RNR is NrdD1, with the later evolution of NrdD2, possibly through a NrdD3-like intermediate.

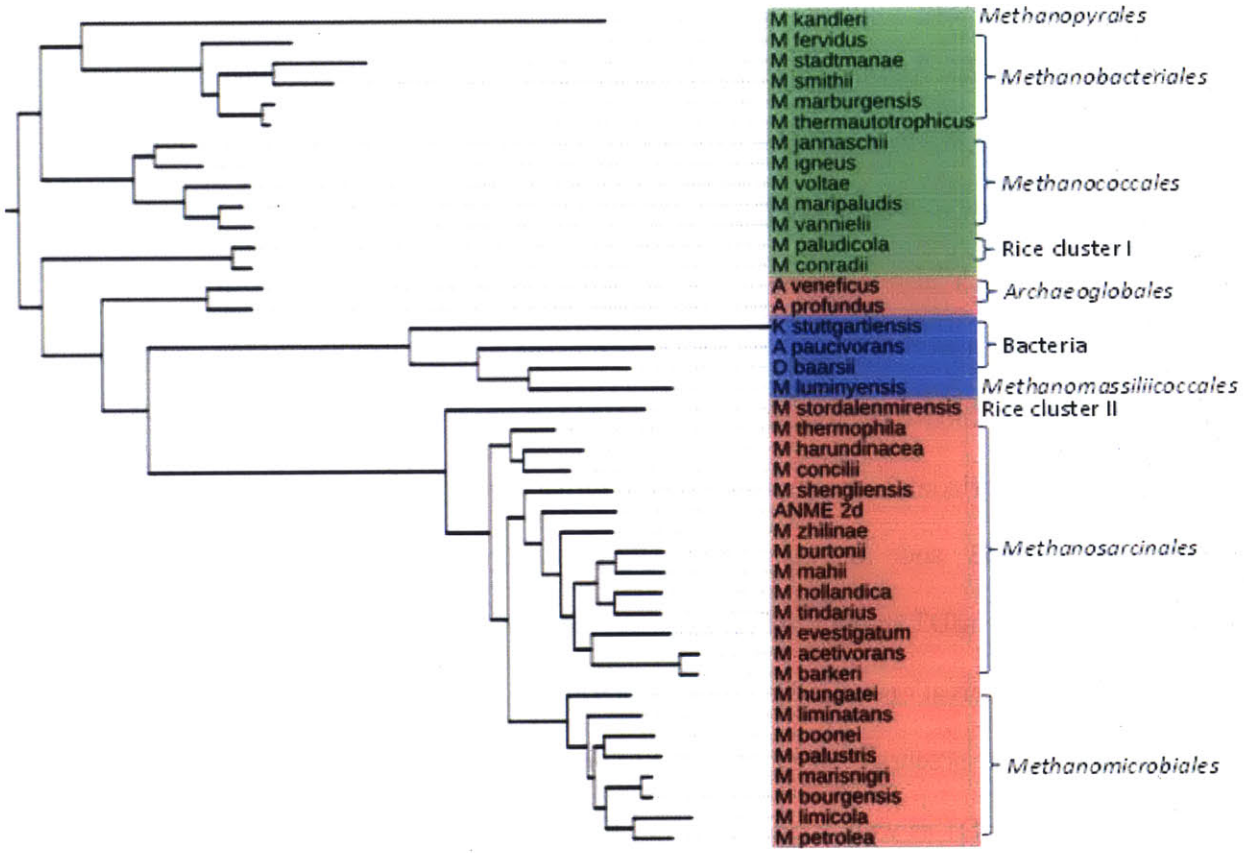


Figure 4.9 Phylogenetic tree of NrdDs in the Euryarchaeota branch, with different archaeal orders labeled. NrdD subtypes are indicated by highlighting (green=NrdD1, blue=NrdD2, red=NrdD3). Constructed using PhyML[27] and plotted using iTOL[28].

4.2.3. A class III-like ancestor for RNR

Phylogenetic analysis has had limited value in resolving the identity of the ancestor of RNRs. Analysis is confounded because of widespread inter-domain horizontal gene transfer,[9] and because of the highly divergent primary sequences of the class II and III RNRs, suggesting that most of the phylogenetic signal has been lost in evolutionary time, Instead, arguments have been based on the availability of cofactors and simplicity of different types of radical chemistry.

Despite the partially conserved radical chemistry for nucleotide reduction, the different cofactors used by the three classes of RNR revealed the surprising modularity of the radical-generating unit. The simpler chemical structure of formate, iron sulfur clusters and SAM has been used by Reichard to argue for the primordial nature of the formate-dependent class III RNRs.[75] However, it was also pointed out by Sjöberg that redoxins and AdoCbl are likely ancient molecules present in the RNA-protein world before the advent of RNR.[47] The recruitment of the pre-existing AdoCbl module would thus offer a direct pathway for the evolution of a radical-dependent enzyme from a non-radical precursor.

We now add further arguments supporting the proposal that the ancestor of all RNRs is the class III RNR, by examining the diversity of radical-dependent enzymes to find precedent for the various evolutionary scenarios. Our arguments are based on: **1**) the precedent of a redoxin-dependent class III RNR evolving from a formate-dependent one, but not the converse; **2**) the requirement for a thiyl radical intermediate in all G• radical enzymes but not in all AdoCbl enzymes; **3**) the high diversity of radical SAM enzymes and of G• enzymes, but not of AdoCbl enzymes with the RNR fold; **4**) the variability and less stringent requirements for the active site of the anaerobic class III RNR.

1) The class II and III RNRs differ not only in the cofactor used to generate the thiyl radical, but also in the reductants used for nucleotide reduction. Here we argue on grounds of parsimony that, given the evidence for redoxin-dependent class III RNRs evolving from formate-dependent ones, it unlikely that the ancestral RNR is a redoxin-dependent class II RNR. We analyze two unlikely scenarios below:

i) The first RNR is a redoxin-dependent class II RNR. NrdD1 evolves through an intermediate formate-dependent class II RNR. We consider this unlikely because

there is no evidence for such an enzyme in modern organisms, even in formate-producing organisms like Methanobacteria, Clostridia, Bacteroidetes and Proteobacteria, which contain a formate-dependent class III RNR and redoxin-dependent class II RNR. It would be puzzling for this evolutionary intermediate to have vanished despite the selection pressure required for it to evolve in the first place.

ii) The first RNR is a redoxin-dependent class II RNR. NrdD1 evolves through an intermediate redoxin-dependent class III RNR. We consider this unlikely because of our arguments for a formate-dependent ancestor of class III RNRs.

2) A characteristic that RNRs share with G• enzymes in the use of a thiyl radical for substrate activation.[80] This is true even of the newly discovered non-canonical G• enzyme PhnJ, which is structurally unrelated to RNR,[81] suggesting that the thiyl radical may be a mechanistic necessity for G• enzymes, rather than an evolutionary contingency. The reason why G• is not directly used for substrate activation may be rationalized by higher steric demands compared to the thiyl radical, or a higher reorganization barrier for the α-C on the peptide backbone in changing from sp^2 to sp^3 geometry.

In contrast, the class II RNR is the only known AdoCbl enzyme that uses a thiyl radical for substrate activation, all other AdoCbl enzymes use 5′-dA• for direct generation of the substrate radical, and are based on the structurally unrelated TIM barrel scaffold.[82]

Even for G• and AdoCbl enzymes catalyzing similar reactions, different chemical mechanisms are used. For example, in the G• glycerol dehydratase, dehydration of the substrate radical is thought to be catalyzed by conserved active site acid / base residues, generating a ketyl radical that regenerates the thiyl radical and G•.[83] On the other hand, the AdoCbl-dependent diol

dehydratase catalyzes 1,2-migration of the OH group, generating a more reactive alkyl radical that regenerates the 5′-dA radical.[84] A similar example is the AdoCbl-dependent ethanolamine ammonia lyase, which is thought to catalyze 1,2-migration of the NH_2 group.[85] The analogous mechanism may not be possible for quaternary amine choline, which may rationalize why the choline-TMA lyase reaction is catalyzed by a G• enzyme.[86]

The regeneration of the 5′-dA radical by the thiyl radical in class II RNRs is thought to be thermodynamically uphill and driven forward by the formation of the C-Co^{III} bond.[87] The mechanistic rationale for converting a highly reactive 5′-dA radical to a less reactive thiyl radical is puzzling and counter to the mode of reactivity of other AdoCbl enzyme, thus we suggest that the use of a thiyl radical is an evolutionary relic.

3) The complexity of the activation reaction involving two separate proteins has been an argument against the RNR ancestor being a G• enzyme.[16,47] We argue instead that the radical SAM superfamily,[88] to which the activase NrdG belongs, includes enzymes that catalyze a diverse array of reactions, many of which are more complex than G• generation. Enzymes in this evolutionarily ancient superfamily are involved in the biosynthesis of many cofactors proposed to play a role in the "RNA world". Their substrates include peptides and other macromolecules.[89,90] An example is the anaerobic sulfatase maturase, which was suggested to provide a mechanistic,[91] and perhaps evolutionary, link to G• activating enzymes. The discovery of the non-canonical G• enzyme PhnJ demonstrates that the generation of G• by a radical SAM enzyme evolved independently at least twice.[81] Therefore the evolution of a G• RNR from a non-radical precursor is not inconceivable.

Another possibility is the evolution of class III RNR from a pre-existing G• enzyme. The large cavity formed by the 10 stranded barrel of G• enzymes insulates the G• cofactor and radical

intermediates from the environment, and have evolved to accommodate a large range of substrates.[92] Benzylsuccinate synthase, like NrdD1, catalyzes the reaction of two different substrate molecules.[93] The fold of G• enzymes is also malleable, as illustrated by the extensive remodeling of YjjI, an enzyme present in *E. coli* that has an unknown function. In addition "split" NrdDs and YfiD, described above, demonstrate the modularity of the C-terminal domain, which could conceivably be replaced by other cofactors like AdoCbl. The examples above provide precedent for structural and biochemical diversification of G• enzymes, providing a pathway for the evolution of a G• class III RNR. In contrast, the class II RNR is the only AdoCbl-dependent enzyme with this fold.

4) Even among class III RNRs, there is diversity in active site architecture. This variability, even between the redoxin-dependent NrdD2 and NrdD3, may be tolerated due to the strict anaerobic conditions required by this enzyme. The reaction mechanism of class I and II RNRs is thought to minimize the accumulation of radical intermediates,[87] and their highly conserved active-site architecture with a H-bonding framework involving a conserved Asn,[94] and deeply buried disulfide[95] necessitating the redox relay via a C-terminal Cys pair, may reflect stringent selection due to competing pathways for radical quenching in the presence of O_2. For example, despite the high reactivity of sulfur-based radicals towards O_2,[96] the disulfide anion radical formed on the *E. coli* class Ia RNR has a lifetime of minutes.[97,98] In contrast, a thiosulfuranyl radical accumulates in the *E. coli* class III RNR under conditions of limiting formate,[2] which is stable for several minutes under anaerobic conditions but rapidly quenched upon exposure to air.

The difference in active-site architecture between class II and redoxin-dependent class III RNRs demonstrates that the multiple configurations are able to carry out the redoxin-dependent

chemistry under anaerobic conditions. Given the more stringent selection experienced by class I and II RNRs in the presence of O_2, it is conceivable that the selection pressure drove the evolution of a class I and II active-site architecture from the NrdD2 architecture, but not the reverse.

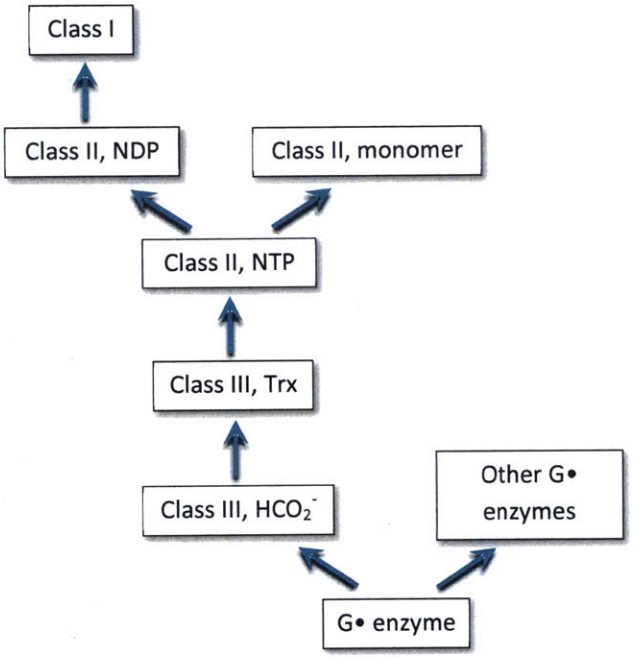


Figure 4.9 Proposed scenario for evolution of RNRs.

4.2.3.1. Conclusions

Based on the arguments above, we favor the scenario similar to the one originally described by Reichard (**Figure 4.9**).[75] We envision the genesis of the formate-dependent class III RNR from a G• enzyme like PFL, in a primordial cell rich with ribonucleotides and formate produced through an ancient C1 metabolism resembling the Wood–Ljungdahl pathway.[99] Adoption of redoxin-dependent chemistry allowed acquisition by organisms with other types of formate-independent metabolism. Oxygenation of the atmosphere drove the adoption of AdoCbl

as the radical generating cofactor. Subsequently, acquisition by a thermophilic archaeon led to the transition from NTPs to NDPs as substrates. From one of the NDP-reducing RNRs evolved the prototype of class I RNR, eliminating the reliance on the costly AdoCbl cofactor. Diversification of metallocofactors driven by metal availability and specific requirements for radical stability and maintenance led to the current assortment of class I RNRs.

4.3. Acknowledgements

We thank Michael A. Funk for identifying sequence motifs involved in binding of NTP or NDP substrates.

4.4. References

(1) Licht, S.; Gerfen, G. J.; Stubbe, J. *Science* **1996**, *271*, 477.

(2) Wei, Y.; Mathies, G.; Yokoyama, K.; Chen, J.; Griffin, R. G.; Stubbe, J. *J. Am. Chem. Soc.* **2014**, *136*, 9001.

(3) Jiang, W.; Xie, J.; Varano, P. T.; Krebs, C.; Bollinger Jr, J. M. *Biochemistry* **2010**, *49*, 5340.

(4) Zhang, Y.; An, X.; Stubbe, J.; Huang, M. *J. Biol. Chem.* **2013**, *288*, 13951.

(5) Cotruvo, J. A.; Stubbe, J. *Annu. Rev. Biochem.* **2011**, *80*, 733.

(6) Blakely, R.; Barker, H. *Biochem. Biophys. Res. Commun.* **1964**, *16*, 391.

(7) Booker, S.; Stubbe, J. *Proc. Natl. Acad. Sci. U. S. A.* **1993**, *90*, 8352.

(8) Lin, A. N.; Ashley, G. W.; Stubbe, J. *Biochemistry* **1987**, *26*, 6905.

(9) Lundin, D.; Gribaldo, S.; Torrents, E.; Sjöberg, B.-M.; Poole, A. M. *BMC evolutionary biology* **2010**, *10*, 383.

(10) Gleason, F. K.; Hogenkamp, H. P. C. *J. Biol. Chem.* **1970**, *245*, 4894.

(11) Hamilton, F. D. *J. Biol. Chem.* **1974**, *249*, 4428.

(12) Torrents, E.; Poplawski, A.; Sjöberg, B.-M. *J. Biol. Chem.* **2005**, *280*, 16571.

(13) Riera, J.; Robb, F. T.; Weiss, R.; Fontecave, M. *Proc. Natl. Acad. Sci. U. S. A.* **1997**, *94*, 475.

(14) Tauer, A.; Benner, S. A. *Proceedings of the National Academy of Sciences* **1997**, *94*, 53.

(15) Jordan, A.; Torrents, E.; Jeanthon, C.; Eliasson, R.; Hellman, U.; Wernstedt, C.; Barbé, J.; Gibert, I.; Reichard, P. *Proceedings of the National Academy of Sciences* **1997**, *94*, 13487.

(16) Lundin, D.; Berggren, G.; Logan, D. T.; Sjöberg, B.-M. *Life* **2015**, *5*, 604.

(17) Larsson, K.-M.; Jordan, A.; Eliasson, R.; Reichard, P.; Logan, D. T.; Nordlund, P. *Nat. Struct. Mol. Biol.* **2004**, *11*, 1142.

(18) Sintchak, M. D.; Arjara, G.; Kellogg, B. A.; Stubbe, J.; Drennan, C. L. *Nat Struct Biol* **2002**, *9*, 293.

(19) Reichard, P. *Science* **1993**, *260*, 1773.

(20) Johnson, J. E.; Webb, S. M.; Thomas, K.; Ono, S.; Kirschvink, J. L.; Fischer, W. W. *Proceedings of the National Academy of Sciences* **2013**, *110*, 11238.

(21) Wei, Y.; Funk, M. A.; Rosado, L. A.; Baek, J.; Drennan, C. L.; Stubbe, J. *Proceedings of the National Academy of Sciences* **2014**, *111*, E3756.

(22) Kengen, S. W. M.; Tuininga, J. E.; de Bok, F. A. M.; Stams, A. J. M.; de Vos, W. M. *J. Biol. Chem.* **1995**, *270*, 30453.

(23) Ronimus, R. S.; Koning, J.; Morgan, H. W. *Extremophiles* **1999**, *3*, 121.

(24) Crooks, G. E.; Hon, G.; Chandonia, J.-M.; Brenner, S. E. *Genome Res.* **2004**, *14*, 1188.

(25) Licht, S. S.; Lawrence, C. C.; Stubbe, J. *J. Am. Chem. Soc.* **1999**, *121*, 7463.

(26) Larsson, K.-M.; Logan, D. T.; Nordlund, P. r. *ACS Chem. Biol.* **2010**, *5*, 933.

(27) Guindon, S.; Lethiec, F.; Duroux, P.; Gascuel, O. *Nucleic Acids Res.* **2005**, *33*, W557.

(28) Letunic, I.; Bork, P. *Nucleic Acids Res.* **2011**, *39*, W475.

(29) Logan, D. T.; Mulliez, E.; Larsson, K. M.; Bodevin, S.; Atta, M.; Garnaud, P. E.; Sjoberg, B. M.; Fontecave, M. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 3826.

(30) Grinberg, I.; Shteinberg, T.; Hassan, A. Q.; Aharonowitz, Y.; Borovok, I.; Cohen, G. *J. Bacteriol.* **2009**, *191*, 1169.

(31) Eliasson, R.; Pontis, E.; Jordan, A.; Reichard, P. *J. Biol. Chem.* **1999**, *274*, 7182.

(32) Sievers, F.; Higgins, D. G. In *Multiple Sequence Alignment Methods*; Springer: 2014, p 105.

(33) Sze, I. S. Y.; McFarlan, S. C.; Spormann, A.; Hogenkamp, H. P. C.; Follmann, H. *Biochem. Biophys. Res. Commun.* **1992**, *184*, 1101.

(34) Torrents, E.; Buist, G.; Liu, A.; Eliasson, R.; Kok, J.; Gibert, I.; Gräslund, A.; Reichard, P. *J. Biol. Chem.* **2000**, *275*, 2463.

(35) Eliasson, R.; Pontis, E.; Sun, X.; Reichard, P. *J. Biol. Chem.* **1994**, *269*, 26052.

(36) Dwivedi, B.; Xue, B.; Lundin, D.; Edwards, R. A.; Breitbart, M. *BMC evolutionary biology* **2013**, *13*, 33.

(37) Andersson, J.; Westman, M.; Hofer, A.; Sjöberg, B.-M. *J. Biol. Chem.* **2000**, *275*, 19443.

(38) Zimanyi, C. M.; Ando, N.; Brignole, E. J.; Asturias, F. J.; Stubbe, J.; Drennan, C. L. *Structure* **2012**, *20*, 1374.

(39) Fairman, J. W.; Wijerathna, S. R.; Ahmad, M. F.; Xu, H.; Nakano, R.; Jha, S.; Prendergast, J.; Welin, R. M.; Flodin, S.; Roos, A. *Nat. Struct. Mol. Biol.* **2011**, *18*, 316.

(40) Smalley, D.; Rocha, E. R.; Smith, C. J. *J. Bacteriol.* **2002**, *184*, 895.

(41) Stairs, C. W.; Roger, A. J.; Hampl, V. *Mol. Biol. Evol.* **2011**, *28*, 2087.

(42) Siedow, A.; Cramm, R.; Siddiqui, R. A.; Friedrich, B. *J. Bacteriol.* **1999**, *181*, 4919.

(43) Jordan, A.; Torrents, E.; Sala, I.; Hellman, U.; Gibert, I.; Reichard, P. *J. Bacteriol.* **1999**, *181*, 3974.

(44) Wagner, A. V.; Schultz, S.; Bomke, J.; Pils, T.; Lehmann, W. D.; Knappe, J. *Biochem. Biophys. Res. Commun.* **2001**, *285*, 456.

(45) Herlemann, D. P. R.; Geissinger, O.; Ikeda-Ohtsubo, W.; Kunin, V.; Sun, H.; Lapidus, A.; Hugenholtz, P.; Brune, A. *Appl. Environ. Microbiol.* **2009**, *75*, 2841.

(46) Brune, A. *Nat. Rev. Microbiol.* **2014**, *12*, 168.

(47) Poole, A. M.; Logan, D. T.; Sjöberg, B.-M. *J. Mol. Evol.* **2002**, *55*, 180.

(48) Borovok, I.; Gorovitz, B.; Schreiber, R.; Aharonowitz, Y.; Cohen, G. *J. Bacteriol.* **2006**, *188*, 2512.

(49) Sjöberg, B.-M.; Torrents, E. *Infect. Immun.* **2011**, *79*, 2663.

(50) Lee, K.-M.; Go, J.; Yoon, M. Y.; Park, Y.; Kim, S. C.; Yong, D. E.; Yoon, S. S. *Infect. Immun.* **2012**, *80*, 1639.

(51) Torrents, E.; Aloy, P.; Gibert, I.; Rodríguez-Trelles, F. *J. Mol. Evol.* **2002**, *55*, 138.

(52) Torrents, E.; Westman, M.; Sahlin, M.; Sjöberg, B.-M. *J. Biol. Chem.* **2006**, *281*, 25287.

(53) Jonna, V. R.; Crona, M.; Rofougaran, R.; Lundin, D.; Johansson, S.; Brännström, K.; Sjöberg, B.-M.; Hofer, A. *J. Biol. Chem.* **2015**.

(54) Wu, C. H.; Jiang, W.; Krebs, C.; Stubbe, J. *Biochemistry* **2007**, *46*, 11577.

(55) Hristova, D.; Wu, C. H.; Jiang, W.; Krebs, C.; Stubbe, J. *Biochemistry* **2008**, *47*, 3989.

(56) Mann, G. J.; Graeslund, A.; Ochiai, E. I.; Ingemarson, R.; Thelander, L. *Biochemistry* **1991**, *30*, 1939.

(57) Roshick, C.; Iliffe-Lee, E. R.; McClarty, G. *J. Biol. Chem.* **2000**, *275*, 38111.

(58) Aravind, L.; Wolf, Y. I.; Koonin, E. V. *J. Mol. Microbiol. Biotechnol.* **2000**, *2*, 191.

(59) Young, P.; Ohman, M.; Xu, M. Q.; Shub, D. A.; Sjöberg, B. *J. Biol. Chem.* **1994**, *269*, 20229.

(60) Berglund, O. *J. Biol. Chem.* **1975**, *250*, 7450.

(61) Jin, J.; Li, Z.-J.; Wang, S.-W.; Wang, S.-M.; Chen, S.-J.; Huang, D.-H.; Zhang, G.; Li, Y.-H.; Wang, X.-T.; Wang, J. *BMC genomics* **2014**, *15*, 793.

(62) Jover, L. F.; Effler, T. C.; Buchan, A.; Wilhelm, S. W.; Weitz, J. S. *Nat. Rev. Microbiol.* **2014**, *12*, 519.

(63) Forterre, P. *Biochimie* **2005**, *87*, 793.

(64) Warren, R. *Annual Reviews in Microbiology* **1980**, *34*, 137.

(65) Berglund, O. *J. Biol. Chem.* **1972**, *247*, 7276.

(66) Bandea, C. I. **2009**.

(67) Ando, N.; Brignole, E. J.; Zimanyi, C. M.; Funk, M. A.; Yokoyama, K.; Asturias, F. J.; Stubbe, J.; Drennan, C. L. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 21046.

(68) Holmgren, A. *J. Biol. Chem.* **1978**, *253*, 7424.

(69) Balsera, M.; Uberegui, E.; Susanti, D.; Schmitz, R. A.; Mukhopadhyay, B.; Schürmann, P.; Buchanan, B. B. *Planta* **2013**, *237*, 619.

(70) Edgell, D. R.; Belfort, M.; Shub, D. A. *J. Bacteriol.* **2000**, *182*, 5281.

(71) Landthaler, M.; Begley, U.; Lau, N. C.; Shub, D. A. *Nucleic Acids Res.* **2002**, *30*, 1935.

(72) Lazarevic, V. *Nucleic Acids Res.* **2001**, *29*, 3212.

(73) Friedrich, N. C.; Torrents, E.; Gibb, E. A.; Sahlin, M.; Sjöberg, B.-M.; Edgell, D. R. *Proceedings of the National Academy of Sciences* **2007**, *104*, 6176.

(74) Freeland, S. J.; Knight, R. D.; Landweber, L. F. *Science* **1999**, *286*, 690.

(75) Reichard, P. *Trends Biochem. Sci.* **1997**, *22*, 81.

(76) Stubbe, J.; Ge, J.; Yee, C. S. *Trends Biochem. Sci.* **2001**, *26*, 93.

(77) Poole, A. M.; Horinouchi, N.; Catchpole, R. J.; Si, D.; Hibi, M.; Tanaka, K.; Ogawa, J. *J. Mol. Evol.* **2014**, *79*, 204.

(78) Nordlund, P.; Reichard, P. *Annu. Rev. Biochem.* **2006**, *75*, 681.

(79) Borrel, G.; O'Toole, P. W.; Harris, H. M. B.; Peyret, P.; Brugère, J.-F.; Gribaldo, S. *Genome Biology and Evolution* **2013**, *5*, 1769.

(80) Eklund, H.; Fontecave, M. *Structure* **1999**, *7*, R257.

(81) Kamat, S. S.; Williams, H. J.; Dangott, L. J.; Chakrabarti, M.; Raushel, F. M. *Nature* **2013**, *497*, 132.

(82) Banerjee, R.; Ragsdale, S. W. *Annu. Rev. Biochem.* **2003**, *72*, 209.

(83) Feliks, M.; Ullmann, G. M. *The Journal of Physical Chemistry B* **2012**, *116*, 7076.

(84) Smith, D. M.; Golding, B. T.; Radom, L. *J. Am. Chem. Soc.* **2001**, *123*, 1664.

(85) Warncke, K.; Canfield, J. M. *J. Am. Chem. Soc.* **2004**, *126*, 5930.

(86) Craciun, S.; Marks, J. A.; Balskus, E. P. *ACS Chem. Biol.* **2014**, *9*, 1408.

(87) Licht, S.; Stubbe, J. In *Comprehensive Natural Products Chemistry*; Barton, S., Nakanishi, K., Meth-Cohn, O., Poulter, C., Eds.; Elsevier Science: New York, 1999; Vol. 5, p 163.

(88) Sofia, H. J.; Chen, G.; Hetzler, B. G.; Reyes-Spindola, J. F.; Miller, N. E. *Nucleic Acids Res.* **2001**, *29*, 1097.

(89) Yan, F.; LaMarre, J. M.; Röhrich, R.; Wiesner, J.; Jomaa, H.; Mankin, A. S.; Fujimori, D. G. *J. Am. Chem. Soc.* **2010**, *132*, 3953.

(90) Anton, B. P.; Saleh, L.; Benner, J. S.; Raleigh, E. A.; Kasif, S.; Roberts, R. J. *Proceedings of the National Academy of Sciences* **2008**, *105*, 1826.

(91) Benjdia, A.; Subramanian, S.; Leprince, J.; Vaudry, H.; Johnson, M. K.; Berteau, O. *FEBS J.* **2010**, *277*, 1906.

(92) Lehtiö, L.; Goldman, A. *Protein Engineering Design and Selection* **2004**, *17*, 545.

(93) Leuthner, B.; Leutwein, C.; Schulz, H.; Hörth, P.; Haehnel, W.; Schiltz, E.; Schägger, H.; Heider, J. *Mol. Microbiol.* **1998**, *28*, 615.

(94) Kasrayan, A.; Persson, A. L.; Sahlin, M.; Sjöberg, B.-M. *J. Biol. Chem.* **2002**, *277*, 5749.

(95) Eriksson, M.; Uhlin, U.; Ramaswamy, S.; Ekberg, M.; Regnström, K.; Sjöberg, B.-M.; Eklund, H. *Structure* **1997**, *5*, 1077.

(96) Reddy, S. G.; Wong, K. K.; Parast, C. V.; Peisach, J.; Magliozzo, R. S.; Kozarich, J. W. *Biochemistry* **1998**, *37*, 558.

(97) Persson, A. L.; Eriksson, M.; Katterle, B.; Pötsch, S.; Sahlin, M.; Sjöberg, B.-M. *J. Biol. Chem.* **1997**, *272*, 31533.

(98) Lawrence, C. C.; Bennati, M.; Obias, H. V.; Bar, G.; Griffin, R. G.; Stubbe, J. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 8979.

(99) Poehlein, A.; Schmidt, S.; Kaster, A.-K.; Goenrich, M.; Vollmers, J.; Thürmer, A.; Bertsch, J.; Schuchmann, K.; Voigt, B.; Hecker, M. *PLoS One* **2012**, *7*, e33439.

Chapter 5:

A ferredoxin-dependent class III ribonucleotide reductase in *Methanosarcina barkeri*

5. A ferredoxin-dependent class III ribonucleotide reductase in *Methanosarcina barkeri*

5.1. Introduction

In many anaerobic bacteria and archaea, the reduction of ribonucleotides to deoxyribonucleotides is carried out by an O_2-sensitive class III ribonucleotide reductase (RNR).[1] The class III RNRs that have been most extensively characterized, from bacteriophage T4 and its host *Escherichia coli*, use formate as the hydrogen donor for nucleotide reduction, oxidizing it to CO_2.[2] We recently reported a second subtype of class III RNR from the bacterium *Neisseria bacilliformis*, which uses the thioredoxin (TrxA) / thioredoxin reductase (TrxB) / NADPH system for nucleotide reduction.[3] Here we describe a third subtype from the methanogenic archaeon *Methanosarcina barkeri*, which uses a reduction system consisting of reduced ferredoxin (Fdx), a [4Fe4S] protein ferredoxin:disulfide reductase (FDR), and a conserved thioredoxin-like protein present in the RNR operon (NrdH), in conjunction with the Fdx-dependent anaerobic metabolism of this organism.

RNRs are essential enzymes present in nearly all cellular organisms and many viruses.[4,5] All RNRs characterized to date are structurally homologous and initiate radical-dependent nucleotide reduction via a transient thiyl radical on a conserved Cys residue on the top face of the ribose in the active site.[6,7] Apart from the class III RNR, there are a further two classes of RNR that differ in the cofactor used to generate this thiyl radical.[8] Class I RNRs use cofactors that require reduced metals (Fe, Mn, and Fe/Mn) and O_2 for their biogenesis, and are present only in aerobic organisms. Class II RNRs use adenosylcobalamin in an O_2-independent reaction and are present in both aerobes and anaerobes. Class III RNRs use an O_2-sensitive glycyl radical (G•) situated in the α protein (NrdD),[9] which is generated by a separate activating enzyme

(NrdG) via radical S-adenosylmethionine (SAM)-$[4Fe4S]^{1+}$ chemistry.[10] The class III RNRs are only found in facultative and obligate anaerobes.

The mechanism of nucleotide reduction has been most extensively studied in the class I and II RNRs, where nucleotide reduction proceeds with the concomitant generation of a disulfide between a pair of conserved Cys residues on the bottom face of the ribose in the active site.[11] Subsequent turnovers require the reduction of this disulfide by a redoxin protein (thioredoxin, glutaredoxin or NrdH).[12-14] The radical-dependent reduction mechanism involves acid / base catalysis by a conserved Glu residue that is essential for turnover in these enzymes.[7]

In contrast, the *E. coli* and bacteriophage T4 class III RNR (subtype NrdD1) use formate as a reductant and have a single Cys on the bottom face.[2,15] Reaction of NrdD1 with nucleotide results in the formation of a thiosulfuranyl radical involving a three-electron bond between this Cys and a conserved Met residue.[16] This radical is thought to be the oxidant of formate and is chemically competent for formation of the deoxynucleotide product.

A recent bioinformatics study of the RNR database (RNRdb) led us to identify a second class III RNR subtype (NrdD2) with distinct active-site residues, suggesting a different mechanism for nucleotide reduction.[3] In NrdD2 the active-site Met is not conserved, and the presence of a pair of bottom face Cys and Glu residues suggested chemistry similar to class I and II RNRs. In the crystal structures of *Thermotoga maritima* obtained independently by us and by Aurelius *et al.*[17], the active site "thiyl radical loop" is displaced from its expected position compared to other RNRs, and both we and Aurelius *et al.* have suggested that these structures may represent an inactive state of the enzyme. Nevertheless, a model constructed based on the crystal structure suggested that the Cys and Glu residues were in a suitable position to carry out the chemistry.[3] Biochemical evidence was provided by the recombinant NrdD2 from *N.*

bacilliformis, which catalyzed nucleotide reduction using the *E. coli* TrxA / TrxB / NADPH system as the electron source.

The type of NrdD present in an organism was observed to correlate with its anaerobic metabolism. The redoxin-dependent NrdD2 is present in diverse non-fermenting bacteria and non-methanogenic archaea, in conjunction with the nearly universal occurrence of redoxins. Among bacteria, the formate-dependent NrdD1 is present in fermenting bacteria where pyruvate-formate lyase catalyzes the conversion of pyruvate to formate and acetyl-coA,[18] providing a source of formate for nucleotide reduction.

Most methanogenic archaea contain either NrdD1 or a third subtype of class III RNR (NrdD3). Although NrdD3 is phylogenetically more closely related to methanogen NrdD1s, its active-site more closely resembles NrdD2, lacking the catalytic Met but containing a pair of bottom face Cys residues. Also, all NrdD3 operons contain a thioredoxin-like protein NrdH, suggesting redoxin-dependent nucleotide reduction. However, unlike other redoxin-dependent RNRs, NrdD3 lacks any active-site residue that could carry out acid / base catalysis, suggesting differences in the mechanism of nucleotide reduction.

Class I methanogens (Methanopyrales, Methanococcales, Methanobacteriales) and Rice cluster I use NrdD1. These organisms live under conditions with low partial pressure of H_2, and formate is thought to be generated reversibly by the F_{420}-dependent formate dehydrogenase. Formate is also used in purine biosynthesis[19] and as a substrate for methanogenesis.[20] Class II methanogens (Methanomicrobiales, Methanosarcinales) use NrdD3. Many of them lack formate dehydrogenase and use formyl-THF for purine biosynthesis. The source of electrons for nucleotide reduction by NrdD3 is unknown.

Although NrdD2 is the most widely distributed class III RNR, it is uncommon in methanogens. We speculate that this may be related to the unique electron carriers used in their energy metabolism: Fdx and coenzyme F_{420}, instead reduced pyridine nucleotides, the substrate of TrxB.[21]

To investigate the source of electrons for nucleotide reduction by NrdD3, we cloned and reconstituted the class III RNR system from *Methanosarcina barkeri*. We carried out a bioinformatics study suggesting that the reducing equivalents for this RNR system originate from Fdx, through the [4Fe4S] protein FDR. We demonstrate that the *M. barkeri* NrdD3 catalyzes the nucleotide reduction in vitro using the recombinant NrdH / FDR / Fdx system (**Figure 5.1**). The distribution of this form of class III RNR is discussed in relation to the different methanogenic pathways occurring in these organisms.

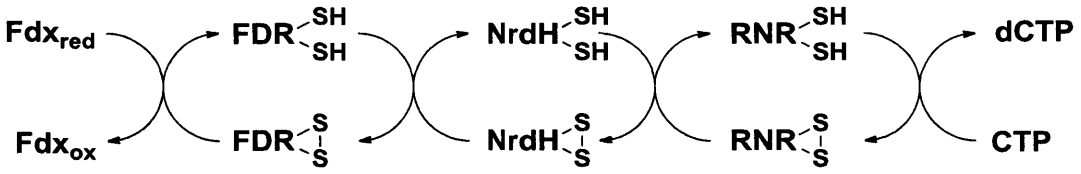


Figure 5.1 Model for Fdx-dependent nucleotide reduction by MbNrdD. Fdx provides the electron source for reduction of FDR, which reduces the redoxin NrdH, which in turn regenerates the active site Cys pair required for nucleotide reduction by NrdD.

5.2. Materials and methods

5.2.1. Materials and general methods

All chemical reagents were purchased from Sigma–Aldrich, unless otherwise indicated. Primers were purchased from Integrated DNA Technologies. UV-vis absorption spectroscopy was performed on an Agilent 8453 Diode Array spectrophotometer or a Varian Cary 3 UV-vis

spectrophotometer. Anaerobic procedures were carried out in a custom-designed MBraun glovebox at 30°C. All solutions and proteins were made anaerobic on a Schlenk line by 3 cycles of evacuation (5 min) followed by flushing with Ar gas (10 min) before being brought into the glovebox. Nucleotides and SAM were brought into the glovebox as lyophilized solids. *Pyrococcus furiosus* genomic DNA was purchased from ATCC. *M. barkeri* (str. Fusaro) genomic DNA was a gift from Prof. William W. Metcalf, University of Illinois.

5.2.2. Cloning of *M. barkeri* genes

The genes were amplified from genomic DNA by PCR using Q5 polymerase (NEB) and the primers (**Table 5.1**), and inserted into their respective linearized plasmids (**Table 5.1**) using a Gibson isothermal assembly kit (NEB) following the manufacturer's protocol.[22] All constructs were confirmed by DNA sequencing by Quintara Biosciences.

The genes for MbNrdH was inserted into pET28a (Novagen) linearized with NdeI and HindIII, to give the plasmid pET28a-MbNrdH. The pET28a plasmid contains an N-terminal His_6-tag followed by a thrombin cleavage site (MGSSHHHHHHSSGLVPRGSH-).

To increase yields and stability of some of the proteins, a new plasmid was constructed to enable expression of proteins with a fusion to *P. furiosus* maltose binding protein, as previously reported.[23] The gene for PfMBP was inserted into pSV272 linearized with SacI and BsaI, to give pPfMBP. The PfMBP fusion contains an N-terminal His_6-tag (MHHHHHHSSGG-).

The genes for MbNrdG and MbFDR1 were inserted into pPfMBP linearized with NdeI and EagI, to give plasmids pPfMBP-MbNrdG and pPfMBP-MbFDR1 respectively.

The gene for MbNrdD was inserted into pET24a (Novagen) linearized with NdeI and HindIII, to give the plasmid pET-MbNrdD. This construct consists of a deletion of the N-

terminal 36 non-conserved residues prior to the ATP cone domain, and installation of a His_4-tag that abrogated unusually tight binding to the TALON affinity resin, followed by a hydrophilic linker found to increase yields of soluble protein (MHHHHGSGSGSGSG-).

The MbNrdD(C219S) mutant was obtained by site-directed mutagenesis of pET24a-MbNrdD using the primer CAAGACCATTTAGCCAGGATTGGGATCTCC (mutation underlined), confirmed by sequencing.

Plasmid	Gene product	Primer	Sequence
pET24a	MbNrdD	Fw	ctttaagaaggagatatacatATGCATCACCATCAcgGCAGCGGAAGTGGAAGCggcAGCGGCT CTGACAACAAGCCAGT
		Rv	GGTGCTCGAGTGCGGCCGCAAGCTtACCTGAGTTCCCTTGGACTGTATCG
pPfMBP	MbNrdG	Fw	CCTGTATTTTCAGGGCGCCcatATGAAAGTAAACTACGCAGGGACTGTCC
		Rv	GCCCGTTTGATCTCGAGTGCGGCCGtCATACAGAGTCTGAGAGCACTTCG
pET28a	MbNrdH	Fw	GCCTGGTGCCGCGCGGCAGCcatATGGCAAAAATAATCATATATACAACGG
		Rv	GGTGCTCGAGTGCGGCCGCAAGCTtACAGGACTTCCTGGATTTTTTCAGGG
pPfMBP	MbFDR1	Fw	CCTGTATTTTCAGGGCGCCcatATGACTGATCATAATGAACTAAAACAG
		Rv	GCCCGTTTGATCTCGAGTGCGGCCGtTAATTCGCTTTGAAGAAAAGTTCGC
pSV272	PfMBP	Fw	caggtctcccatgCATCATCATCATCATCACAGCAGCGGCGGAAAAGTTGTTATTTGGC
		Rv	cagagctcccTCCTTGCATGTTGTTAAGG

Table 5.1 Primers used for amplification of genes by PCR.

5.2.3. Expression and purification of proteins

5.2.3.1. Protein expression

The plasmids were separately transformed into BL21 (DE3) codon plus (RIL) cells (Stratagene), grown on LB-agar plates with 50 μg/mL kanamycin (Kan) and 30 μg/mL chloramphenicol (Cm). A single colony was inoculated into 5 mL starter culture of LB (50 μg/mL Kan and 30 μg/mL Cm in all growths), grown at 37°C until saturated (12 h), and transferred into 200 mL of LB. For expression of MbNrdD, which contains a putative Zn binding site, 50 μM of $ZnSO_4$ was added to the medium.[24] The cultures were grown at 37°C with shaking at 200 rpm. At OD600 ~0.8, the temperature was decreased to 25°C (or 30°C for FDR1) and

IPTG (Promega) was added to a final concentration of 0.1 mM. After 12 h (or 4 h for FDR1), cells were harvested by centrifugation (4,000 × g, 10 min, 4°C). Typical yield was ~5 g of cell paste per L.

5.2.3.2. Standard procedure: used for purification of MbNrdG and MbFDR1

After harvesting, the cells were suspended in 5 mL of lysis buffer (50 mM Tris-HCl pH 8, 1 mM phenylmethylsulfonyl fluoride (PMSF), 0.2 mg/mL lysozyme, 0.03% Triton X, 1 uL of Benzonase (Novagen)). The cell suspension was frozen at -80°C, then thawed and incubated at room temperature for 40 min to allow for lysis. 15 mL of buffer A (20 mM Tris-HCl pH 7.5, 5 mM β-mercaptoethanol (BME)) containing 1.3% streptomycin sulfate was added. The mixture was shaken for an additional 10 min, and the precipitated DNA was removed by centrifugation (20,000 × g, 10 min, 4°C). Solid $(NH_4)_2SO_4$ was then added to 60% saturation. The solution was shaken for an additional 10 min, and the precipitated protein was isolated by centrifugation (20,000 × g, 10 min, 4°C), yielding a ~1 g cell pellet.

The pellet was dissolved in 20 mL of buffer B (20 mM Tris pH 7.5, 5 mM BME, 0.2 M KCl) and incubated with 2 mL of TALON resin (Clontech) with shaking for 30 min. The column was then packed (0.8 × 4 cm) and washed with 10 column volumes (CV) of buffer B. Protein was eluted with 5 CV of buffer B containing 150 mM imidazole. The eluted protein was precipitated with solid $(NH_4)_2SO_4$ to 60% saturation, and isolated by centrifugation (20,000 × g, 10 min, 4°C). The pellet was dissolved in 0.5 mL of buffer B and desalted using a Sephadex G-25 column (1.5 cm × 8.5 cm, 15 mL), pre-equilibrated with buffer C (20 mM Tris pH 7.5, 50 mM KCl, 5% glycerol, 1 mM DTT). The eluted protein was concentrated to ~200 μL by

ultrafiltration (Amicon YM-30), frozen in aliquots in liquid N_2 and stored at -80°C. For yields, extinction coefficients and SDS PAGE gels, see **Figure 5.7**.

5.2.3.3. Purification of MbNrdH

The lower molecular weight MbNrdH was purified by modification of the standard procedure described above. After harvesting, the cells were suspended in 5 mL of 20 mM Tris pH 7.5, followed by lysis by three freeze-thaw cycles. The subsequent steps were identical to those described above, except that BME was omitted from all buffers, and protein precipitation was carried out using $(NH_4)_2SO_4$ to 70% saturation. For yields, extinction coefficients and SDS PAGE gels, see **Figure 5.7**.

5.2.3.4. Purification of MbNrdD

MbNrdD required high concentrations of salt and / or glycerol for stability, and was purified by modification of the standard procedure described above. After lysis, streptomycin sulfate precipitation and ammonium sulfate precipitation, and binding to the TALON resin, the column was washed with 5 CV of buffer B, then 10 CV of buffer B containing 2 M NaCl to remove bound DNA, then 5 CV of buffer B. Protein was eluted with 5 CV of buffer D (20 mM Tris pH 7.5, 300 mM KCl, 20% glycerol) containing 150 mM imidazole and 5 mM BME. The eluted protein was precipitated with solid $(NH_4)_2SO_4$ to 60% saturation, and isolated by centrifugation (20,000 × g, 10 min, 4°C). The pellet was dissolved in 0.5 mL of buffer D and desalted using a Sephadex G-25 column (1.5 cm × 8.5 cm, 15 mL), pre-equilibrated with buffer D containing 1 mM DTT. The eluted protein was concentrated to ~300 μM by ultrafiltration

(Amicon YM-30), frozen in aliquots in liquid N_2 and stored at -80°C. For yields, extinction coefficients and SDS PAGE gels, see **Figure 5.7**.

5.2.4. Reconstitution of [4Fe4S] clusters of MbNrdG and FDR1

The procedure was carried out in a glovebox. Solutions of Na_2S and of $Fe(NH_4)_2(SO_4)_2$ in water (100 mM) were freshly prepared in the glovebox. Solutions of the respective proteins (~200 μM) were made anaerobic on a Schlenk line and brought into the glovebox. A solution of DTT (dithiothreitol, 1 M) was added to 10 mM, followed by ordered addition of the solution of Na_2S (5 eq.) and $Fe(NH_4)_2(SO_4)_2$ (5 eq.). The mixture was incubated for 12 h at 4°C. EDTA (ethylenediaminetetraacetic acid, 5 eq.) was then added and the solution was desalted by repeated dilution with 20 mM Tris pH 7.5 and concentration by ultrafiltration (Amicon YM-30). The final material contained 3.3 and 3.6 atoms of Fe per MbNrdG and FDR1 peptide respectively determined by the ferrozine assay.[25]

5.2.5. Generation of the MbNrdD G•

In a 1.5 mL polypropylene Eppendorf tube, a 50 μL mixture of MbNrdD (40 μM), MbNrdG (40 μM), SAM (1 mM), Bicine potassium salt pH 7.5 (10 mM) and diaminoacridine hydrochloride (10 μM), KCl (300 mM), glycerol (20%) was placed 1 m away from a fluorescent lamp in the glovebox at 30°C for 1 h. This mixture was used directly for assays without further purification. For inspection by X-band EPR spectroscopy, the reaction was carried out on a 200 μL scale in an EPR tube sealed with a rubber stopper. The solution was quenched in liquid N_2 immediately after removal from the glovebox. The amount of G• in the solution was determined by comparing the EPR signal intensity to that of a $CuSO_4$ standard.[26] A typical yield of 0.1 radicals per MbNrdD polypeptide was reproducibly obtained.

5.2.6. X-band EPR spectroscopy

Continuous wave X-band EPR spectra were recorded at 77 K in the MIT Department of Chemistry Instrumentation Facility on a Bruker ESP-300 X-band spectrometer equipped with a quartz finger Dewar filled with liquid N_2. Experimental conditions were as follows: microwave frequency, 9.45 GHz; modulation amplitude, 0.15 mT; modulation frequency, 100 kHz; time constant, 5.12 ms; scan time, 41.9 s; microwave power, 20 μW.

5.2.7. Activity assay for dCTP formation by MbNrdD using DTT as the electron source

The assay mixture contained in 50 μL: MbNrdD (10 μM, ~1 μM G•), MbNrdH (10 μM), ATP (0.5 mM), 5-[^{3}H]-CTP (0.5 mM, 3730 cpm/nmol), DTT (2 mM) in assay buffer (50 mM Tris pH 7.5, 200 mM KCl, 10 mM $MgSO_4$) and was incubated at 30°C. Aliquots (10 μL) were removed at 30 s intervals and quenched with 2% perchloric acid (10 μL). Subsequent to neutralization and removal of the phosphates using calf intestine alkaline phosphatase (Roche), dCTP formation was analyzed by the method of Steeper and Steuart.[27] One unit of activity is equivalent to 1 nmol of dCTP/min. The specific activity of MbNrdD is 55 U/mg NrdD protein (~0.83 s^{-1} per G•).

5.2.8. Activity assay for dCTP formation by MbNrdD using Ti(III) citrate as the electron source

Reproducible activity was obtained by initiating the reaction in the following manner: to a 32.5 μL mixture containing ATP, 5-[^{3}H]-CTP, Ti(III) citrate and assay buffer was added 5 μL of MbFDR1, followed immediately by addition of a 12.5 μL mixture containing MbNrdD and MbNrdH. The final assay mixture contained in 50 μL: MbNrdD (10 μM, ~1 μM G•), MbNrdH

(10 μM), MbFDR1 (10 μM), ATP (0.5 mM), 5-[^{3}H]-CTP (0.5 mM, 3730 cpm/nmol), Ti(III) citrate (1 mM) in assay buffer and was incubated at 30°C. Aliquots (10 μL) were removed at 30 s intervals and quenched with 2% perchloric acid (10 μL). Workup is as described above. The specific activity of MbNrdD under these conditions was 41 U/mg NrdD protein (~0.61 s^{-1} per G•).

5.2.9. Stoichiometry of Ti(III) citrate consumption and dCTP production

The assay mixture was divided into 10 μL aliquots containing: MbNrdD (10 μM, ~1 μM G•), MbNrdH (2 μM), MbFDR1 (2 μM), dATP (0.1 mM), 5-[^{3}H]-CTP (0.5 mM, 3730 cpm/nmol) and varying amounts of Ti(III) citrate (0-0.6 mM) in assay buffer. The reaction was initiated as described above and incubated at 30°C for 2 h to allow for complete consumption of Ti(III). Workup of the samples was carried out as described above to quantify dCTP formed.

5.2.10. Assay for cytosine (Cyt) formation by MbNrdD(C219S)

The assay mixture contained in 50 μL: MbNrdD(C219S) (10 μM, ~1 μM G•), ATP (0.5 mM), 5-[^{3}H]-CTP (0.5 mM, 3730 cpm/nmol) in assay buffer (50 mM Tris pH 7.5, 200 mM KCl, 10 mM $MgSO_4$) and was incubated at 30°C. Aliquots (10 μL) were removed at 2, 4 and 8 min and quenched with 2% perchloric acid (10 μL). Formation of Cyt was analyzed by passing the mixture through an anion exchange column to remove the nucleoside triphosphates as previously described.[16]

5.3. Results

5.3.1. Generation of the MbNrdD G•.

Initial attempts to purify MbNrdD were confounded by the instability of the protein, which aggregated over time and bound tightly to various chromatographic resins and to DNA. We later found that this instability and non-specific binding could be overcome by addition of 200-300 mM of KCl and / or 20% glycerol to the buffers used for chromatography, storage and assay buffers. To generate active MbNrdD for biochemical studies, we incubated MbNrdD with MbNrdG and SAM in the presence of the diaminoacridine / bicine photoreduction system, resulting in the generation of a radical with a characteristic doublet EPR signal shown in **Figure 5.2**, consistent with its assignment as G•. MbNrdD stored in buffer containing 20 mM Tris pH 7.5, 300 mM KCl and 20% glycerol was stable at 4°C in the glovebox for several weeks, during which a yield of ~0.1 G• per NrdD peptide was reproducibly obtained by EPR.

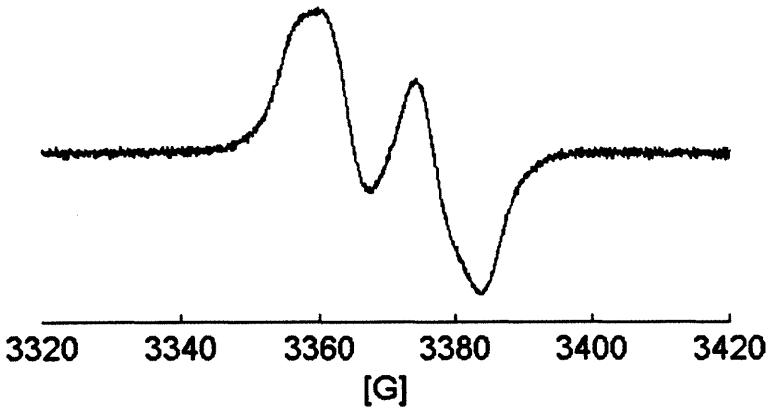


Figure 5.2. Spectrum of the MbNrdD G• (40 μM NrdD peptide).

5.3.2. NrdH is required for nucleotide reduction by MbNrdD using DTT as an electron source

In the previously characterized redoxin-dependent *N. bacilliformis* NrdD2,[3] nucleotide reduction was proposed to occur with the concomitant oxidation of a pair of conserved active-site Cys residues, through chemistry analogous to that proposed for the class I and II RNRs.[7] This pair of Cys residues is also present in the same location in the primary sequence of NrdD3, suggesting similar redoxin-dependent chemistry. All NrdD3s that we have examined occur adjacent to a thioredoxin-like protein NrdH, making it a likely candidate for the electron donor for NrdD3.

To establish whether NrdH can support nucleotide reduction by NrdD3, we carried out assays using MbNrdD (0.1 G•/α), MbNrdH and DTT as an electron source, and the results are summarized in **Table 5.2**. MbNrdD was active for reduction of CTP to dCTP with ATP as an effector, but much less active for CDP reduction (~10% of the activity for CTP reduction, **Table 5.2**). This promiscuous activity was also previously observed for *N. bacilliformis* NrdD2.[3] Catalytic activity was dependent on the presence of G• and NrdH (**Table 5.2**). As with *N. bacilliformis* NrdD2, DTT alone fails to efficiently reduce NrdD in the absence of NrdH. Formate failed to produce any dCTP. Thus the experiments demonstrate that MbNrdH is required for nucleotide reduction by MbNrdD with DTT as the electron source.

CTP-reduction activity was enhanced 2.5-fold by ATP and 2-fold by dATP (**Table 5.2**). MbNrdD has an annotated N-terminal ATP cone domain that controls the activity of many RNRs by binding the activator (ATP) or the inactivator (dATP). However, in MbNrdD, this domain lacks several key residues for nucleotide binding and is predicted to be inactive. Thus in

MbNrdD, both ATP and dATP would be predicted to bind to the specificity site and activate nucleotide reduction. The degree of activation is smaller than observed in *E. coli* NrdD (5-fold enhancement of CTP reduction by ATP).[28]

TrxB sequences are not highly conserved among methanogens containing NrdD3, and it was noted that the *Methanosarcina* TrxBs are more closely related to that of sulfate-reducing bacteria than to that of other methanogens.[29] Since bacterial and archaeal TrxBs are known to be highly specific for their cognate thioredoxins,[30] we hypothesize that the reductant for NrdH in these organisms is not TrxB.

Reaction conditions	Activity (U/mg)
Complete (CTP, ATP)	55 ± 3
- SAM (no G• formed)	N.D.
-NrdH	N.D.
-DTT	N.D.
-DTT + formate (10 mM)	N.D.
-ATP	22 ± 1
-ATP, + dATP (0.1 mM)	47 ± 2
-CTP, + CDP (0.5 mM)	6 ± 1

Table 5.2 Requirements for dCTP formation by MbNrdD using DTT as the electron source. N.D. = activity not detected, < 10 turnovers per G• over 20 min.

5.3.3. Identification of FDR as a potential reductant for NrdH

We had previously noted the correlation between the energy metabolism of an organism with the subtype of class III it contains.[3] The diverse mechanisms for energy conservation in methanogens have been extensively studied and are becoming increasingly well understood,[21] and therefore we examined the distribution of class III RNRs among the metabolically diverse Methanosarcinales for clues regarding the source of electrons for reduction of NrdH (**Table 5.3**).

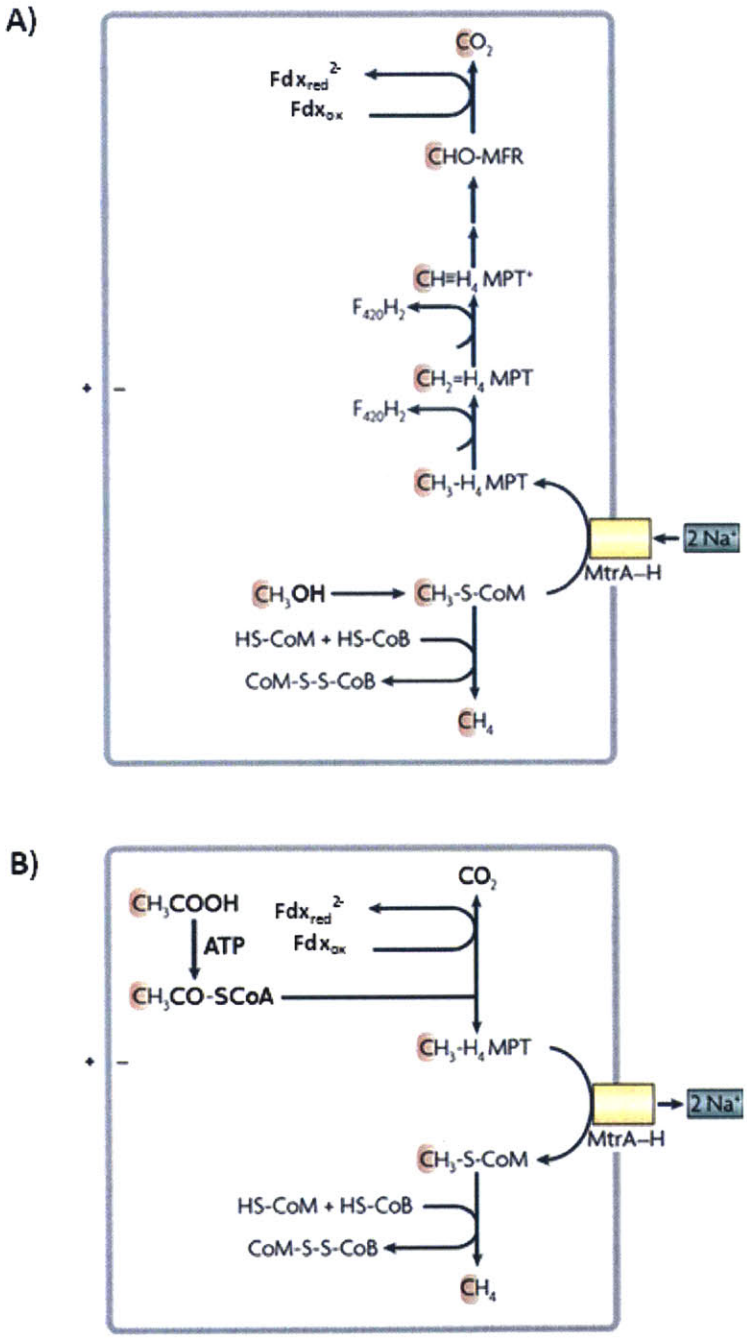


Figure 5.3 Pathways for methanogenesis from methanol or acetate. A) Methylotrophic methanogenesis generates 4 electrons in the form of $F_{420}H_2$ and 2 electrons in the form of reduced Fdx. B) Aceticlastic methanogenesis produces 2 electrons in the form of reduced Fdx.

M. barkeri and other organisms in the order Methanosarcinales generate methane from from methanol or methylamines (methylotrophic methanogenesis, **Figure 5.3A**), or from acetate (aceticlastic methanogenesis, **Figure 5.3B**). These pathways differ in the electron carriers used in energy conservation: the breakdown of each molecule of acetate produces 2 electrons in the form of reduced Fdx from the carbon monoxide dehydrogenase-acetyl coA synthase system (CODH-ACS); while the oxidation of each methyl group produces 4 electrons in the form of $F_{420}H_2$ from the oxidation of methyl- and methylene tetrahydromethanopterin (H_4MPT) and 2 electrons in the form of reduced Fdx from formylmethanofuran dehydrogenase (FMF-DH).[31]

Most members, including aceticlastic *Methanosaeta* and *Methanosarcina* species have NrdD3 and a class II RNR (**Table 5.3**). However, several obligate methylotrophs (*Methanococcoides*, *Methanosalsum*, *Methanolobus*) contain NrdD3 and also NrdD2 but no class II RNR (**Table 5.3**). Because of the precedent of class II RNRs and NrdD2 using the TrxA / TrxB / NADPH system for nucleotide reduction, and because NADPH can be produced by F_{420}-NADP reductase from $F_{420}H_2$ generated in the methylotrophic pathway,[32] the non-orthologous replacement of the class II RNR with NrdD2 may suggest that they fulfill the common function of NADPH-dependent nucleotide reduction. Therefore we hypothesized that NrdD3 may use an alternative reductant: reduced Fdx.

We also carried out a BLAST search for NrdH within the genomes of methanogens and found that is most closely related to the thioredoxin-like protein Mrx of unknown function, which occurs adjacent to a [4Fe4S] protein FDR, which has recently been structurally and biochemically characterized.[33] Proteins of this family previously characterized from plants and cyanobacteria catalyze the reduction of thioredoxin using electrons from reduced ferredoxin.[34]

Since FDR and TrxB are the only two cytosolic proteins with chemical precedent for reduction of Trx-like proteins, we hypothesized that FDR is the direct reductant of NrdH.

FDRs have been classified into several subtypes according to their primary sequence,[35] and two (which we name FDR1 and FDR2) are broadly distributed in Methanosarcinales (**Table 5.3**). Methylotrophs contain FDR1, which is located adjacent to FMF-DH, a source of reduced Fdx. Aceticlastic methanogens contain FDR2, which has a C-terminal rubredoxin fusion and occurs adjacent to Mrx.[33] Metabolically versatile *Methanosarcina* species contain both. Because of its distribution, we hypothesize that the source of reduced Fdx for FDR2 is CODH.

Organism	Metabolism	RNRs	FDRs
Methanosarcina	Versatile	NrdD3, NrdJ	FDR1 FDR2
Methanococcoides *Methanosalsum* *Methanolobus*	Methylotrophic	NrdD3, NrdD2	FDR1
Methanosaeta	Aceticlastic	NrdD3, NrdJ	FDR2

Table 5.3 Occurrence of RNRs and FDRs in Methanosarcinales.

5.3.4. FDR1 supports nucleotide reduction by MbNrdD using Ti(III) citrate as an electron source

To establish whether a FDR can provide electrons for nucleotide reduction, we carried out assays using MbNrdD (0.1 G•/α), NrdH, FDR1 and Ti(III) citrate as an electron source (see **Figure 5.1**), and the results are summarized in **Table 5.4**. Catalytic activity for nucleotide reduction with the Ti(III) / FDR1 / NrdH system was comparable to the DTT / NrdH system, and was dependent on the presence of both NrdH and FDR1 (**Table 5.4**). The number of dCTPs formed per Ti(III) in the reaction mixture is 0.45 (**Figure 5.4**), suggesting a 1:2 stoichiometry, in

agreement with the proposal that the reducing equivalents are provided by the Ti(III) / FDR1 / NrdH system (**Figure 5.1**).

Reaction conditions	Activity (U/mg)
Complete (CTP, ATP)	41 ± 1
-NrdH	N.D.
-FDR1	N.D.

Table 5.4 Requirements for dCTP formation by MbNrdD using Ti(III) citrate as the electron source. N.D. = activity not detected, < 10 turnovers per G• over 20 min.

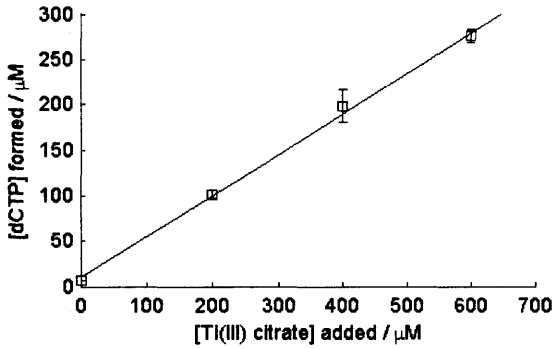


Figure 5.4 Amount of 5-[^{3}H]-dCTP formed after incubation of MbNrdD with 5-[^{3}H]-CTP, dATP, NrdH, FDR1 and limiting amounts of Ti(III) citrate at 30°C for 2 h. Stoichiometry of dCTP produced per Ti(III) added is 0.45. The concentration of MbNrdD G• in the reaction is ~1 μM.

5.3.5. MbNrdD(C219S) mutant generates Cyt but not dCTP

In *E. coli* NrdA (EcNrdA), which has served as a model system for investigation of the mechanism of nucleotide reduction, C225 is believed to be involved in protonation of the 2′-OH leaving group and formation of a disulfide concomitant with nucleotide reduction (**Chapter 2 Introduction**). As a preliminary investigation into the role of the bottom face cysteine, the C219S mutant was constructed. Reaction with substrate (CTP) and allosteric effector (ATP) leads to time-dependent formation of Cyt, a known breakdown product of the 3′-keto-dC, with an initial rate of 0.046 s^{-1} (**Figure 5.5**). dCTP is not detected and the reaction is unaffected by the presence of NrdH and DTT.

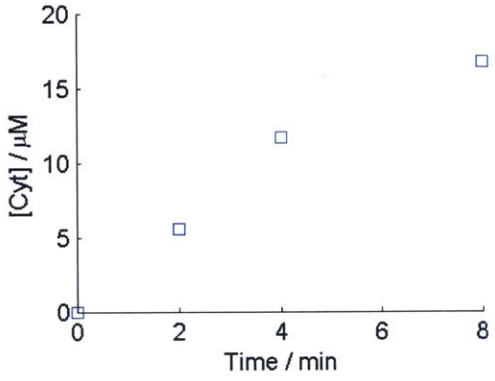


Figure 5.5 Time-dependent formation of 5-[^{3}H]-Cyt upon incubation of MbNrdD(C219S) with 5-[^{3}H]-CTP and ATP. The initial rate is 0.046 s^{-1}. The concentration of MbNrdD G• in the reaction is ~1 μM.

5.4. Discussion

The role of low potential Fdx's in anaerobic energy conservation and in driving anaerobic biosynthetic reactions is becoming increasingly appreciated.[36] Our experiments with the *M. barkeri*, class III RNR suggest another role for Fdx providing the electron source for ribonucleotide reduction. Although we have yet to demonstrate that Fdx is competent for providing electrons for nucleotide reduction by MbNrdD, the ability of Fdx to provide electrons for glutathione disulfide reduction by FDR2 was recently reported.[33]

The cloning of this RNR system was motivated by the observation that the type of class III RNR correlates with the energy metabolism in methanogens, which we now rationalize based on the redox gradient used by the organism for energy conservation. In all methanogens, the CoM-CoB heterodisulfide, produced concomitantly with CH_4 generation, serves as the electron sink in redox processes involved in energy metabolism. However, the electron source differs between organisms. Class I methanogens containing the formate-dependent NrdD1 utilize

shallow redox gradients, with H_2 or formate as electron carriers.[20,37] Class II methanogens containing the putatively Fdx-dependent NrdD3 utilize steep redox gradients, with Fdx playing a larger role as an electron carrier in pathways involving cytochromes. For example, Fdx is thought to be the sole cytosolic electron carrier for aceticlastic methanogenesis.[31,38]

Apart from the unique electron source for nucleotide reduction, the configuration of active site residues in MbNrdD is also remarkable. In EcNrdA, which has served as a model system for investigation of the mechanism of nucleotide reduction, E441 is believed to play a role in both the first and second half reactions of RNR (**Chapter 2 Introduction**). Analysis of the MbNrdD sequence shows a lack of conserved acid / base residues in the active site, suggesting that the mechanism may be different. In the first half reaction of EcNrdA, the first irreversible step is thought to be elimination of H_2O, catalyzed by the deprotonation of the 3′-OH by E441 and protonation of the 2′-OH by C225, forming a neutral ketyl radical (**Figure 5.6A**). In MbNrdD, a possible mechanism involves the reversible formation of a cation radical intermediate, catalyzed by the protonation of the 2′-OH by C219 (**Figure 5.6B**). Mutation of C219 in the MbNrdD(C219S) mutant is expected to make the elimination of OH^- leaving group much less favorable, slowing down the subsequent irreversible reaction forming 3′-keto-dC (**Figure 5.6C**). This is one possible explanation for the slow rate of 3′-keto-dC formation by MbNrdD(C219S), which is 5% the rate of dCTP formation by the wild type MbNrdD, although more thorough characterization of this mutant is required.

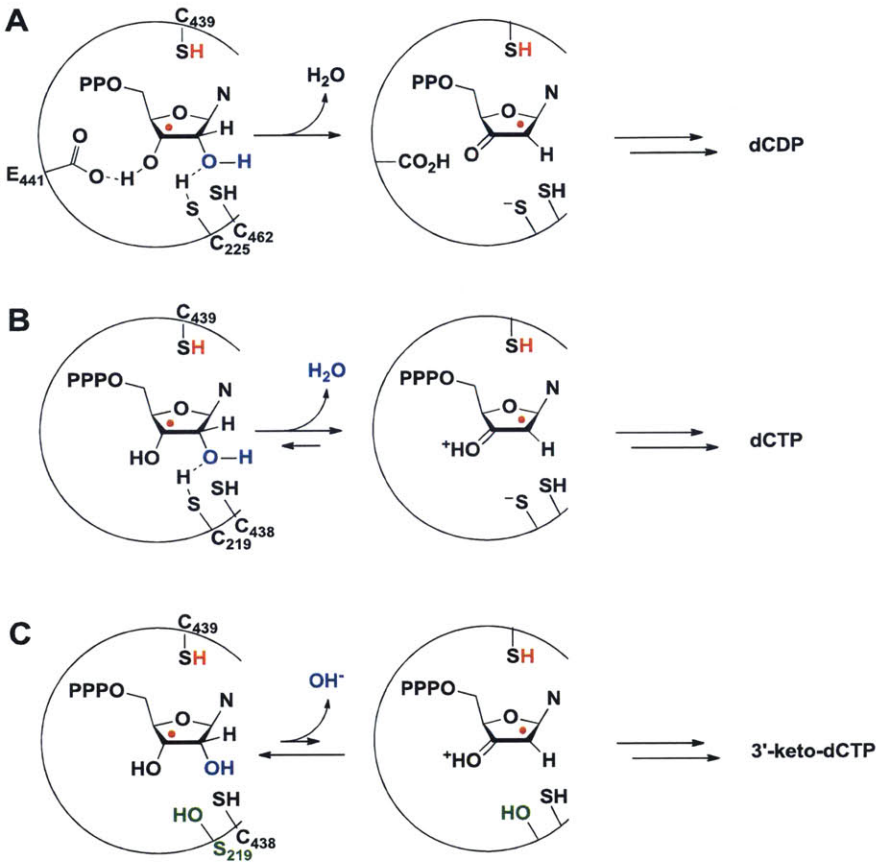


Figure 5.6 Possible mechanisms for elimination of the 3′-OH in A) EcNrdA, B) MbNrdD and C) MbNrdD(C219S).

5.5. Acknowledgements

We are grateful to Prof. William W. Metcalf for the gift of *M. barkeri* genomic DNA.

5.6. References

(1) Reichard, P. *J. Biol. Chem.* **1993**, *268*, 8383.

(2) Mulliez, E.; Ollagnier, S.; Fontecave, M.; Eliasson, R.; Reichard, P. *Proc. Natl. Acad. Sci. U. S. A.* **1995**, *92*, 8759.

(3) Wei, Y.; Funk, M. A.; Rosado, L. A.; Baek, J.; Drennan, C. L.; Stubbe, J. *Proceedings of the National Academy of Sciences* **2014**, *111*, E3756.

(4) Nordlund, P.; Reichard, P. *Annu. Rev. Biochem.* **2006**, *75*, 681.

(5) Hofer, A.; Crona, M.; Logan, D. T.; Sjöberg, B.-M. *Crit. Rev. Biochem. Mol. Biol.* **2012**, *47*, 50.

(6) Licht, S.; Gerfen, G. J.; Stubbe, J. *Science* **1996**, *271*, 477.

(7) Licht, S.; Stubbe, J. In *Comprehensive Natural Products Chemistry*; Barton, S., Nakanishi, K., Meth-Cohn, O., Poulter, C., Eds.; Elsevier Science: New York, 1999; Vol. 5, p 163.

(8) Stubbe, J. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 2723.

(9) Sun, X.; Ollagnier, S.; Schmidt, P. P.; Atta, M.; Mulliez, E.; Lepape, L.; Eliasson, R.; Graslund, A.; Fontecave, M.; Reichard, P.; Sjöberg, B. M. *J. Biol. Chem.* **1996**, *271*, 6827.

(10) Gambarelli, S.; Luttringer, F.; Padovani, D.; Mulliez, E.; Fontecave, M. *ChemBioChem* **2005**, *6*, 1960.

(11) Mao, S. S.; Holler, T. P.; Yu, G. X.; Bollinger, J. M., Jr.; Booker, S.; Johnston, M. I.; Stubbe, J. *Biochemistry* **1992**, *31*, 9733.

(12) Holmgren, A. *J. Biol. Chem.* **1979**, *254*, 3672.

(13) Jordan, A.; Åslund, F.; Pontis, E.; Reichard, P.; Holmgren, A. *J. Biol. Chem.* **1997**, *272*, 18044.

(14) Avval, F. Z.; Holmgren, A. *J. Biol. Chem.* **2009**, *284*, 8233.

(15) Logan, D. T.; Andersson, J.; Sjöberg, B.-M.; Nordlund, P. *Science* **1999**, *283*, 1499.

(16) Wei, Y.; Mathies, G.; Yokoyama, K.; Chen, J.; Griffin, R. G.; Stubbe, J. *J. Am. Chem. Soc.* **2014**, *136*, 9001.

(17) Aurelius, O.; Johansson, R.; Bågenholm, V.; Lundin, D.; Tholander, F.; Balhuizen, A.; Beck, T.; Sahlin, M.; Sjöberg, B.-M.; Mulliez, E. *PLoS One* **2015**, *10*, e0128199.

(18) Knappe, J.; Sawers, G. *FEMS Microbiol. Lett.* **1990**, *75*, 383.

(19) White, R. H. *J. Bacteriol.* **1997**, *179*, 3374.

(20) Costa, K. C.; Wong, P. M.; Wang, T.; Lie, T. J.; Dodsworth, J. A.; Swanson, I.; Burn, J. A.; Hackett, M.; Leigh, J. A. *Proceedings of the National Academy of Sciences* **2010**, *107*, 11050.

(21) Thauer, R. K.; Kaster, A.-K.; Seedorf, H.; Buckel, W.; Hedderich, R. *Nat. Rev. Microbiol.* **2008**, *6*, 579.

(22) Gibson, D. G.; Young, L.; Chuang, R.-Y.; Venter, J. C.; Hutchison, C. A.; Smith, H. O. *Nat Meth* **2009**, *6*, 343.

(23) Fox, J. D.; Routzahn, K. M.; Bucher, M. H.; Waugh, D. S. *FEBS Lett.* **2003**, *537*, 53.

(24) Luttringer, F.; Mulliez, E.; Dublet, B.; Lemaire, D.; Fontecave, M. *J. Biol. Inorg. Chem.* **2009**, *14*, 923.

(25) Fish, W. *Methods Enzymol.* **1988**, *158*, 357.

(26) Malmström, B. G.; Reinhammar, B.; Vänngård, T. *Biochim. Biophys. Acta (BBA)-Bioenergetics* **1970**, *205*, 48.

(27) Steeper, J.; Steuart, C. *Anal. Biochem.* **1970**, *34*, 123.

(28) Eliasson, R.; Pontis, E.; Sun, X.; Reichard, P. *J. Biol. Chem.* **1994**, *269*, 26052.

(29) McCarver, A. C.; Lessner, D. J. *FEBS J.* **2014**, *281*, 4598.

(30) Arnér, E. S.; Holmgren, A. *Eur. J. Biochem.* **2000**, *267*, 6102.

(31) Ferry, J. G. *FEMS Microbiol. Rev.* **1999**, *23*, 13.

(32) Eirich, L. D.; Dugger, R. S. *Biochimica et Biophysica Acta (BBA)-General Subjects* **1984**, *802*, 454.

(33) Kumar, A. K.; Kumar, R. S. S.; Yennawar, N. H.; Yennawar, H.; Ferry, J. G. *Biochemistry* **2015**.

(34) Staples, C. R.; Ameyibor, E.; Fu, W.; Gardet-Salvi, L.; Stritt-Etter, A.-L.; Schürmann, P.; Knaff, D. B.; Johnson, M. K. *Biochemistry* **1996**, *35*, 11425.

(35) Balsera, M.; Uberegui, E.; Susanti, D.; Schmitz, R. A.; Mukhopadhyay, B.; Schürmann, P.; Buchanan, B. B. *Planta* **2013**, *237*, 619.

(36) Kaster, A.-K.; Moll, J.; Parey, K.; Thauer, R. K. *Proceedings of the National Academy of Sciences* **2011**, *108*, 2981.

(37) Thauer, R. K.; Kaster, A.-K.; Goenrich, M.; Schick, M.; Hiromoto, T.; Shima, S. *Annu. Rev. Biochem.* **2010**, *79*, 507.

(38) Welte, C.; Deppenmeier, U. *Biochimica et Biophysica Acta (BBA)-Bioenergetics* **2014**, *1837*, 1130.

5.7. Appendix

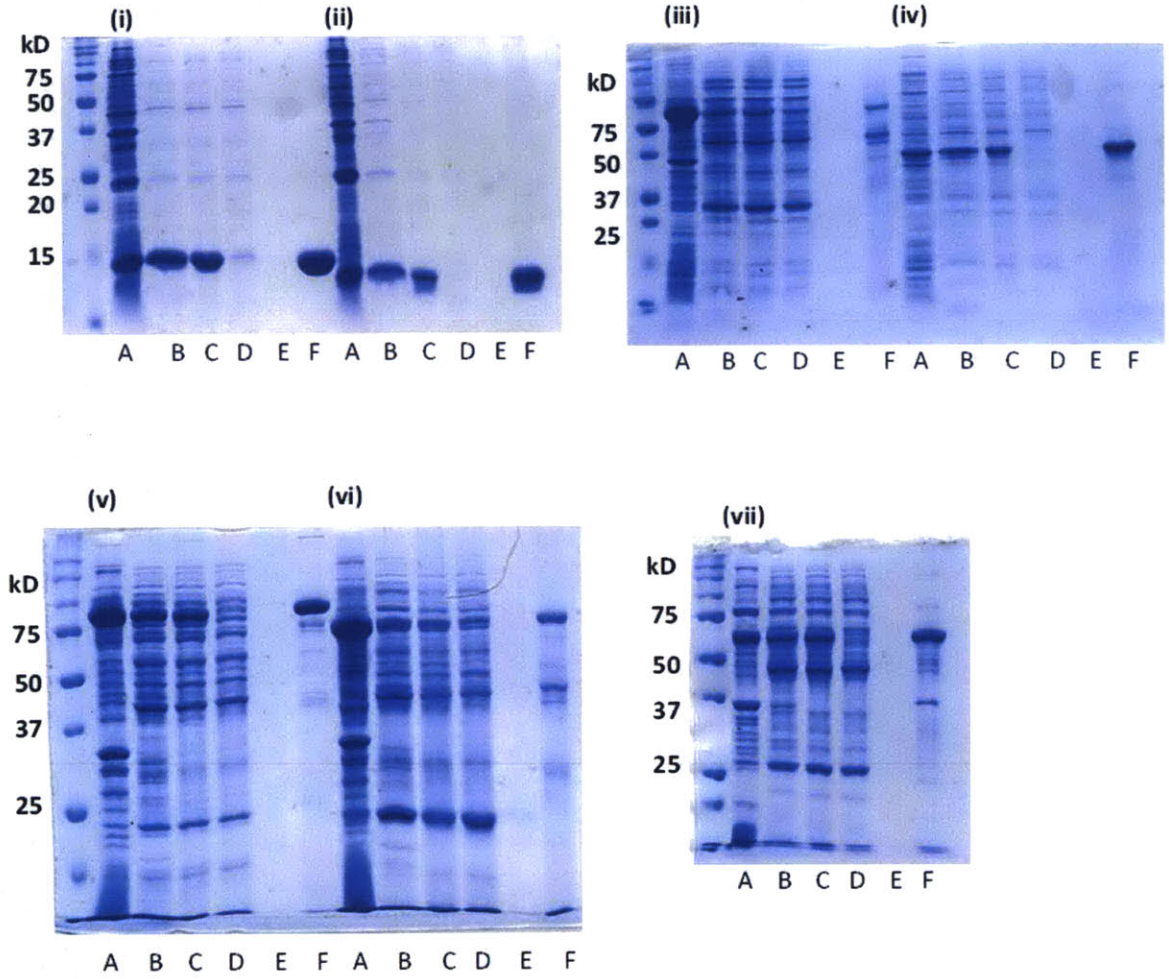


Figure 5.7 Purification of recombinant *M. barkeri* proteins (SDS/PAGE 10% gels, molecular weight, extinction coefficient at 280 nm and yield in mg / g cells is given). (i) MbNrdH (11.5 kD, ε_{280} = 1615 $M^{-1}cm^{-1}$, 20 mg / g); (ii) MbTrxA (11 kD, ε_{280} = 15595 $M^{-1}cm^{-1}$, 20 mg / g); (iii) PfMBP-MbFDR2 (66.4 kD, ε_{280} = 113680 $M^{-1}cm^{-1}$, 2 mg / g); (iv) MbTrxB (33.5 kD, ε_{280} = 18910 $M^{-1}cm^{-1}$, 10 mg / g); (v) MbNrdD (90.5 kD, ε_{280} = 100220 $M^{-1}cm^{-1}$, 15 mg / g); (vi) PfMBP-MbNrdG (76.2 kD, ε_{280} = 106120 $M^{-1}cm^{-1}$, 10 mg / g); (vii) PfMBP-MbFTR1 (58 kD, ε_{280} = 102220 $M^{-1}cm^{-1}$, 5 mg / g). Lanes: (A) streptomycin sulfate precipitate; (B) streptomycin sulfate supernatant; (C) ammonium sulfate precipitate; (D) TALON column flow-through; (E) TALON column wash; and (F) TALON column eluted protein.