Ribonucleotide Reduction and Redox Regulation in Archaea: Surprising Twists on a Common Theme

By

Hector Hugo Hernandez

B.A., Chemistry (2000)

University of South Florida

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

June 2008

Signature of Author _____
Department of Chemistry
June 10, 2008

Certified By _____
Catherine L. Drennan
Associate Professor of Chemistry and Biology
Thesis Supervisor

Accepted By _____
Robert W. Field
Chairman, Department Committee on Graduate Students

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

Professor John M. Essigmann

Committee Chairman

Professor Catherine L. Drennan

Thesis Supervisor

Professor Stuart S. Licht

Committee Member

Ribonucleotide Reduction and Redox Regulation in Archaea: Surprising Twists on a Common Theme

By
Hector Hugo Hernandez

Submitted to the Department of Chemistry on June 10, 2008 in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in Biological Chemistry

ABSTRACT

To study the nucleotide reduction system in Archaea, we have expressed the *Archaeoglobus fulgidus* gene encoding the adenosylcobalamin (AdoCbl) dependent ribonucleotide reductase (*afu*RNR) protein. Initial characterization of the recombinant *afu*RNR indicates that it reduces CDP to dCDP in an AdoCbl dependent manner. One astonishing finding is the capacity of *afu*RNR to reduce both diphosphate and triphosphate nucleosides to their respective deoxynucleosides. This is the first instance of a RNR displaying dual substrate selectivity.

Our investigation of redox regulation focused on thioredoxin reductase (TrR) and thioredoxin (Tr). In addition to providing the reducing equivalents for many RNRs, TrRs are involved in maintaining the intracellular redox environment. This process typically involves the transfer of electrons from NADPH to Trs. Here the characterization of a putative TrR (Ta0984) and Tr (Ta0866) from *Thermoplasma acidophilum* are presented. To explore the apparent inability of *ta*TrR to use NADPH or NADH as a reductant, we carried out a complete electrochemical characterization, which ruled out redox potential as the source of this non-reactivity. A 2.35 Å resolution structure of *ta*TrR, also presented here, shows that despite the overall structural similarity to the well-characterized TrR from *E. coli*, the "NADPH binding pocket" is not conserved. *E. coli* TrR residues implicated in NADPH binding, H175, R176, R177, and R181 have been substituted with E185, M186, Y187, and M191 in the *ta* protein. Thus, we have identified a Tr and TrR protein system from *T. acidophilum* for which the TrR shares overall structural and redox properties with other TrRs, but lacks the appropriate binding motif to use the standard NADPH reductant. Our discovery of a TrR that does not use NADPH provides a new twist in redox regulation.

Further, four *A. fulgidus* genes encoding four Trs (*afu*Tr 1, *afu*Tr 2, *afu*Tr 3, and *afu*Tr 4) proteins have been identified and their proteins expressed. Electrochemical characterization via protein-film voltammetry shows that the *afu*Trs reduction potentials are -32 mV for *afu*Tr 1, -301 mV for *afu*Tr 2, -287 mV for *afu*Tr 3, and -309 mV for *afu*Tr 4. This set of reduction potentials found in this Archaeon represents the greatest range in a specific organism.

Thesis Supervisor: Catherine L. Drennan

Title: Associate Professor of Chemistry and Biology

ACKNOWLEDGEMENTS

I would like to thank many people for their contributions to my tenure as a graduate student, and their help towards successful completion of the work described here. When I arrived in Cambridge in June 2000, I did not know what to expect, but I was sure I was in for the time of my life. I felt I had, after a much crooked road, finally arrived home. I still feel the thrill I felt when I walked under the Building 7 dome at 77 Massachusetts Avenue. Now, eight years later, the time has come to close this chapter of my life and begin a new.

First, I want to thank my advisor, Professor Cathy Drennan, for providing an environment in which I have been able to grow as a researcher and scientist. Her guidance, assistance, and advice were invaluable, and have made this work possible. I have gained a deeper appreciation for structural biology and broadened my interests in science largely due to her encouragement. Cathy supported me through both exhilarating and ominous times in my life. She showed me how exciting science is and even more important, how exciting it can be to be able to share it with others.

Most of all, I want to thank her for giving me the liberty to explore life as a graduate student to the fullest of my capabilities. It was because of her encouragement, and understanding that I have matured as a scientist and a scholar, gaining invaluable insight into the workings of an academic and research institution.

I also want to thank Professor Stuart Licht for his mentoring and help. I am thankful for the conversation we have had over the years regarding my projects and the perplexing results that I obtained. On more than one occasion he has provided insightful comments and opinions which allowed me to develop and sharpen my skills as an observer of scientific phenomena.

To Professor John Essigmann, I can not thank you enough for your friendship and encouragement over these last eight years. You have been my scientific father. Now the boy is a grown man. I do not know what else to say, except thank you.

To the many Professors who have guided me through these last years at MIT, thank you for your support. I need to mention Professor Steve Lerman for his counsel and assistance, first, in his role as Chair of the Faculty in the Faculty Policy Committee, and

later as Dean for Graduate Students. To Professor Rafael Bras, your words always resound with me. I will not forget them. You will be missed at MIT. To Professor Colin Stultz, I did it, this is just the beginning.

To all my friends over the years at MIT: You all rock! Will, I look forward to many dinners and particularly to Saturday morning flying. Jarrett, thanks for stealing me away late at night from lab for a salad at CCF. Peter, I promise I will have more sunflower seeds and JW Black for you. Eric, thank for being such a great lab mate and a room mate for all those years. Zari, you can stop busting my chops now (actually not yet). Brian, Cali is not the same without you. Charles, I will miss you being up two floors and our computer room discussions. Joost, keep the seat warm and the beer cold for me. Lauren, I miss that smile. Mike and Xu, thanks for being there and keeping me sane and warm with goodies. To Penny S. I can't wait to see you soon. It has been to long. Roberto and Sarah, thanks and I hope Cali treats you well. Danya, it is your turn to go through this now. Kyle, keep the hops coming. To the '24' crew. When is Jack coming back? Jess, are you going to get a county accent now? Guido, I am coming to see you soon. Luis, Curra, and Lola, two more months and then Granada. Tzanko, thanks for your friendship over the years. Ramon, nos vemos en Barcelona!

To all the members of the Drennan Lab over the years, it has been a pleasure to be in the lab with you. Good luck with your endeavors. To the members of the Essigmann lab, thanks for everything. To the members of the Licht Lab, it was great to have you next door.

To the USF Crew. We've come a long way baby! Ron, always a kool kat. It was good to see you, Jen, and the girls. Chris thanks for all the memories. I am sure there are more to come. Ainsley, I do not know how I could have gotten through without you sitting there next to me. I hope it wasn't too painful. I miss the Taco Bell runs, lunch at Bennigan's, and the afternoon run to the Yuengling brewery.

To the Cuba crew, Ivette, Juani, and Yohan, thanks for showing me what it means to want something enough that you are willing to give everything up for it. Oh that the world would have such passion.

To Lily, I am lucky to have you here with me. This is just the beginning. To Mr. J. Scott Burns, and Mrs. Cynthia Burns, thank you for taking me in and making me feel

welcome and loved in your family. To Matu, you are an amazing woman. Thank you for raising such a beautiful family.

To my grandparents, Thomas Martinez and Rafaela Martinez. They left their home, family, and friends in Cuba to provide a better life for their children. Abuelita, te extraño. To my sister, Danery Ela Foster, thanks for taking care of me and being there over the years. To my Mother, Nerida Hernandez Martinez Perez Alfonso, you are the reason I am here. Even in the darkest of times you believed in me. Thank you for the principles you instilled in me. I remember what you told me once, "They can take everything away from you except what is in your head." Well, madre mia, I have filled it with great things.

- Research is an engaging and exciting, but cruel mistress -

TABLE OF CONTENTS

ABBREVIATIONS

AdoCbl	Adenosylcobalamin 5'-deoxyadenosylcobalamin
AdPentCbl	adeninylpentylcobalamin
afu	Archaeoglobus fulgidus
ATCC	American Tissue Culture Center
BSA	bovine serum albumin
CIP	calf-intestine alkaline phosphatase
DTNB	5,5′-dithionitrobenxoic acid
DTT	1,4-dithio-DL-threitol
E	extinction coefficient
Ec	*Escherichia coli*
E.coli	*Escherichia coli*
EDTA	ethylenediaminetetraacetic acid
EPG	edge-plane pyrolytic graphite
F	the Farraday constant
FAD	flavin adenine dinucleotide
$FADH_2$	flavin adenine dinucleotide, reduced form
FO	flavin oxidizing conformation
FR	flavin reducing conformation
Grx	glutaredoxin
GS	glutathione
GrR	glutathione reductase
HEPES	N-2-hydroxyethyl-piperazine-N'-2-ethanesulfonic acid
IPTG	isopropylthio-β-D-galactoside
Kan	kanamycin
kb	kilo base pairs
KPi	potassium phosphate
LB	Luria Bertani broth
LIC	ligation independent cloning
L.leichmannii	*Lactobacillus leichmannii*
min	minute
MW	molecular weight
n	the number of electrons transferred in a redox process
n_{app}	the apparent number of electrons involved in a redox process
NAD^+	nicotinamide adenine dinucleotide, oxidized form
NADH	nicotinamide adenine dinucleotide, reduced form
$NADP^+$	nicotinamide adenine dinucleotide phosphate, oxidized form
NADPH	nicotinamide adenine dinucleotide phosphate, reduced form
NaOAc	sodium acetate
Ni-NTA	nickel nitrilotriacetic acid-agarose gel
OD	optical density
ORF	open reading frame
PAGE	polyacrylimide gel electrophosesis
PCR	polymerase chain reaction

PEG	polyethylene glycol
PMSF	phenil methyl sulfonyl fluoride
PFV	protein film voltammetry;
R	the ideal gas constant.
RNR	ribonucleotide reductase
RTPR	ribonucleoside triphosphate reductase from *L.leichmannii*
s	second
SafO	safranine O
SDS	sodium dodecyl sulfate
T	temperature
ta	*Thermoplasma acidophilum*
Tr	thioredoxin
TrR	thioredoxin reductase
TRIS-HCl	tris(hydroxymethyl)-aminomethane hydrochloride
wt	wild-type
XO	xanthine oxidase

LIST OF FIGURES

LIST OF TABLES

LIST OF SCHEMES

This page intentionally left blank

Chapter 1:

Introduction to the Ribonucleotide Reduction, Thioredoxin Reductase, and Thioredoxin Pathways

INTRODUCTION

Nucleotide reduction and redox regulation are essential cellular processes. DNA synthesis depends on a balanced supply of deoxynucleotides pools (1, 2). Loss of deoxynucleotide equilibrium increases mutation rates in DNA replication and can lead to cell death (3). Redox regulation in the cell is also intimately tied to cell survival. The cytoplasm of the cell is generally a reducing environment due to the high concentration of reductants such as the NADPH dependent thioredoxin reductase (TrR) / thioredoxin (Tr) system. Cellular metabolism and integrity are maintained by carefully balancing the redox state of all components for optimal overall function.

While nucleotide reduction and redox regulation are key processes in all organisms, this thesis focuses on ribonucleotide reductases, thioredoxin reductases, and thioredoxins from archaea. Archaea are the most obscure of life's three domains. The members of this domain were first differentiated from bacteria based on rRNA sequence patterns. The first archaea were isolated from high temperature, high salinity, and acidity environments. Because of their ability to grow in such inhospitable environments, it has been postulated that they contain vestiges of the first biochemical pathways. Recently, with the advent of genomic sequencing, the secrets held in their genomes have been revealed, yielding a trove of biological and evolutionary information. One surprising observation has been how many of the features contained therein are specific to archaea.

Due to their extreme environmental growth conditions, there is interest in elucidating the manner in which archaea have evolved to cope with, and maintain cellular integrity. Two underlying questions are (1) how do archaea preserve genomic integrity; (2) what mechanisms are used to regulate and maintain the redox environment inside the cell?

This introduction chapter will (1) provide a brief introduction to the classes of RNRs, (2) discuss the accepted regulation, (3) substrate preference, (4) and reaction mechanism of RNR, (5) give an overview of RNR structures, (6) discuss the role of the TrR / Tr system as a hydrogen donor to RNRs; (7) introduce the reader to features of archaea; and (8) finish with the discussion of the current knowledge of RNR / TrR / Tr systems in archaea.

RIBONUCLEOTIDE REDUCTASES

Ribonucleotide reductases (RNRs) are enzymes that are present in all living organisms. These enzymes catalyze the reduction of nucleotides to deoxynucleotides at the 2' ribose site thus providing the building blocks for the synthesis and repair of DNA. Activation and regulation of RNRs assures the maintenance of the proper ratios of deoxynucleotides in the cell. Imbalances in deoxynucleotide levels increase mutagenesis and are detrimental to the cell (Recently reviewed in (4-6).

The reaction carried out by RNRs is of interest to many fields of research. Evolutionary biologists are drawn to the enzyme because of its proposed principal role in the transformation from an RNA-based world to a DNA-based world (7-9). The exquisite allosteric control of RNRs, which allow these enzymes to catalyze the reduction of four unique substrates all in one active site amazes enzymologists, whereas the similarity of the core structural elements between the seemingly unrelated amino acid sequences of RNRs enthralls structural biochemists (10). Epidemiologists are interested in these enzymes because of the possibility they hold for the development of unique system-specific inhibitors of cancer, and organism-based diseases such as malaria and tuberculosis (11-13).

RNRs need a hydrogen donor to re-reduce the disulfide bond of the active site cystine to two thiols, the latter of which contribute reducing equivalents in the synthesis of deoxynucleotides. Class I and class II RNRs contain a C-terminal CxxC motif, referred to as the c-terminal tail, which 'swings' into the active site barrel and reduces the oxidized active site cystine during every round of catalysis. In turn, the c-terminal tail disulfide is reduced by thioredoxin for the next round of catalysis. Regeneration of Tr is mediate using by TrR and NADPH. The thioredoxin reductase (TrR) and thioredoxin (Tr) system from *E. coli* was the first system isolated as a hydrogen donor for RNRs, and to date, this is the best characterized TrR / Tr system(14, 15).

CLASSES OF RIBONUCLEOTIDE REDUCTASES

Currently, RNRs are classified into three different classes (7) characterized by the cofactor that initiates activation of the thiyl radical in the active site (Scheme 1.1 and Table 1.1). The three RNR classes are reviewed below:

Class I RNRs – The di-iron tyrosyl radical class I RNR prototype was first isolated and characterized in the early 1960s from *E. coli (16)*. Humans and higher order organisms also contain a class I RNR. These enzymes are oligomers with the form $(\alpha_2\beta_2)_n$. In *E. coli* the α_2 dimer is designated R1 and is the catalytic portion of the molecule, whereas the β_2 dimer is designated R2 and is responsible for the radical generation. The R2 monomer contains a μ-oxo-bridged diferric cluster and a tyrosine (Y122) and generates a stable tyrosine radical. During catalysis, the radical species is transferred over 35 Å to the active site in the R1 dimer (17, 18). Class I R1 structure from *E. coli* was solved in 1994 (19), whereas a class I R2 structure from *E. coli* was solved in 1990 (20) (Figure 1.1). To date, no structure of a R1R2 complex from class I exists.

Class II RNRs - Studies with the organism *L. leichmannii* yielded the first adenosylcobalamin-dependent RNR (RTPR) (21-24). Class II RNRs can be found in bacteria, archaea, and one eukaryote. In this class, AdoCbl is the radical-generating cofactor (24). The ability to use a small molecule to initiate the active site radical instead of a separate protein sets class II RNRs apart from the other classes. A class II monomeric and a dimeric structure have been solved (10, 25) (Figure 1.2).

Class III RNRs - Class III RNRs were discovered in *E. coli* grown under anaerobic conditions in 1989 (26) and have been found in all anaerobic organisms to date. Class III RNRs are α_2 dimers, which are activated by a β activase protein that contains an 4Fe-4S cluster and utilizes *S*-adenosylmethionine (AdoMet) to generate a glycyl radical located on the catalytic subunit (27-30). A class III α_2 structure from bacteriophage T4 was recently solved (Figure 1.3) (31).

ALLOSTERIC REGULATION OF RIBONUCLEOTIDE REDUCTASES

A unique feature of RNRs is that one enzyme catalyzes the reduction of all four nucleotides to their corresponding deoxynucleotides and that this substrate specificity is regulated by the binding of deoxynucleoside triphosphates allosteric effectors to a distal site.

Early work on *E. coli* RNR identified two allosteric sites in the R1 subunit that regulate activity at the active site (see Scheme 1.2) (32). These sites have been identified

as the activity site and the specificity site. The activity site is found at the N-terminus of the class Ia RNRs and binds either ATP or dATP, with ATP stimulating overall activity and dATP inhibiting catalytic activity (32). The specificity site, which is common to class Ia and class Ib RNRs, binds dGTP, dTTP, dATP, and ATP (32). Binding of dGTP stimulates the reduction of ADP to dADP, binding of dTTP induces the reduction of GDP to dGDP, and finally binding of ATP or dATP to the specificity site promotes the reduction of CDP or UDP to dCDP and dUDP (Scheme 1.3).

Subsequent work with *Salmonella typhimurium* lead to the isolation of a second class I RNR, which differs from the classic RNR in that it is missing the N-terminus amino acids comprising the nucleotide binding domain involved in regulating the overall activity of class I RNRs (described above) (33). The *S. typhimurium* and *E. coli* class I RNRs differ only in this respect. These observations lead to the separation of these enzymes into class Ia for the originally isolated RNR and class Ib for the subsequently isolated RNR. There is also a recently discovered class Ib RNR that uses a Fe-Mn cofactor for radical generation (34).

The class II RNRs display a similar relationship between the binding of effectors at the specifity site and the reduction of substrates at the activity site (35). No operational activity site has been reported for a class II RNR.

Class III RNRs show similar effector-dependent substrate specificity as class Ia RNRs. Early experiments suggested two separate allosteric binding sites, but these sites exhibited slightly different specificity than the *E. coli* class Ia RNR. Class III RNR effector binding experiments show that these RNRs can bind two dATP molecules and only one ATP molecule (36). Also, unlike class Ia RNRs, when dATP binds to either allosteric binding site, overall activity is inhibited. These sites have been called the purine and pyrimidine binding sites. When ATP binds the pyrimidine site, the enzyme reduces pyrimidines, and when dGTP or dTTP bind the purine site, the enzyme reduces ATP and GTP (31, 36).

SUBSTRATE PREFERENCE FOR NDP VERSUS NTP

All class I and most class II RNRs use NDPs as substrates and class III RNRs use NTPs as substrates. RTPR and two other class II RNRs, the RNR from the cyanobacteria

Anabaena sp. Strain PCC7120 and the RNR from the eukaryote *Euglena gracilis*, are the only characterized class II enzymes that use NTPs as substrates (21, 37, 38). The class III RNRs use NTPs as substrates.

The structural source for this preference is not understood. Superposition of the RTPR and *T. maritima* active site barrels do not reveal obvious structural differences that would explain the substrate preference (please see discussion in Chapter 2 for more detail).

MECHANISM OF THE RIBONUCLEOTIDE REACTION

Mechanistic studies performed by Professor JoAnne Stubbe's laboratory on *E. coli* and *L. leichmannii* RNRs suggested that these seemingly distantly related enzymes utilized the same catalytic mechanism (Scheme 1.4) (39). Recent structural work confirmed that even though there is little conserved amino acid homology between the classes of RNRs; there are highly conserved structural features that regulate specificity and activity of RNRs (10, 19).

Although generated by different cofactors, once the radical is initiated, the reduction of nucleotides proceeds in a similar manner for all RNRs (Scheme 1.4, Step 1). The first step of the reaction is the abstraction of a hydrogen atom from the 3` position of the substrate (2). The 2` hydroxyl group is protonated by one of the redox active cysteines on the protein (3), followed by the loss of water, creating a radical-cation intermediate (4). The radical-cation intermediate undergoes a two step one electron reduction by the redox active cysteines, allowing the formation of the deoxy-2` carbon of the pentose ring and the formation of a disulfide on the active site cysteines (5). This step is followed by the return of the initial hydrogen to the 3` position and the regeneration of the thiyl radical located on the finger loop (40).

STRUCTURES OF RIBONUCLEOTIDE REDUCTASES

There are structures of the *E. coli* apo RNR and the RNR complexed with an effector and a substrate (Fig 1.1) (19, 41). The major difference in the crystallization conditions between the two structures is that in the apo form, the active site cysteines were oxidized, but in the complexed structure, they were fully reduced. The authors

suggest that the in order for substrate to bind, the cysteines in the active site need to be reduced (41). However, this statement contradicts the biochemical data. Although the structure provided the first glimpse of the location of effector and substrate binding, it does not explain the allosteric regulation of the specificity site and of active site.

The recent structure of RTPR gave a glimpse of how a monomeric RNR can be allosterically regulated. In class I and class II dimeric RNRs, the allosteric binding site is located at the ends of a 4-helix bundle, which composes the dimeric interface (two helices from each monomer) (colored magenta in Figures 1.1 and 1.2). The effector binding pocket is capped by a loop that locks in the effector (Figure 1.1A, inset). In the monomeric RTPR structure, the effector binding helix bundle and loop have been structurally conserved by what could be described as two 'cut and paste' events (10). The recent structure of a class II RNR from *Thermotoga maritima* in complex with substrate and effectors was the first structure of a dimeric class II RNR (Figure 1.2B). Unfortunately, the engineered construct used to elucidate the *tm*RNR structure lacked 182 C-terminal residues (residues 645-827) which bind AdoCbl.

Although all three classes of RNR have low sequence homology, the structure of the central core, where catalysis occurs is highly conserved (Figure 1.4) (10). The core of all three classes of RNR is composed of a 10 strand α/β-barrel with two parallel five stranded β-sheets oriented in anti-parallel fashion. The barrel harbors the active site cysteine that forms the thiyl radical at the tip of the 'finger loop'. The barrel also contains the cysteines which provide reducing equivalents during catalysis. The spacing between the cysteine on the finger loop (C462 in *E. coli*)and the second cysteine (C439 in *E. coli*) responsible for providing reducing equivalents appears to be an important marker for identifying the class of RNR from sequence information alone. For class II RNRs the conserved motif is NPC(G/A)EX_8C(N/V)L (Figure 1.5) (42). This motif is of variable distance in class I RNRs and is not seen in class III RNRs. Structural comparison of these motif loops shows that the motif loop in class I RNRs threads itself out the bottom of the barrel and interacts with the same loop from the other monomer. In class II structures, this loop extends to the end of the barrel Figure 1.6)

The significance of this conserved motif in class II RNRs is not clear. This sequence is not involved in AdoCbl binding or in substrate or effector binding and

recognition. Structural alignments of RNRs show that it is outside of the active site barrel, where there is little conserved structure between the classes of RNRs. As seen in the class I structure, there is no reason why this loop could not be of variable length (Figure 1.6). The high level of conserved amino acids in this motif suggests that it may play a role in a yet to be characterized cellular interaction.

Even though the current structural snapshots of different RNRs provide much information about how and where substrate and effectors bind, we lack a complete understanding of how RNRs are regulated. The large conformational changes proposed to occur during catalysis have yet to be captured in crystallographic studies. The non-physiological conditions of the crystallization of RNRs (pH 4.5-5.5) leads to questions regarding the ionic interactions seen in the crystal structures between effectors and protein. There is still much work to be done to unravel the structural rearrangements needed to transmit information between the effectors bound at the specificity site and the selection of substrate at the activity site.

THIOREDOXINS AND THIOREDOXIN REDUCTASES

The reduction of the RNR active site disulfide is accomplished using two different mechanisms. Class I and class II RNRs contain two active site cysteines, which contribute reducing equivalents in the synthesis of deoxynucleotide building blocks for DNA synthesis. The active site disulfide in class I and class II RNRs is reduced during every round of catalysis through cysteines in its c-terminal tail. These c-terminal cysteines are in turn reduced by the NADPH dependent Tr / TrR / NADPH system or glutathione system (43), whereas the cysteines in class III RNRs are reduced by formate (44).

Trs are the major disulfide reductase family responsible for maintaining intracellular proteins in their reduced state (45). Trs are small enzymes composed of about 120 amino acids and are characterized by the active site sequence CxxC (46). Tr is reduced by electrons from NADPH via TrR (Scheme 1.5). Tr and TrR proteins are ubiquitous from archaea to eukaryote species (47).

TrR is an FAD-containing enzyme with a general structure that is composed of a nucleotide binding domain and an FAD binding domain. The Tr / TrR system from *E.*

coli was the first isolated hydrogen donor for ribonucleotide reductase (RNR) and is the best characterized Tr / TrR system. The structures of *E. coli* Tr / TrR have been solved by crystallography (Figure 1.7 and 1.8) (48, 49). Biochemical work on the TrR / Tr system suggests that TrR is a highly flexible protein which undergoes large conformational changes during catalysis. The crystal structure of *E. coli* TrR provides evidence for this change in conformation (49). This structure of TrR shows that the cysteines which participate in the reduction of Tr are buried deep in the interface of between the FAD binding domain and the NADPH binding domain and that the NADPH binding site is located 17 Å from the isoalloxazine ring. The authors proposed a catalytic model in which in the redox active cysteines are located near the FAD (FO conformation) for re-reduction, then there is a 66° conformational change of the enzyme which brings the NADPH is near the FAD (FR conformation) for hydride transfer from the NADPH to the FAD moiety (Figure 1.7) (50). This large change was captured in the x-ray structure of the *E. coli* TrR / Tr complex (Figure 1.8) (51).

Recent studies reveal that the TrR / Tr system is involved in regulation of many cellular functions (52). Besides being a hydrogen donor for RNRs, Trs have been implicated as having a role, for example, in reduction of H_2O_2, protein repair by methionine sulfoxide reduction, regulation of chloroplast photosynthesis enzyme, redox regulation of transcription factor, and regulation of apoptosis (43, 53, 54). Cellular redox regulation is a complex interconnected system, and understanding the regulation of the TrR / Tr system is essential to examining the flux of reducing equivalents through cellular pathways.

ARCHAEA

Seminal work performed by Woese and Fox elucidated the third branch of life, the archaea (55). Their study of small subunit ribosomal RNA sequences (SSU rRNA) led to the isolation and characterization of SSU rRNA from *Methanobacterium thermoautotrophicum ΔH*, a methanogen (56). Subsequent characterization of SSU rRNA from several other methanogens gave rise to the identification of a new kingdom of life. Originally, archaea were thought to reside only in extreme environments, but they have now been found in all ecological niches (57). Most archaea organisms are difficult

to cultivate and are recalcitrant to genetic manipulation due to cell wall structure and natural resistance to common antibiotics (58)

Sequencing efforts have produced 52 annotated archaeal genomes with a further 35 genomes whose sequencing and annotation is currently being completed (NCBI Entrez Genome Project http://www.ncbi.nlm.nih.gov/sites/entrez?db=genomeprj). Analysis of the completed genomes reveals that there is much to be learned from these organisms. Archaeal genomes have a large number of proteins that are unique to archaea and for which no biological role can be assigned (> 50% in some organisms). Almost all of the proteins that can be assigned to metabolic pathway homologues are awaiting biochemical and biophysical characterization.

There has been growing interest in the study of archaea. They exhibit unique molecular signatures that have expanded the known range of biochemical reactions and are capable of using a wide assortment of molecules to provide the metabolic energy for cellular reactions (59, 60). Their ability to inhabit and thrive under the most extreme environmental conditions suggests novel pathways that allow for specialized survival mechanisms. Biofilms and community structures of archaea are changing how scientists view interactions between biological systems (61, 62). The exchange of signaling molecules between species, gene expression profiles, and lateral gene transfer plays an integral role in the endurance of archaeal systems. In addition, they hold evolutionary records of key biologically relevant pathways (63, 64). The study of chemical processes mediated by archaea potentially could yield a host of biological novelties.

ARCHAEAL RIBONUCLEOTIDE REDUCTASE SYSTEMS

Archaeal RNR systems present an exciting and unique opportunity for the study of DNA damage and repair processes. Many archaeal organisms thrive under conditions of extreme temperature, pH, and pressure. The biochemical processes involved in maintaining the fidelity of DNA replication and repair is still being elucidated (Recently reviewed in (65-67). Regretfully, there has been little progress in elucidating the role RNRs play in this new environment because organisms are extremely difficult to grow and manipulate. Recently, genomic data has allowed for the mining of extremophilic genomes for putative RNR enzymes.

Analysis of the 53 completed archaeal genomes reveal that there are a total of 79 annotated RNRs, with 20 organisms containing 2 or more RNRs (NCBI Entrez Genome Project) (Table 1.2). There are 14 class I RNR subunits (only 4 subunits correspond to the α protein), 32 class II RNRs, and 33 class III RNRs (this number does not include the activating protein). One organism, *Methanocorpusculum labreanum Z*, does not contain an annotated RNR. It would be interesting to reexamine this genome to look for a possible analogue of an RNR.

How these organisms maintain and regulate the nucleotide pool responsible for DNA damage and repair remains a mystery. As the characterization of these organisms progresses, it will be exciting to see if their RNRs exhibit the same substrate preference and allosteric regulation patterns seen in other RNRs, or if we will see the emergence of novel substrate preferences and regulation. Biophysical and biochemical analysis of archaeal RNRs is important to understanding the role these enzymes play in the maintenance and regulation of nucleotide reduction.

The RNR from *T. acidophilum (ta*RNR*)* is the only archaeal RNR to have been characterized (9, 68). This RNR is a dimeric class II RNR which exhibits allosteric effector patterns similar to those seen in *E. coli* class Ib RNRs (68). *ta*RNR exhibits a preference for nucleoside diphosphate substrates (9). The temperature optimum for the reduction of substrates is 55 °C (9). To expand our knowledge of archaeal RNRs, we propose to characterize a RNR from the archaeon *Archaeoglobus fulgidus*.

THIOREDOXIN REDUCTASE AND THIOREDOXINS IN ARCHAEA

Disulfide bond reduction has been studied in archaea in the context of protein disulfide oxoreductases (69, 70). While there has been extensive effort undertaken to understand prokaryote and eukaryote TrR / Tr systems, there is a paucity of work completed on archaeal systems.

Genomic analysis of archaea reveals that of the 53 genomes, 13 do not contain sequences which correspond to putative TrRs. The remaining 40 genomes contain a total of 63 TrR sequences with 11 of them holding 2 or more putative TrRs (Table 1.3).

Compared to RNRs and TrRs, Tr genes are abundant. There are a total of 114 putative Trs in the annotated genomes. There are 10 organisms which do not contain

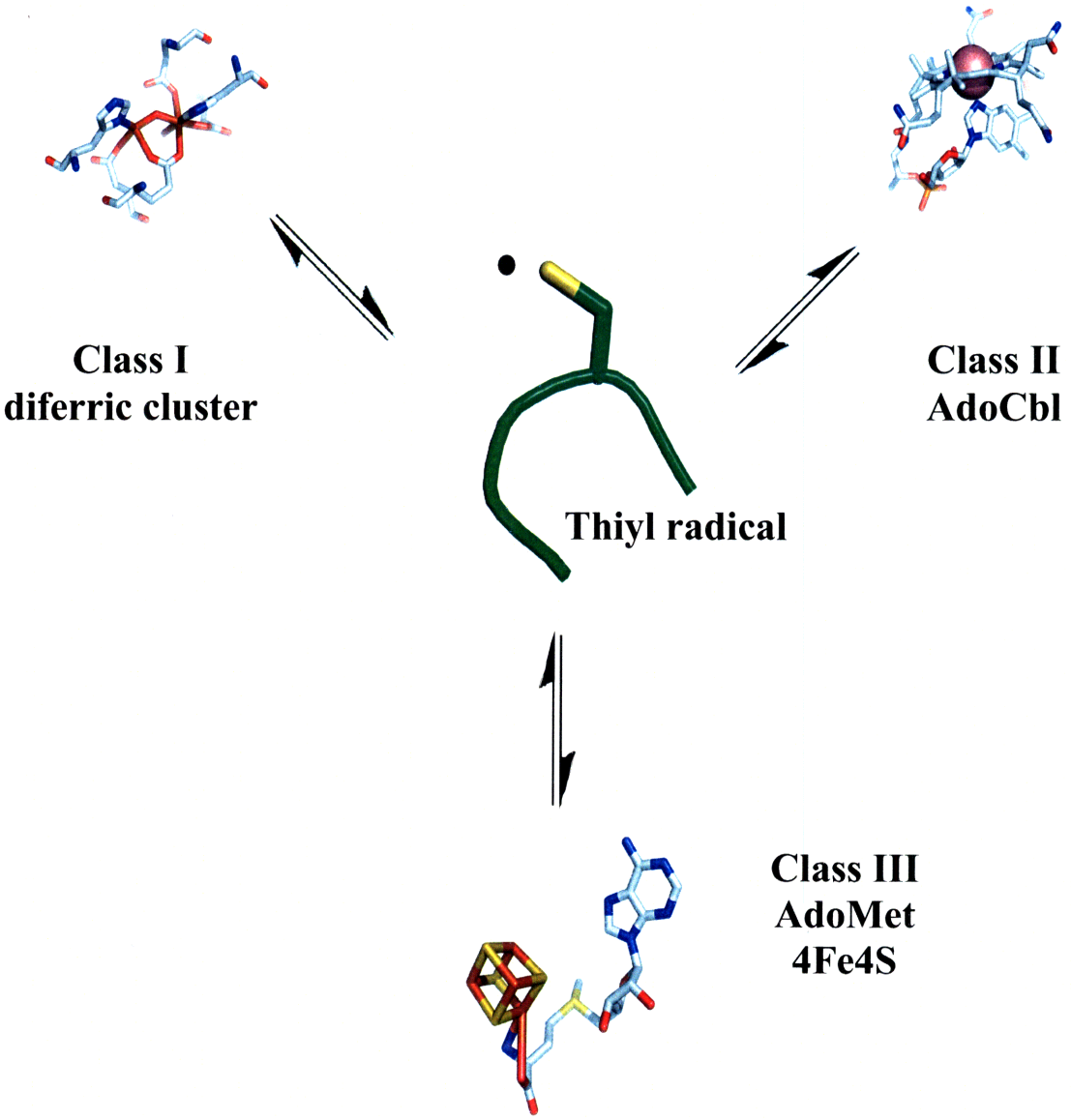
putative Trs. Of the remaining genomes, 27 have two or more Trs with 12 of them having four or more proteins each (Table 1.4). The organism with the most Trs, *Haloarcula marismortui ATCC 43049*, has a total of 8 Trs. Deciphering why there are so many Trs in one organism may help us understand how archaea have adapted to live under environmentally stressful conditions.

Genome analysis has provided a wealth of information about redox regulation players in archaea. To date biochemical analysis has been performed on TrR / Tr from *Aeropyrum pernix* (71), a TrR from *Pyrococcus horioshii* (72, 73), and the TrR / Tr from *Sulfolobus solfataricus* (74-77).

There is much to be done before a consensus is reached about whether pathways involving RNRs, TrRs, and Trs are different in archaea than in other kingdoms. The breadth of extreme environments in which archaea thrive suggests that there may be alternate regulation patterns which provide preference of substrate in RNRs. The lack of identifiable TrRs in some archaea suggests alternative pathways to providing reducing equivalents in the cell. The study of important enzyme families in archaea provides an opportunity to investigate how organisms have adapted to living at the extreme edges of life.

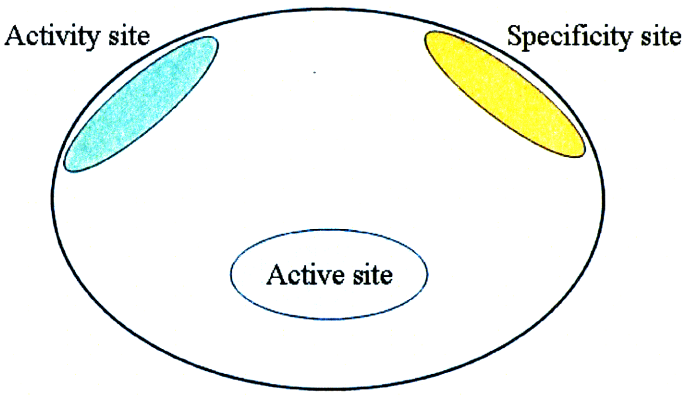
This thesis investigates the RNR (*afu*RNR), four Trs (Tr-1-4) from *Archaeoglobus fulgidus,* and the TrR (*ta*TrR) from *Thermoplasma acidophilum*. The second chapter covers the initial characterization of the recombinant *afu*RNR. Here we show that *afu*RNR reduces CDP and CTP to their corresponding deoxynucleotides in an AdoCbl dependent manner. In the third chapter we report the structural and biochemical characterization of *ta*TrR and our finding that *ta*TrR does not use NADPH as a source of reducing equivalents. We finish with the biochemical and electrochemical characterization of the four *afu*Trs in chapter 4.

Scheme 1.1 Metal-dependent activation of the thiyl radical of RNRs. The three classes of RNRs are identified by the metal cofactor which initiates the abstraction of the hydrogen atom, thus generating the radical at the active site cysteine. In the strictly aerobic Class I RNRs, the radical is generated at the diferric cluster in the R2 subunit and is translated through the R1 protein to the active site. In Class II RNRs, AdoCbl undergoes homolytic cleavage creating a 5'-deoxy adenosyl radical which abstracts the hydrogen atom from the active site cysteine. Class III RNRs function under anaerobic conditions and use AdoMet to generate a stable glycyl radical. This radical in turn abstracts the hydrogen atom from the active site cysteine to initiate the catalytic cycle.

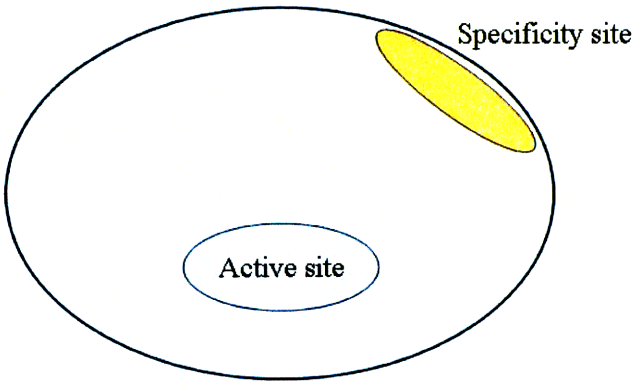


Scheme 1.2 Model of allosteric regulation for Class I and Class II RNRs. Only one subunit monomer is shown in scheme. (A) Class Ia RNRs and some Class II RNRs have an N-terminus activity site which regulates overall reduction of nucleotides by RNR. When ATP is bound at the activity site, overall activity is up regulated, but when dATP is bound, enzyme activity is down regulated. (B) Class Ib and most Class II RNRs do not possess the N-terminus activity site. In all RNRs, the specificity site regulates which nucleotide is reduced to its corresponding deoxynucleotide in the active site. See Scheme 1.3 for details of specificity regulation.

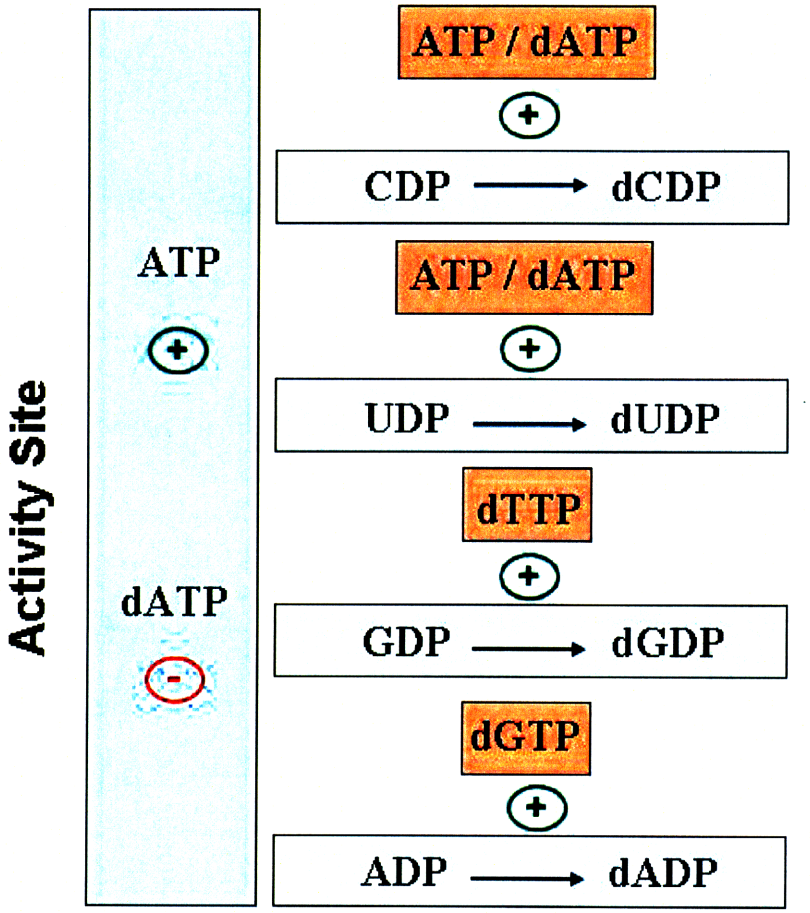
A



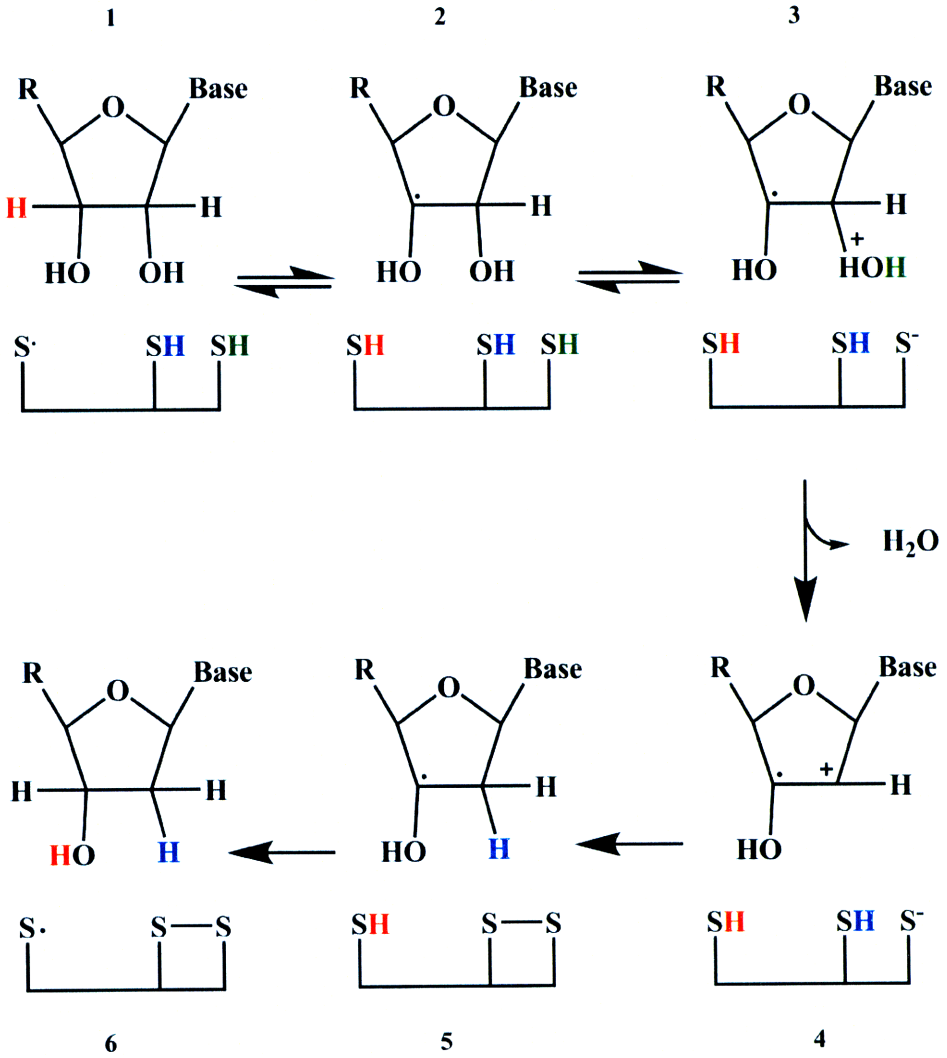
B



Scheme 1.3 A schematic representation of the classic pattern of regulation as exemplified in Class I and Class II RNRs. In Class Ia RNRs, overall activity is up regulated when ATP is bound and down regulated when dATP is bound. The specificity effectors are in brown boxes and substrates are in light grey boxes. When ATP or dATP is bound at the specificity site, RNRs preferentially convert CDP or UDP to dCDP or dUDP. The binding of dTTP at the specificity site preferentially reduces GDP to dGDP and binding of dGTP to specificity site reduces ADP to dADP. In Class II RTPR the substrates are nucleotide triphosphates instead of nucleotide diphosphates, but the enzyme regulation follows the same pattern.

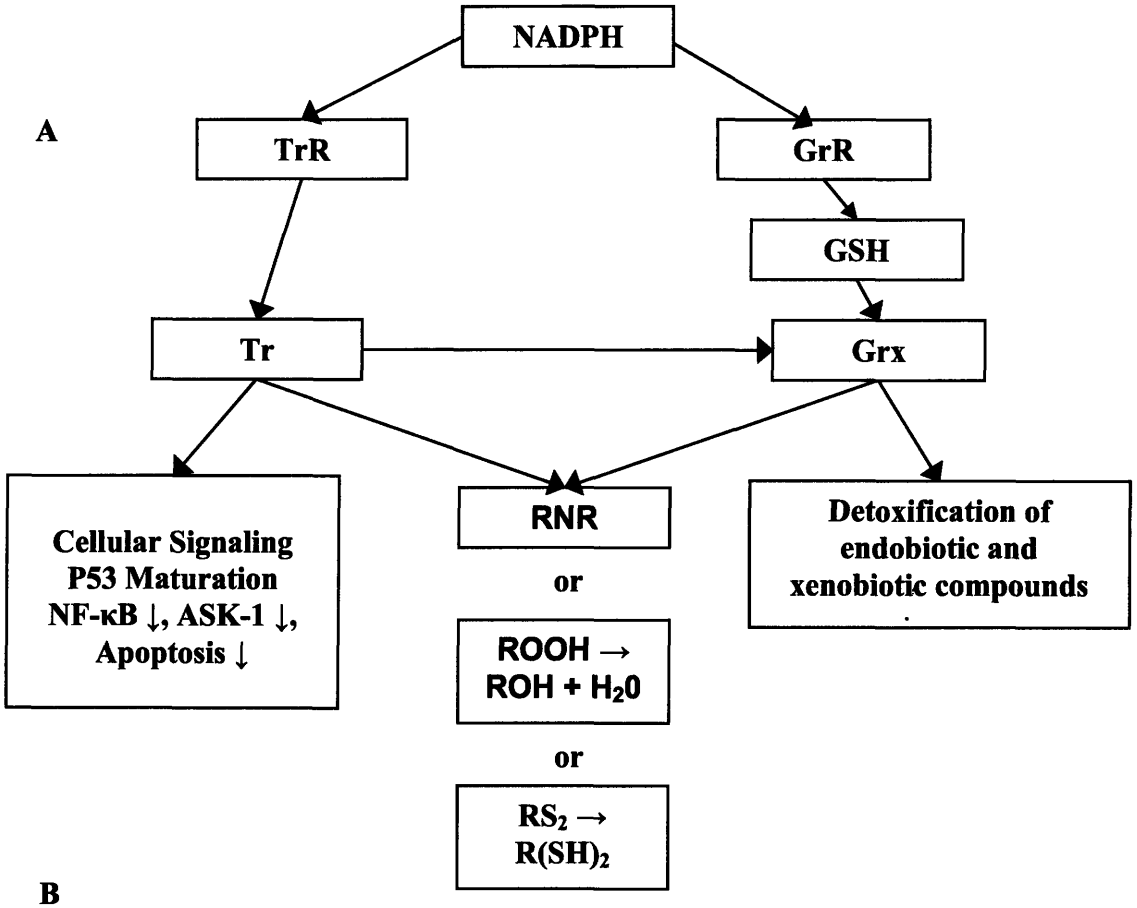


Scheme 1.4 A minimal reaction mechanism for RNRs. In Class III enzymes, the distal cysteine is lacking and the enzyme uses formate as the reductant. The first step of the reaction is the abstraction of a hydrogen atom from the 3` position of the substrate (2). The 2` hydroxyl group is protonated by one of the redox active cysteines on the protein (3), followed by the loss of water, creating a radical-cation intermediate (4). The radical-cation intermediate undergoes a two step one electron reduction by the redox active cysteines, allowing the formation of the deoxy-2` carbon of the pentose ring and the formation of a disulfide on the active site cysteines (5). This is followed by the return of the initial hydrogen to the 3` position and the regeneration of the thiyl radical located on the finger loop.



Scheme 1.5 NADPH participation in cellular redox regulation. (A) The NADPH-dependent redox regulatory pathways. (B) Mechanism of NADPH dependant cellular reduction by the thioredoxin reductase system.

A



B

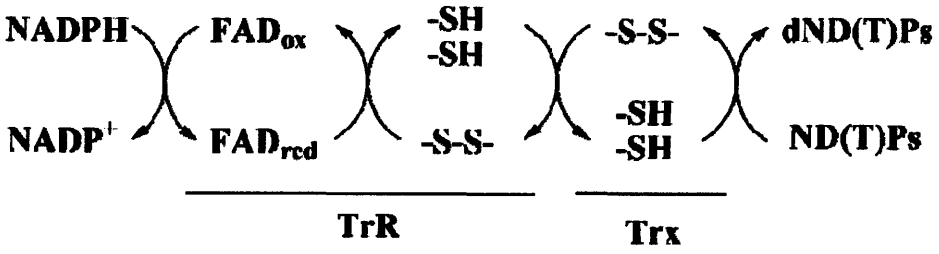


Table 1.1 Overview of the classes of RNRs. Table lists key features of classes of RNRs.

	Class			
	Ia	**Ib**	**II**	**III**
Distribution	Archaea / Bacteria / Eukaryotes	Bacteria / Archaea	Archaea / Bacteria / Eukaryotes	Archaea / Bacteria
Anaerobic / Aerobic	Aerobic	Aerobic	Anaerobic / Aerobic	Anaerobic
Protein Structure	$\alpha_2\beta_2$	$\alpha_2\beta_2$	α / α_2	$\alpha_2\beta$
Radical Generator	Fe-O-Fe	Fe-O-Fe	AdoCbl	4Fe-4S / AdoMet
Catalytic Radical	Thiyl	Thiyl	Thiyl	Thiyl
Reductant	Redoxin-(SH_2)	Redoxin-(SH_2)	Redoxin-(SH_2)	Formate

Table 1.2 RNRs in completed archaea genomes. Table consists of all completed and annotated Archaea genomes as of June 2008. Table includes name of organisms, what archaea subdivision they are assigned, the organism reference sequence, the protein locus tag, what class the particular RNR falls under, and the optimal growth temperature for the organisms. Only one RNR, the *Thermoplasma acidophilum* RNR, has been previously characterized. All others are putative RNRs annotated by sequence homology modeling at the NCBI Entrez Genome Project. Some organisms have multiple RNRs.

Organism	Group	RefSeq	Locus Tag	Class	Growth T (°C)
Aeropyrum pernix K1	Crenarchaeota	NC_000854.2	APE_2062.1	II	90-95
Archaeoglobus fulgidus DSM 4304	Euryarcaeota	NC_009917.1	AF1664	II	83
Caldivirga maquilingensis IC-167	Crenarchaeota	NC_009954.1	Cmaq_0639	II	85
Candidatus Korarchaeum cryptofilum OPF8	Other Archaea	NC_010482.1	Kcr_0134	II	N/A
Candidatus methanoregula boonei 6A8	Euryarchaeota	NC_009712.1	Mboo_0022 Mboo_1709	II III	37
Haloarcula marismortui ATCC 43049	Euryarchaeota	NC_006396.1	rrnAC1622	II	40-50
Halobacterium salinarum R1	Euryarchaeota	NC_010364.1	OE3328R OE4345R OE4346R	II Ia – β Ia – α	50
Halobacterium sp. NRC-1	Euryarchaeota	NC_002607.1	VNG2383G VNG2384G	Ia – β Ia – α	42
Haloquadratum walsbyi DSM 16790	Euryarchaeota	NC_008212.1	HQ2859A	III	37-40
Hyperthermus butylicus DSM 5456	Crenarchaeota	NC_008818.1	COG1328F Hbut_0854	III II	95-106

Organism	Group	RefSeq	Locus Tag	Class	Growth T (°C)
Ignicoccus hospitalis KIN4/I	Crenarchaeota	NC_009776.1	Igni_0036	III	> 80
Metallosphaera sedula DSM 5348	Crenarchaeota	NC_009440.1	Msed_1716	II	70
Methanobrevibacter smithii ATCC 35061	Euryarchaeota	NC_009515.1	Msm_1383	III	37-40
Methanocaldococcus jannaschii DSM 2661	Euryarchaeota	NC_000909.1	MJ0832	II	85
Methanococcoides burtonii DSM 6242	Euryarchaeota	NC_007955.1	Mbur_2125 Mbur_2140	III III	23.4
Methanococcus aeolicus Nankai-3	Euryarchaeota	NC_009635.1	Maeo_1041	III	42
Methanococcus maripaludis C5	Euryarchaeota	NC_009135.1	MmarC5_1471	III	24-25
Methanococcus maripaludis C6	Euryarchaeota	NC_009975.1	MmarC6_0728	III	25-40
Methanococcus maripaludis C7	Euryarchaeota	NC_009637.1	MmarC7_1227	III	25-40
Methanococcus maripaludis S2	Euryarchaeota	NC_005791.1	MMP0227	III	25-40
Methanococcus vannielii SB	Euryarchaeota	NC_009634.1	Mevan_1241	III	30
Methanocorpusculum labreanum Z	Euryarchaeota	NC_008942.1	N / A	N / A	37
Methanoculleus marisnigri JR1	Euryarchaeota	NC_009051.1	Memar_0306 Memar_0320 Memar_1998	III III II	21-25
Methanopyrus kandleri AV19	Euryarchaeota	NC_003551.1	MK0901	III	98
Methanosaeta thermophila PT	Euryarchaeota	NC_008553.1	Mthe_1264 Mthe_1314	III II	>60
Methanosarcina acetivorans C2A	Euryarchaeota	NC_003552.1	MA0072 MA1665	III II	35-40
Methanosarcina	Euryarchaeota	NC_007355.1	Mbar_A1037	III	35-40

Organism	Group	RefSeq	Locus Tag	Class	Growth T (°C)
barkeri str. Fusaro			Mbar_A3615	II	
Methanosarcina mazei Go1	Euryarchaeota	NC_003901.1	MM_1367 MM_2830	III II	30-40 30-40
Methanosphaera stadtmanae DSM 3091	Euryarchaeota	NC_007681.1	Msp_0254	III	36-40
Methanospirillum hungatei JF-1	Euryarchaeota	NC_007796.1	Mhun_2811	III	37
Methanothermobacter thermautotrophicus str. Delta H	Euryarchaeota	NC_000916.1	MTH652 MTH1539	II III	65-70
Nanoarchaeum equitans Kin4-M	Nanoarchaeota	NC_005213.1	NEQ345	III	>80
Natronomonas pharaonis DSM 2160	Euryarchaeota	NC_007426.1	NP3346A NP4434A NP5066A NP6166A NP6168A	Ib – α Ib – β Ib – β Ia – β Ia – β	37-40
Nitrosopumilus maritimus SCM1	Crenarchaeota	NC_010085.1	Nmar_1627	II	37
Picrophilus torridus DSM 9790	Euryarchaeota	NC_005877.1	COG0209F	II	60
Pyrobaculum aerophilum str. IM2	Crenarchaeota	NC_003364.1	PAE3155	II	100
Pyrobaculum arsenaticum DSM 13514	Crenarchaeota	NC_009376.1	Pars_1670	II	>80
Pyrobaculum calidifontis JCM 11548	Crenarchaeota	NC_009073.1	Pcal_1797	II	90-95
Pyrobaculum islandicum DSM 4184	Crenarchaeota	NC_008701.1	Pisl_0226 Pisl_0525	III II	100
Pyrococcus abyssi	Euryarchaeota	NC_000868.1	PAB2337	III	103

Organism	Group	RefSeq	Locus Tag	Class	Growth T (°C)
GE5			PAB1057	II	
Pyrococcus furiosus DSM 3638	Euryarchaeota	NC_003413.1	PF0440 PF1971	II III	100
Pyrococcus horikoshii OT3	Euryarchaeota	NC_000961.1	PH0021 PH0363	III II	98
Staphylothermus marinus F1	Crenarchaeota	NC_009033.1	Smar_1080	III	92
Sulfolobus acidocaldarius DSM 639	Crenarchaeota	NC_007181.1	Saci_1112 Saci_1353 Saci_2188 Saci_2212 Saci_2220	Ia – β II Ia – β Ia – β Ia – β	70-75
Sulfolobus solfataricus P2	Crenarchaeota	NC_002754.1	SSO0929 SSO2498	II Ia – β	85
Sulfolobus tokodaii str. 7	Crenarchaeota	NC_003106.2	ST1259	II	80
Thermococcus kodakarensis KOD1	Euryarchaeota	NC_006624.1	TK1736 TK2298	II III	85
Thermofilum pendens Hrk 5	Crenarchaeota	NC_008698.1	Tpen_0335	III	88
Thermoplasma acidophilum DSM 1728	Euryarchaeota	NC_002578.1	Ta1475	II	59
Thermoplasma volcanium GSS1	Euryarchaeota	NC_002689.2	TVN0092	II	60
Thermoproteus neutrophilus V24Sta	Crenarchaeota	NC_010525.1	Tneu_0378 Tneu_1608	III II	85
ncultured methanogenic archaeon RC-I	Euryarchaeota	NC_009464.1	RCIX705	III	35-40

Table 1.3 TrRs in completed archaea genomes. Table consists of all completed and annotated archaea genomes as of June 2008. Table includes name of organisms, the archaea subdivision to which they are assigned, the organism reference sequence, and the protein locus tag. All protein, except for the TrRs from *Aeropyrum pernix*, *Pyrococcus horioshii*, and *Sulfolobus solfataricus*, are putative TrRs annotated by sequence homology at the NCBI Entrez Genome Project. Some organisms have multiple TrRs.

Organism	Group	RefSeq	Locus Tag
Aeropyrum pernix K1	Crenarchaeota	NC_000854.2	APE_1061.1
Archaeoglobus fulgidus DSM 4304	Euryarcaeota	NC_009917.1	AF1554
Caldivirga maquilingensis IC-167	Crenarchaeota	NC_009954.1	Cmaq_0211
Candidatus Korarchaeum cryptofilum OPF8	Other Archaea	NC_010482.1	Kcr_1054
Candidatus methanoregula boonei 6A8	Euryarchaeota	NC_009712.1	N / A
Haloarcula marismortui ATCC 43049	Euryarchaeota	NC_006396.1	rrnAC1129 rrnAC1470 rrnAC3545 pNG7078
Halobacterium salinarum R1	Euryarchaeota	NC_010364.1	OE2805R OE6204F OE7085F
Halobacterium sp. NRC-1	Euryarchaeota	NC_002607.1	VNG6074G VNG6452G VNG7055
Haloquadratum walsbyi DSM 16790	Euryarchaeota	NC_008212.1	HQ1592A HQ1652A HQ2445A
Hyperthermus butylicus DSM 5456	Crenarchaeota	NC_008818.1	Hbut_0180
Ignicoccus hospitalis KIN4/I	Crenarchaeota	NC_009776.1	N / A
Metallosphaera sedula DSM 5348	Crenarchaeota	NC_009440.1	Msed_0632
Methanobrevibacter smithii ATCC 35061	Euryarchaeota	NC_009515.1	Msm_034
Methanocaldococcus jannaschii	Euryarchaeota	NC_000909.1	MJ1536

Organism	Group	RefSeq	Locus Tag
DSM 2661			
Methanococcoides burtonii DSM 6242	Euryarchaeota	NC_007955.1	Mbur_0101
Methanococcus aeolicus Nankai-3	Euryarchaeota	NC_009635.1	Maeo_0766
Methanococcus maripaludis C5	Euryarchaeota	NC_009135.1	N / A
Methanococcus maripaludis C6	Euryarchaeota	NC_009975.1	N / A
Methanococcus maripaludis C7	Euryarchaeota	NC_009637.1	N / A
Methanococcus maripaludis S2	Euryarchaeota	NC_005791.1	MMP0959
Methanococcus vannielii SB	Euryarchaeota	NC_009634.1	N / A
Methanocorpusculum labreanum Z	Euryarchaeota	NC_008942.1	N / A
Methanoculleus marisnigri JR1	Euryarchaeota	NC_009051.1	N / A
Methanopyrus kandleri AV19	Euryarchaeota	NC_003551.1	MK1561
Methanosaeta thermophila PT	Euryarchaeota	NC_008553.1	Mthe_1150
Methanosarcina acetivorans C2A	Euryarchaeota	NC_003552.1	MA1368
Methanosarcina barkeri str. Fusaro	Euryarchaeota	NC_007355.1	Mbar_A28987
Methanosarcina mazei Go1	Euryarchaeota	NC_003901.1	MM_2353
Methanosphaera stadtmanae DSM 3091	Euryarchaeota	NC_007681.1	Msp_1413
Methanospirillum hungatei JF-1	Euryarchaeota	NC_007796.1	N / A
Methanothermobacter thermautotrophicus str. Delta H	Euryarchaeota	NC_000916.1	MTH708
Nanoarchaeum equitans Kin4-M	Nanoarchaeota	NC_005213.1	N / A
Natronomonas pharaonis DSM 2160	Euryarchaeota	NC_007426.1	NP1384A NP1388A NP2680A NP3544A
Nitrosopumilus maritimus SCM1	Crenarchaeota	NC_010085.1	Nmar_0672
Picrophilus torridus DSM 9790	Euryarchaeota	NC_005877.1	PTO0431 PTO0734
Pyrobaculum aerophilum str. IM2	Crenarchaeota	NC_003364.1	PAE1725 PAE2744
Pyrobaculum arsenaticum DSM 13514	Crenarchaeota	NC_009376.1	Pars_0776

Organism	Group	RefSeq	Locus Tag
Pyrobaculum calidifontis JCM 11548	Crenarchaeota	NC_009073.1	Pcal_0626
Pyrobaculum islandicum DSM 4184	Crenarchaeota	NC_008701.1	Pisl_1704
Pyrococcus abyssi GE5	Euryarchaeota	NC_000868.1	PAB0500
Pyrococcus furiosus DSM 3638	Euryarchaeota	NC_003413.1	PF1422
Pyrococcus horikoshii OT3	Euryarchaeota	NC_000961.1	PH1426
Staphylothermus marinus F1	Crenarchaeota	NC_009033.1	N / A
Sulfolobus acidocaldarius DSM 639	Crenarchaeota	NC_007181.1	Saci_0029 Saci_093 Saci_1169 Saci_2144
Sulfolobus solfataricus P2	Crenarchaeota	NC_002754.1	SSO222 SSO2416 SSO2765
Sulfolobus tokodaii str. 7	Crenarchaeota	NC_003106.2	ST0438 ST0546 ST2133
Thermococcus kodakarensis KOD1	Euryarchaeota	NC_006624.1	TK2100
Thermofilum pendens Hrk 5	Crenarchaeota	NC_008698.1	N / A
Thermoplasma acidophilum DSM 1728	Euryarchaeota	NC_002578.1	Ta0984
Thermoplasma volcanium GSS1	Euryarchaeota	NC_002689.2	TVN0470 TVN1129
Thermoproteus neutrophilus V24Sta	Crenarchaeota	NC_010525.1	N / A
ncultured methanogenic archaeon RC-I	Euryarchaeota	NC_009464.1	RRC321 RRC232

Table 1.4 Trs in completed archaea genomes. Table consists of all completed and annotated Archaea genomes as of June 2008. Table includes name of organisms, the archaea to subdivision they are assigned, the organism reference sequence, and the protein locus tag. All protein, except for the Trs from *Aeropyrum pernix* and *Sulfolobus solfataricus*, are putative Trs annotated by sequence homology at the NCBI Entrez Genome Project. Some organisms have multiple Trs.

Organism	Group	RefSeq	Locus Tag
Aeropyrum pernix K1	Crenarchaeota	NC_000854.2	APE_0641
			APE_1722
			APE_2273.1
Archaeoglobus fulgidus DSM 4304	Euryarcaeota	NC_009917.1	AF0711
			AF0769
			AF1284
			AF2144
Caldivirga maquilingensis IC-167	Crenarchaeota	NC_009954.1	Mboo_0072
			Mb00_0097
			Mboo_1798
			MBoo_2203
Candidatus Korarchaeum cryptofilum OPF8	Other Archaea	NC_010482.1	Kcr_0134
Candidatus methanoregula boonei 6A8	Euryarchaeota	NC_009712.1	Mboo_0022
			Mboo_1709
Haloarcula marismortui ATCC 43049	Euryarchaeota	NC_006396.1	rrnAC1371
			rrnAC2110
			rrnAC2141
			rrnAC2458
			rrnAC3126
			rrnB027
			pNG6029
			pNG7124
Halobacterium salinarum R1	Euryarchaeota	NC_010364.1	OE1252R
			OE1277F
			OE2862F
			OE4648F

Organism	Group	RefSeq	Locus Tag
			OE6204F
			OE7084F
Halobacterium sp. NRC-1	Euryarchaeota	NC_002607.1	VNG6073G
			VNG6453G
			VNG7055
Haloquadratum walsbyi DSM 16790	Euryarchaeota	NC_008212.1	HQ1323A
			HQ1614A
			HQ1629A
			HQ2511A
			HQ2557A
Hyperthermus butylicus DSM 5456	Crenarchaeota	NC_008818.1	Hbut_0700
			Hbut_1122
Ignicoccus hospitalis KIN4/I	Crenarchaeota	NC_009776.1	N / A
Metallosphaera sedula DSM 5348	Crenarchaeota	NC_009440.1	Msed_0776
			Msed_1614
Methanobrevibacter smithii ATCC 35061	Euryarchaeota	NC_009515.1	Msm_083
Methanocaldococcus jannaschii DSM 2661	Euryarchaeota	NC_000909.1	MJ0307
Methanococcoides burtonii DSM 6242	Euryarchaeota	NC_007955.1	Mbur_0227
			Mbur_0334
			Mbur1670
			Mbur_1851
			Mbur_2325
Methanococcus aeolicus Nankai-3	Euryarchaeota	NC_009635.1	Maeo_0589
			Mbur_0334
			Maeo_0589
			Mbur_0334
			Mbur1670
			Mbur_1851
			Mbur_1851
Methanococcus maripaludis C5	Euryarchaeota	NC_009135.1	N / A
Methanococcus maripaludis C6	Euryarchaeota	NC_009975.1	MmarC6_1697
Methanococcus maripaludis C7	Euryarchaeota	NC_009637.1	N / A

Organism	Group	RefSeq	Locus Tag
Methanococcus maripaludis S2	Euryarchaeota	NC_005791.1	MMP1635
Methanococcus vannielii SB	Euryarchaeota	NC_009634.1	N / A
Methanocorpusculum labreanum Z	Euryarchaeota	NC_008942.1	N / A
Methanoculleus marisnigri JR1	Euryarchaeota	NC_009051.1	Memar_0227 Memar_1658
Methanopyrus kandleri AV19	Euryarchaeota	NC_003551.1	N / A
Methanosaeta thermophila PT	Euryarchaeota	NC_008553.1	Mthe_0711 Mthe_1133
Methanosarcina acetivorans C2A	Euryarchaeota	NC_003552.1	MA09651 MA3212 MA3702 MA4254
Methanosarcina barkeri str. Fusaro	Euryarchaeota	NC_007355.1	Mbar_A1859 Mbar_A2283 Mbar_A2310 Mbar_A2577 Mbar_A3617
Methanosarcina mazei Go1	Euryarchaeota	NC_003901.1	MM_0436 MM_0706 MM_2079 MM_2354
Methanosphaera stadtmanae DSM 3091	Euryarchaeota	NC_007681.1	Msp_0103 Msp_0820
Methanospirillum hungatei JF-1	Euryarchaeota	NC_007796.1	Mhun_0815 Mhun_2493 Mhun_3175
Methanothermobacter thermautotrophicus str. Delta H	Euryarchaeota	NC_000916.1	MTH807
Nanoarchaeum equitans Kin4-M	Nanoarchaeota	NC_005213.1	N / A
Natronomonas pharaonis DSM 2160	Euryarchaeota	NC_007426.1	NP1310A NP1448A NP1606A NP3114A

Organism	Group	RefSeq	Locus Tag
			NP3704A NP3914A NP4120A
Nitrosopumilus maritimus SCM1	Crenarchaeota	NC_010085.1	Nmar_0210 Nmar_0230
Picrophilus torridus DSM 9790	Euryarchaeota	NC_005877.1	PTO0121
Pyrobaculum aerophilum str. IM2	Crenarchaeota	NC_003364.1	N / A
Pyrobaculum arsenaticum DSM 13514	Crenarchaeota	NC_009376.1	Pars_1573
Pyrobaculum calidifontis JCM 11548	Crenarchaeota	NC_009073.1	Pcal_1702
Pyrobaculum islandicum DSM 4184	Crenarchaeota	NC_008701.1	Pisl_0610
Pyrococcus abyssi GE5	Euryarchaeota	NC_000868.1	N / A
Pyrococcus furiosus DSM 3638	Euryarchaeota	NC_003413.1	PF1457
Pyrococcus horikoshii OT3	Euryarchaeota	NC_000961.1	N / A
Staphylothermus marinus F1	Crenarchaeota	NC_009033.1	Smar_0032 Smar_0093
Sulfolobus acidocaldarius DSM 639	Crenarchaeota	NC_007181.1	Saci_0041 Saci_1485 Saci_1823
Sulfolobus solfataricus P2	Crenarchaeota	NC_002754.1	SSO0368 SSO2232
Sulfolobus tokodaii str. 7	Crenarchaeota	NC_003106.2	STS171 ST2123
Thermococcus kodakarensis KOD1	Euryarchaeota	NC_006624.1	TK0838
Thermofilum pendens Hrk 5	Crenarchaeota	NC_008698.1	Tpen_0425 Tpen_1129
Thermoplasma acidophilum DSM 1728	Euryarchaeota	NC_002578.1	Ta0866
Thermoplasma volcanium GSS1	Euryarchaeota	NC_002689.2	TVN0777

Organism	Group	RefSeq	Locus Tag
Thermoproteus neutrophilus V24Sta	Crenarchaeota	NC_010525.1	Tneu_1511
ncultured methanogenic archaeon RC-I	Euryarchaeota	NC_009464.1	RCIX1080 RCIX1079 RCIX360 LRC226 LRC131

Figure 1.1 Structures of Class Ia RNR from *E. coli*. (A) Ribbon representation of R1 dimer showing N-terminal activity site in blue (blue oval), specificity site in magenta (magenta oval), and dimeric interface 4-helix bundle in cyan (PDB 4r1r). The structure has GDP bound in the active site (black oval) and dTTP in the specificity site (magenta oval). Inset highlights the placement of the effector atop the dimer interface. (B) Ribbon representation of R2 dimer (PDB 1rib). Orange circles indicate diferric sites in R2 dimer. Red indicates α-helices, yellow indicates β-strands, and green indicates loop regions. Models are not to scale.

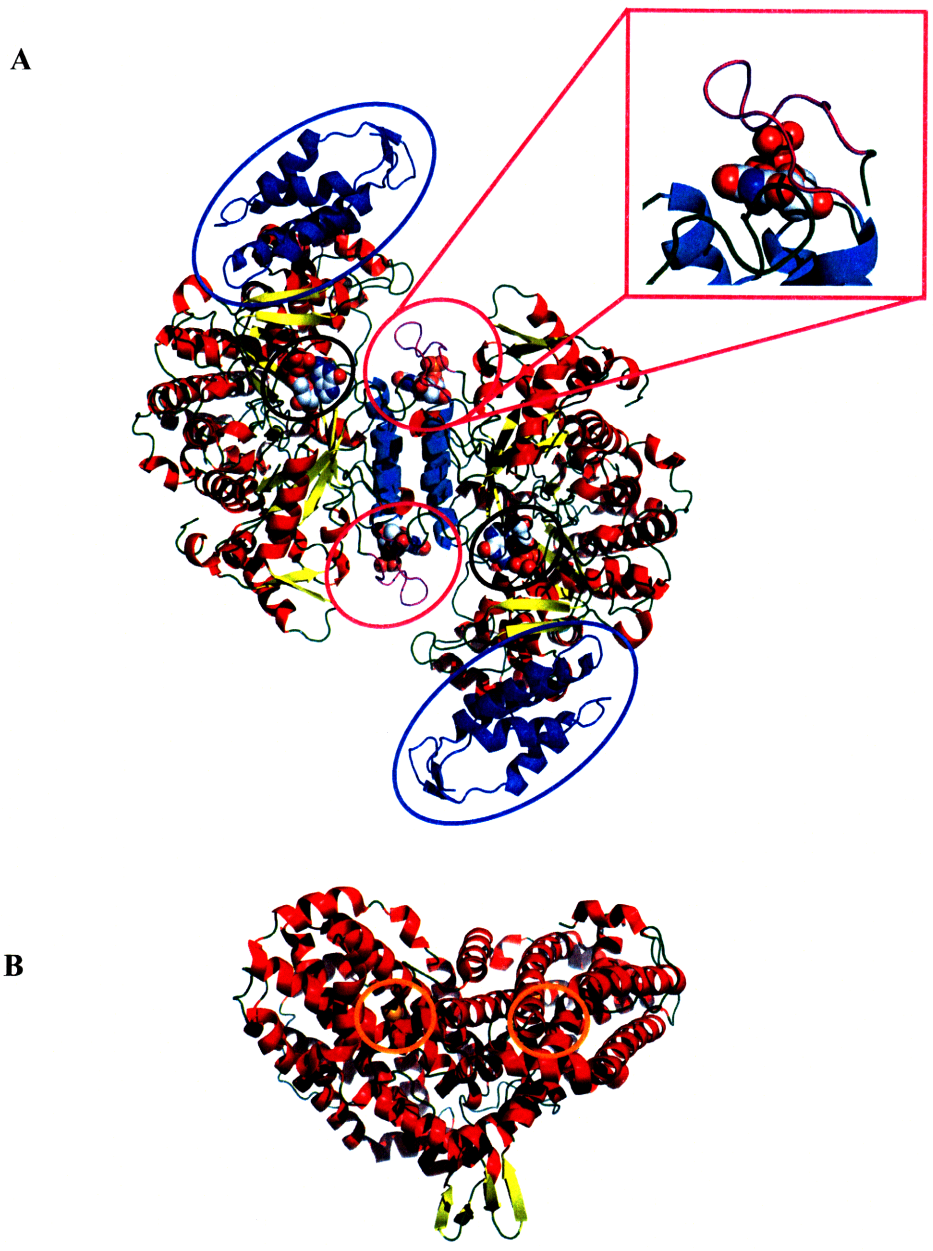


Figure 1.2 Structures of Class II RNRs. (A) Ribbon representation of RTPR monomer showing the conserved specificity site in magenta, and 4-helix bundle in cyan (PDB 1111). (B) Ribbon representation of the Class II RNR from *Thermotoga maritima* with the 4-helix bundle dimeric interface in cyan and the specificity site in magenta (PDB 1xjg). Unfortunately, the *T. maritima* structure is missing the 182 C-terminal amino acids. Red indicates α-helices, yellow indicates β-strands, and green indicates loop regions. Active site is circled in black and specificity site is highlighted by magenta oval. Models are not to scale.

A

B

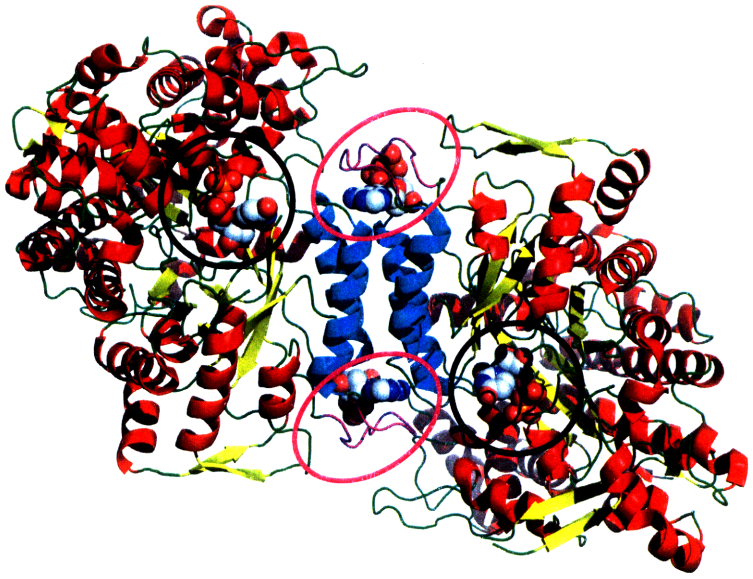


Figure 1.3 Structure of the bacteriophage T4 Class III RNR (PDB 1h7a). Ribbon representation of the Class III RNR showing the specificity site in magenta (magenta circle), and dimeric interface 4-helix bundle in cyan. Red indicates α-helices, yellow indicates β-strands, and green indicates loop regions.

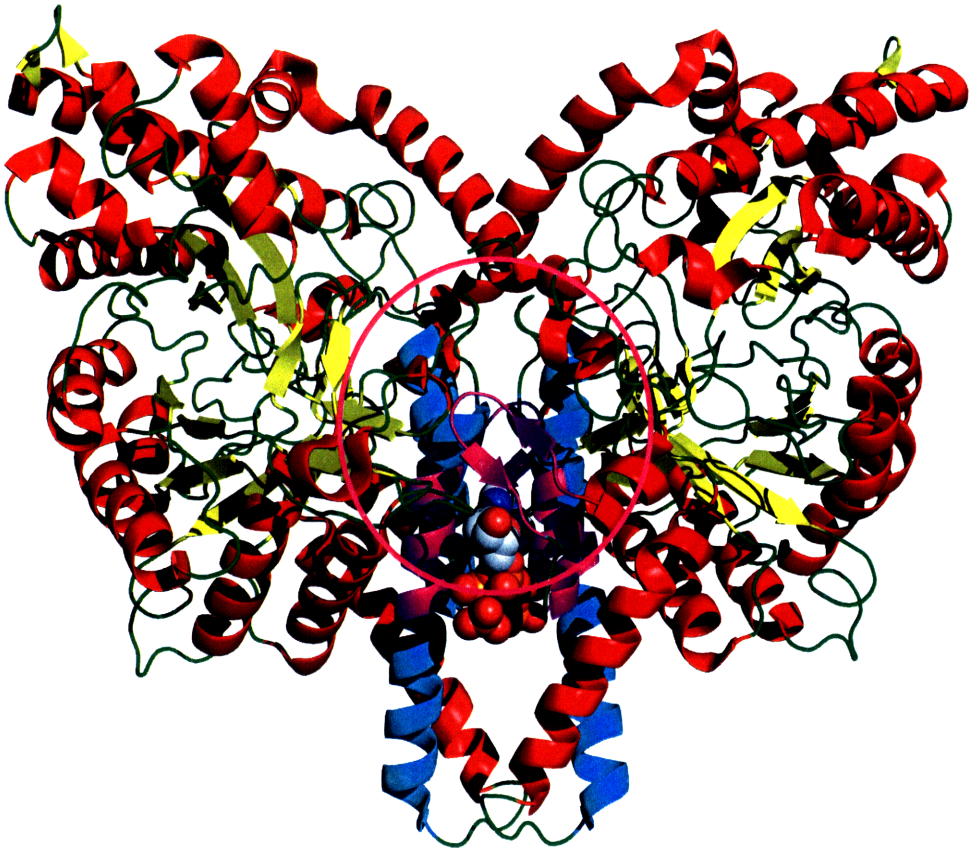


Figure 1.4 Ribbon diagram of the structurally conserved RNR β-barrel. Top and side views of the 10-strand barrel active site. The top view of the barrel shows the finger loop which contains the active site cysteine. (A) *E. coli* RNR, (B) RTPR (C) bacteriophage T4 RNR.

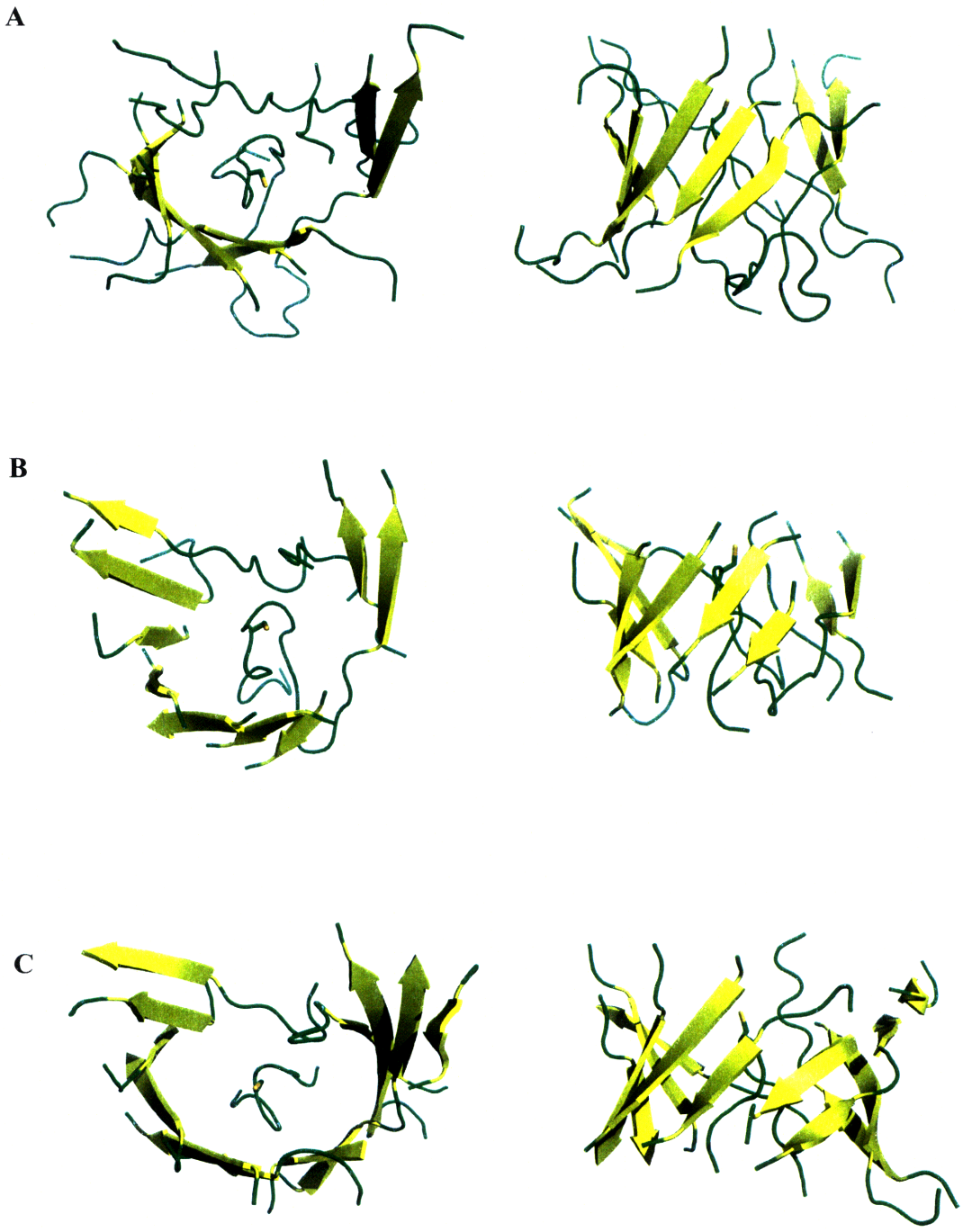


Figure 1.5 Sequence alignment of Class II RNRs showing the conserved motif, NPC(G/A)EX$_8$C(N/V)L, between the active thiyl radical cysteine (C436) and one of the redox active cysteines, (C447). The two cysteines are highlighted by a star at the bottom of the alignment. The spacing between these two cysteines varies between 12 and 15 amino acids for Class I RNRs.

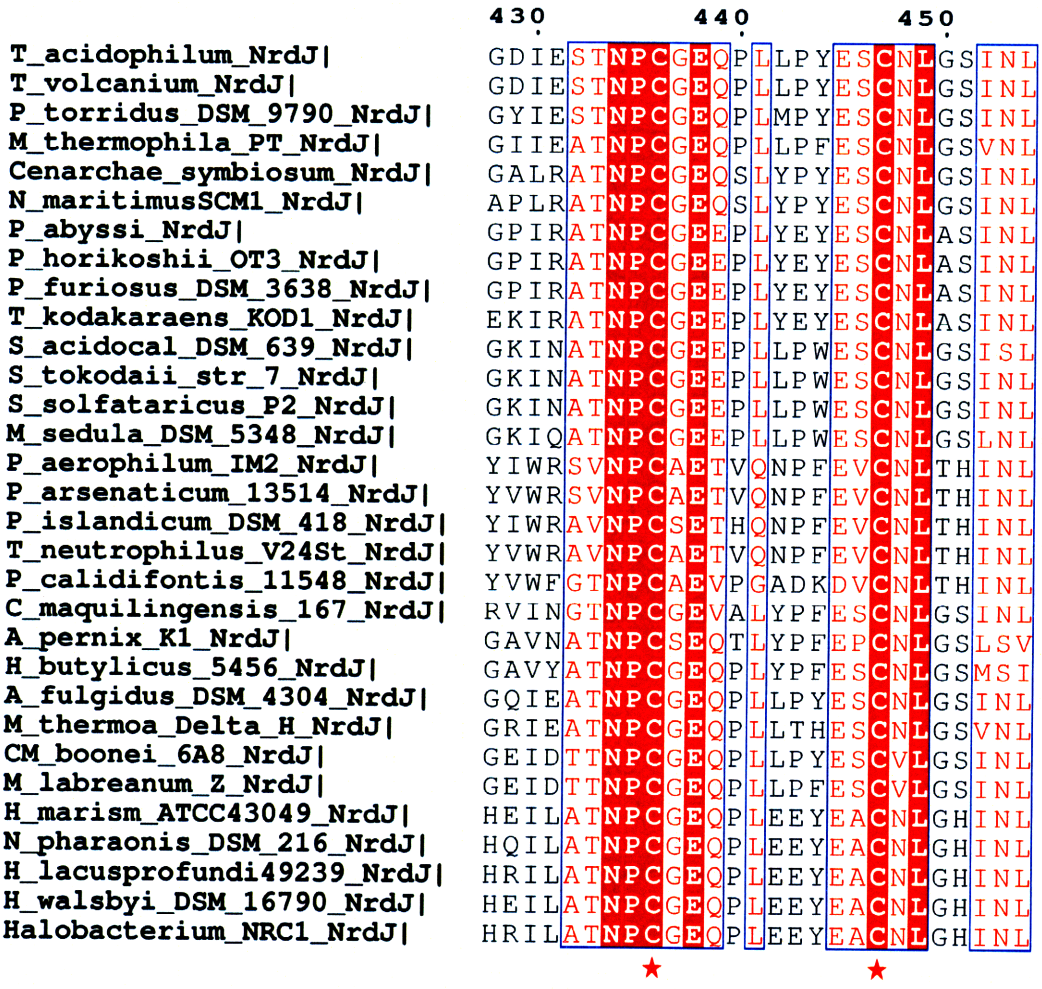


Figure 1.6 Structural locations of the conserved class-dependent motifs. (A) The class I ($Cx_{22}C$ in *E. coli*) motif (black oval mapped onto the *E. coli* structure) comprises residues which extend from the barrel to the dimeric interface. (B) The class II $Cx_{10}C$ motif (black circle) ends at the bottom of the *T. maritima* barrel. The dimer interface is colored in cyan, the effector binding loop is colored in magenta, α-helices are colored in red, β-sheets are colored in yellow, and the motif loops are colored in orange.

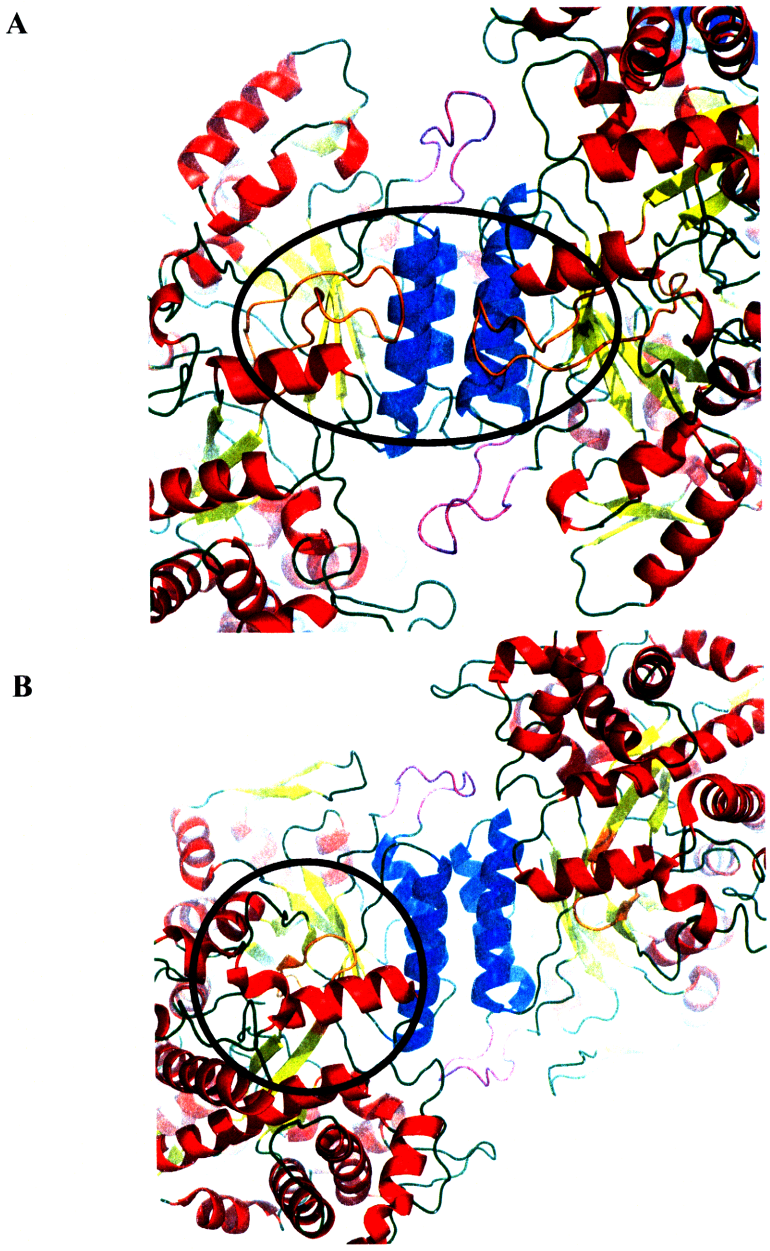


Figure 1.7 Structures of the FO and FR form of TrR from *E. coli* (PDB 1tdf (FO) and 1f6m (FR)). (A) Structure of the FR conformation of *E. coli* TrR. (B) Structure of the FO conformation of *E. coli* TrR. The circle highlights the movement in the NADPH binding domain relative to the FAD binding domain. The α-helices are colored in red and β-sheets are colored in yellow. Figures are not to scale.

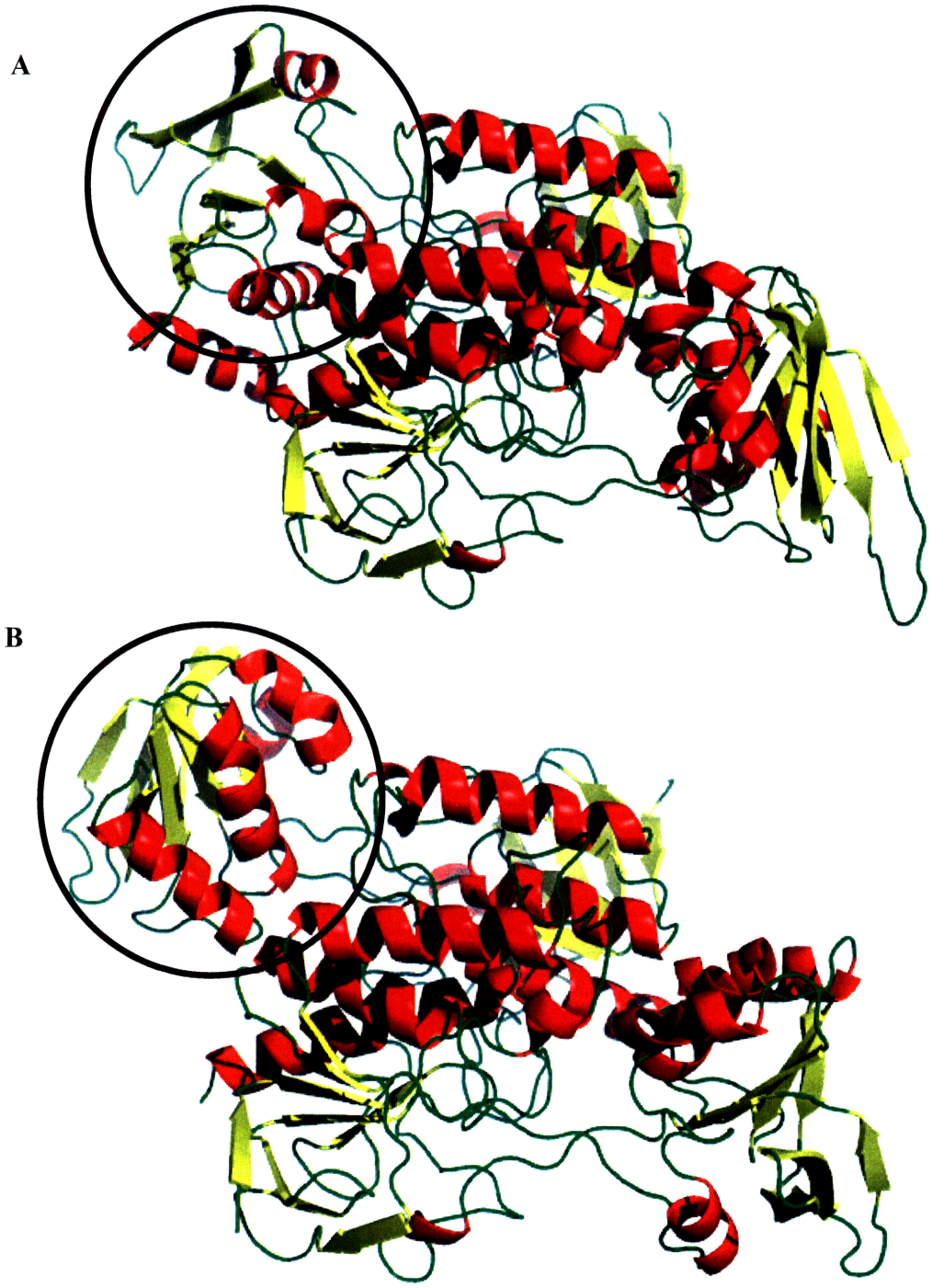


Figure 1.8 Structure of the oxidized form of Tr from *E. coli* (PDB 2trx). The disulfide is highlighted in the black circle. The α-helices are colored in red and β-sheets are colored in yellow.

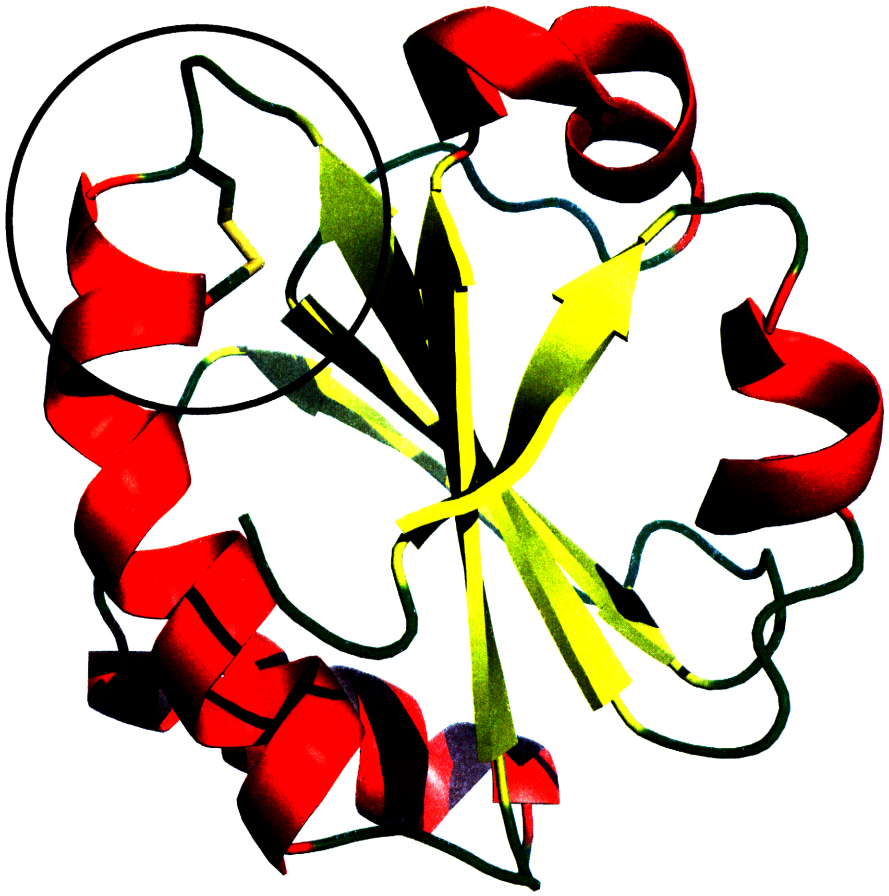
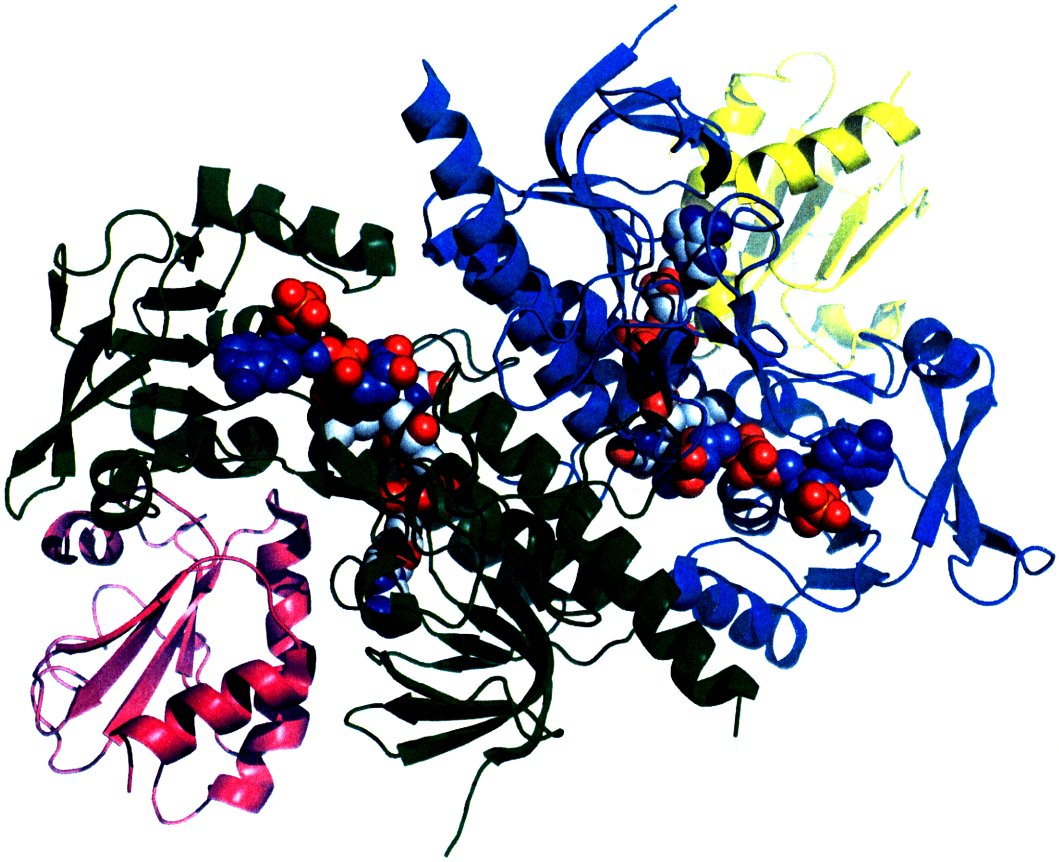


Figure 1.9 Structure of the complex of Tr and TrR from *E. coli* (PDB 1f6m) with the TrR dimer is colored cyan and green and the Tr monomers are yellow and magenta. The structure was solved with FAD and $AADP^+$, an NADPH analogue.



REFERENCES

1. Mathews, C. K.,(2006) DNA precursor metabolism and genomic stability. *Faseb J* 20, (9), 1300-1314.

2. Mathews, C. K., Ji, J.,(1992) DNA precursor asymmetries, replication fidelity, and variable genome evolution. *Bioessays* 14, (5), 295-301.

3. Meuth, M.,(1989) The molecular basis of mutations induced by deoxyribonucleoside triphosphate pool imbalances in mammalian cells. *Exp Cell Res* 181, (2), 305-316.

4. Herrick, J., Sclavi, B.,(2007) Ribonucleotide reductase and the regulation of DNA replication: an old story and an ancient heritage. *Mol Microbiol* 63, (1), 22-34.

5. Nordlund, P., Reichard, P.,(2006) Ribonucleotide reductases. *Annu Rev Biochem* 75, 681-706.

6. Kolberg, M., Strand, K. R., Graff, P., Andersson, K. K.,(2004) Structure, function, and mechanism of ribonucleotide reductases. *Biochim Biophys Acta* 1699, (1-2), 1-34.

7. Stubbe, J.,(2000) Ribonucleotide reductases: the link between an RNA and a DNA world? *Curr Opin Struct Biol* 10, (6), 731-736.

8. Stubbe, J., Ge, J., Yee, C. S.,(2001) The evolution of ribonucleotide reduction revisited. *Trends Biochem Sci* 26, (2), 93-99.

9. Tauer, A., Benner, S. A.,(1997) The B12-dependent ribonucleotide reductase from the archaebacterium Thermoplasma acidophila: an evolutionary solution to the ribonucleotide reductase conundrum. *Proc Natl Acad Sci U S A* 94, (1), 53-58.

10. Sintchak, M. D., Arjara, G., Kellogg, B. A., Stubbe, J., Drennan, C. L.,(2002) The crystal structure of class II ribonucleotide reductase reveals how an allosterically regulated monomer mimics a dimer. *Nat Struct Biol* 9, (4), 293-300.

11. Cerqueira, N. M., Fernandes, P. A., Ramos, M. J.,(2007) Ribonucleotide reductase: a critical enzyme for cancer chemotherapy and antiviral agents. *Recent Patents Anticancer Drug Discov* 2, (1), 11-29.

12. Shao, J., Zhou, B., Chu, B., Yen, Y.,(2006) Ribonucleotide reductase inhibitors and future drug design. *Curr Cancer Drug Targets* 6, (5), 409-431.

13. Pereira, S., Fernandes, P. A., Ramos, M. J.,(2004) Theoretical study of ribonucleotide reductase mechanism-based inhibition by 2'-azido-2'-deoxyribonucleoside 5'-diphosphates. *J Comput Chem* 25, (2), 227-237.

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

15. Orr, M. D., Vitols, E.,(1966) Thioredoxin from Lactobacillus leichmannii and its role as hydrogen donor for ribonucleoside triphosphate reductase. *Biochem Biophys Res Commun* 25, (1), 109-115.

16. Reichard, P.,(1962) Enzymatic synthesis of deoxyribonucleotides. I. Formation of deoxycytidine diphosphate from cytidine diphosphate with enzymes from Escherichia coli. *J Biol Chem* 237, 3513-3519.

17. Sjoberg, B. M., Reichard, P.,(1977) Nature of the free radical in ribonucleotide reductase from Escherichia coli. *J Biol Chem* 252, (2), 536-541.

18. Ehrenberg, A., Reichard, P.,(1972) Electron spin resonance of the iron-containing protein B2 from ribonucleotide reductase. *J Biol Chem* 247, (11), 3485-3488.

19. Uhlin, U., Eklund, H.,(1994) Structure of ribonucleotide reductase protein R1. *Nature* 370, (6490), 533-539.

20. Nordlund, P., Sjoberg, B. M., Eklund, H.,(1990) Three-dimensional structure of the free radical protein of ribonucleotide reductase. *Nature* 345, (6276), 593-598.

21. Blakley, R. L., Barker, H. A.,(1964) Cobamide stimulation of the reduction of ribotides to deoxyribotides in Lactobacillus leichmannii. *Biochem Biophys Res Commun* 16, (5), 391-397.

22. Vitols, E., Brownson, C., Gardiner, W., Blakley, R. L.,(1967) Cobamides and ribonucleotide reduction. V. A kinetic study of the ribonucleoside triphosphate reductase of Lactobacillus leichmannii. *J Biol Chem* 242, (13), 3035-3041.

23. Hogenkamp, H. P., Ghambeer, R. K., Brownson, C., Blakley, R. L., Vitols, E.,(1968) Cobamides and ribonucleotide reduction. VI. Enzyme-catalyzed hydrogen exchange between water and deoxyadenosylcobalamin. *J Biol Chem* 243, (4), 799-808.

24. Sando, G. N., Blakley, R. L., Hogenkamp, H. P., Hoffmann, P. J.,(1975) Studies on the mechanism of adenosylcobalamin-dependent ribonucleotide reduction by the use of analogs of the coenzyme. *J Biol Chem* 250, (22), 8774-8779.

25. Larsson, K. M., Jordan, A., Eliasson, R., Reichard, P., Logan, D. T., Nordlund, P.,(2004) Structural mechanism of allosteric substrate specificity regulation in a ribonucleotide reductase. *Nat Struct Mol Biol* 11, (11), 1142-1149.

26. Fontecave, M., Eliasson, R., Reichard, P.,(1989) Oxygen-sensitive ribonucleoside triphosphate reductase is present in anaerobic Escherichia coli. *Proc Natl Acad Sci U S A* 86, (7), 2147-2151.

27. Ollagnier, S., Mulliez, E., Gaillard, J., Eliasson, R., Fontecave, M., Reichard, P.,(1996) The anaerobic Escherichia coli ribonucleotide reductase. Subunit structure and iron sulfur center. *J Biol Chem* 271, (16), 9410-9416.

28. Ollagnier, S., Mulliez, E., Schmidt, P. P., Eliasson, R., Gaillard, J., Deronzier, C., Bergman, T., Graslund, A., Reichard, P., Fontecave, M.,(1997) Activation of the anaerobic ribonucleotide reductase from Escherichia coli. The essential role of the iron-sulfur center for S-adenosylmethionine reduction. *J Biol Chem* 272, (39), 24216-24223.

29. Tamarit, J., Mulliez, E., Meier, C., Trautwein, A., Fontecave, M.,(1999) The anaerobic ribonucleotide reductase from Escherichia coli. The small protein is an activating enzyme containing a [4fe-4s](2+) center. *J Biol Chem* 274, (44), 31291-31296.

30. Tamarit, J., Gerez, C., Meier, C., Mulliez, E., Trautwein, A., Fontecave, M.,(2000) The activating component of the anaerobic ribonucleotide reductase from Escherichia coli. An iron-sulfur center with only three cysteines. *J Biol Chem* 275, (21), 15669-15675.

31. Logan, D. T., Andersson, J., Sjoberg, B. M., Nordlund, P.,(1999) A glycyl radical site in the crystal structure of a class III ribonucleotide reductase. *Science* 283, (5407), 1499-1504.

32. Brown, N. C., Reichard, P.,(1969) Role of effector binding in allosteric control of ribonucleoside diphosphate reductase. *J Mol Biol* 46, (1), 39-55.

33. Jordan, A., Pontis, E., Atta, M., Krook, M., Gibert, I., Barbe, J., Reichard, P.,(1994) A second class I ribonucleotide reductase in Enterobacteriaceae:

characterization of the Salmonella typhimurium enzyme. *Proc Natl Acad Sci U S A* 91, (26), 12892-12896.

34. Jiang, W., Yun, D., Saleh, L., Barr, E. W., Xing, G., Hoffart, L. M., Maslak, M. A., Krebs, C., Bollinger, J. M., Jr.,(2007) A manganese(IV)/iron(III) cofactor in Chlamydia trachomatis ribonucleotide reductase. *Science* 316, (5828), 1188-1191.

35. Chen, A. K., Bhan, A., Hopper, S., Abrams, R., Franzen, J. S.,(1974) Substrate and effector binding to ribonucleoside triphosphate reductase of Lactobacillus leichmannii. *Biochemistry* 13, (4), 654-661.

36. Eliasson, R., Pontis, E., Sun, X., Reichard, P.,(1994) Allosteric control of the substrate specificity of the anaerobic ribonucleotide reductase from Escherichia coli. *J Biol Chem* 269, (42), 26052-26057.

37. Gleason, F. K., Olszewski, N. E.,(2002) Isolation of the gene for the B12-dependent ribonucleotide reductase from Anabaena sp. strain PCC 7120 and expression in Escherichia coli. *J Bacteriol* 184, (23), 6544-6550.

38. Torrents, E., Trevisiol, C., Rotte, C., Hellman, U., Martin, W., Reichard, P.,(2006) Euglena gracilis ribonucleotide reductase: the eukaryote class II enzyme and the possible antiquity of eukaryote B12 dependence. *J Biol Chem* 281, (9), 5604-5611.

39. Stubbe, J.,(1990) Ribonucleotide reductases: amazing and confusing. *J Biol Chem* 265, (10), 5329-5332.

40. Ashley, G. W., Harris, G., Stubbe, J.,(1986) The mechanism of Lactobacillus leichmannii ribonucleotide reductase. Evidence for 3' carbon-hydrogen bond cleavage and a unique role for coenzyme B12. *J Biol Chem* 261, (9), 3958-3964.

41. Eriksson, M., Uhlin, U., Ramaswamy, S., Ekberg, M., Regnstrom, K., Sjoberg, B. M., Eklund, H.,(1997) Binding of allosteric effectors to ribonucleotide reductase protein R1: reduction of active-site cysteines promotes substrate binding. *Structure* 5, (8), 1077-1092.

42. Borovok, I., Kreisberg-Zakarin, R., Yanko, M., Schreiber, R., Myslovati, M., Aslund, F., Holmgren, A., Cohen, G., Aharonowitz, Y.,(2002) Streptomyces spp. contain class Ia and class II ribonucleotide reductases: expression analysis of the genes in vegetative growth. *Microbiology* 148, (Pt 2), 391-404.

43. Arner, E. S., Holmgren, A.,(2000) Physiological functions of thioredoxin and thioredoxin reductase. *Eur J Biochem* 267, (20), 6102-6109.

44. Mulliez, E., Ollagnier, S., Fontecave, M., Eliasson, R., Reichard, P.,(1995) Formate is the hydrogen donor for the anaerobic ribonucleotide reductase from Escherichia coli. *Proc Natl Acad Sci U S A* 92, (19), 8759-8762.

45. Lillig, C. H., Holmgren, A.,(2007) Thioredoxin and related molecules--from biology to health and disease. *Antioxid Redox Signal* 9, (1), 25-47.

46. Martin, J. L.,(1995) Thioredoxin--a fold for all reasons. *Structure* 3, (3), 245-250.

47. Williams, C. H., Arscott, L. D., Muller, S., Lennon, B. W., Ludwig, M. L., Wang, P. F., Veine, D. M., Becker, K., Schirmer, R. H.,(2000) Thioredoxin reductase two modes of catalysis have evolved. *Eur J Biochem* 267, (20), 6110-6117.

48. Holmgren, A., Soderberg, B. O., Eklund, H., Branden, C. I.,(1975) Three-dimensional structure of Escherichia coli thioredoxin-S2 to 2.8 A resolution. *Proc Natl Acad Sci U S A* 72, (6), 2305-2309.

49. Waksman, G., Krishna, T. S., Williams, C. H., Jr., Kuriyan, J.,(1994) Crystal structure of Escherichia coli thioredoxin reductase refined at 2 A resolution. Implications for a large conformational change during catalysis. *J Mol Biol* 236, (3), 800-816.

50. Mulrooney, S. B., Williams, C. H., Jr.,(1997) Evidence for two conformational states of thioredoxin reductase from Escherichia coli: use of intrinsic and extrinsic quenchers of flavin fluorescence as probes to observe domain rotation. *Protein Sci* 6, (10), 2188-2195.

51. Lennon, B. W., Williams, C. H., Jr., Ludwig, M. L.,(2000) Twists in catalysis: alternating conformations of Escherichia coli thioredoxin reductase. *Science* 289, (5482), 1190-1194.

52. Holmgren, A., Johansson, C., Berndt, C., Lonn, M. E., Hudemann, C., Lillig, C. H.,(2005) Thiol redox control via thioredoxin and glutaredoxin systems. *Biochem Soc Trans* 33, (Pt 6), 1375-1377.

53. Powis, G., Kirkpatrick, D. L.,(2007) Thioredoxin signaling as a target for cancer therapy. *Curr Opin Pharmacol* 7, (4), 392-397.

54. Lemaire, S. D., Michelet, L., Zaffagnini, M., Massot, V., Issakidis-Bourguet, E.,(2007) Thioredoxins in chloroplasts. *Curr Genet* 51, (6), 343-365.

55. Woese, C. R., Fox, G. E.,(1977) The concept of cellular evolution. *J Mol Evol* 10, (1), 1-6.

56. Balch, W. E., Magrum, L. J., Fox, G. E., Wolfe, R. S., Woese, C. R.,(1977) An ancient divergence among the bacteria. *J Mol Evol* 9, (4), 305-311.

57. Robertson, C. E., Harris, J. K., Spear, J. R., Pace, N. R.,(2005) Phylogenetic diversity and ecology of environmental Archaea. *Curr Opin Microbiol* 8, (6), 638-642.

58. Lange, M., Ahring, B. K.,(2001) A comprehensive study into the molecular methodology and molecular biology of methanogenic Archaea. *FEMS Microbiol Rev* 25, (5), 553-571.

59. Faber, K., Kroutil, W.,(2005) New enzymes for biotransformations. *Curr Opin Chem Biol* 9, (2), 181-187.

60. Spormann, A. M., Widdel, F.,(2000) Metabolism of alkylbenzenes, alkanes, and other hydrocarbons in anaerobic bacteria. *Biodegradation* 11, (2-3), 85-105.

61. Hartzell, P. L., Millstein, J., LaPaglia, C.,(1999) Biofilm formation in hyperthermophilic Archaea. *Methods Enzymol* 310, 335-349.

62. Reysenbach, A. L., Shock, E.,(2002) Merging genomes with geochemistry in hydrothermal ecosystems. *Science* 296, (5570), 1077-1082.

63. Pikuta, E. V., Hoover, R. B., Tang, J.,(2007) Microbial extremophiles at the limits of life. *Crit Rev Microbiol* 33, (3), 183-209.

64. Ramette, A., Tiedje, J. M.,(2007) Biogeography: an emerging cornerstone for understanding prokaryotic diversity, ecology, and evolution. *Microb Ecol* 53, (2), 197-207.

65. Kelman, Z., White, M. F.,(2005) Archaeal DNA replication and repair. *Curr Opin Microbiol* 8, (6), 669-676.

66. Grogan, D. W.,(2004) Stability and repair of DNA in hyperthermophilic Archaea. *Curr Issues Mol Biol* 6, (2), 137-144.

67. Majernik, A. I., Jenkinson, E. R., Chong, J. P.,(2004) DNA replication in thermophiles. *Biochem Soc Trans* 32, (Pt 2), 236-239.

68. Eliasson, R., Pontis, E., Jordan, A., Reichard, P.,(1999) Allosteric control of three B12-dependent (class II) ribonucleotide reductases. Implications for the evolution of ribonucleotide reduction. *J Biol Chem* 274, (11), 7182-7189.

69. Ladenstein, R., Ren, B.,(2006) Protein disulfides and protein disulfide oxidoreductases in hyperthermophiles. *Febs J* 273, (18), 4170-4185.

70. Pedone, E., Limauro, D., Bartolucci, S.,(2008) The machinery for oxidative protein folding in thermophiles. *Antioxid Redox Signal* 10, (1), 157-169.

71. Jeon, S. J., Ishikawa, K.,(2002) Identification and characterization of thioredoxin and thioredoxin reductase from Aeropyrum pernix K1. *Eur J Biochem* 269, (22), 5423-5430.

72. Kashima, Y., Ishikawa, K.,(2003) A hyperthermostable novel protein-disulfide oxidoreductase is reduced by thioredoxin reductase from hyperthermophilic archaeon Pyrococcus horikoshii. *Arch Biochem Biophys* 418, (2), 179-185.

73. Kashima, Y., Ishikawa, K.,(2003) Alkyl hydroperoxide reductase dependent on thioredoxin-like protein from Pyrococcus horikoshii. *J Biochem* 134, (1), 25-29.

74. Arcari, P., Masullo, L., Masullo, M., Catanzano, F., Bocchini, V.,(2000) A NAD(P)H oxidase isolated from the archaeon Sulfolobus solfataricus is not homologous with another NADH oxidase present in the same microorganism. Biochemical characterization of the enzyme and cloning of the encoding gene. *J Biol Chem* 275, (2), 895-900.

75. Grimaldi, P., Ruocco, M. R., Lanzotti, M. A., Ruggiero, A., Ruggiero, I., Arcari, P., Vitagliano, L., Masullo, M.,(2008) Characterisation of the components of the thioredoxin system in the archaeon Sulfolobus solfataricus. *Extremophiles*.

76. Ruggiero, A., Ruocco, M. R., Grimaldi, P., Arcari, P., Masullo, M., Zagari, A., Vitagliano, L.,(2005) Crystallization and preliminary X-ray crystallographic analysis of Sulfolobus solfataricus thioredoxin reductase. *Acta Crystallogr Sect F Struct Biol Cryst Commun* 61, (Pt 10), 906-909.

77. Ruocco, M. R., Ruggiero, A., Masullo, L., Arcari, P., Masullo, M.,(2004) A 35 kDa NAD(P)H oxidase previously isolated from the archaeon Sulfolobus solfataricus is instead a thioredoxin reductase. *Biochimie* 86, (12), 883-892.

Chapter 2:

Characterization of ribonucleotide reductase from *Archaeoglobus fulgidus*

INTRODUCTION

Ribonucleotide reductases (RNRs) are present in all living organisms. Using radical-based chemistry, they catalyze the reduction of nucleotides to deoxynucleotides, providing the building blocks for DNA replication and repair. Due to their importance, they have been the subjects of several recent reviews (1-3).

Despite their central function, RNRs are not evolutionarily conserved in terms of substrate preference (ribonucleoside di- versus tri-phosphates), metallocofactors, or quaternary structure and sequence. RNRs are named based on substrate preferences: ribonucleoside diphosphate reductases (RDPRs) and ribonucleoside triphosphates reductases (RTPRs). They are also divided into classes based on the cofactor requirements (Scheme 1.1). Class I RNRs, found in eukaryotes, bacteria, bacteriophages, and viruses, utilize a diferric-tyrosyl radical. Class II RNRs, found in bacteria, bacteriophages, algae, and Archaea, utilize coenzyme B_{12} (adenosylcobalamin, AdoCbl). Class III RNRs, found in anaerobic bacteria and bacteriophages, utilize a 4Fe-4S cluster and *S*-adenosylmethionine to generate a glycyl radical. While Class I and III RNRs tend to be higher order oligomeric enzymes with the minimal composition of $\alpha_2\beta_2$ and $\alpha_2\beta$ respectively, Class II RNRs can be monomeric (α) or dimeric (α_2). In terms of substrate preference, all Class I and most Class II RNRs use nucleoside diphosphates (NDPs) as substrates, but there are some Class II enzymes that use nucleoside triphosphates (NTPs) as substrates, and no previously characterized RNR uses both NDPs and NTPs. Class III RNRs use NTPs as substrates. The structural basis for this phosphate preference is unclear (4).

Despite these variations, the overall enzyme mechanisms appear to be conserved (5). For each class of RNR, the metallocofactor is used to generate a thiyl radical in the active site of the α-subunit. While the mechanism of that radical generation may be different, once the radical is generated, the reduction of nucleotides proceeds in a similar manner (Figure 2.1) (6). The first step of the reaction is the abstraction by the thiyl radical of the hydrogen from the 3' position of the substrate. The subsequent multiple step radical rearrangement results in the loss of the 2' hydroxyl group in the form of water. For class III RNR, reducing equivalents for the reaction are provided by oxidation of formate (7). In class I and II RNRs, reducing equivalents are supplied through active site disulfide formation. Re-reduction of the active site disulfide occurs via two

cysteine residues of the C-terminal tail. In turn, the C-terminal tail disulfide is re-reduced in a thioredoxin (Tr)/thioredoxin reductase (TrR)/NADPH-dependent reaction (8, 9).

A unique feature of RNRs is that one enzyme catalyzes the reduction of all four nucleotides to their corresponding deoxynucleotides, and that this substrate specificity is regulated by the binding of deoxynucleoside triphosphates effectors to a distal site on the α-subunits. Somehow, this distant effector binding changes the catalytic site conformation and thus the preference for each ribonucleotide substrate. For the limited number of RNRs studied, the rules of regulation (Scheme 2.1) are similar, but more experiments on a diverse group of RNRs are needed.

In addition to the allosteric regulation of specificity, some RNRs are subject to allosteric regulation of activity. An activity site, found at the N-terminus of Class Ia RNRs, binds either ATP to stimulate overall activity and dATP to inhibit catalytic activity. Some Class I RNRs do not have this site and are referred to as Class Ib enzymes. For a few Class II RNRs, the sequence for an 'activity site' exists but no allosteric regulation of activity has been experimentally demonstrated (10).

Despite the importance of RNR enzymes, there are a relatively small number of them that have been well characterized, and some of the best-studied are from closely related organisms, e.g. *E. coli* and *Salmonella typhimurium*. Most of the work has focused on Class I RNRs while Class II and III remain relatively unexamined. To better understand this family of enzymes, we have cloned and characterized a putative class II RNR from an extremophilic Archaeal organism. *Archaeoglobus fulgidus*, first isolated from a geothermally heated sea floor at Vulcano Island, Italy (11), is an anaerobic motile aquatic autotrophic and an organic sulfate / sulfite respirer whose optimal growth temperature is 83 °C. We find that *A. fulgidus* RNR (*afu*RNR) is a dimeric Class II RNR, which surprisingly shows evidence of activity for both nucleoside diphosphate and nucleoside triphosphate substrates. Thus for this extreme environment, we find the first example of an RNR which exhibits dual selectivity for the phosphorylation state of its substrate.

MATERIALS AND METHODS

Materials PfuTurbo™ DNA Polymerase was obtained from Stratagene (Santa Clara, CA). Genomic DNA was obtained from the American Tissue Culture collection (ATCC). Restriction enzymes, calf intestine phosphatase (CIP), Klenow polymerase fragment, and TAQ polymerase, were purchased from New England Biolabs (NEB, Brunswick, MA). Deoxynucleotides were purchased from Amersham Biosciences (Piscataway, NJ). TOP10, NovaBlue, Rosetta competent cells, and pET28 vector, were purchased from Novagen (San Diego, CA). Zero Blunt™ TOPO™ PCR Cloning Kit was purchased from Invitrogen (Carlsbad, CA). Primers were purchased from Integrated DNA Technologies (IDT) (Coralville, IA). Thrombin (bovine) was purchased from CalBiochem (San Diego, CA). Qiaquick gel extraction minicolumn and Qiagen Miniprep Kits were purchased from Qiagen. SDS-PAGE gels were obtained from Bio-Rad Laboratories (Hercules, CA). All other chemicals used were purchased from Sigma-Aldrich (St. Louis, MO). DNA sequencing was performed at MGH Sequencing Facility and at the MIT Center for Cancer Research Biopolymers Lab (Cambridge, MA). RTPR purified by the Booker method was a kind gift from Professor JoAnne Stubbe at MIT (12).

Analysis of Archaeoglobus fulgidus genome. The genome of *Archaeoglobus fulgidus* (RefSeq NC_000917, GenBank AE000782) was analyzed using the web interface at the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov) for a putative adenosylcobalamin (AdoCbl) ribonucleotide reductase (RNR) open reading frame (ORF). An ORF encoding a 752 amino acid protein, AF1664 was identified by sequence homology to known AdoCbl dependent RNRs using PSI-BLAST (13). Protein sequences were aligned using the program ClustalW (14) and the output was rendered using the web interface ESPript 2.2 (15) (Figure 2.2 and 2.3).

Cloning of afuRNR. The *afu*RNR gene was amplified by the polymerase chain reaction (PCR) using genomic DNA from a cell suspension obtained from ATCC and the primers 5'-ATC GCA TAT GAA GCT CAA TCG TAC TGC TGA AAC T-3' and 5'-TGC TCG AGT TCA ACC GCA CCT CGT AAA CCC GCA G-3' for the putative *afu*RNR gene (restriction sites are underlined). PCR products were cloned into the expression vector pET28a using *Nde*I and *Xho*I restriction enzymes. The sequence of the pET28a-*afu*RNR construct was verified by dideoxy sequencing at the MIT Biopolymers Sequencing Facility. The resulting construct, pET28a-*afu*RNR, was transformed into Rosetta™ (DE3)pLysS cells for facile protein expression.

Purification of recombinant afuRNR. Rosetta™ (DE3)pLysS cells transformed with the pET28a-*afu*RNR construct were used to inoculate TB media supplemented with 20 μg ml^{-1} kanamycin and 34 μg ml^{-1} chloroamphenicol. The culture was grown at 37 ºC to an OD_{600} of 0.8, prior to induction with 0.8 mM isopropylthio-ß-D-galactoside (IPTG) for 4 hrs. The cells were harvested by centrifugation, resuspended in lysis buffer (10 mM TRIS-HCl pH 8.0, 150 mM NaCl, 10 mM imidazole, 1 mM PMSF), and disrupted by sonication using a Digital Sonifier 250 (Branson, CT, USA). Insoluble material was removed by ultracentrifugation at 40,000 x g using a Beckman Avanti J-25 centrifuge (Beckman, CA, USA). The supernatant was then subjected to heating at 70 °C for 45 min, followed by centrifugation at 40,000 x g. The clarified supernatant was loaded directly onto a Ni-NTA column, washed with 1 M NaCl buffer, and eluted with buffer containing 10 mM TRIS-HCl pH 8.0, 150 mM NaCl, 200 mM imidazole. Eluted fractions were analyzed by an SDS-PAGE 4-20% gel and fractions containing protein that migrated at the size of the predicted *afu*RNR sequence were pooled and concentrated using an Amicon stirred ultrafiltration cell (Millipore, MA, USA).

The eluted protein was dialyzed (35 kDa MW cutoff) overnight in 20 mM HEPES, pH 7.5, 20 mM NaCl, 0.5 mM EDTA, 1 mM 1,4-dithio-DL-threitol (DTT) with thrombin (0.05 unit ml^{-1}) to remove the N-terminal poly-His-tag. The resulting cleavage products were loaded onto an SEC-200 high-prep gel-filtration size exclusion column from Amersham Biosciences (Piscataway, NJ, USA), which was pre-equilibrated with storage buffer (10 mM HEPES, pH 7.0, 150 mM NaCl). Fractions were analyzed via SDS-PAGE (Figure 2.4) and those containing protein with an approximate MW ~84,000 kDa were pooled and concentrated using Centriprep YM-10 filter from Amicon (Millipore, Bedford, MA, USA). Protein concentration was determined using a Cary 300 spectrophotometer (Varian, CA, USA) and the theoretical extinction coefficient (ε_{280} = 94,450 M^{-1} cm^{-1}), obtained from ProtParam (ExPaSy web server).

RNR radioactive enzyme assays. RTPR assays were conducted as described by Booker and Stubbe (12) with minor modifications as described below. Briefly each assay mixture contained 25 mM HEPES (pH 7.5), 4 mM EDTA, 1 M NaOAc, 30 mM DTT, 1 mM [2-^{14}C]-CTP (specific activity 1.5 x 10^3 x 3.0 x10^3 cpm/nmol), 8 μM AdoCbl, and 0.48 nmol of RTPR. The final volume was 700 μL. The reaction mixture was incubated in light protective tubes for 5 min at 37

°C, and an aliquot (100 µL) containing everything except AdoCbl was removed at the 0 time point. Each assay was initiated by the addition of AdoCbl, and at various time points, a 100-µL aliquot was removed and quenched in 200 µL of 2% perchloric acid. The contents were neutralized with 180 µL of 0.2 N NaOH and 50 µL of 0.5 M Tris-HC1, pH 8.5, 1.0 mM EDTA. Subsequent to the addition of 10 units of calf intestine phosphatase (CIP), the reaction was incubated at 37 "C for 1.5 h. To each reaction vial 30 µL of carrier cytosine and dC (120 nmol) were added, and the reaction was diluted to 1.5 mL with H_2O. Each reaction was loaded onto 0.75 X 7 cm AGl-X2 columns (borate form, 50-100 mesh) prepared by the method of Steeper and Steuart (1970) (16). Each column was washed with 12 mL of H_2O, and a 1-mL portion of the eluate was subjected to scintillation counting. Changes to the described assay include an increase in the final volume of 700 µL (compared to 510 µL), and that the assays were performed in light protective 1.5 ml eppendorf tubes instead of a dark room. The light protective tube assay was developed to minimize the experimental hazards of working at such high temperatures.

For assays of *afu*RNR, the light protective tubes were also used. In addition, experiments were performed at 83 °C (instead of 37 °C) and the reactions were neutralized by rapid quenching in liquid nitrogen instead of 2 % perchloric acid. Briefly, an assay mixture contains the following: 0.06 – 4.9 mM [^{14}C]-CDP (S.A. = 1.5 x 10^3 – 3.0 x 10^3 cpm/nmol), 25 mM HEPES pH 7.5 (instead of 50 mM HEPES), 10 mM $MgCl_2$, 40-80 µM *afu*RNR, 1.4 – 34 mM DTT and 0.74 to 43 µM AdoCbl in a final volume of 700 µL. The assay was performed as follows. The mixture containing each reagent except AdoCbl was incubated at 83 °C for 5 min. For the initial time point, a 100 µL aliquot was removed and quenched in liquid nitrogen. The reaction was started by the addition of AdoCbl to the mixture. At different time points, a 100 µL aliquot was removed and quenched in liquid nitrogen. After all reactions were stopped, alkaline phosphate buffer (NEB Buffer 3 – 50 mM TRIS-HCl pH 7.9, 100 mM NaCl, 10 mM $MgCl_2$, 1 mM DTT) was added followed by addition of 10 U calf intestine alkaline phosphatase (final volume = 250 µL). The samples were then incubated at 37 °C for 1 hour. Finally, products were separated using the boronate column method and each 1 mL sample of flow through was analyzed using a scintillation counter (16). In this chapter, 1 unit (U) of enzyme activity is defined as 1 nmol substrate converted per min (1 U = 1 nmol / min) (See Scheme 2.2).

Stability of AdoCbl in light protective tubes. To verify that AdoCbl was stable in the light protective 1.5 mL eppendorf tubes used in the modified RTPR and *afu*RNR assays, we subjected AdoCbl to visible light exposure in both clear 1.5 mL tubes and 1.5 mL light protective tubes. Experimental mixtures containing 25 mM HEPES, 10 mM $MgCl_2$, and 39 μM AdoCbl were placed into clear 1.5 ml eppendorf and light protective 1.5 ml eppendorf tubes at 25 °C. A 100 μL aliquot was removed from each tube and a UV/visible spectrum was recorded from 800 to 240 nm using a Cary 300 spectrophotometer (Figure 2.5). Aliquots were removed every 2 min for the duration of 10 min with a final timepoint taken at 20 min and subsequent spectra recorded from 600 to 240 nm.

AdoCbl has been reported to undergo a photolysis reaction in the presence of visible light (17, 18). The instability of AdoCbl is marked by a change in the UV/visible spectra of the coenzyme. The formation of product from the light induced photolysis of AdoCbl can be followed by monitoring the visible spectra changes between 300 and 600 nm. In Figure 2.5A, AdoCbl present in the clear 1.5 mL tube and in the standard RNR assay mixture displays light-induced Co-C bond homolysis followed by conversion into the product hydroxocobalamin (19). There is an appearance of a peak at 352 nm, and the well defined isosbestic points at 335, 367 nm, and 482 nm indicate a single reaction was accountable for the spectral changes. The final spectrum is very similar to the spectra of hydroxocobalamin (20, 21).

In contrast, the AdoCbl RNR reaction mixture contained within the light protective tube exhibited no measurable change in the absorbance spectra over the duration of the experiment (Figure 2.5B). In this experiment, there is no development of the peak at 352 nm indicating that there is no light induced photolysis of AdoCbl (see Figures 2.5A).

Controls to verify stability of substrates at 83 °C. Before assays were conducted with a thermophilic RNR enzyme, controls were carried out to determine the stability of substrates at the organism's temperature optima of 83 °C.

Spectra of DTT in 1.0 mL of buffer (25 mM HEPES, pH 7.5, 10 mM $MgCl_2$) was collected at 83 °C to check the stability of DTT at this temperature. The total concentration of DTT was 30.0 mM. UV/vis spectra were collected from 600 to 240 nm using a Cary 300 spectrophotometer (Figure 2.6). An initial aliquot of 100 μL was collected at 25 °C and the

sample incubated at for 5 min at 83 °C. Additional aliquots were taken at 15 min and 30 min after the first 83 °C time point and subsequent spectra recorded from 600 to 240 nm.

The change in absorption spectra between a reduced and oxidized DTT can be monitored at 283 nm (ε_{283} = 273 M^{-1} cm^{-1}) as was reported by Cleland in 1964 (22). The spectrum of 30 mM DTT in RNR reaction buffer at 25 °C indicates that there is some oxidation (A_{283} = 0.38) present in the DTT solution (Figure 2.6). As the temperature is increased to 83 °C, DTT exhibits an increase in oxidation as seen by the increase in absorbance at 283 nm over the 30 min duration of the experiment (A_{283} = 0.68). In RNR assay buffer, there is an 8.33% oxidation of DTT over the duration of the experiment. Although there is oxidation of DTT at 83 °C in 30 min, the assay is only 7.5 min long.

The stability of AdoCbl in 1.0 mL of buffer (25 mM HEPES, pH 7.5, 10 mM $MgCl_2$) at 83 °C in the presence and absence of DTT was monitored by UV/visible spectra. The total concentrations of DTT and AdoCbl were 30 mM and 47 μM respectively. UV/vis spectra were recorded from 600 to 240 nm (Figure 2.7). The spectra of AdoCbl in RNR reaction buffer (Figure 2.7) or AdoCbl plus DTT in RNR assay buffer (Figure 2.7B) do not show changes within the absorbance range of AdoCbl. Absorbance differences around 290 nm (Figure 2.7B) can be attributed to the changes in DTT oxidation state as described above (Figure 2.6).

The stability of substrates CDP and CTP at 83 °C was analyzed using HPLC. A 1.0 mL reaction solution was prepared consisting of 1.0 mM CDP or 1.0 mM CTP in either 25 mM HEPES, pH 8.0, 20 mM $MgCl_2$ or 50 mM Tris-HCl pH 8.0, 20 mM $MgCl_2$. These solutions were incubated at 83 °C for 5 min. An initial 100 μL aliquot was collected and flash frozen in liquid nitrogen and lyophilized. Subsequent aliquots were collected at 5 min intervals, treated in the same manner as described above, for a total of 20 min. Lyophilized samples were resuspended in 100 μL HPLC running buffer A (described below).

Chromatographic separation of nucleotides was achieved using an Alltima C18 column (Alltech Inc., Nicholasville, KY). All analytical columns were 250 x 4 mm, with a 5-μm particle size. Products were separated from substrate using an isocratic buffer system consisting of 50% buffer A (10% acetonitrile and 90% H_2O) and 50% buffer B (10% acetonitrile and 90% 1.5 M ammonium acetate (pH 7)). The flow rate was 1.0 mL/min, and UV spectra were recorded at 260 nm.

This method shows that CDP is stable in HEPES (RNR assay buffer) over the course of 20 min (Figure 2.8A). However in Tris-HCl buffer CDP degradation occurs fairly rapidly, resulting in a >40% loss of CDP over the course of the experiment (Figure 2.8B).

For CTP, over the course of a 20 min experiment in HEPES (RNR assay buffer) at 83 °C, there is only a slight change in the relative ratios of CTP, CDP, and CMP in the sample between the initial timepoint at 5 min 83 °C (Figure 2.9A) and the 25 min time point at 83 °C (Figure 2.9B). While some of the CMP and CDP in the 5 min time point is likely due to temperature dependent degradation, some samples of CTP from our supplier (Sigma-Aldrich) do have CMP and CDP contaminants. In contrast, when Tris-HCl is used in place of HEPES, the degradation of CTP is substantially increased. After incubation at 83 °C for five min (T = 5), it is apparent that a significant amount of CTP has degraded to CDP and CMP (Figure 2.10A) compared to that of CTP in HEPES buffer (Figure 2.9A). At the 25 min, all of the CTP has degraded to CDP and CMP (Figure 2.10B). Due to the temperature sensitivity of the Tris-HCl buffer, all assays of RNR were conducted with HEPES b0uffer.

HPLC analysis of CDP afuRNR reaction products. An *af*uRNR reaction consisting of 1.0 mM CDP was carried out as described above. An initial 100 µL aliquot was collected and flash frozen in liquid nitrogen and lyophilized to denature protein. The reaction was allowed to proceed for a total of 25 min, and a 100 µL aliquot was collected and treated as described above. Lyophilized samples were resuspended in 100 µL HPLC running buffer A.

Chromatographic separation of nucleotides was achieved using an Alltima C18 column (Alltech Inc., Nicholasville, KY). All analytical columns were 250 x 4 mm, with a 5-µm particle size. The mobile phase consisted of Buffer A (10 mM tetrabutylammonium hydroxide, 10 mM KH_2PO_4, 0.25% methanol, pH 7.0) and buffer B (2.8 M tetrabutylammonium hydroxide, 100 mM KH_2PO_4, 30% methanol, pH 5.5). All mobile phase buffers were filtered and degassed. Separation of nucleotides was achieved with the Alltima C18 column using a 95 min gradient 0% B, at 20 min 10% B, at 25 min 30% B, at 75 min 55% B, and at 95 min 100% B with a continuous gradient between concentrations. All separations were performed at the rate of 1.0 ml/min.

Single base overhang primer extension assay. Three 17-mer oligonucleotides (5`-GTT CTG AAA ATG GTC GC-3`, 5`- CTT CTG AAA ATG GTC GC-3`, and 5`-TTT CTG AAA ATG GTC GC-3`) and a complementary 16-mer oligonucleotide (5`-GCG ACC ATT TTC AGA A-3`) were obtained from IDT Inc., and purified on a 20% PAGE gel. Next, each 17-mer oligonucleotide was phosphorylated at the 5'-end with 1.5 mM ^{32}P-ATP (Perkin Elmer, MA, USA), in a 1× T4 polynucleotide kinase buffer, using 20 units of T4 polynucleotide kinase (NEB, MA, USA) in a volume of 20 μL. The mixture was incubated at 37 °C for 60 min. The reaction was quenched by heating at 90 °C for 5 min and then desalted using Micro Bio-Spin chromatography columns containing Bio-Gel P-6 (Bio-Rad Laboratories) to remove unincorporated ATP. The phosphorylated 17-mers were then incubated with equal amounts of the 16-mer complementary strand in a 20 μL solution containing 50 mM Tris-HCl, pH 7.6, 10 mM $MgCl_2$, and 1.0 mM DTT. To allow for strand annealing, the mixture was brought to 90 °C for 5 min and was then slowly cooled to room temperature before finally cooling to 4 °C.

*afu*RNR assays were performed as described above but with 1.0 mM CTP, GTP, or ATP as substrates, 25 mM HEPES pH 7.5, 10 mM $MgCl_2$, 0.65 μM *afu*RNR, 20 mM DTT, and 8.0 μM AdoCbl except as noted in figure legend. Each reaction was run for 10 – 25 min and then quenched by rapid cooling on ice. A polymerase extension assay was conducted using 10 units of the Klenow polymerase (New England BioLabs) for 60 min at 37 °C. The sample was supplemented with denaturing buffer and was heated at 90 °C for 5 min to stop the polymerase reaction and to denature the oligomers. This step was followed by rapid chilling on ice, and separation on a 20% polyacrylamide gel with 7.0 M urea, and then visualization of gel using phosphoimagery (See Scheme 2.3).

The reactions using RTPR were conducted in an assay mixture composed of 1.0 mM CTP, GTP, or ATP, 25 mM HEPES pH 7.5, 10 mM $MgCl_2$, 1 M NaOAc, 0.85 μM RTPR, 20 mM DTT and 8.0 μM AdoCbl for 20 min. The polymerase extension assay and subsequent separation and visualization of reaction were conducted as described above.

RESULTS

Cloning, purification, and characterization of afuRNR. Using the *A. fulgidus* genome database, a putative ribonucleotide reductase homologue was identified from a ORFs (Accession no. AF1664). Comparison of this sequence with all known protein sequences using PSI-BLAST

revealed sequence similarities with other characterized AdoCbl dependent Class II RNRs (Figure 2.2). The AF1644 gene encodes a protein of 752 amino acids with a predicted molecular mass of 84,456 Da (Figure 2.3). The deduced amino acid sequence suggests that *afu*RNR does not contain an N-terminal ATP activation domain found in the *T. acidophilum* RNR (10); however, it does contain a conserved $Cx_{10}C$ sequence identified as the classic Class II AdoCbl dependent RNR motif (23). This ORF was subsequently amplified by PCR, cloned, sequenced, and compared to the sequence in the *A. fulgidus* genome database.

Purified *afu*RNR protein yields absorption spectra with absorbance maxima at 279 nm. Expression and subsequent purification yields ~28 mg from a 1 liter culture. The molecular mass of the protein is estimated to be approximately 84 kDa by SDS-PAGE (Figure 2.4). This molecular mass is in agreement with values predicted by the gene analysis. The native molecular mass is determined to be about 160 kDa by gel-filtration using a Sephacryl SEC-200 HR 16/60 column. These results suggest that the native state of *afu*RNR is a homodimer, similar to other reported Class II RNRs (10).

Biochemical characterization of afuRNR. To examined whether *afu*RNR is a Class II enzyme in that it requires AdoCbl to generate the thiyl radical we carried out activity assays in the presence and absence of AdoCbl. In the presence of AdoCbl, *afu*RNR displays enzyme-dependent reduction of CDP to dCDP. The rate of reduction increases with an increase in *afu*RNR concentration (Figure 2.11, red line and blue line). In the absence of AdoCbl, there is no measurable conversion of CDP to dCDP (Figure 2.11, black line). The $MgCl_2$ dependence of *afu*RNR was also investigated. In the absence of $MgCl_2$, there is CDP reduction, but activity increases to maximum activity between 1.00 and 10.0 mM $MgCl_2$ (Figure 2.12). The temperature profile of *afu*RNR enzyme activity fits the reported optimal growth conditions of 83 °C for *A. fulgidus* (Figure 2.13). The activity of the enzyme drops significantly at 90 °C and is abolished at 100 °C. These results show that *afu*RNR is a Class II RNR whose optimal activity is at 83 °C and with a concentration of $MgCl_2$ greater than 1.0 mM.

The *Km* for CDP, AdoCbl, and DTT were determined (see Figures 2.14-2.16). Table 2.1 compares the results for *afu*RNR with the published results for RTPR and the only other Archaeal thermophilic RNR; *T. acidophilum* RNR. The biochemical characteristics of *afu*RNR

are comparable to those previously published, with *Kms* for substrate in mM range and the *Km* for AdoCbl in the μM range.

Substrate phosphorylation state preference. The preference of *afu*RNR for CMP, CDP, and CTP was investigated using a radioactive assay. As is true for all RNR enzymes, there is no discernable activity with CMP (data not shown). However, with CDP and CTP as the substrate, *afu*RNR displays reductive activity (Figure 2.17, blue and magenta lines). This result is surprising since all characterized RNRs have preference for either a diphosphate or a triphosphate, but do not accept both. As a control, the radiolabeled CDP and CTP substrates were used with RTPR, an RNR which uses only triphosphates as substrates. As seen in figure 2.17, RTPR exhibits activity when CTP is used as substrate (black line) but there is no measurable activity when CDP is used (red line).

Because this result is surprising, we wanted to confirm that CDP is in fact being converted to dCDP and that the observed enzyme activity is not due to a CTP contaminant. Thus the identity of the product of the *afu*RNR reaction with CDP as substrate was determined via HPLC. The HPLC method separates CDP (Figure 2.18A) and dCDP (Figure 2.18B) as shown by the standard spectra. In addition, CTP can be easily distinguished form CDP or dCDP as it elutes with a peak separation which is greater than 20 min (Figure 2.18C). Using this method, we were able to verify that dCDP is being produced by *afu*RNR, confirming that CDP is a substrate and that activity with CDP is not due to a CTP contamination (Figure 2.18F, UV/visible spectra; 2.18G, radioactive trace). Unfortunately, we were not able to use HPLC to confirm that CTP is a substrate, since it was not possible to resolve the CTP substrate from the product (data not shown). The lack of definitive identification of CTP reaction products lead to the development of a primer extension assay to allow for the identification of the products of the CTP reduction reaction.

The general scheme of the primer extension assay is given in Scheme 2.3. The assay exploits the ability of a polymerase to incorporate deoxynucleoside triphosphates into a radiolabeled primer across from the complementary template with a corresponding base pair overhang. The radiolabeled extended primer is then separated from non-extended substrate primer on a PAGE gel. Since the Klenow polymerase fails to incorporate ND(T)Ps and only incorporates their deoxynucleotide counterpart (24-26), this assay tests for RNR activity when

deoxynucleotides are not provided. Further, to discriminate between the di- versus triphosphate nucleoside products of the RNR reaction, this assay takes advantage of the discretion of a polymerase for dNTP over a dNDP.

The results of a primer extension assay with the products from the *afu*RNR nucleotide reduction assay are shown in Figure 2.19A. When NTPs are used as substrates for the RNR reaction, the dNTP products are incorporated by the polymerase into the primer and we observe the extension of the primer by one base. If NDPs are substituted in the RNR reaction, the polymerase can not incorporate the dNDPs and as such, there is no primer extension and no separation on the PAGE gel (Figure 2.19A).

If the activity of the polymerase demonstrated in Figure 2.19 is due to a small contaminant of dNTP in the NTP samples or if it is due to low level reactivity of the polymerase with NTPs themselves, there should be no positive correlation between the extent of the RNR assay and the amount of 17-mer formed by the subsequent primer extension. All time points should show the same amount of 17-mer or less if the containment is being degraded at the high temperatures. To investigate this issue of time dependence, the *afu*RNR assay was run for 180 s with products removed at various time points. We find a clear dependence on the extent of *afu*RNR assay and the amount of 17-mer formed (Figure 2.20A). When the *afu*RNR assay is allowed to proceed for longer time periods more 17-mer is produced, providing evidence that the polymerase activity is due to the time-dependent formation of dNTP by *afu*RNR and not to a contamination. When no *afu*RNR is added in the presence of NTPs, no extension is observed (Figure 2.19B, C). For comparison with *afu*RNR, we carried out a similar time-dependent reaction using RTPR (Figure 2.20B). We find the complete conversion of radiolabeled primer to the corresponding 17mer occurring after the RTPR assay is run for 120 s. Although the primer extension reaction using *afu*RNR-produced products does not proceed to completion like the one using RTPR produced products, comparison of both results along with the results observed in Figure 2.19A provide strong evidence that *afu*RNR can reduce NTPs to their corresponding dNTPs. We believe that the lack of completion in the case of *afu*RNR may be due to temperature dependent degradation or dCTP since that assay was carried out at 83 °C compared to 37 °C for RTPR.

DISCUSSION

Due to their central function in DNA replication, RNRs have been targeted in anti-parasitic, anti-viral and anti-tumor therapies. While RNRs have proven to be successful chemotherapeutic targets (27), progress with anti-parasitic therapies has been slower. Although one could imagine that class and species specific differences could be exploited to develop targeted RNR inhibitors, the lack of biochemical and structural data on the Class II and III enzymes has impeded their development. In addition, the small number of characterized RNRs from any class has left questions unanswered about basic properties of these enzymes such as substrate preference for NDPs versus NTPs, and whether the rules of allosteric regulation are in fact common to all RNRs. Here we have gone to the extremes of life in studying an RNR from *A. fulgidus*, a sulfur-reducing thermophilic Archaeon that grows between 60 °C and 90 °C. By characterizing this enzyme from this extremophile we are probing the question of which features of RNRs are truly conserved.

We identified ORF (AF1664) in the *A. fulgidus* genome as a putative Class II RNR based on the presence of a highly conserved 10-amino acid sequence block connecting active site cysteine residues (C322 – C333 in *T. maritima*; 408 – 419 in *L. leichmannii* and 436 – 4475 *A. fulgidus*). While the identity of all residues in this motif is not completely conserved, the spacing between cysteines is absolute (23). In Class I RNRs, the distance between the active site cysteines is more variable. Since these connecting residues in both classes comprise a loop that runs on the bottom of β-barrel structure, the rationale for this conservation is not clear (see Chapter 1 for more discussion). However, this motif does appear to be a good predictor as we show here that AF1664 is in fact a Class II RNR (Figure 2.22), dependent on AdoCbl for activity (Figure 2.11s). As with many of the characterized Class II RNRs, size exclusion chromatography suggests that this Class II enzyme is a dimer.

Due to the high growth temperature of the *A. fulgidus* organism (60-90 °C), assaying this RNR is non-trivial. First it was necessary to establish buffer conditions under which substrates and cofactors are stable at high temperatures. While Tris-HCl buffer did not protect substrates from degradation at high temperatures, HEPES buffer did (Figure 2.9). To improve reproducibility, we modified the assay to use liquid nitrogen to quench this high-temperature reaction rather than ice, which did not cool the reaction quickly enough. We also used light-protective tubes to prevent light-induced cleavage of AdoCbl, rather than performing the

reactions in the dark as was done previously (12). Controls with clear and light-protective tubes show that photolysis is not an issue when the light-protective tubes are used (Figure 2.5). In our hands, reproducibility was dramatically increased when assays are carried out using light-protective tubes outside of a darkroom. This assay improvement could be utilized when studying non-thermophilic Class II enzymes.

With this modified assay in hand, the biochemical properties of *afu*RNR were investigated. The activity of *afu*RNR is enhanced 3 fold in the presence of $MgCl_2$ concentration (Figure 2.12), and in terms of temperature-dependence of activity, *afu*RNR displays very little measurable activity at 30 °C, which increases slowly until the temperature reaches 60 °C. The enzyme exhibits the greater than 60% activity between 70 °C and 90 °C with highest activity measured at 83 °C. At 100 °C the activity is abrogated (Figure 2.13). The 20 °C range of activity of *afu*RNR is similar to other RNRs, allowing for deoxynucleotide production under a range of environmental stresses. Interestingly, the enzyme displays higher CDP reduction in absence of allosteric effectors than in their presence (data not shown). While this seems to contradict the rules of allosteric regulation of substrate specificity in which effectors enhance activity with particular substrates, more work must be done before conclusions are drawn. Because *afu*RNR's activity without effectors is comparable to the activity of other RNRs (Table 2.1), all assays were performed without effectors present. It should be noted that other RNRs are assayed without effectors. In one case, sodium acetate is used to enhance activity instead of adding effectors to the reaction mix (28).

*afu*RNR's unusual behavior with effectors was only the first observation of nonstandard activity. The investigations of the substrate preference of *afu*RNR towards nucleotides lead to surprising results. As expected, there was no activity with CMP as substrate. But when CTP was substituted for CDP in the reduction assay, the enzyme exhibited reduction activity. When the kinetic parameters were analyzed, the results were even more perplexing as the K_ms for both CDP and CTP are similar. The K_m of *afu*RNR for CDP is 0.92 mM and the K_m for CTP is 1.29 mM (Figure 2.21). The cellular concentrations of nucleotides in *A. fulgidus* have not been measured, but assuming that they were within the concentration range of other organisms (For *E. coli*: ATP = 3.1 mM, ADP = 0.84 mM, CTP = 0.28 mM, CDP = 0.07 mM, GTP = 0.47 mM, and GDP = 0.20 mM), both CDP and CTP should be substrates for *afu*RNR (29).

One other thermophilic RNR was initially reported to be active with both NDP as well as NTPs, but upon closer examination, the activity with NTPs was found to be due to temperature dependent hydrolysis leading to formation of NDPs (30). Our control experiments however show that in HEPES buffer, CTP does not hydrolyze significantly although it does hydrolyze more than CDP. To rule out that small amounts of CDP in the CTP sample are responsible for the measured activities, we set about to verify that dCTP is in fact produced by *afu*RNR. For completeness, we also confirmed that CTP contamination of CDP samples is not the source of the activity with that nucleotide. For the latter, we used an HPLC analysis that clearly distinguishes dCDP from CDP and CTP (Figure 2.18), allowing us to verify dCDP production. Unfortunately, this method failed to resolve dCTP as a product of the *afu*RNR reaction. However, due to the greater instability of triphosphate nucleotides at 83 °C, it is impossible to tell if the failure to identify dCTP was due to degradation or to a lack of reactivity with CTP. Regrettably, the RNR assay requires the removal of all the deoxynucleoside phosphates, making it impossible to distinguish if the substrate acted upon by the enzyme is a diphosphate or a triphosphate nucleoside.

The inability to confirm by HPLC the product of *afu*RNR activity with CTP led to the development of a primer extension assay (See Scheme 2.3). Since a primer can only be extended by a polymerase if dNTP is available, primer extension serves as proof of dNTP production. Negative control experiments with dNDP confirm the lack of primer extension with deoxynucleoside diphosphates (Figure 2.19A), while a positive control with dNTP-producing RTPR shows primer extension when an RNR is used as a source of deoxynucleoside triphosphates. When *afu*RNR is reacted with CTP, the primer is extended, while no primer extension is observed with CDP. To rule out that a contaminant of dCTP in the CTP sample could lead to a false positive, we ran a no-enzyme control in the presence of CTP and the absence of *afu*RNR, the primer is not extended confirming that primer extension is *afu*RNR dependent. This experiment serves as the best evidence to date that an RNR is active with both NDPs and NTPs.

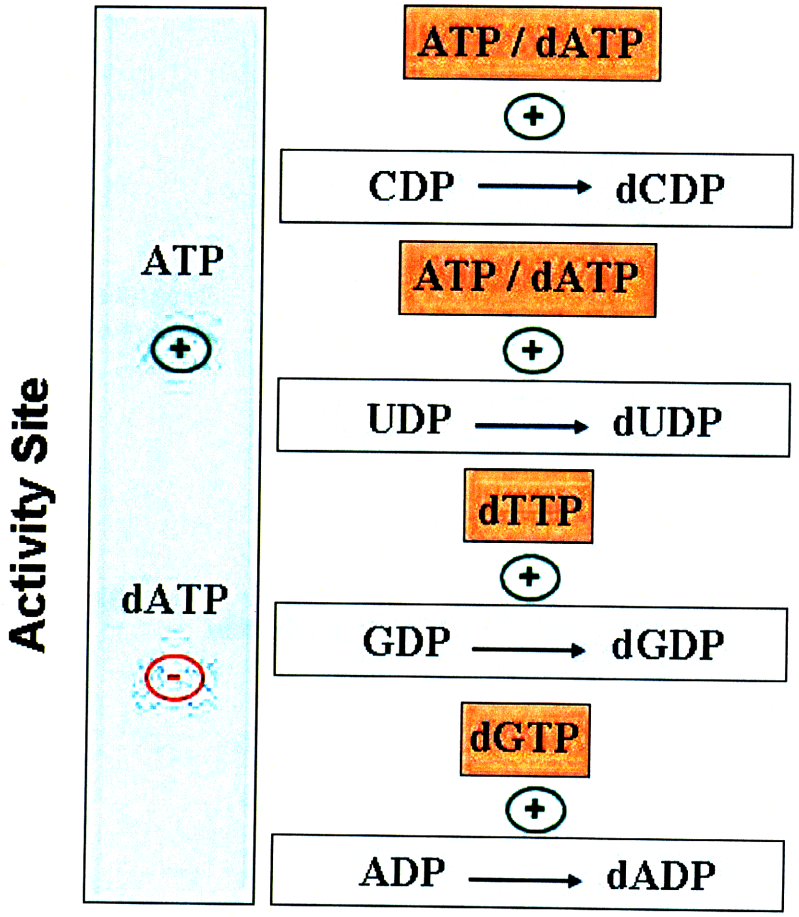
In light of the primer extension result for *afu*RNR, it is interesting to think about the molecular mechanism among RNR enzymes for determining substrate preference for NDPs and NTPs. When the structure of Class II RTPR was determined (4), the active site of this NTP-

specific RNR was compared with that of NDP-specific Class I RNR from *E. coli* in the hope of identifying the structural differences responsible for this variation in substrate specificity. However, there were no obvious structural differences to explain specificity, no extra positively charged residues to bind the extra phosphate, no extra pocket that could be detected. With the phosphate binding regions so similar, it is unclear why either enzyme does not bind either substrate. When the second structure of a Class II RNR became available, this topic was revisited (31). Unlike the Class II enzyme from *L. leichmannii* (RTPR), the Class II enzyme from *T. maritima* uses NDPs as substrates (10). Again, structural comparisons showed similar active sites with no clear structural rationale to explain this specificity (Figure 2.23). It should be noted that all structures currently available are in an open, non-catalytic state. Presumably, the structure of a closed, catalytic RNR will show variations in the substrate phosphate binding area that will explain the specificity. However, given the similarities in the size and in amino acid content in these active sites, the idea of a dual specificity RNR is plausible from a structural perspective.

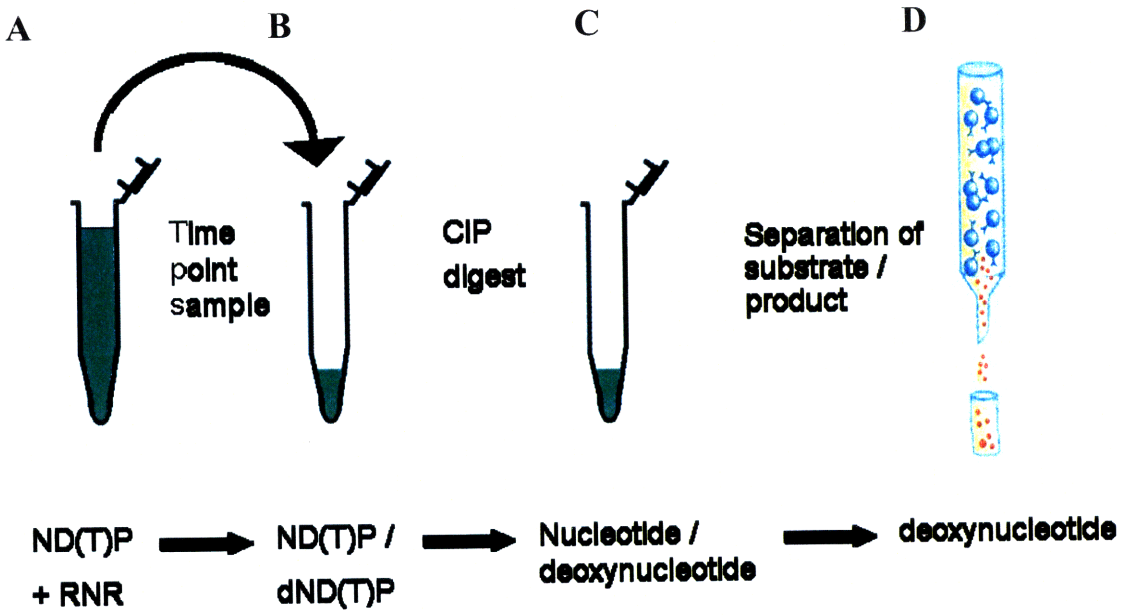
While there is nothing inconsistent with the idea of a dual specificity RNR from a structural perspective, the cellular perspective is harder to rationalize. Maintenance of the levels of dNTPs is essential to the cell's survival (32). Too much or too little of any given dNTP is harmful (33, 34). For the majority of organisms that use Class I or Class II RNRs, dNDP are produced, and the regulation of dNTP levels occurs via nucleotide kinases (35). How the correct balance of dNTPs can be maintained in the cellular context where the one RNR enzyme produces both dNDPs and dNTPs is an open question. A genomic investigation of *A. fulgidus* shows that it contains all the standard nucleotide kinases to convert nucleotide diphosphates to nucleotide triphosphates. However, these proteins have not been characterized from this organism and may enforce a regulatory mechanism that is unique. Another possibility is that the cellular regulation of this RNR will control dNTP balance. Given that the allosteric regulation of *afu*RNR appears to vary from the norm, it is possible that this Class II enzyme has evolved such that its use of NDP substrates versus NTP substrates is highly regulated, and that either NDPs or NTPs is rarely, if ever, used under *in vivo* conditions. Perhaps instability of dNTPs due to high temperature is a problem for the organism as well as for the scientist studying the organism's enzymes, and that as a result, *A. fulgidus* will have nonstandard concentrations of nucleotides. Further work will be necessary to resolve these questions.

These studies of an RNR from an extremophile challenge some of the paradigms in the RNR field. While it is clear that we still have much to learn about this fascinating and medically-important enzyme family, this work suggests the importance of expanding our studies of these enzymes to cover additional evolutionary ground. If we wish to exploit species and class dependent differences between RNRs for drug design purposes, we need more data about RNRs from outlying organisms. At the extremes, there promise to be some surprises.

Scheme 2.1 A schematic representation of the classic pattern of regulation as exemplified in Class I and Class II RNRs. In Class Ia RNRs, overall activity is up regulated when ATP is bound and down regulated when dATP is bound. The specificity effectors are in brown boxes and substrates are in light grey boxes. When ATP or dATP is bound at the specificity site, RNRs preferentially convert CDP or UDP to dCDP or dUDP. The binding of dTTP at the specificity site preferentially reduces DGP to dGDP and binding of dGTP to specificity site reduces ADP to dADP. In Class II RTPR the substrates are nucleotide triphosphates instead of nucleotide diphosphates, but overall, enzyme regulation follows the same pattern.

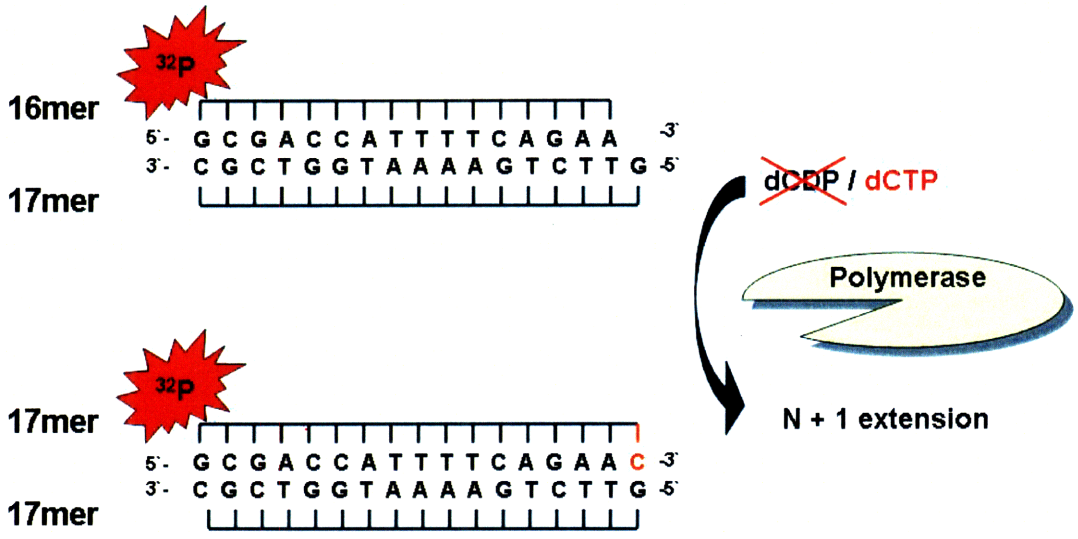


Scheme 2.2 Modified RNR ^{14}C radioactive nucleotide reduction assay. (A) Instead of carrying out the assay in a dark room to avoid light induced C-Co homolysis of AdoCbl cofactor, the reaction mixture is contained in a 1.5 ml light protective eppendorf tube. (B) At each timepoint a 100 µL sample is removed from the 700 µL reaction mixture and quenched in liquid nitrogen. (C) Each sample was then subjected to CIP digest to remove all 5`-phosphates. (D) Nucleoside substrates were separated from deoxynucleoside substrates using the borate affinity column method. A 1 mL aliquot of elution is mixed with 9 mL of scintillation fluid and subject to scintillation counting.



Scheme 2.3 Scheme of primer extension assay. (A) Polymerase lacking a 3` to 5` exonuclease activity extends the ^{32}P radiolabeled primer by inserting the corresponding base pair. (B) Reaction mixture is denatured and separated by 20% denaturing PAGE. Bands are then visualized on a phosphoimagery.

A



B

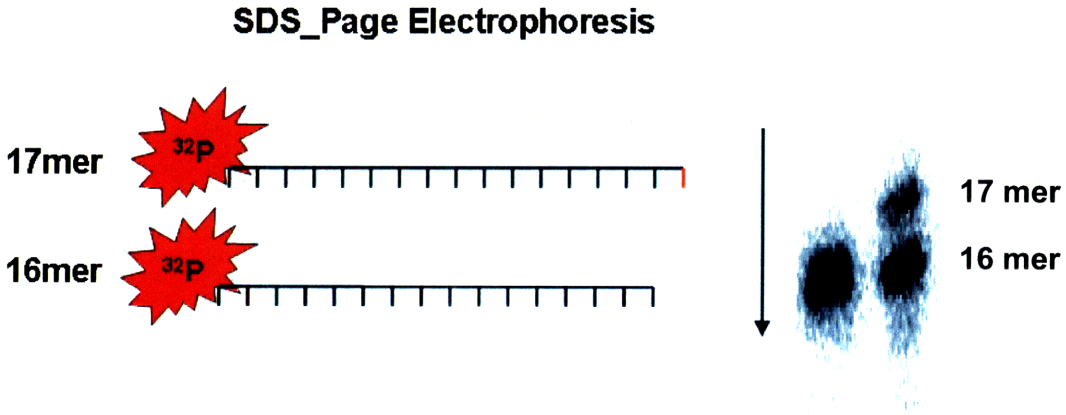


Table 2.1 Kinetic parameters of *A. fulgidus* ribonucleotide reductase, *L. leichmannii* ribonucleoside triphosphate reductase (36) and *T. acidophilum* ribonucleotide diphosphate reductase (37). The assays were carried out as described under experimental procedures, and the kinetic data were derived from Michaelis-Menten equation. Mean values of at least three independent measurements, together with standard deviations, are given.

	Substrate	**Km**	**Effector**	**REF**
A. fulgidus	AdoCbl	11 ± 7.6 μM	N/A	This work
A. fulgidus	CDP	0.92 ± 0.55 mM	N/A	This work
A. fulgidus	CTP	1.29 ± 0.14 mM	N/A	This work
A. fulgidus	DTT	4.18 ± 0.78 mM	N/A	This work
L. leichmannii	AdoCbl	0.30 μM	1.0 mM CTP / 1 mM dATP	(36)
L. leichmannii	CTP	0.13 mM	1 mM dATP	(36)
L. leichmannii	ATP	0.22 mM	1 mM dGTP	(36)
T. acidophilum	ADP	0.64 mM	100 μM AdoCbl / 1 mM dGTP	(37)

Figure 2.1 Reaction of RNRs. The thiyl radical generated by the diferric tyrosine radical (Class I), AdoCbl (Class II), or glycyl radical (Class III) abstracts a hydrogen atom from the substrate initiating a multi-step rearrangement which results in the net loss of the 2′ hydroxyl of the nucleoside moiety. The reducing equivalents are provided in Class I and II RNRs by the oxidation of two active site cysteines to a disulfide bond, while in Class III RNRs, the reducing equivalents are provided by the oxidation of formate. Finally, the thiyl radical is regenerated and the product, either a dNDP or dNTP is formed (Figure modified from (4)).

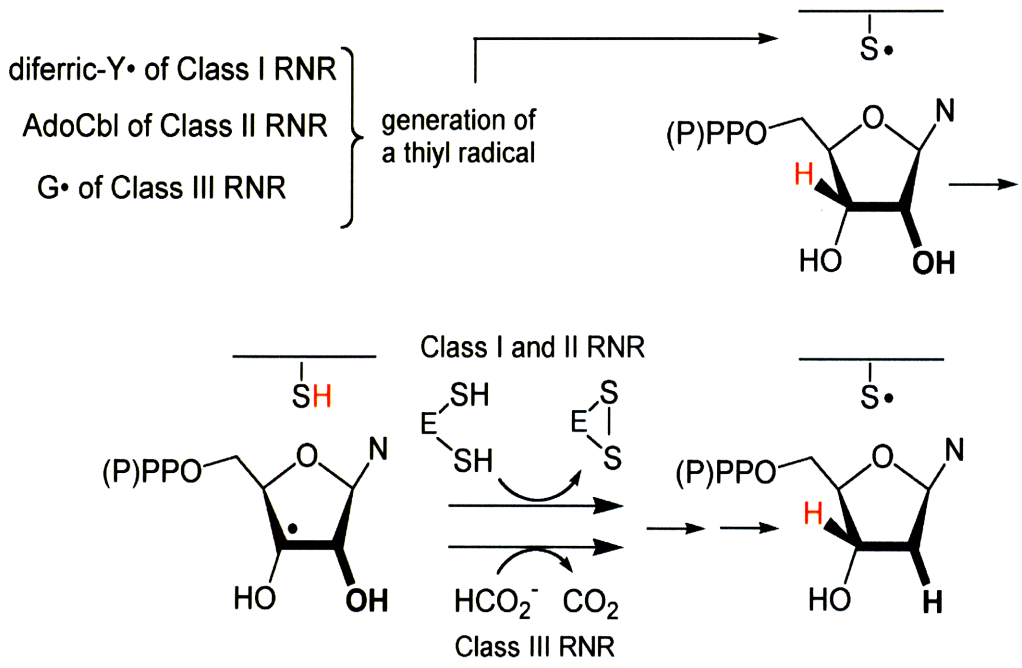


Figure 2.2 Structural alignment of *afu*RNR and two characterized thermophilic dimeric Class II RNR sequences from PSI-BLAST (13). Sequences were aligned using the program ClustalW (14) and the output was rendered using the web interface ESPript 2.2 (15). Sequences are *A. fulgidus* (gi:6626247), *T. acidophilum* (gi:30407158), and *T. maritima* (gi:12057205). White letters in red blocks are strictly conserved residues. Numbers correspond to each organism's amino acid sequence. Red letters in blue boxes are more than 50% conserved residues. Top line is structural elements of the *T. maritima* structure (PDB 1XJE). Numbering of core barrel elements is according to RTPR structure (4)

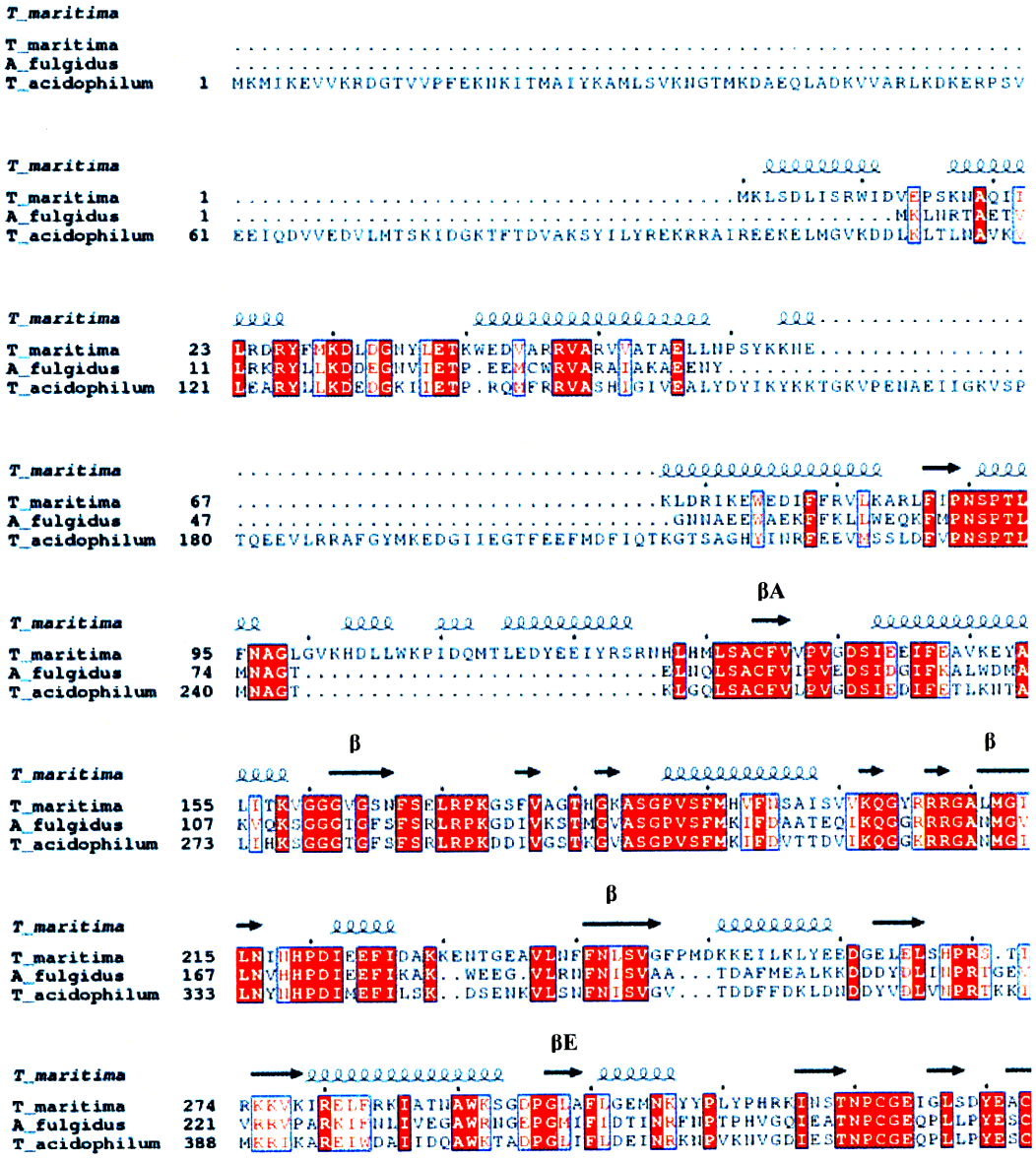


Figure 2.2 (cont.) Alignment of selected dimeric Class II AdoCbl dependent RNR sequences

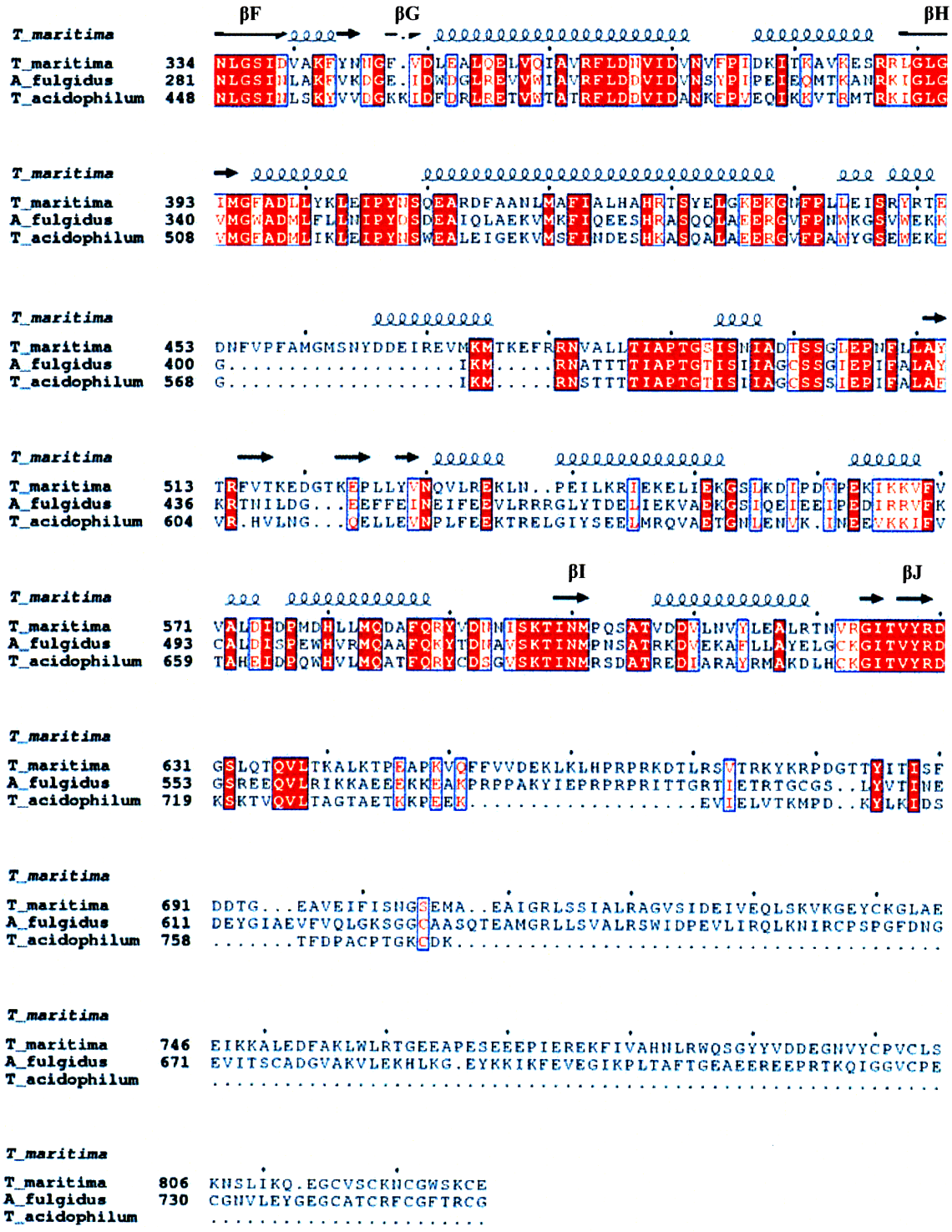


Figure 2.3 DNA and corresponding amino acid sequence of ORF AF1664 from *A. fulgidus* encoding for a putative RNR. DNA sequence was verified at the MIT Biopolymers Laboratory. (black = nucleotide sequence, green = corresponding 1-letter amino acid sequence; numbers on left correspond to nucleotide numbers)

```
  1  atgaagctca atcgtactgc tgaaactgtg ctcagaaaga ggtatcttct caaggatgat
       m  k  l  n  r  t  a  e  t  v  l  r  k  r  y  l  l  k  d  d

 61  gaaggtaacg taattgagac tccagaagag atgtgctgga gggttgccag ggcaattgca
       e  g  n  v  i  e  t  p  e  e  m  c  w  r  v  a  r  a  i  a

121  aaggctgagg agaattacgg aaacaatgct gaagagtggg ctgaaaagtt cttcaagctc
       k  a  e  e  n  y  g  n  n  a  e  e  w  a  e  k  f  f  k  l

181  ctctgggagc agaagttcat gcccaactcc ccaacgctga tgaatgcggg gacggagctg
       l  w  e  q  k  f  m  p  n  s  p  t  l  m  n  a  g  t  e  l

241  aatcagcttt ctgcatgctt cgtcatcccc gttgaggact ccattgacgg cattttcaag
       n  q  l  s  a  c  f  v  i  p  v  e  d  s  i  d  g  i  f  k

301  gccctctggg atatggcgaa ggttcaaaag agcgggggag ggacgggatt cagcttttca
       a  l  w  d  m  a  k  v  q  k  s  g  g  g  t  g  f  s  f  s

361  aggctgagac cgaagggaga tatagtaaaa tccacaatgg gcgttgcaag cggaccggta
       r  l  r  p  k  g  d  i  v  k  s  t  m  g  v  a  s  g  p  v

421  agcttcatga aaatatttga tgctgcgacg gagcagataa agcagggagg gaggagaaga
       s  f  m  k  i  f  d  a  a  t  e  q  i  k  q  g  g  r  r  r

481  ggggcaaaca tgggcgttct gaacgtccac cacccggaca ttgaggaatt cataaaggcg
       g  a  n  m  g  v  l  n  v  h  h  p  d  i  e  e  f  i  k  a

541  aaatgggagg agggggtgct gagaaacttc aacatcagcg ttgccgcaac cgacgccttc
       k  w  e  e  g  v  l  r  n  f  n  i  s  v  a  a  t  d  a  f

601  atggaagctt tgaagaaaga tgatgattac gacctcatca accccagaac gggggaggtt
       m  e  a  l  k  k  d  d  d  y  d  l  i  n  p  r  t  g  e  v

661  gtaaggagag ttcccgccag aaagatattc aatttgattg tggagggagc atggcgcaac
       v  r  r  v  p  a  r  k  i  f  n  l  i  v  e  g  a  w  r  n
```

Figure 2.3 (cont.) DNA and corresponding amino acid sequence of ORF AF1664 from *A. fulgidus*

```
721  ggagagccgg gaatgatttt catcgacacc ataaaccgct tcaaccccac tccgcacgtg
       g  e  p  g  m  i  f  i  d  t  i  n  r  f  n  p  t  p  h  v

781  ggacagattg aagcaacaaa cccctgcgga gagcaaccgc tcctgcccta cgaatcctgc
       g  q  i  e  a  t  n  p  c  g  e  q  p  l  l  p  y  e  s  c

841  aaccttggca gcataaacct cgcaaagttc gttaaagatg gggaaatcga ctgggatggg
       n  l  g  s  i  n  l  a  k  f  v  k  d  g  e  i  d  w  d  g

901  ctcagggagg ttgtctggat tgccgtgagg tttttggatg acgttatcga tgtcaacagc
       l  r  e  v  v  w  i  a  v  r  f  l  d  d  v  i  d  v  n  s

961  taccccattc ccgaaatcga gcagatgacc aaagccaacc gcaaaatagg tctcggagtt
       y  p  i  p  e  i  e  q  m  t  k  a  n  r  k  i  g  l  g  v

1021 atgggctggg ctgacatgct ctttctgctc aacatcccct acgacagcga tgaagccatc
       m  g  w  a  d  m  l  f  l  l  n  i  p  y  d  s  d  e  a  i

1081 cagcttgctg agaaggtgat gaagttcatt caggaggaaa gtcacagggc ttcgcagcag
       q  l  a  e  k  v  m  k  f  i  q  e  e  s  h  r  a  s  q  q

1141 ctggcggagg agaggggtgt tttccccaac tggaagggca gcgtgtggga gaagaaggga
       l  a  e  e  r  g  v  f  p  n  w  k  g  s  v  w  e  k  k  g

1201 ataaagatga ggaacgcaac gacaacaaca atcgcaccaa cgggaacgat aagcattatc
       i  k  m  r  n  a  t  t  t  t  i  a  p  t  g  t  i  s  i  i

1261 gctggctgtt caagcggaat agagccgatt ttcgctttag catacaagag aacgaacatt
       a  g  c  s  s  g  i  e  p  i  f  a  l  a  y  k  r  t  n  i

1321 cttgacggag aagagttctt cgagataaat gagatattcg aggaagtttt aaggagaaga
       l  d  g  e  e  f  f  e  i  n  e  i  f  e  e  v  l  r  r  r

1381 gggctgtaca cagatgagct gattgagaaa gttgccgaaa aaggcagcat tcaggaaata
       g  l  y  t  d  e  l  i  e  k  v  a  e  k  g  s  i  q  e  i

1441 gaggagattc ccgaggacat aaggagggtc ttcaagtgcg ctctcgacat ctctccagag
       e  e  i  p  e  d  i  r  r  v  f  k  c  a  l  d  i  s  p  e
```

Figure 2.3 (cont.) DNA and corresponding amino acid sequence of ORF AF1664 from *A. fulgidus.*

```
1501  tggcatgtga gaatgcaggc ggccttccag aagtacacgg acaatgcggt cagcaaaacc
        w  h  v   r  m  q   a  a  f  q   k  y  t   d  n  a   v  s  k  t

1561  ataaatatgc cgaactctgc tacgcggaaa gacgttgaga aggctttcct gctggcatac
        i  n  m   p  n  s   a  t  r  k   d  v  e   k  a  f   l  l  a  y

1621  gagcttggat gtaagggcat aaccgtttac agagatggca gcagagagga gcaggttctc
        e  l  g   c  k  g   i  t  v  y   r  d  g   s  r  e   e  q  v  l

1681  agaataaaga aggctgagga ggagaagaag gaggcgaagc cgagaccccc tgccaagtac
        r  i  k   k  a  e   e  e  k  k   e  a  k   p  r  p   p  a  k  y

1741  atcgagccga ggccgaggcc gaggataacc accggcagaa cgattgagac cagaacaggc
        i  e  p   r  p  r   p  r  i  t   t  g  r   t  i  e   t  r  t  g

1801  tgcggttccc tgtatgtgac catcaacgag gacgagtacg ggattgccga agtcttcgtg
        c  g  s   l  y  v   t  i  n  e   d  e  y   g  i  a   e  v  f  v

1861  cagcttggaa aaagcggagg ctgcgctgcg tcgcagacag aggcaatggg gaggctgctt
        q  l  g   k  s  g   g  c  a  a   s  q  t   e  a  m   g  r  l  l

1921  tcagtcgcgc ttaggagttg gattgaccct gaggtgctga taaggcagct caagaacata
        s  v  a   l  r  s   w  i  d  p   e  v  l   i  r  q   l  k  n  i

1981  cgctgtccct cgccaggatt cgacaacgga gaggtcataa catcctgcgc ggatggagtg
        r  c  p   s  p  g   f  d  n  g   e  v  i   t  s  c   a  d  g  v

2041  gcaaaggttc tcgaaaagca cctaaaggga gagtacaaga aaataaagtt cgaagttgag
        a  k  v   l  e  k   h  l  k  g   e  y  k   k  i  k   f  e  v  e

2101  ggaatcaagc ccctcacggc cttcaccggc gaggctgagg agagagagga gccgaggacg
        g  i  k   p  l  t   a  f  t  g   e  a  e   e  r  e   e  p  r  t

2161  aagcagattg gcggagtgtg tcctgaatgc gggaatgttc ttgagtacgg ggagggttgc
        k  q  i   g  g  v   c  p  e  c   g  n  v   l  e  y   g  e  g  c

2221  gcaacctgca ggttctgcgg gtttacgagg tgcggttga
        a  t  c   r  f  c   g  f  t  r   c  g  -
```

Figure 2.4 SDS-PAGE (4-20% gradient (w/v)) gel of the isolated *afu*RNR from the overexpression of *E. coli* Rosetta™ (DE3) pLysS strain. Lanes 1 and 10 are molecular weight markers; lane 2 is Ni-NTA flow through; lane 3 is elution buffer A (10 mM TRIS-HCl pH 8.0, 150 mM NaCl, 10 mM imidazole) wash; lane 4 is elution buffer B (10 mM TRIS-HCl pH 8.0, 1.0 M NaCl, 200 mM imidazole) wash; lane 5 is elution buffer C wash (10 mM TRIS-HCl pH 8.0, 150 mM NaCl, 30 mM imidazole); lanes 6-9 are elution buffer D wash (10 mM TRIS-HCl pH 8.0, 150 mM NaCl, 200 mM imidazole).

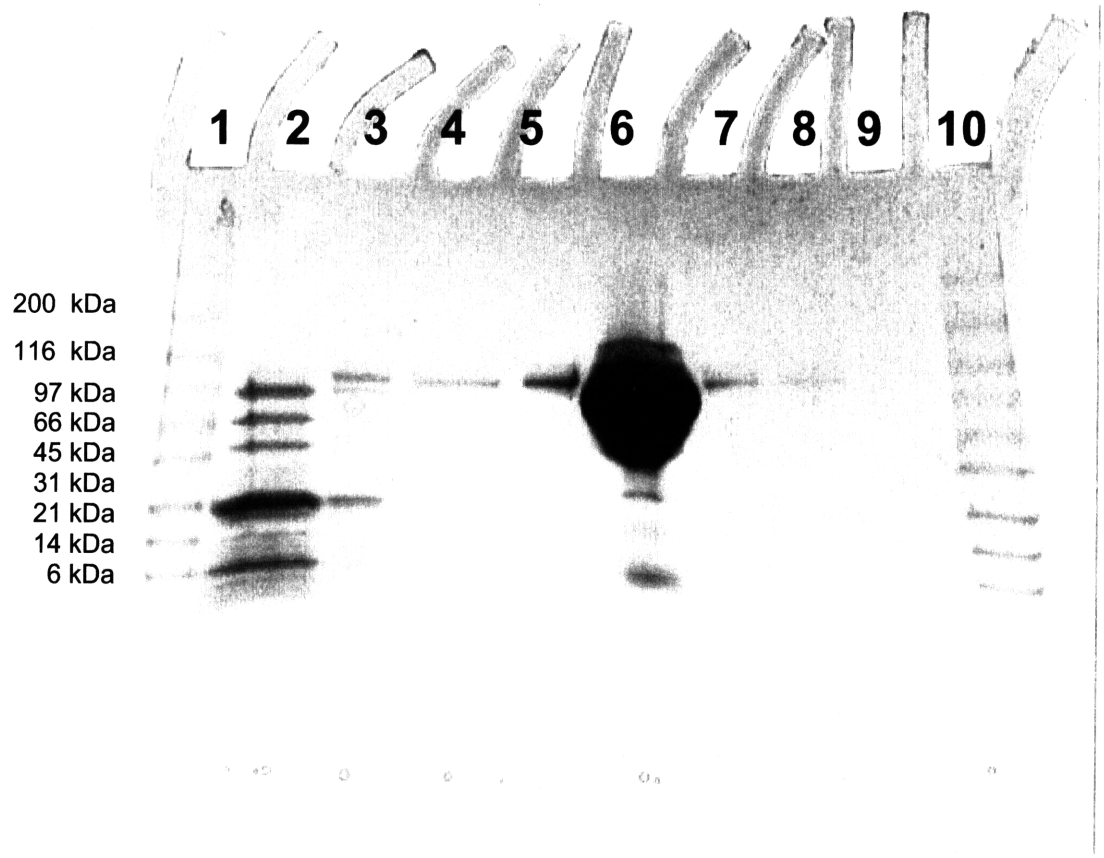


Figure 2.5 Examination of the ability of 'light protective' eppendorf tubes to prevent Co-C bond homolysis in (A) clear 1.5 mL eppendorf tube and in (B) light protective 1.5 mL eppendorf reaction tubes (USA Scientific, Ocala, FL). Aliquots were collected at 0 (black line), 2 min (red line), 4 min (blue line), 6 min (aqua line), 8 min (magenta line), 10 min (green line), and 20 min (brown line). Arrows indicate the direction of change in absorbance as AdoCbl is converted to hydroxocobalamin. The formation of hydroxocobalamin is a result of light induced Co-C homolysis followed by H_2O binding. There was no appreciable change in absorbance from the sample of AdoCbl in the light protective reaction tubes over the time of RNR activity assay.

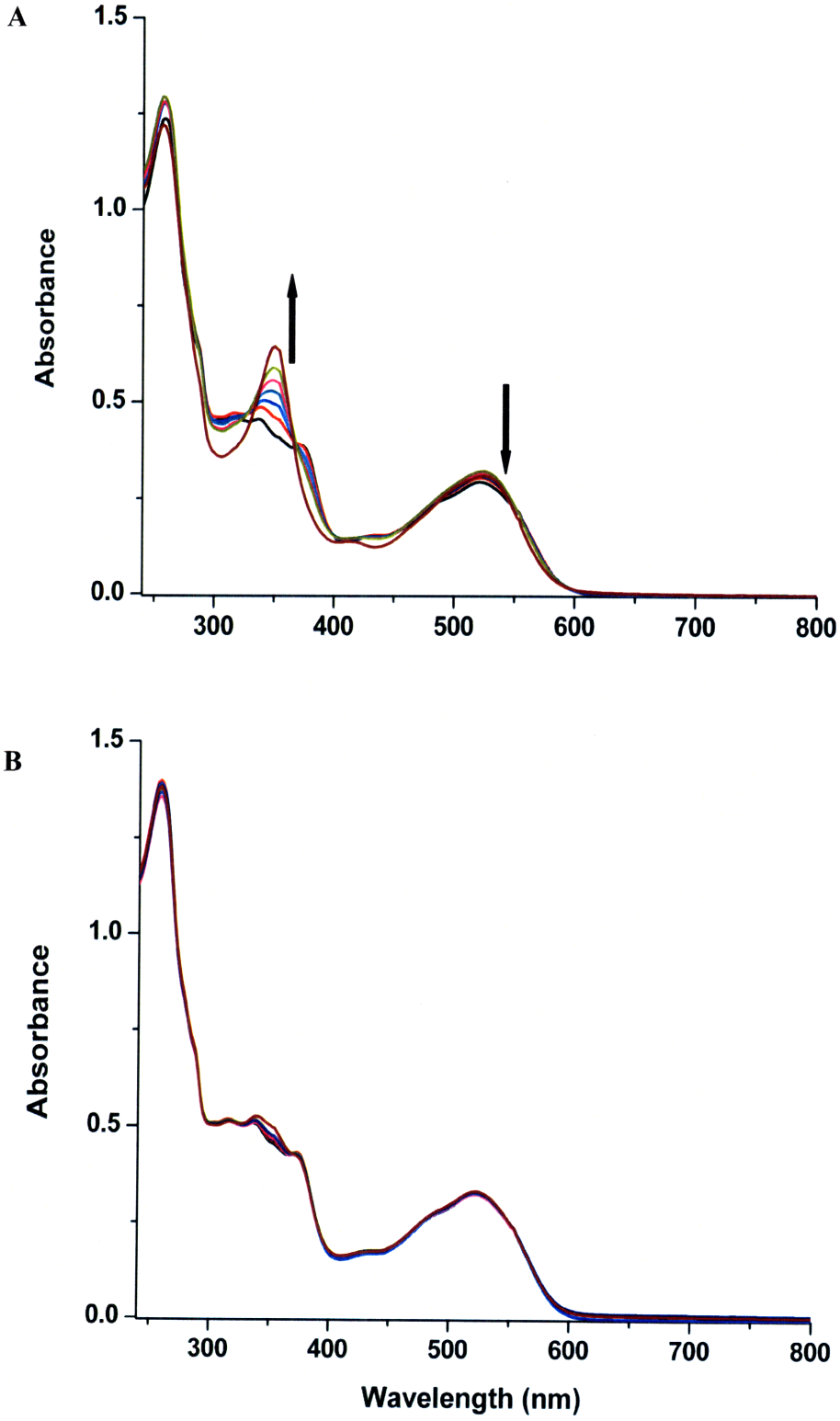

A
1.5
1.0
0.5
0.0
Absorbance
300
400
500
600
700
800
B
1.5
1.0
0.5
0.0
Absorbance
300
400
500
600
700
800
Wavelength (nm)

Figure 2.6 Stability of DTT at 83 ºC in RNR Buffer. An initial UV/visible spectra of DTT (30 mM) was collected at 25 ºC in RNR buffer (50 mM HEPES, 20 mM $MgCl_2$) prior to raising the temperature to 83 ºC. After a 5 min equilibration period, spectra were obtained every 5 min for a total of 30 min. Each spectrum was as follows: 25 ºC (red line), 83 ºC after 5 min incubation at 83 ºC (blue), 83 ºC at 20 min (aqua), and 83 ºC at 30 min (magenta). Black line is buffer blank. Oxidation of DTT can be monitored by measuring the increase in absorbance at 283 nm.

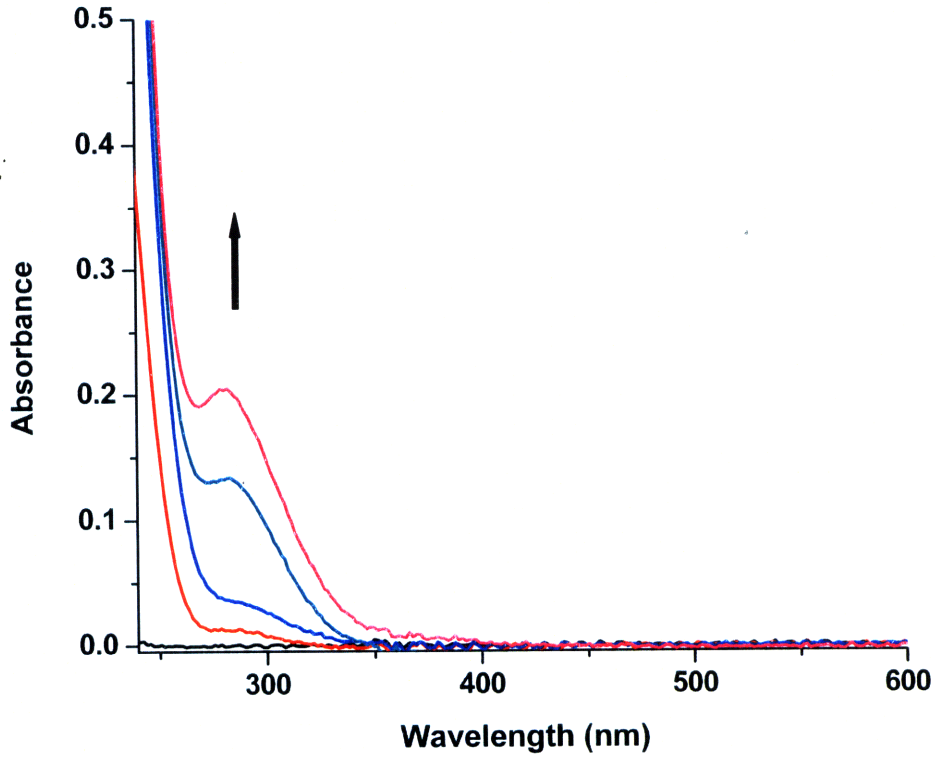
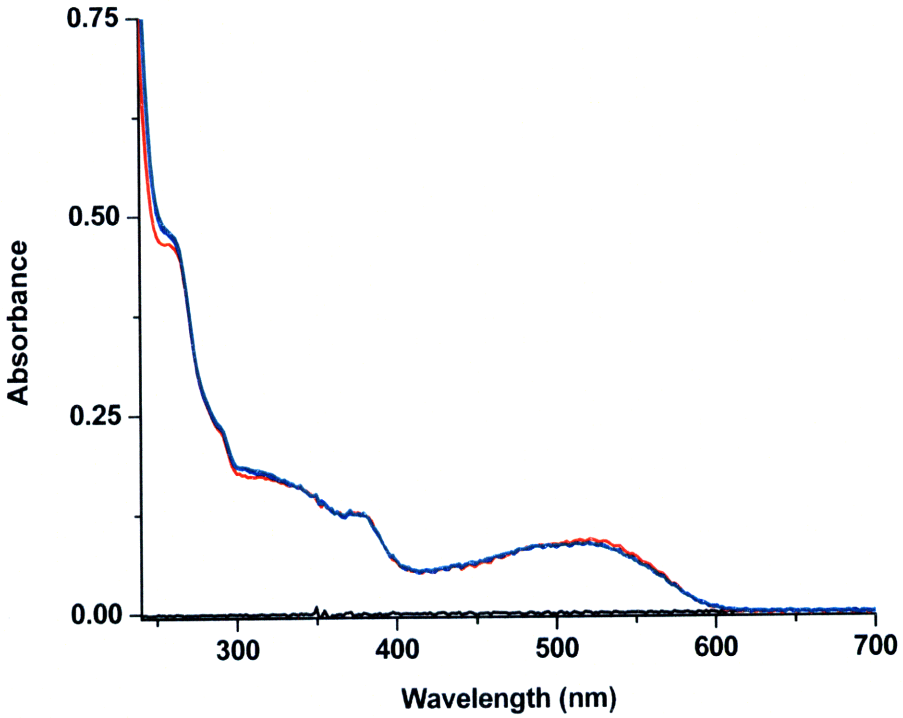


Figure 2.7 Stability of AdoCbl at 83 °C. (A) Spectral recordings of AdoCbl in light protective reaction tubes in RNR reaction buffer (25 mM HEPES, pH 8.0, 20 mM $MgCl_2$) at 83 °C (A) without DTT and (B) with 30.0 mM DTT. Samples were incubated at 83 °C for 5 min and spectra were collected for a total of 10 min. Graph displays timepoint at 0 (red line), 5 (blue line), and 10 (aqua line) min. In B, the absorbance in the 290-310 nm wavelength range can be attributed to the DTT spectra (See Figure 2.6). Black line is buffer blank.

A


0.75
0.50
0.25
0.00
Absorbance
300
400
500
600
700
Wavelength (nm)

B

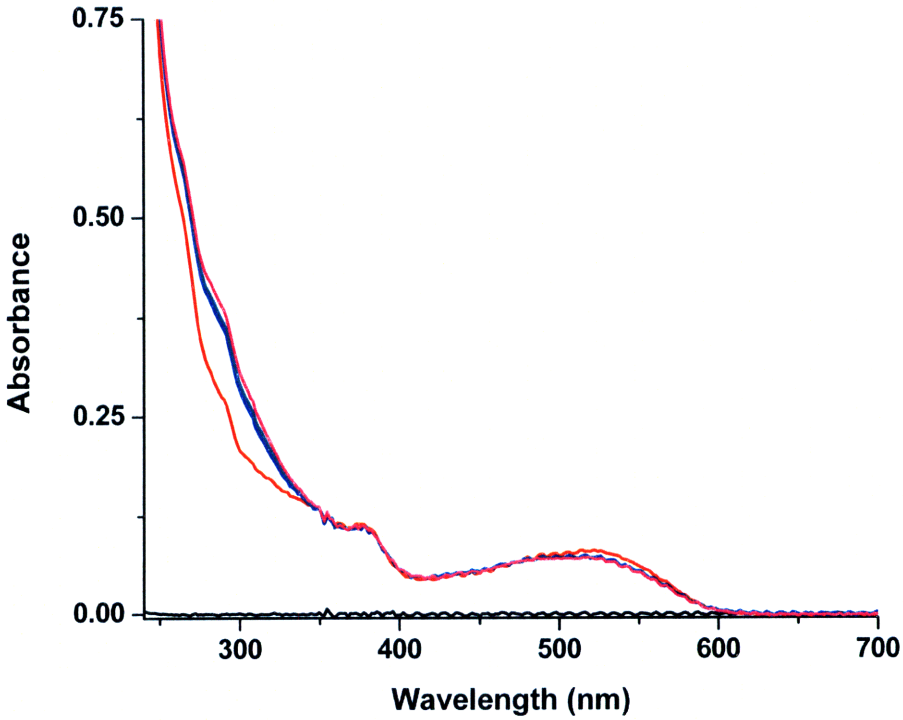

0.75
0.50
0.25
0.00
Absorbance
300
400
500
600
700
Wavelength (nm)

Figure 2.8 Stability test of CDP in HEPES and Tris-HCl buffer at 83 °C. CDP added to both the RNR reaction buffer (25 mM HEPES, pH 8.0, 20 mM $MgCl_2$), and Tris-HCl reaction buffer (50 mM Tris-HCl, pH 8.0, 20 mM $MgCl_2$) for 20 min at 83 °C. (A) HPLC elution profile of CDP in HEPES RNR reaction buffer. (B) HPLC elution profile of CDP in Tris-HCl buffer (50 mM Tris-HCl, pH 8.0, 20mM $MgCl_2$). Over course of the reaction, CDP was observed to decompose into CDP and CMP.

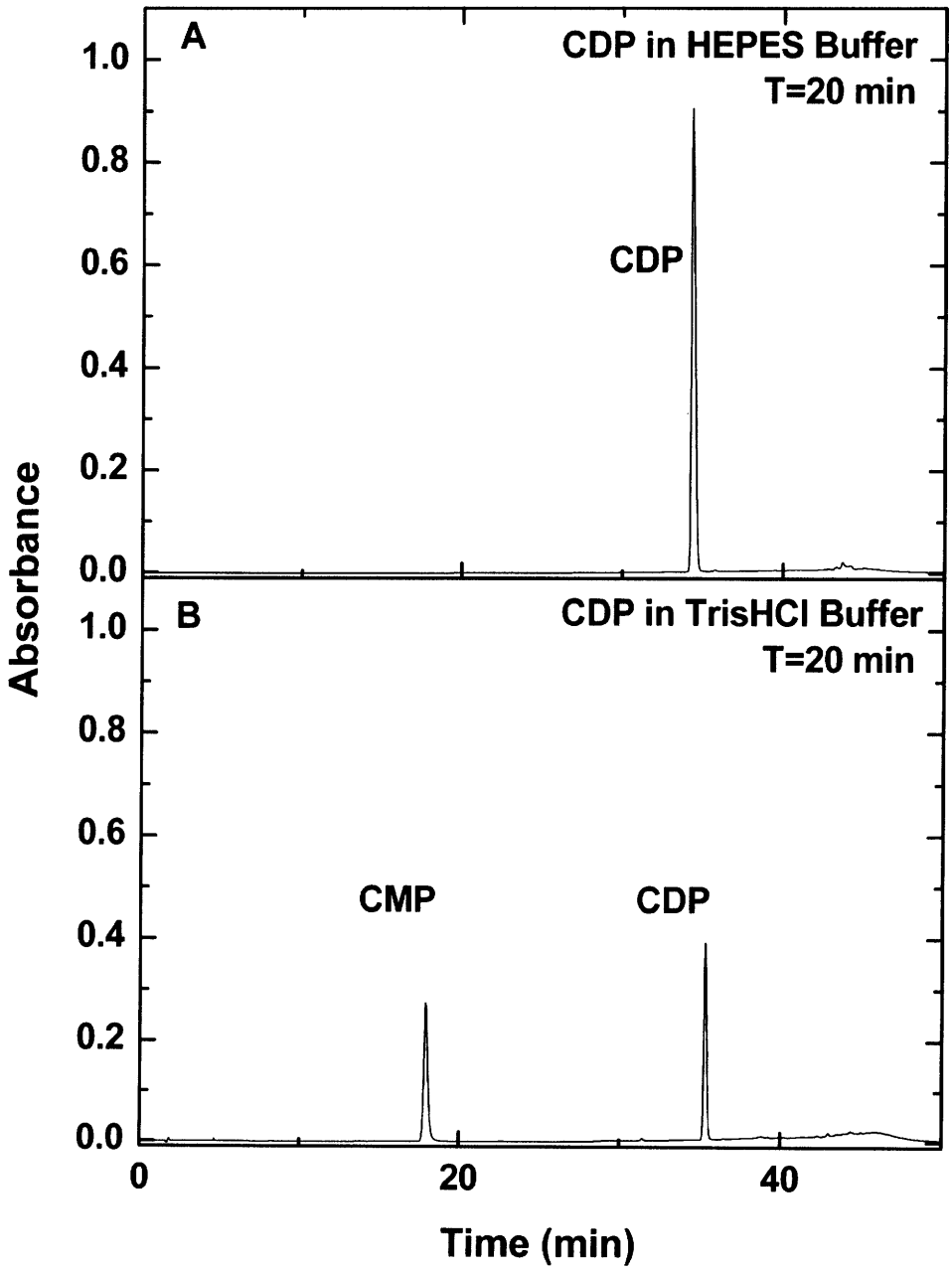


Figure 2.9 Stability test of CTP in HEPES RNR reaction buffer at 83 °C. (A) Initial HPLC elution profile of CTP in HEPES RNR reaction buffer after 5 min incubation at 83 °C to mimic RNR assay conditions. (B) HPLC elution profile of CTP in HEPES reaction buffer (25 mM HEPES, pH 8.0, 20 mM $MgCl_2$) at 25 min. Over course of the 20min, CTP exhibits slight change in ratios of peaks.

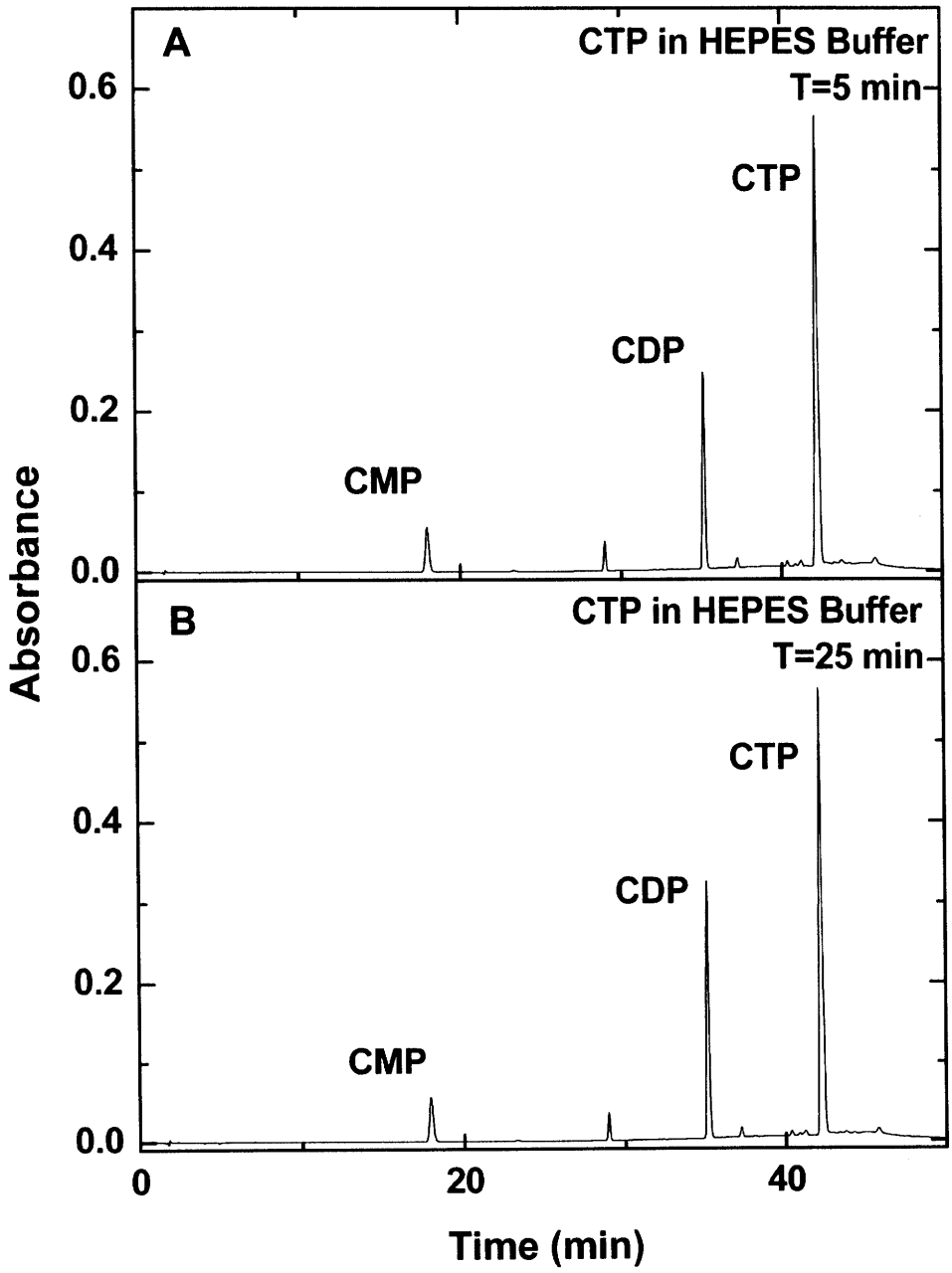


Figure 2.10 Stability test of CTP in Tris-HCl buffer at 83 °C. (A) Initial HPLC elution profile of CTP in Tris-HCl reaction buffer (50 mM Tris-HCl, pH 8.0, 20 mM $MgCl_2$) after 5 min incubation at 83 °C to mimic RNR assay conditions. (B) HPLC elution profile of CTP in Tris-HCl reaction buffer at 25 min. There is substantial degradation of CTP after the 5 min reading, and by 25 min all of the CTP has been decomposed into CDP and CMP.

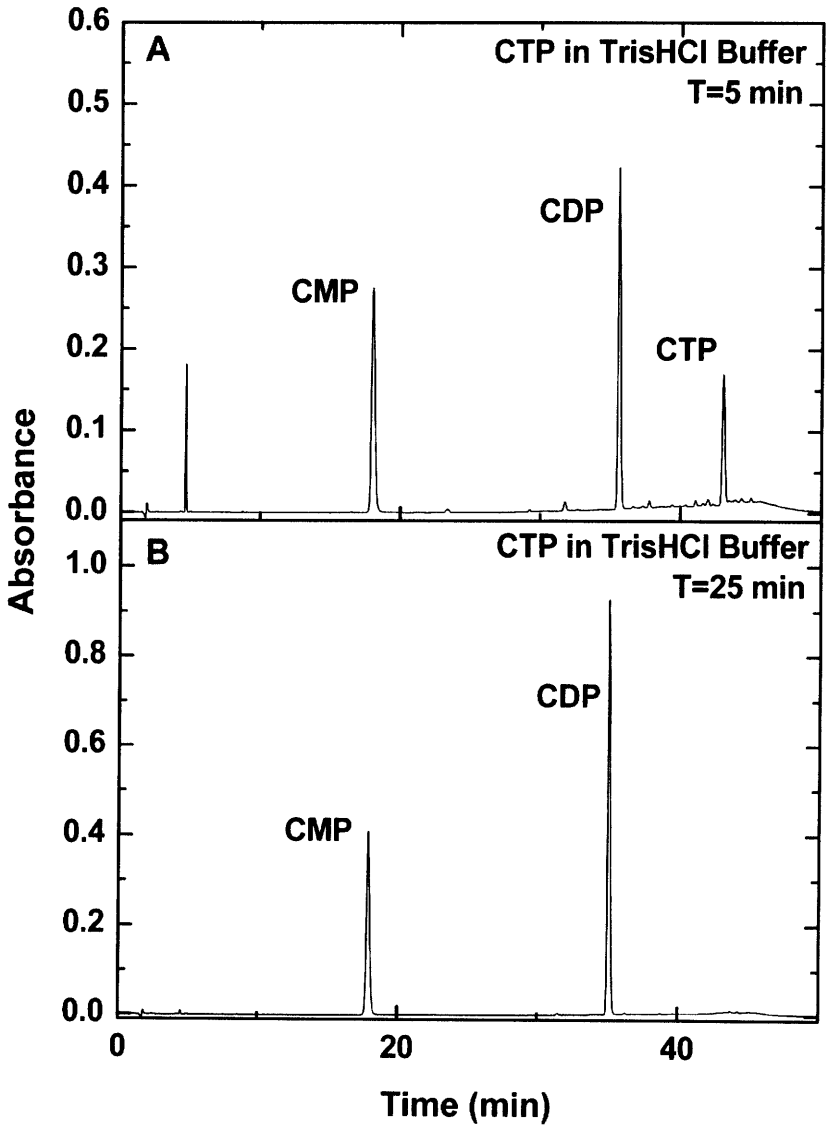


Figure 2.11 The RNR from *A. fulgidus* is AdoCbl dependent. No observable reduction of CDP by *afu*RNR in the absence (black line and squares) of AdoCbl, but *afu*RNR, in the presence of AdoCbl, reduces CDP to dCDP in an enzyme dependent manner (1 mM [2-^{14}C]-CDP (S.A. = 1.5 x 10^3 – 3.0 x 10^3 cpm/nmol), 25 mM HEPES pH 7.5, 10 mM $MgCl_2$, 0.40 μM (red line) or 0.80 μM *afu*RNR (blue line), 9.23 mM DTT and 0.85 μM AdoCbl).

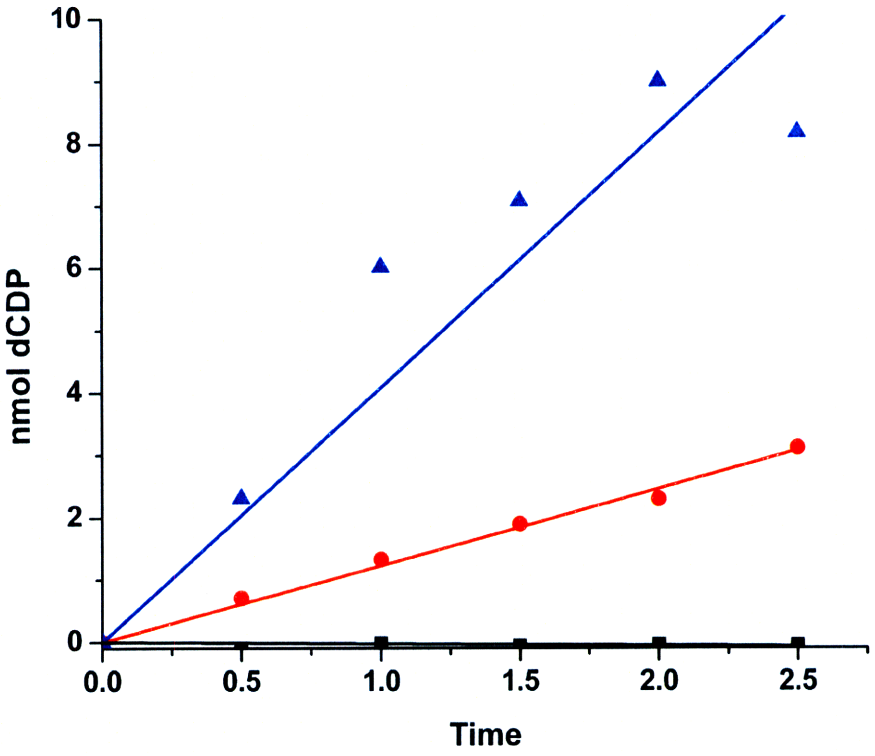


Figure 2.12 Dependence of *afu*RNR on $MgCl_2$ concentration. The 700 µL reaction mixture contained 1.0 mM [2-^{14}C]-CDP (S.A. = 1.5 x 10^3 – 3.0 x 10^3 cpm/nmol), 20 mM HEPES, pH 7.5, 0.40 µM *afu*RNR (30 µL of 3.5 µg/µL), 30 mM DTT 29 µM AdoCbl and 0.5 – 15.0 mM $MgCl_2$ and was carried out at 83 °C.

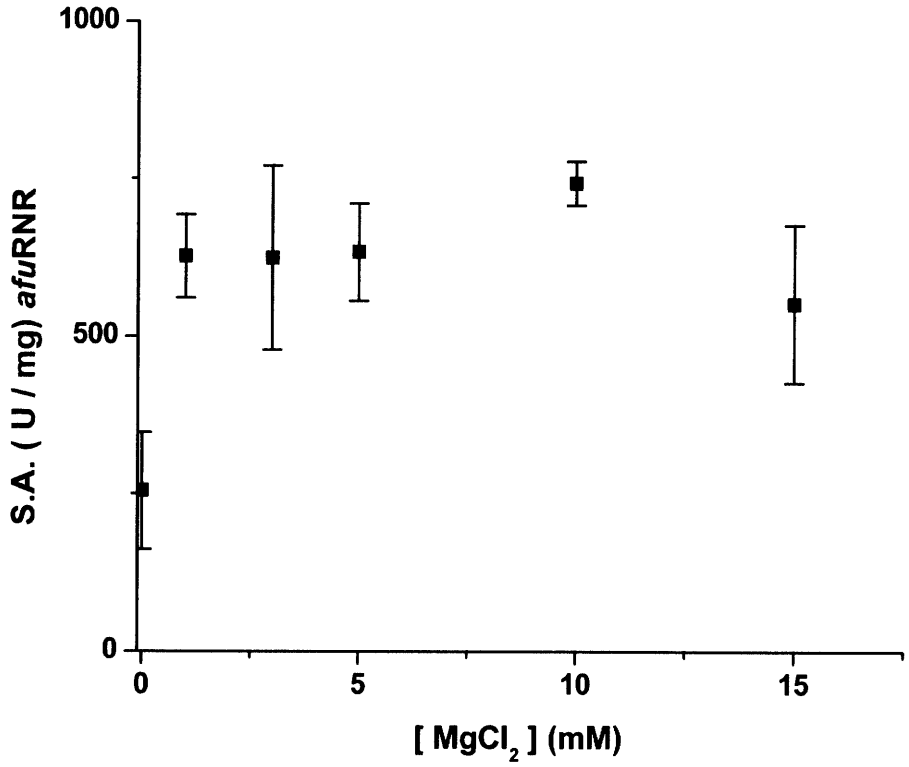


Figure 2.13 Temperature dependence of *afu*RNR activity. The *afu*RNR displays maximum activity at 83 °C. This is in agreement with the optimal growth temperature of *A. fulgidus* which is reported at 83 °C (38). The 700 µL assay mixture contained 1.0 mM [2-^{14}C]-CDP (S.A. = 1.5 x 10^3 – 3.0 x 10^3 cpm/nmol), 25 mM HEPES pH 7.5, 10 mM $MgCl_2$, 0.40 µM *afu*RNR (30 µL of 3.5 µg/µL), 29 mM DTT, and 43.4 µM AdoCbl.

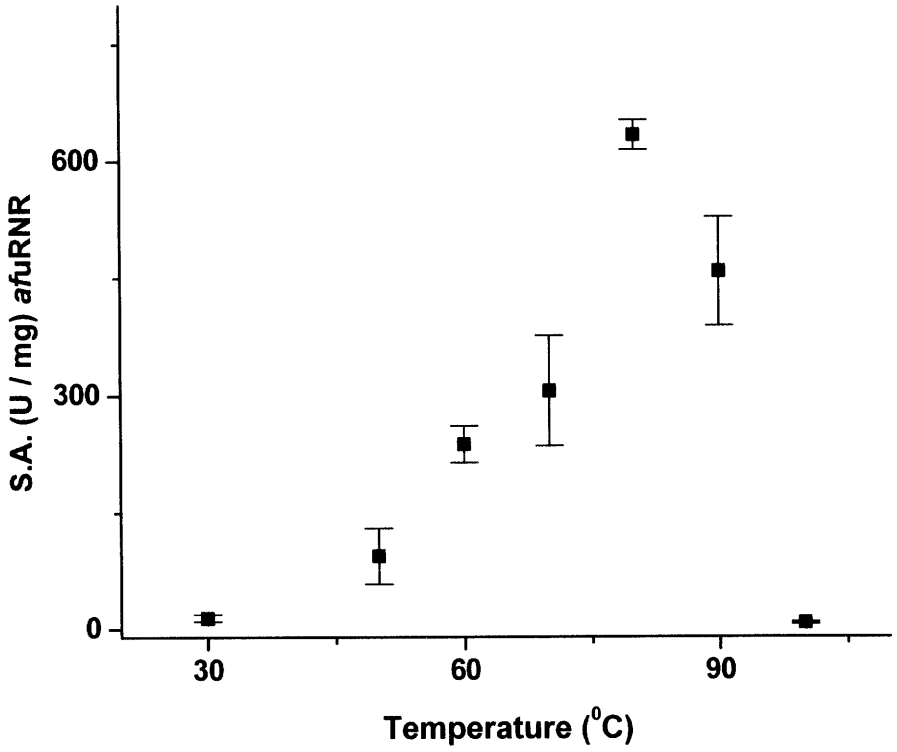


Figure 2.14 *Km* Determination of *afu*RNR for CDP. The aparent *Km* for CDP is 0.92 ±0.16 mM. The 700 µL assay mixture contained 0.06 – 4.9 mM [2-^{14}C]-CDP (S.A. = 1.5 x 10^3 – 3.0 x 10^3 cpm/nmol), 25 mM HEPES pH 7.5, 10 mM $MgCl_2$, 0.40 µM *afu*RNR (30 µL of 3.5 µg/µL), 30 mM DTT and 35.0 µM AdoCbl at 83 ºC.

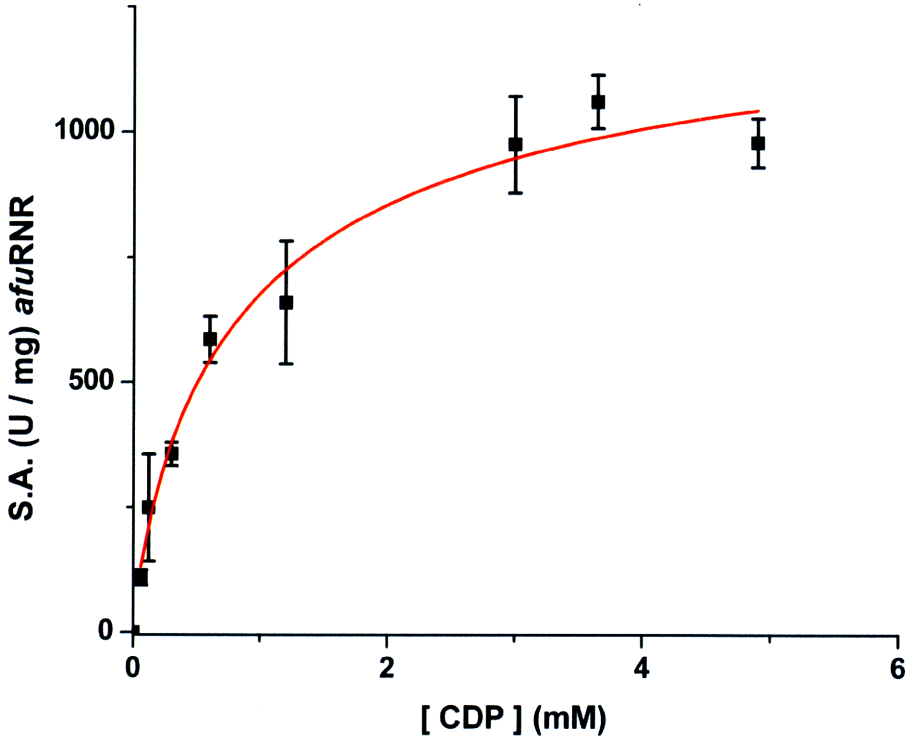


Figure 2.15 *Km* Determination of *afu*RNR for AdoCbl. The aparent *Km* for AdoCbl is 11 ±7.6 µM. The assay mixture contained 1.0 mM [2-^{14}C] CDP (S.A. = 1.5 x 10^3 – 3.0 x 10^3 cpm/nmol), 25 mM HEPES pH 7.5, 10 mM $MgCl_2$, 0.40 µM *afu*RNR (30 µL of 3.5 µg/µL), 29 mM DTT and 0.74 – 43 µM AdoCbl at 83 ºC.

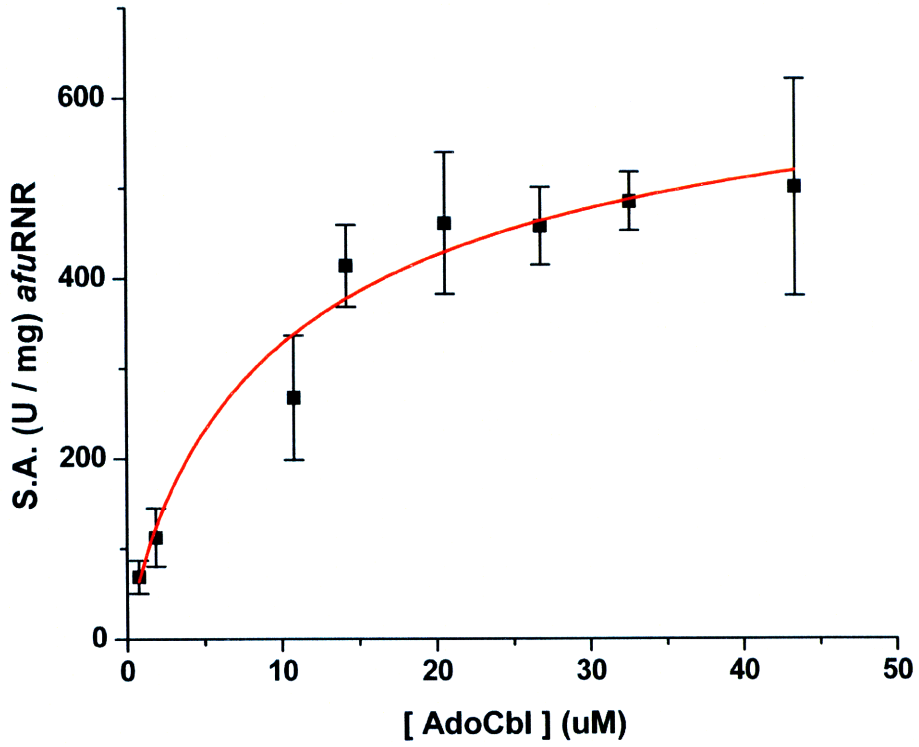


Figure 2.16 *Km* Determination of *afu*RNR for DTT. The aparent *Km* for DTT is 4.18 ±0.78 mM. The assay mixture contained 1.0 mM [2-^{14}C]-CDP (S.A. = 1.5 x 10^3 – 3.0 x 10^3 cpm/nmol), 25 mM HEPES pH 7.5, 10 mM $MgCl_2$, 0.40 µM *afu*RNR (30 µL of 3.5 µg/µL), 26 µM AdoCbl, and 0 – 36 mM DTT at 83 ºC.

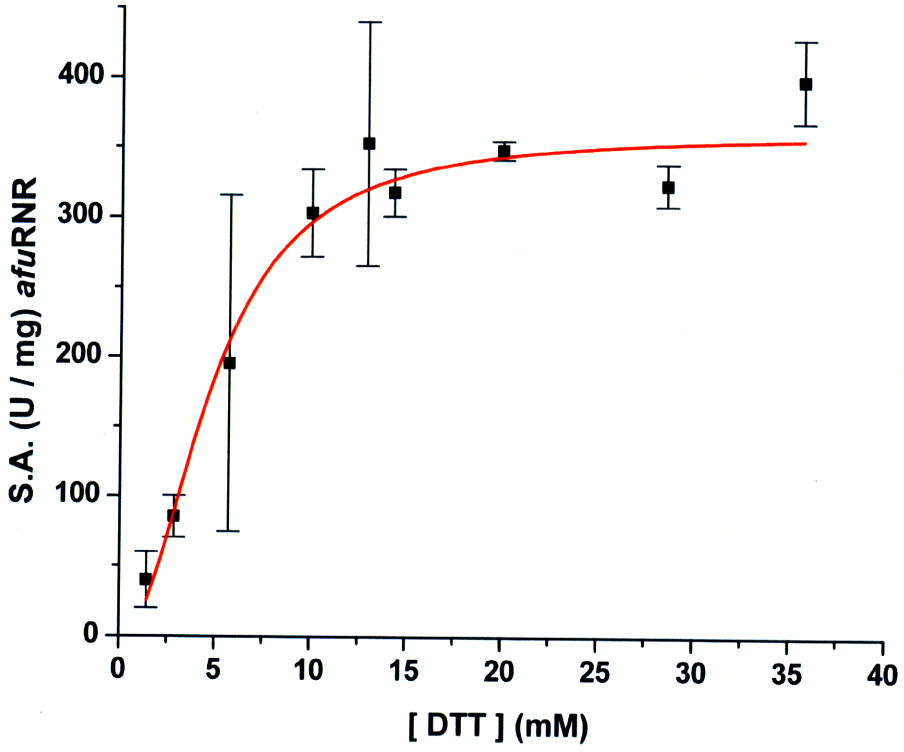


Figure 2.17 Reduction of CDP and CTP by *afu*RNR and RTPR. *afu*RNR reduces both CDP (blue line) and CTP (magenta line) to their corresponding deoxynucleotides under RNR reaction buffer conditions. In contrast, RTPR reduces CTP (black line) to dCTP but does not reduce CDP (red line) to dCDP. A 700 µL *afu*RNR assay mixture contained 1.0 mM [2-^{14}C]-CDP or [2-^{14}C]-CTP (S.A. = 1.5 x 10^3 – 3.0 x 10^3 cpm/nmol), 25 mM HEPES pH 7.5, 10 mM $MgCl_2$, 0.40 µM *afu*RNR, 26 µM AdoCbl, and 35 mM DTT at 83 °C. A 700 µL RTPR assay mixture contained 1.0 mM [2-^{14}C]-CDP or [2-^{14}C]-CTP (S.A. = 1.5 x 10^3 – 3.0 x 10^3 cpm/nmol), 25 mM HEPES pH 7.5, 1.0 M NaOAc, 10 mM $MgCl_2$, 0.48 µM RTPR, 8.0 µM AdoCbl, and 30 mM DTT at 37 °C.

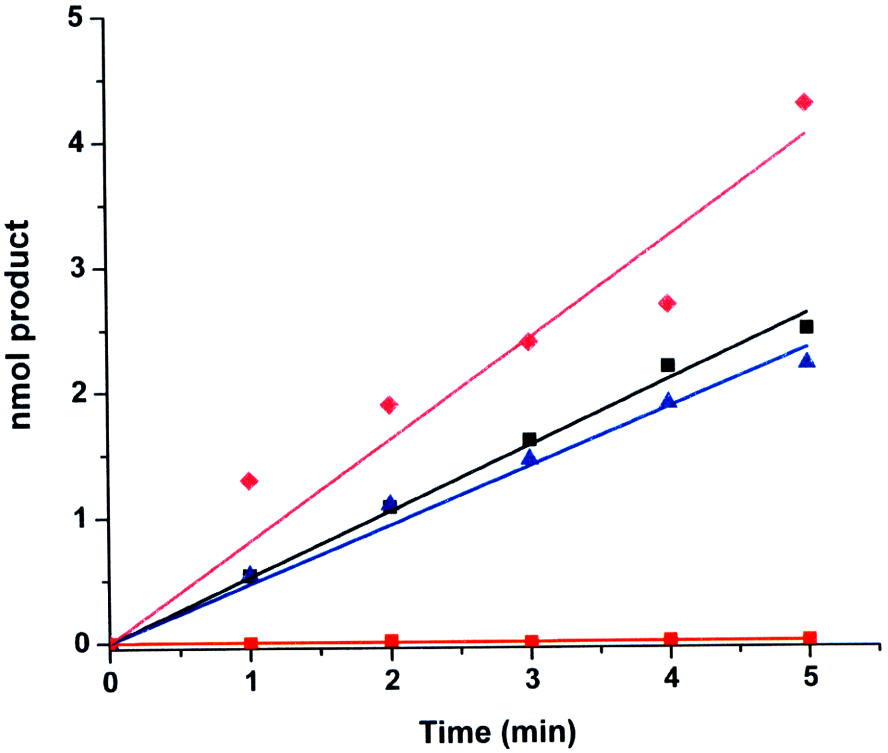


Figure 2.18 HPLC analysis of *afu*RNR CDP reaction. (A) UV/visible spectra of CDP standard elution profile. (B) UV/visible spectra of dCDP standard elution profile. (C) UV/visible spectra of CTP standard elution profile. (D) UV/visible spectra of *afu*RNR reaction mixture before addition of AdoCbl. (E) Radiogram of *afu*RNR reaction mixture before addition of AdoCbl. (F) UV/visible spectra of *afu*RNR reaction after 20 min. (G) Radiogram of endpoint of *afu*RNR reaction after 20 min.

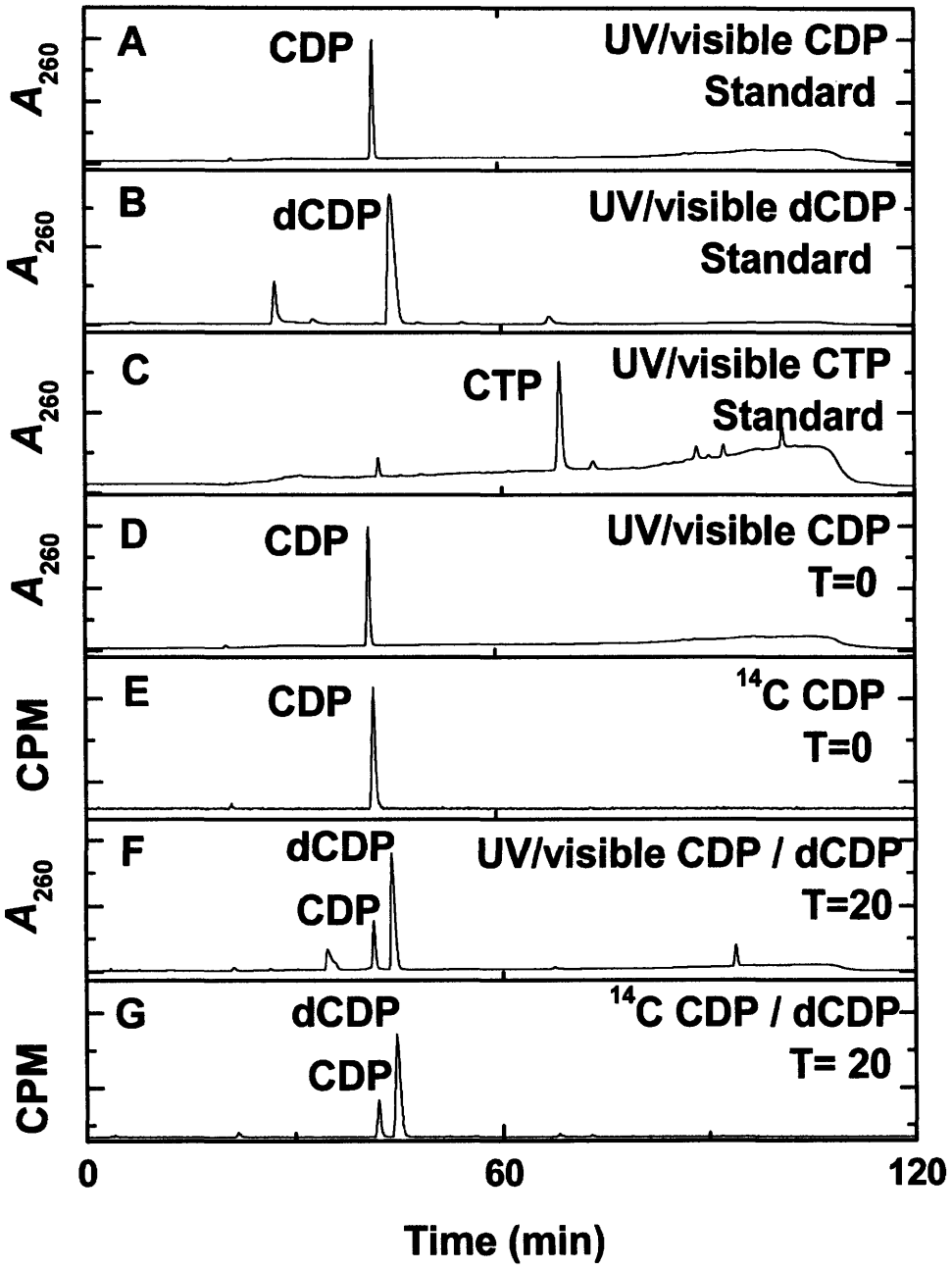
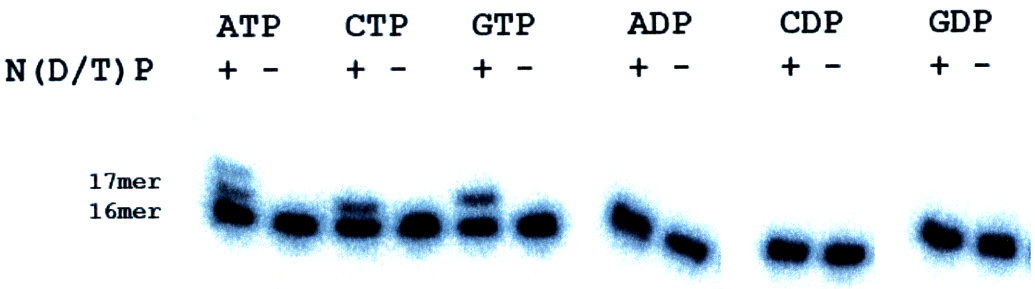
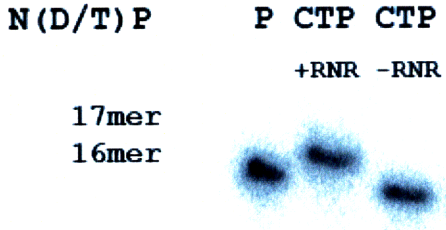

A
CDP
UV/visible CDP
Standard
B
dCDP
UV/visible dCDP
Standard
C
CTP
UV/visible CTP
Standard
D
CDP
UV/visible CDP
T=0
E
CDP
14C CDP
T=0
F
dCDP
CDP
UV/visible CDP / dCDP
T=20
G
dCDP
CDP
14C CDP / dCDP
T= 20
A260
CPM
0
60
120
Time (min)

Figure 2.19 Primer extension assay measuring incorporation of dNTPs produced by *afu*RNR reaction. (A) Reactions were run in the presence and absence of each NTP or NDP as control for nucleotide insertion to its complementary base. The 1+ (18-mer) primer extension seen in the ATP product extension is a known artifact of the Klenow fragment polymerase attributed to high concentration of dATP. (B) Primer extension is *afu*RNR dependent. Due to the longer incubation time (25 min) and the higher *afu*RNR concentration (300 μM) all of the 16mer is converted to 17mer. The lane marked 'P' is the primer control. (C) The incubation of NDPs, dNDPs, or NTPs in RNR assay buffer mixture in the absence of *afu*RNR does not lead to primer extension by the Klenow fragment polymerase.

A


ATP CTP GTP ADP CDP GDP
N(D/T)P + - + - + - + - + - + -
17mer
16mer

B


N(D/T)P P CTP CTP
+RNR -RNR
17mer
16mer

C

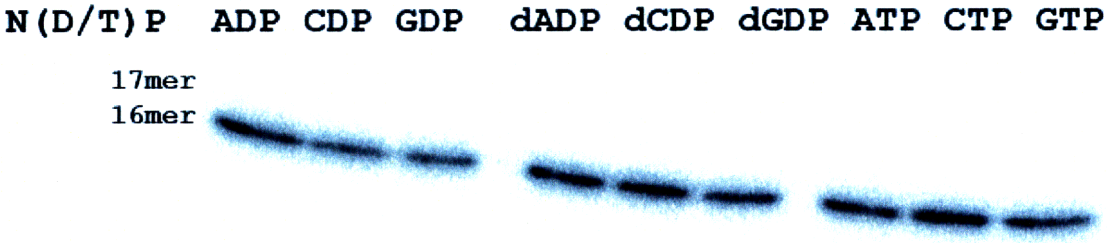
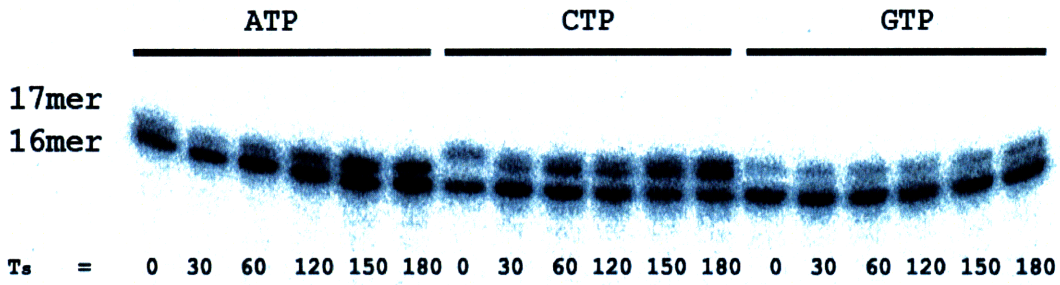

N(D/T)P ADP CDP GDP dADP dCDP dGDP ATP CTP GTP
17mer
16mer

Figure 2.20 Primer extension assay showing incorporation of dNTPs. (A) Time course of *afu*RNR dNTP products incorporated into complementary strand using Klenow fragment polymerase. (B) Time course of RTPR dNTP products incorporated into complementary strand by Klenow fragment polymerase.

A



B

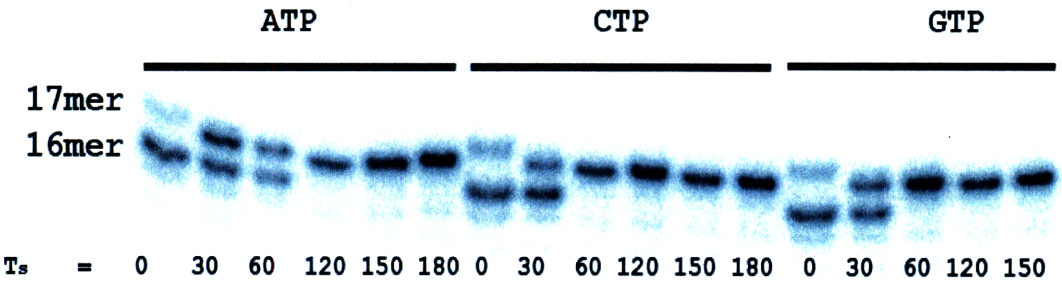


Figure 2.21 *Km* Determination of *afu*RNR for CTP. The apparent *Km* for CTP is 1.3 ±0.14 mM. The 700 µL reaction mixture contained 0.07 – 5.25 mM [^{14}C]-CTP (S.A. = 2.0 x 10^3 – 3.0 x 10^3 cpm/nmol), 25 mM HEPES pH 7.5, 10 mM $MgCl_2$, 0.46 µM *afu*RNR, 30 mM DTT and 35 µM AdoCbl at 83 ºC.

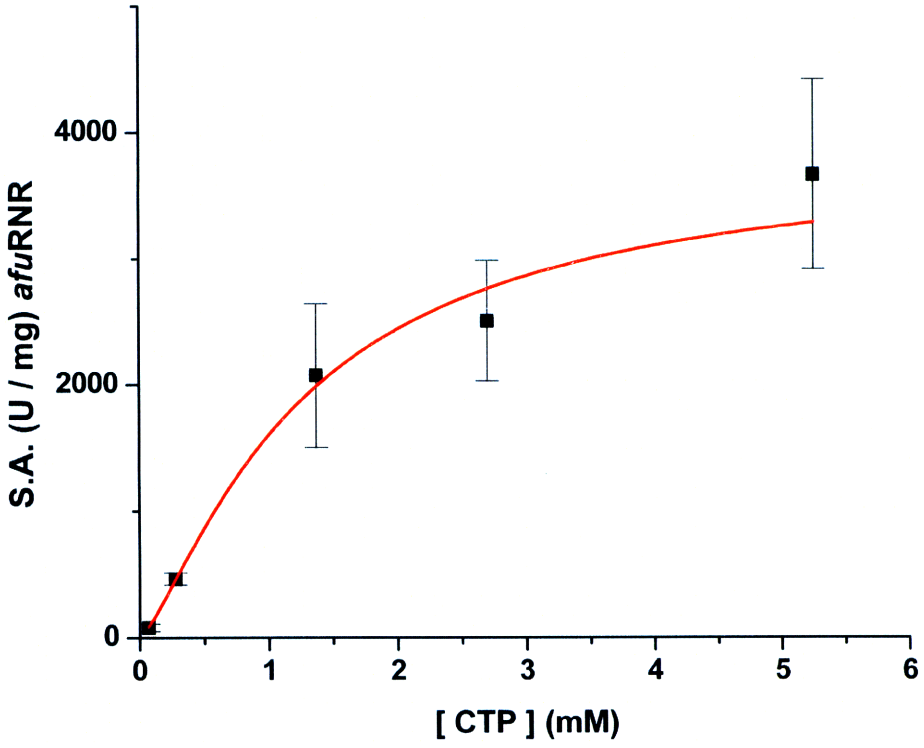


Figure 2.22 Sequence alignment of Class II RNRs showing the conserved motif, NPC(G/A)EX$_8$C(N/V)L, between the active thiyl radical cysteine (C436) and one of the redox active cysteines, (C447). The two cysteines are highlighted by a star at the bottom of the alignment. The spacing between these two cysteines varies between 12 and 15 amino acids for Class I RNRs.

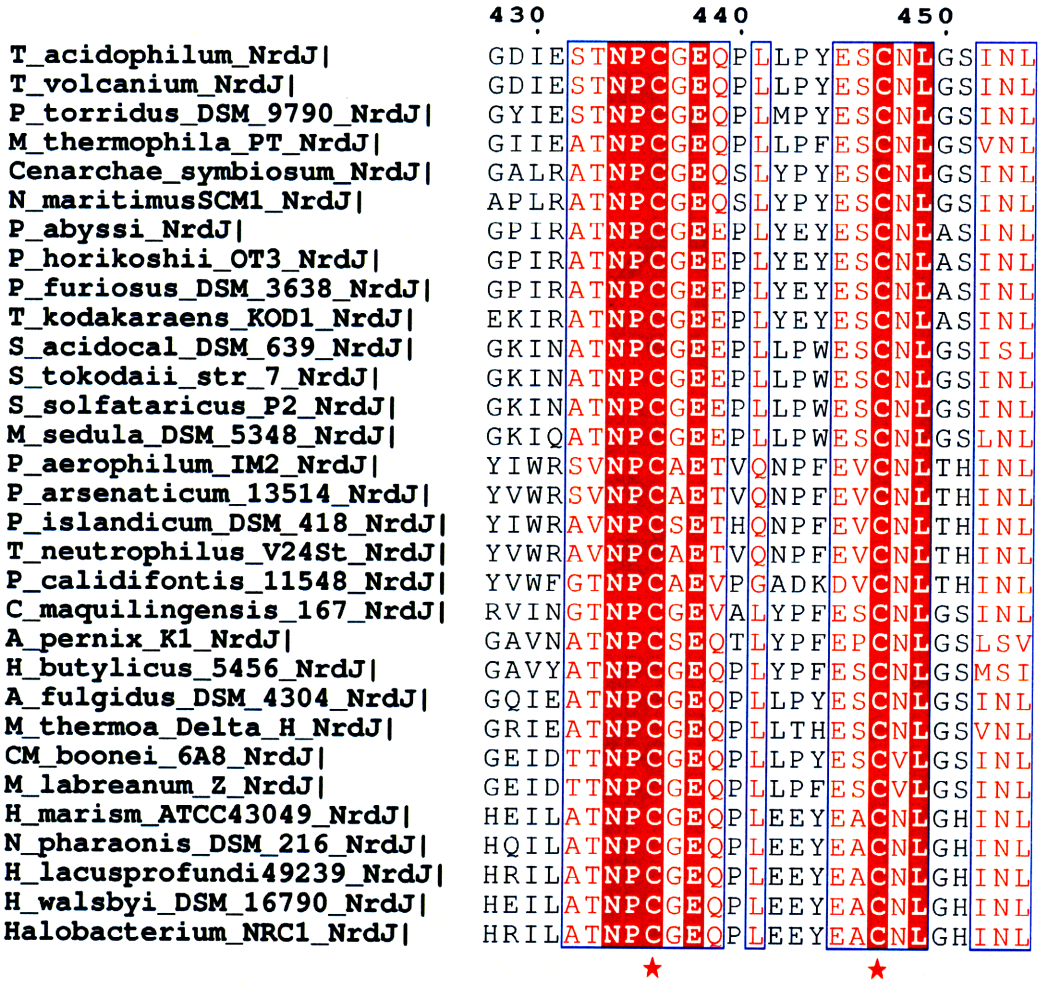
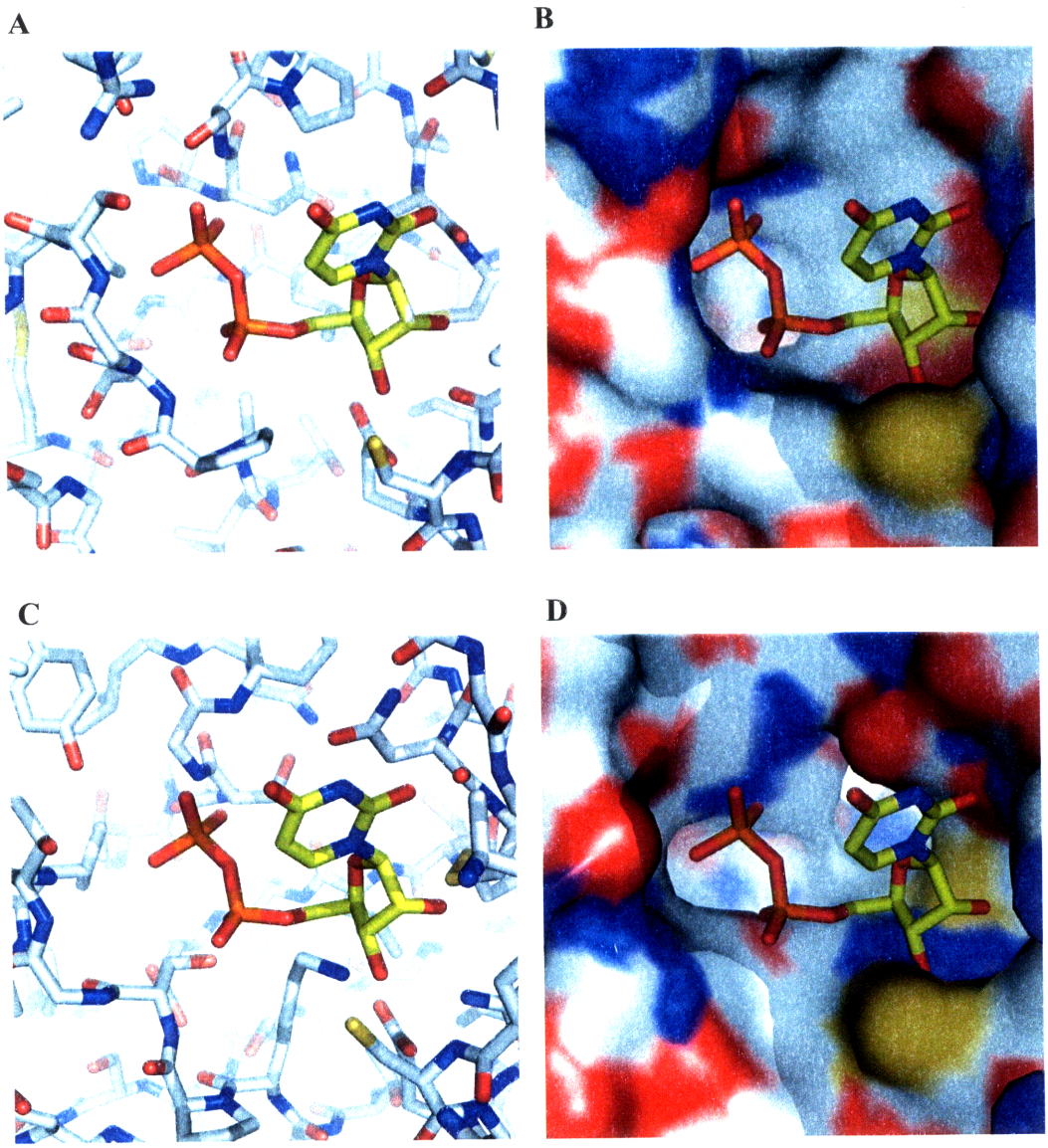


Figure 2.23 Stick and surface representation of *T. maritima* and RTPR active site. (A) Stick representation of the active site of *T. maritima* with GDP co-crystallized in the active site plus residues within 4 Å of substrate. (B) Surface representation of *T. maritima* active site with substrate. (C) Stick representation of the active site of RTPR with GDP modeled in the active site and residues within 4 Å of substrate. (D) Surface representation of RTPR active site with modeled substrate. Colors correspond to: red for oxygen, dark blue for nitrogen, and grey for carbon. The active site cysteine is colored in Yellow in the space fill model.



REFERENCES

1. Nordlund, P., Reichard, P.,(2006) Ribonucleotide reductases. *Annu Rev Biochem* 75, 681-706.

2. Kolberg, M., Strand, K. R., Graff, P., Andersson, K. K.,(2004) Structure, function, and mechanism of ribonucleotide reductases. *Biochim Biophys Acta* 1699, (1-2), 1-34.

3. Stubbe, J.,(2000) Ribonucleotide reductases: the link between an RNA and a DNA world? *Curr Opin Struct Biol* 10, (6), 731-736.

4. Sintchak, M. D., Arjara, G., Kellogg, B. A., Stubbe, J., Drennan, C. L.,(2002) The crystal structure of class II ribonucleotide reductase reveals how an allosterically regulated monomer mimics a dimer. *Nat Struct Biol* 9, (4), 293-300.

5. Stubbe, J.,(1990) Ribonucleotide reductases: amazing and confusing. *J Biol Chem* 265, (10), 5329-5332.

6. Ashley, G. W., Harris, G., Stubbe, J.,(1986) The mechanism of Lactobacillus leichmannii ribonucleotide reductase. Evidence for 3' carbon-hydrogen bond cleavage and a unique role for coenzyme B12. *J Biol Chem* 261, (9), 3958-3964.

7. Mulliez, E., Ollagnier, S., Fontecave, M., Eliasson, R., Reichard, P.,(1995) Formate is the hydrogen donor for the anaerobic ribonucleotide reductase from Escherichia coli. *Proc Natl Acad Sci U S A* 92, (19), 8759-8762.

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

9. Orr, M. D., Vitols, E.,(1966) Thioredoxin from Lactobacillus leichmannii and its role as hydrogen donor for ribonucleoside triphosphate reductase. *Biochem Biophys Res Commun* 25, (1), 109-115.

10. Eliasson, R., Pontis, E., Jordan, A., Reichard, P.,(1999) Allosteric control of three B12-dependent (class II) ribonucleotide reductases. Implications for the evolution of ribonucleotide reduction. *J Biol Chem* 274, (11), 7182-7189.

11. Achenbach-Richter, L., Stetter, K. O., Woese, C. R.,(1987) A possible biochemical missing link among archaebacteria. *Nature* 327, (6120), 348-349.

12. Booker, S., Stubbe, J.,(1993) Cloning, sequencing, and expression of the adenosylcobalamin-dependent ribonucleotide reductase from Lactobacillus leichmannii. *Proc Natl Acad Sci U S A* 90, (18), 8352-8356.

13. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., Lipman, D. J.,(1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25, (17), 3389-3402.

14. Higgins D., T. J., Gibson T.Thompson J.D., Higgins D.G., Gibson T.J,(1994) CLUSTAL W: improving the sensitivity of progressivemultiple sequence alignment through sequence weighting,position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.*, (22), 4673-4680.

15. Gouet, P., Courcelle, E., Stuart, D. I., Metoz, F.,(1999) ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics* 15, (4), 305-308.

16. Steeper, J. R., Steuart, C. D.,(1970) A rapid assay for CDP reductase activity in mammalian cell extracts. *Anal Biochem* 34, 123-130.

17. Barker, H. A., Smyth, R. D., Wawszkiewicz, E. J., Lee, M. N., Wilson, R. M.,(1958) Enzymic preparation and characterization of an alpha-L-beta-methylaspartic acid. *Arch Biochem Biophys* 78, (2), 468-476.

18. Barker, H. A., Smyth, R. D., Weissbach, H., Munch-Petersen, A., Toohey, J. I., Ladd, J. N., Volcani, B. E., Wilson, R. M.,(1960) Assay, purification, and properties of the adenylocobamide coenzyme. *J Biol Chem* 235, 181-190.

19. Weissbach, H., Ladd, J. N., Volcani, B. E., Smyth, R. D., Barker, H. A.,(1960) Structure of the adenylcobamide coenzyme: degradation by cyanide, acid, and light. *J Biol Chem* 235, 1462-1473.

20. Kaczka, E., Wolf, D. E., Folkers, K.,(1949) Identification of crystalline vitamin B12a. *J Am Chem Soc* 71, (4), 1514.

21. Kaczka, E. A., Wolf, D. E., Kuehl, F. A., Jr., Folkers, K.,(1950) Vitamin B12: reactions of cyano-cobalamin and related compounds. *Science* 112, (2909), 354-355.

22. Cleland, W. W.,(1964) Dithiothreitol, a New Protective Reagent for Sh Groups. *Biochemistry* 3, 480-482.

23. Borovok, I., Kreisberg-Zakarin, R., Yanko, M., Schreiber, R., Myslovati, M., Aslund, F., Holmgren, A., Cohen, G., Aharonowitz, Y.,(2002) Streptomyces spp. contain class Ia and class II ribonucleotide reductases: expression analysis of the genes in vegetative growth. *Microbiology* 148, (Pt 2), 391-404.

24. Kornberg, A.,(1957) Enzymatic synthesis of deoxyribonucleic acid. *Harvey Lect* 53, 83-112.

25. Lehman, I. R., Bessman, M. J., Simms, E. S., Kornberg, A.,(1958) Enzymatic synthesis of deoxyribonucleic acid. I. Preparation of substrates and partial purification of an enzyme from Escherichia coli. *J Biol Chem* 233, (1), 163-170.

26. Benkovic, S. J., Cameron, C. E.,(1995) Kinetic analysis of nucleotide incorporation and misincorporation by Klenow fragment of Escherichia coli DNA polymerase I. *Methods Enzymol* 262, 257-269.

27. Metro, G., Fabi, A., Russillo, M., Papaldo, P., De Laurentiis, M., Ferretti, G., Pellegrini, D., Nuzzo, C., Graziano, V., Vici, P., Introna, M., Felici, A., Cognetti, F., Carlini, P.,(2008)

Taxanes and gemcitabine doublets in the management of HER-2 negative metastatic breast cancer: towards optimization of association and schedule. *Anticancer Res* 28, (2B), 1245-1258.

28. Blakley, R. L.,(1978) Ribonucleoside triphosphate reductase from Lactobacillus leichmannii. *Methods Enzymol* 51, 246-259.

29. Traut, T. W.,(1994) Physiological concentrations of purines and pyrimidines. *Mol Cell Biochem* 140, (1), 1-22.

30. Jordan, A., Torrents, E., Jeanthon, C., Eliasson, R., Hellman, U., Wernstedt, C., Barbe, J., Gibert, I., Reichard, P.,(1997) B12-dependent ribonucleotide reductases from deeply rooted eubacteria are structurally related to the aerobic enzyme from Escherichia coli. *Proc Natl Acad Sci U S A* 94, (25), 13487-13492.

31. Larsson, K. M., Jordan, A., Eliasson, R., Reichard, P., Logan, D. T., Nordlund, P.,(2004) Structural mechanism of allosteric substrate specificity regulation in a ribonucleotide reductase. *Nat Struct Mol Biol* 11, (11), 1142-1149.

32. Thelander, L., Reichard, P.,(1979) Reduction of ribonucleotides. *Annu Rev Biochem* 48, 133-158.

33. Mathews, C. K., Ji, J.,(1992) DNA precursor asymmetries, replication fidelity, and variable genome evolution. *Bioessays* 14, (5), 295-301.

34. Meuth, M.,(1989) The molecular basis of mutations induced by deoxyribonucleoside triphosphate pool imbalances in mammalian cells. *Exp Cell Res* 181, (2), 305-316.

35. Mathews, C. K.,(2006) DNA precursor metabolism and genomic stability. *Faseb J* 20, (9), 1300-1314.

36. Vitols, E., Brownson, C., Gardiner, W., Blakley, R. L.,(1967) Cobamides and ribonucleotide reduction. V. A kinetic study of the ribonucleoside triphosphate reductase of Lactobacillus leichmannii. *J Biol Chem* 242, (13), 3035-3041.

37. Tauer, A., Benner, S. A.,(1997) The B12-dependent ribonucleotide reductase from the archaebacterium Thermoplasma acidophila: an evolutionary solution to the ribonucleotide reductase conundrum. *Proc Natl Acad Sci U S A* 94, (1), 53-58.

38. Stetter, K. O., Lauerer, G., Thomm, M., Neuner, A.,(1987) Isolation of Extremely Thermophilic Sulfate Reducers: Evidence for a Novel Branch of Archaebacteria. *Science* 236, (4803), 822-824.

This page intentionally left blank

Chapter 3:

Thioredoxin reductase from *Thermoplasma acidophilum*: a new twist on redox regulation

INTRODUCTION

Tight regulation of the cellular redox environment is crucial for cell survival. Cellular viability relies on the reduction of metabolites for energy production, on the oxidation of small molecules to assemble structural scaffolds, and on the maintenance of the proper disulfide state in the extracellular or intracellular environments. An imbalance in redox regulation can trigger a cascade of genetic responses, the most extreme of which can lead to cell death (Reviewed in (1-3)).

The principal protein systems that act as reductants are NADPH-dependent thioredoxin reductase (TrxR) / thioredoxin (Trx), and glutathione reductase (GrR) / glutathione (GS) / glutaredoxin (Grx) (Figure 3.1A) (1, 2, 4, 5). NADPH reduces the FAD of TrxRs, which in turn reduces Trxs through disulfide exchange (Figure 3.1B). The TrxR / Trx system is responsible for providing reducing equivalents for many cellular processes including the reduction of ribonucleotide reductase (RNR) during the conversion of nucleotides to deoxynucleotides (6). This reaction is essential for maintaining the appropriate levels of deoxynucleotides in the cell for DNA repair, and for cellular replication (7, 8). The significance of the redox regulation of these and other biochemical pathways has lead to renewed interest in anti-bacterial or anti-tumorgenic therapies that are targeted to disrupt redox regulation (1, 9, 10).

Here we study putative TrxR and Trx from *Thermoplasma acidophilum*, an organism first isolated from a self-heating coal refuse pile at the Friar Tuck mine in southwest Indiana. This organism lacks a cell wall and optimal growth occurs at 59 °C and at pH 1-2 (11). Despite the organism's ability to grow at low pH, the cytoplasmic pH of *T. acidophilum* was reported to be between 5.5 and 6.9 (12, 13). The 1.56 Mb genomic sequence of *T. acidophilum* encoding 1,509 open reading frames (ORFs) has been reported (14), allowing us to identify a putative TrxR (Ta0984) and Trx (Ta0866).

We report the cloning, expression, and purification of gene products Ta0984 and Ta0866, and show that Ta0866 exhibits thioredoxin-like activity, and that Ta0984 is able to reduce *T. acidophilum* Trx (*ta*Trx), establishing its identity as a thioredoxin reductase. An accompanying paper describes the full electrochemical characterization of *T. acidophilum* TrxR (*ta*TrxR). While the redox potential is within the expected range for TrxRs, we find that the standard TrxR reducing agents NADPH or NADH do not reduce

the *ta*TrxR flavin. The X-ray structure of *ta*TrxR to 2.35 Å resolution presented here reveals a major modification of the classic TrxR NADPH binding site, suggesting that *T. acidophilum* may utilize a novel reductant.

EXPERIMENTAL PROCEDURES

Materials. The organism *Thermoplasma acidophilum* (ATCC number 27656) was obtained from the American Tissue Culture Center. TOPO and TOPO Zero Blunt PCR cloning kits were obtained from Invitrogen (Carlsbad, CA). NovaBlue competent cells, Rosetta™ (DE3) pLysS competent cells, and pET28a plasmid were purchased from Novagen (Madison, WI, USA). *pfu*Turbo DNA Polymerase and a PCR Optimization Kit were obtained from Stratagene (Cedar Creek, TX, USA). Primers were obtained from Integrated DNA Technologies (Coralville, IA, USA). Nickel nitrilotriacetic acid (Ni-NTA) was purchased from Qiagen (Valencia, CA, USA). Restriction enzymes were purchased from New England BioLabs (Beverly, MA, USA). Kanamycin, chloroamphenicol, TRIS-HCl, ethylenediaminetetraacetic acid (EDTA), and phenylmethylsulfonylfluoride (PMSF) were obtained from Sigma-Aldrich (St. Louis, MO, USA).

Genomic analysis of Thermoplasma acidophilum genome. The genome of *Thermoplasma acidophilum DSM1728* (RefSeq NC_002578, GenBank AL139299) was analyzed using the web interface at the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov) for putative thioredoxin reductase (TrxR) and thioredoxin (Trx) open reading frames (ORFs). An ORF encoding a 319 amino acid protein, Ta0984, and an ORF encoding for a 113 amino acid protein, Ta0866, were identified by sequence homology to known TrxRs and Trxs using PSI-BLAST (15). Protein sequences were aligned using the program ClustalW (16) and the output was rendered using the web interface ESPript 2.2 (17) (Figure 3.2 and 3.3).

Cloning of taTrxR and taTrx. The *ta*TrxR and *ta*Trx genes were amplified by the polymerase chain reaction (PCR) using genomic DNA from a cell suspension and the primers 5'- A TGC CAT ATG GCA TAT GGA ATT TAA CCT GCA TGC AGT A-3' and 5'-TCA GCC AAG CTT CAT TTT TTG GAT ATA GAA TCA-3' for *ta*TrxR and 5' A TGC CAT ATG GCA TAT GAA GAA TTA TAT GGG CTG CG-3' and 5'-TAG CGG CCG CTC AGA CGG AAT CGC AGG GA-3' for *ta*Trx (restriction sites are underlined).

PCR products were cloned into the expression vector pET28a using NdeI and HindIII restriction enzymes for *ta*TrxR and NdeI and NotI for *ta*Trx. The sequence of the pET28a-*ta*TrxR and pET28a-*ta*Trx cDNA was confirmed at the MIT Biopolymers Sequencing Facility (Figures 3.4 and 3.5). After confirmation of the correct sequences, the resulting constructs were transformed into Rosetta™ (DE3)pLysS cells for facile protein expression.

Purification of recombinant taTrxR and taTrx. Rosetta™ (DE3)pLysS transformed with either pET28a-*ta*TrxR or pET28a-*ta*Trx constructs were used to inoculate TB media supplemented with 20 μg mL^{-1} kanamycin and 34 μg mL^{-1} chloroamphenicol. The culture was grown at 37°C to an OD_{600} of 0.8 and induced with 0.8 mM isopropylthio-ß-D-galactoside (IPTG) (not in materials)for 4 hrs. The cells were harvested by centrifugation, resuspended in lysis buffer (10 mM TRIS-HCl pH 8.0, 150 mM NaCl, 10 mM imidazole, 1 mM PMSF), and disrupted by sonication using a Digital Sonifier 250 (Branson, CN, USA). Insoluble material was removed by centrifugation at 40,000 x g using a Beckman Avanti J-25 centrifuge (Beckman, PA, CA, USA). The supernatant was subjected to heating at 70°C for 45 min, and then centrifuged at 40,000 x g. The clarified supernatant was loaded onto a Ni-NTA column, washed with high salt buffer, and eluted with buffer containing 10 mM TRIS-HCl pH 8.0, 150 mM NaCl, 200 mM imidazole. Eluted fractions were analyzed by an SDS-PAGE 4-20% gel and those fractions containing proteins with comparable size to *ta*TrxR or *ta*Trx were concentrated using an Amicon (Millipore, MA, USA) ultrafiltration stir cell.

The eluted protein was dialyzed overnight in 20 mM HEPES, pH 7.5, 20 mM NaCl, 0.5 mM EDTA, 1 mM 1,4-dithio-DL-threitol (DTT) with thrombin (0.05 unit mL^{-1}) to remove the N-terminal poly-His-tag. The digested protein fraction was loaded onto an SEC-30 high-prep size exclusion column from Amersham Biosciences (Piscataway, NJ, USA), which was pre-equilibrated with storage buffer (10 mM HEPES, pH 7.0, 20 mM NaCl). Fractions containing protein of either MW ~34 kDa (*ta*TrxR) or MW ~13 kDa (*ta*Trx) were concentrated using Centriprep YM-3 filter from Amicon (Millipore, Bedford, MA, USA). Protein purity was analyzed by SDS-PAGE 4-20% gel (Figure 3.6) and protein concentration was determined from the theoretical extinction coefficient at 280 nm (23,040 M^{-1} cm^{-1} for *ta*TrxR and 16,500 $M^{-1}cm^{-1}$ for *ta*Trx) (ProtParam, ExPaSy

web server). A Cary 300 spectrophotometer (Varian, CA, USA) was used to analyze the purified *ta*TrxR for flavin incorporation.

Temperature stability of NADPH and NADH. NADPH (0.22 μM) and NADH (0.26 μM) were analyzed for temperature dependent degradation in 50 mM TRIS-HCl, pH 7.5, 150 mM NaCl (Figure 3.7). Initial UV-visible spectra were collected at 20 °C from 800 to 240 nm. The temperature was increased to 80° C and the reaction was allowed to equilibrate for 5 min before full spectra were collected. The reaction was cooled to 20° C and then another full spectrum was collected. To measure time dependent degradation, NADPH or NADH was heated to the desired temperature (20 °C, 60 °C, and 80 °C) and a reading was collected at 340 nm in 5 min increments for 45 min. Degradation of NADPH or NADH was inferred from the decrease in absorbance at 340 nm normalized to 100 % of the absorbance for the sample at time zero. Experiments were performed in triplicate.

Activity of taTrx measured by insulin and DTT assays. Purified *ta*Trx was assayed for disulfide reductase activity by the insulin reduction method (18). The standard assay mixture contained 1.25 mM bovine insulin, 100 mM potassium phosphate, pH 7.0, 2 mM EDTA and *ta*Trx (43 μM, 66 μM, or 70 μM). The reaction was initiated by the addition of 1 mM DTT. An increase in absorbance at 650 nm was monitored at 25 °C.

Purified oxidized *ta*Trx was also assayed for its ability to be reduced by *ec*TrxR using the 5,5'-dithiobis(2-nitrobenzoate (DTNB) method (19). Briefly, this colorimetric assay measures the Trx dependant reduction of DTNB to 5-thio-2-nitrobenzoic acid (TNB), which absorbs strongly at 405-414 nm, using TrxR and NADPH as electron donors. The reaction was performed using 3.8 μM *ec*TrxR, in 50 mM TRIS-HCl, pH 8.0, 2 mM EDTA, 0.5 mM NADPH, and 0.6 mM DTNB. Each assay was initiated by the addition of *ta*Trx (5 – 60 μM) to reaction mixture and the conversion of DTNB to TNB was monitored by the increase in absorbance at 412 nm.

Titration of taTrxR and ecTrxR with NADPH and NADH. *ta*TrxR (43.3 μM) was titrated with NADPH or with NADH (0.5 to 5.0 mole NAD(P)H / mole FAD) under anaerobic conditions at 59° C in 50 mM TRIS, pH 7.5, 150 mM NaCl. Spectra were collected from 800 nm to 240 nm to monitor the reduction of the enzyme bound flavin by NADPH or NADH. *ec*TrxR (40.5 μM) was titrated with NADPH or with NADH (0.5 to 5.0 mol

NADPH or NADH / mol FAD) under anaerobic conditions at 25 °C in 50 mM TRIS, pH 7.5, 150 mM NaCl. Spectra for NADPH reduction of *ec*TrxR were recorded as described above. The reduction of bound FAD was monitored by observing the change in absorbance at 457 nm over 45 min. For *ta*TrxR, spectra were collected every minute and for *ec*TrxR spectra were collected every 0.1 min for initial 5 min and then every 5 min thereafter.

Ability of taTrxR to reduce taTrx. To verify that the reduced *ta*TrxR was capable of passing electrons to *ta*Trx, the xanthine/xanthine oxidase reduction system, described by Massey and co-workers, was used (20). After complete reduction of a sample of *ta*TrxR, the re-oxidation of *ta*TrxR (4 μM) was monitored spectroscopically in the visible region following the addition of *ta*Trx or *ec*Trx (8 μM). The *ta*TrxR reduction system was as follows: in a 1 cm path length cuvette stirred continuously and held at 25° C in an MBraun Labmaster glove box, a solution of 100mM phosphate buffer (pH 7.5) with 1 μM benzyl viologen, 250 μM xanthine and 4 μM *ta*TrxR, an aliquot of xanthine oxidase was added to a final concentration of ~ 0.25 nM, to initiate the reaction. This solution was monitored in the visible region of the spectrum, using an SI Photonics spectrophotometer contained within the glove box, over the course of 1.5 hours, upon which complete reduction of the *ta*TrxR was achieved. Oxidized Trx was then added to a final concentration of 8 μM, and the spectral features were monitored.

Analysis of T. acidophilum genome for alternative redox regulation pathways. To investigate if *T. acidophilum* uses alternative characterized reduction pathways, the *T. acidophilum* genome was mined for known pathways by using NCBI BLAST (15) to compare known proteins involved in the formation of glutathione or of coenzyme F420 to the ORFs of *T. acidophilum.* Pathway members were identified using ExPASy PATHWAY service (http://ca.expasy.org/cgi-bin/lists?pathway.txt).

Crystallization and data collection. *ta*TrxR was crystallized by incubating 2.0 μL of protein solution (65 mg mL^{-1} *ta*TrxR in 10 mM HEPES, pH 7.5 and 20 mM NaCl) and 2.0 μL of precipitant solution (0.28 M $MgCl_2$, 0.1 M Bis-Tris, pH 5.5, 25% (w/v) PEG 3350) in a sitting drop at room temperature. Crystals had rhombohedral geometry, and their bright yellow color confirmed the incorporation of flavin adenine dinucleotide (FAD) by the enzyme. Crystals appeared cubic and grew to approximately 100 μm per

side in 3 to 4 days. These crystals were soaked for 30 seconds in a cryo-protectant solution of 0.28 M $MgCl_2$, 0.1 M Bis-Tris, pH 5.5, 25% (w/v) PEG 3350, 25% PEG 400, before cryo-cooling in liquid nitrogen.

All diffraction data were collected at 100 K at the Advanced Light Source (Berkeley, CA), beam line 5.0.2. Data were processed and scaled using DENZO and SCALEPACK from the HKL (21) program suite (See Table 1). The crystals belong to the I23 (I-centered cubic) space group ($a = b = c = 165.95$ Å), with one homodimer per asymmetric unit.

Molecular replacement solution and refinement. The crystal structure of the thioredoxin reductase was solved by the method of molecular replacement. Initial phases were obtained from a full length poly-alanine model of thioredoxin reductase from *Escherichia coli* (PDB code 1TDE) (22). The solution was found using PHASER (23) with two search models consisting of the flavin binding domain and the NADPH-binding domain of the *E. coli* structure, utilizing data from a resolution range of 4.0 – 12.0 Å. PHASER was first used to find a solution for the FAD-binding domain of *ta*TrxR (residues 1 – 126 and 250 – 319). This solution was fixed and a second round of molecular replacement was performed to find the NADPH-binding domain (residues 127 – 249).

Refinement was carried out in CNS (24) with rigid body refinement followed by simulated annealing. Manual fitting of amino acid residues was done using XFIT (25). Subsequent rounds of refinement included simulated annealing, energy minimization, and B-factor refinement without sigma cutoff. Topology and parameter files for the FAD moiety were obtained from HIC-UP (26) (PDB code 1N1P). Water molecules were added with CNS and were manually checked against $2F_o$-F_c and F_o-F_c electron density maps. A composite omit map was prepared in CNS and used to confirm the assignment of model atoms. The final model contains a biologically relevant dimer in the asymmetric unit, which consists of residues 14-319 for molecule 'A' of 319 and residues 13-319 of 319 for molecule 'B', one flavin molecule per monomer and 114 water molecules.

RESULTS

Cloning, purification, and characterization of taTrxR and taTrx. Using the *T. acidophilum* genome database, we identified two ORFs (Accession no. Ta0984 and

Ta0866): one encoding a putative thioredoxin reductase and the other one encoding a thioredoxin homologue. The Ta0984 gene encodes a protein of 319 amino acids with a predicted molecular mass of 34116 Da (Figure 3.4). The deduced amino acid sequence contains 28% identity (48% similar) to the *ec*TrxR (Figure 3.2). The Ta0866 gene encodes a 133 amino acid protein (Figure 3.5) with a CxxC disulfide-reducing motif and displays 34% identity (58% similarity) to *ec*Trx (Figure 3.3). These ORFs have been amplified by PCR, cloned, sequenced, and compared to the sequence in the *T. acidophilum* genome database.

The purified *ta*TrxR yields visible absorption spectra typical for flavoproteins with absorbance maxima at 380 and 460 nm. Freshly purified protein has an A_{280} / A_{460} ratio of between 7.5 and 8.0, in agreement with the full incorporation of FAD as previously observed for *ec*TrxR (27). Expression and subsequent purification yields ~18.5 mg from a 1 liter culture. The molecular mass of the protein is estimated to be approximately 34 kDa by SDS-PAGE (Figure 3.6A). This molecular mass is in agreement with values predicted by the gene analysis. The native molecular mass is determined to be about 70 kDa by gel-filtration using a Sephacryl SEC-75 HR 16/60 column. These results suggest that the native state of *ta*TrxR is a homodimer, similar to *ec*TrxR (28). Expression of *ta*Trx yields a protein of 13.2 kDa as analyzed by SDS PAGE (Figure 3.6B). Subsequent gel-filtration using a Sephacryl SEC-30 HR 16/60 column reveals a native molecular mass of ~13 kDa, suggesting that the *ta*Trx is a monomer in its native state.

Lack of activity of taTrxR with NADPH or NADH. All previously characterized TrxRs are reduced by NADPH (Reviewed in (2, 29, 30)). For *ec*TrxR, reaction of NADPH with the FAD cofactor occurs rapidly (2,000 turnovers FAD^{-1} min^{-1}) with a K_m for NADPH of 1.2 μM (31). While there is no measurement of K_m for NADH with *ec*TrxR, it has been estimated to be 400 fold higher than that of NADPH (28). Therefore, it is surprising that titration of *ta*TrxR with concentrations of up to 5 fold molar excess NADPH does not reduce its FAD cofactor, showing no spectral change in the visible part of the spectra (Figure 3.8A). For comparison, titration of *ec*TrxR with NADPH results in a rapid reduction of the bound FAD, decreasing the absorbance at 460 nm (Figure 3.8C). Titration of *ta*TrxR with NADH shows minimal reduction of FAD as indicated by the

small change in the shoulder at 460 nm (Figure 3.8B). Although NADPH is the preferred substrate for *ec*TrxR, NADH does reduce the flavin to some extent in this protein (Figure 3.8D). The reduction is, in fact, greater than that observed for *ta*TrxR. Compared to the rapid reduction of *ec*TrxR by NADPH (turnover rate is 2,000 FAD^{-1} min^{-1}) (31), the changes observed for *ta*TrxR appear too slow and too small to be physiologically relevant (Figure 3.8E). The inability of NADPH or NADH to rapidly reduce *ta*TrxR is not due to instability of these components at 59 °C, the physiological temperature of *T. acidophilum*. Analysis of the time-dependent degradation of NADPH and NADH at 59 °C shows that there is negligible degradation at 5 min under assay conditions, not enough to explain this lack of activity (Figure 3.7). At the higher temperature of 80 °C, degradation is increased, but even then less than 10% of NADPH or NADH is degraded during the first five minutes of the assay, and only 30% is degraded by 45 minutes (Figure 3.7).

Disulfide reductase activity of taTrx. The lack of measurable activity between *ta*TrxR and NADPH prevented the use of a coupled assay for *ta*Trx activity, requiring us to turn to other methods to characterize the *ta*Trx. The reduction of insulin disulfides can be monitored by following the increase in turbidity at 650 nm due to the precipitation of the free B-chain (18). The recombinant *ta*Trx reduced insulin in a protein concentration-dependent manner (Figure 3.9). To investigate whether *ta*Trx is able be reduced by *ec*TrxR we performed the DTNB assay using NADPH, *ec*TrxR, and *ta*Trx. *ta*Trx is reduced by *ec*TrxR, with an estimate of the apparent K_m of 14.4 (±7.0) μM for the *ec*TrxR (Figure 3.10). This K_m value is higher than for the *ec*TrxR / *ec*Trx system (K_m of 2.8 μM) (31), which is not surprising in that the two proteins (*ec*TrxR and *ta*Trx) did not evolve to interact. These experiments confirm that the *ta*Trx is a thioredoxin-like protein capable of performing disulfide exchange.

Thioredoxin reductase activity of taTrxR. To verify that *ta*TrxR can catalyze a disulfide exchange with *ta*Trx, we pre-reduced *ta*TrxR using a modified xanthine / xanthine oxidase system (see methods) at 25 °C; after complete reduction of the *ta*TrxR flavin, the reduced enzyme was treated with oxidized *ta*Trx. Upon the introduction of two equivalents of *ta*Trx, the *ta*TrxR was rapidly oxidized, achieving complete re-oxidation of the FAD within one min (Figure 3.11). This experiment indicates that *ta*TrxR is a biologically relevant TrxR, which does not use NADPH as a reductant.

Overall Structure. The structure of *ta*TrxR was solved to 2.35 Å resolution with the physiological dimer in the asymmetric unit and an R and R_{free} of 22.6% and 26.8%, respectively (See Table 1 for statistics). Each *ta*TrxR monomer consists of an FAD-binding domain (residues 14-125 and 251-319) and an "NADPH binding domain" (residues 126-250) (Figure 3.12A). Each domain of *ta*TrxR is quite similar to the equivalent domain of *ec*TrxR, with an r.m.s.d. of 1.23 Å for the 166 Cα of the FAD-binding domain, and an r.m.s.d of 1.17 Å for 121 Cα of the "NADPH-binding domains" (Figure 3.12B and 3.12C). The *ta*TrxR structure has one FAD molecule bound in a similar conformation as the *ec*TrxR structure (Figure 3.13), and with similar protein-FAD interactions (Figure 3.14).

*ec*TrxR is known to undergo a large conformational change during catalysis, rotating between a "flavin oxidizing" form (FO) and a "flavin reducing" form (FR) (22, 32, 33). In the FR form, the FAD domain is positioned such that NADPH can reduce the FAD by hydride transfer (Figure 3.12A). In the FO form, the FAD domain is positioned above the disulfide such that the $FADH_2$ can be oxidized by the disulfide. Structures of both conformations are available for the *E. coli* enzyme: the structure of wild type *ec*TrxR represents the FO state, and a structure of a complex between *ec*TrxR and *ec*Tr represents the FR state. Interestingly, while all other structures of uncomplexed TrxRs are in the FO state, the *ta*TrxR structure is of the FR form. The *ta*TrxR FR conformation is very similar to that of *ec*TrxR in the FR state with an r.m.s.d. of 1.30 Å for chain A and 1.73 Å for chain B when all the C□ atoms are superimposed. As with the *ec*TrxR FR structure, the *ta*TrxR FAD domain is positioned near the putative NADPH binding site, and away from the catalytic cysteines (Cys145 and Cys148), which are oxidized with an S-S distance of 2.04 Å. A discussion of the conformational dynamics of TrxRs is available in the accompanying paper.

Major Differences in the Domain that Typically Binds NADPH. Structural superposition of the "NADPH binding domains" from *ec*TrxR and *ta*TrxR reveal that the backbone of residues 184 - 191 of the *ta*TrxR structure is constrained into a slightly different conformation than is found in *ec*TrxR. This different loop conformation appears to result from the presence of a proline (P188) in the *ta*TrxR sequence (Figure 3.15C), and not the presence or absence of nucleotide in the structure (Figure 3.17). This

variation in loop conformation is subtle and is unlikely to explain the lack of affinity of NADPH for the *ta* enzyme. In contrast, while the "NADPH-binding domains" for *ta*TrxR and *ec*TrxR display an overall similar fold, close examination reveals several highly conserved residues in all bacterial TrxRs that are not conserved in the *ta*TrxR sequence. The *ec*TrxR structure, which was co-crystallized with $NADP^+$, shows that the NADPH 2′-phosphate is surrounded by a histidine (H175) and three arginines (R176, R177, and R181) creating a positively charged pocket that complements the negatively charged phosphate (Figures 3.15A, 3.18A, and 3.16A). The lower activity (400-fold) for NADH compared to NADPH (28) is presumably due to the loss of interactions between the cofactor and residues that form this positively charged pocket (Figure 3.18C). In *ta*TrxR, none of the residues that form this pocket are conserved. The classic NADPH binding motif 'VxxxHRRDxxRA' is replaced with 'VxxxEYMPxxMC' in *ta*TrxR (Figure 3.15D). *ec*TrxR's histidine 175, which points directly toward the 2′-phosphate of the NADPH, is substituted with a glutamic acid (E185) in *ta*TrxR, and the three arginine residues (R176, R177, and R181) are replaced by a tyrosine (Y186) and two methionines (M187 and M191) (Figures 3.15B and 3.15C). These substitutions dramatically change the electrostatics of this protein surface (Figures 3.18 and 3.16). In a model of NADH bound to *ta*TrxR, an unfavorable interaction between E185 and the 2′-phosphate of NADPH is relieved (Figure 3.18D), but neither NADPH nor NADH would be expected to bind with any kind of reasonable affinity, consistent with the lack of rapid reduction of FAD by these molecules. Analysis of all known bacterial and archaeal genomes shows that only one other organism, the close cousin to *T. acidophilum*, *Thermoplasma volcanicum*, contains an identical 'VxxxEYMPxxMC' sequence. All others use a variant of the Arg-rich 'VxxxHRRDxxRA' motif.

DISCUSSION

The protein derived from Ta0984 is a thioredoxin reductase. Our structural and biochemical studies reveal a protein with a FAD cofactor and an overall similar fold to the *ec*TrxR. However, unlike all other characterized TrxRs, *ta*TrxR is not reduced by NADPH. This lack of reactivity is not due to redox potential differences between TrxRs (See accompanying paper). The redox potential of *ta*TrxR is similar to that of *ec*TrxR (-

305 vs. -270 mV) (34), and when reduced with a xanthine / xanthine oxidase system, *ta*TrxR is able to react stoichiometrically with *ta*Trx.

Examination of the *ta*TrxR structure suggests that the lack of activity with NADPH is due to amino acid substitutions as compared to the *E. coli* enzyme. The "NADPH-binding site" of *ta*TrxR does not appear to be designed to bind NADPH, with neutral and negatively charged amino acids replacing the canonical Arg-rich motif. The "NADPH-binding site" is also structurally constrained by P188. Sequence analysis shows that only one other organism, the closely related *Thermoplasma volcanicum*, shares *ta*TrxR's unusual sequence.

The lack of a NADPH-dependent TrxR / Trx reduction system in *T. acidophilum* led us to investigate whether this organism might rely heavily on GrR / Grx systems instead of on thioredoxins. Analysis of the *T. acidophilum* genome, however, revealed no GrR / Grx system, nor enzymes for the biosynthesis of glutaredoxin (35). Our extensive data mining and sequence alignments have not produced any near or distantly related GrR / Grx proteins from this organism.

We also considered whether *T. acidophilum* uses NADH or NADPH in other cellular processes. Our genomic analysis reveals a putative NAD kinase, the enzyme responsible for the conversion of NAD to NADP, and several putative NADH and NADPH utilizing enzymes, which suggests that *T. acidophilium* does use these cofactors. To date, there are only a few characterized enzymes in *T. acidophilum* that use NADH (aldohexose dehydrogenase) (37) or NADPH (glyderaldehyde dehydrogenase) (36) or both NADH and NADPH as substrates (38). However, the use of NADPH or NADH by any *T. acidophilum* enzyme, in addition to the presence of the NAD kinase gene, indicates that availability of these cofactors is not the reason for the observed variation in the TrxR protein.

The above data challenges the paradigm shown in Figure 1A. Our data blocks the arrow to the left, and the lack of GrR / Grx blocks the arrow to the right. Given this, we revisited the *T. acidophilum* genome to search again for other putative TrxR sequences and found none. Unless there is a TrxR with a very different sequence, we have cloned the only one in *T. acidophilum*.

If NADPH does not reduce this *ta*TrxR, and this TrxR is the only TrxR in *T. acidophilum*, and there are no obvious GrR / Grx candidates, there must be some small molecule that can provide reducing equivalents or the cytoplasmic environment would quickly oxidize and the organism would expire. One such molecule, coenzyme F420, has been shown to be a physiological reductant in methane reducing thermophilic organisms (Reviewed in (39-41)). However, genomic analysis of *T. acidophilum* does not reveal any of the enzymes involved in the biosynthesis of F420. The lack of molecular pathways for the biosynthesis of coenzyme F420, along with the lack of other F420-dependent enzymes in *T. acidophilum*, seems to indicate that coenzyme F420 is not a reductant of *ta*TrxR. Of course, we cannot rule out the possibility that F420 is generated by a different set of enzymes in this organism.

Conclusions. To date, there is no well characterized example of an organism that uses a reductant other than NADPH to maintain its intracellular redox environment. Here we present evidence that one organism, *T. acidophilum*, which does not contain a Grx system, uses an alternative method to reduce the bound FAD of the TrxR system. This work lays the foundation for a biochemical search for the molecular redox partners of *T. acidophilum* TrxR.

Table 3.1. Crystallographic and refinement data statistics for *ta*TrxR structure.

Dataset	**Native**
Wavelength (Å)	1.2710
Beamline	ALS 5.0.2
Space Group	I23
a=b=c (Å)	165.95
Resolution (Å)	50. – 2.35
R_{sym} (%)	6.9 (52.1)[a]
No. of Observations	434,549
Unique Observation	31,586
I/□I	32.9 (7.6)
Completeness (%)	99.9 (100.0)
Refinement Statistics	
Resolution Limits (Å)	50. – 2.35
# Unique Reflections	31,123
# Reflections - Test Set	1,824
R_{work} (%)[b]	23.4
R_{free} (%)[c]	26.7
Wilson B factor (Å)	68.3
B factor Chain A (Å)	55.4
B factor Chain B (Å)	62.1
B factor FAD 1 (Å)	49.2
B factor FAD 2 (Å)	59.4
B factor Waters (Å)	68.0
Solvent Content (%)	56.4
# protein atoms	2272 (2282)[d]
# water atoms	115
# non-protein atoms (FAD)	106
Bond length deviation (Å)	0.007
Bond angle deviation (°)	1.10
Ramachandran Plot	
Residues in most favored regions (%)	93.4
Res in allowed region (%)	6.1
Res in disallowed region (%)	0.5

[aP]highest resolution shell is 2.35 Å to 2.43Å

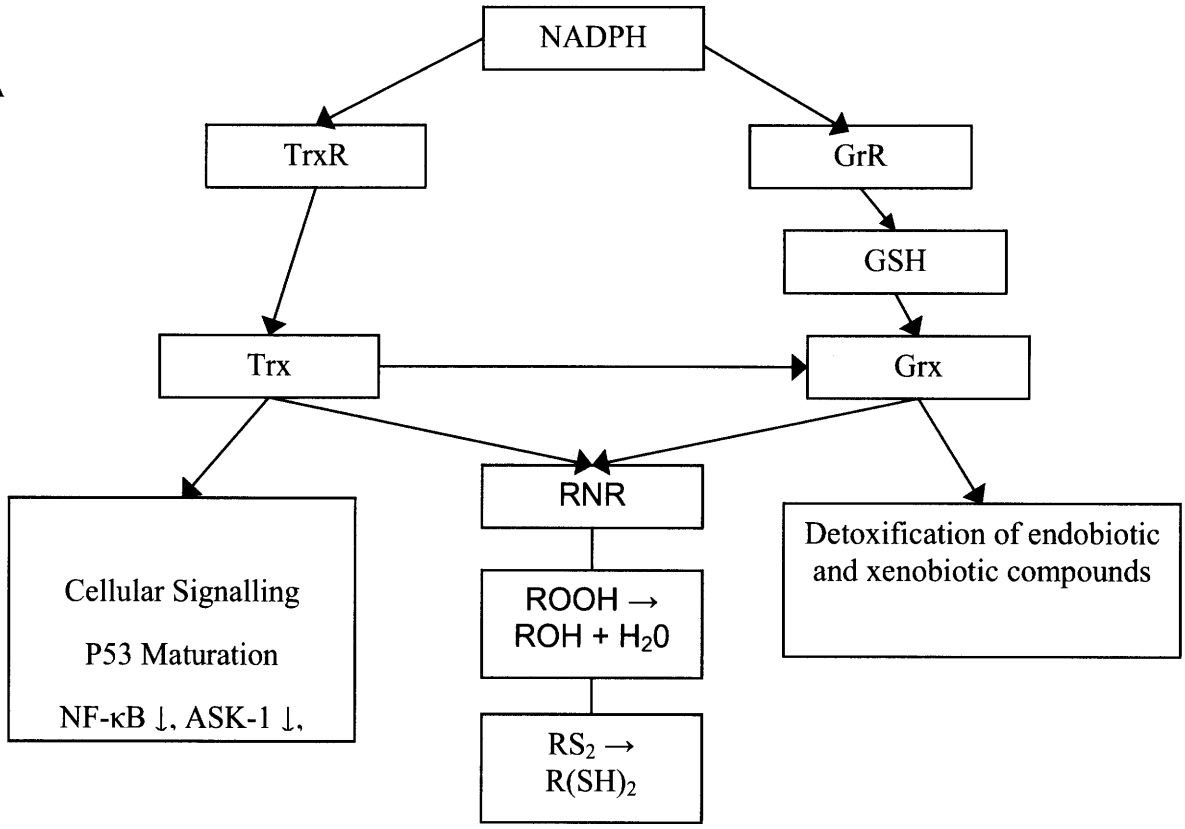
[PbP]$RB_{workB} = \Sigma||Fobs| - |Fcalc|| / \Sigma|Fobs|$

[PcP]RB_{freeB} calculated using 5% of the total reflections, which were not used in refinement

[PdP]Number of atoms in chain A and chain B

Figure 3.1. NADPH participation in cellular redox regulation. (A) The NADPH-dependent redox regulatory pathways. (B) Mechanism of NADPH dependant cellular reduction by the thioredoxin reductase system.

A



B

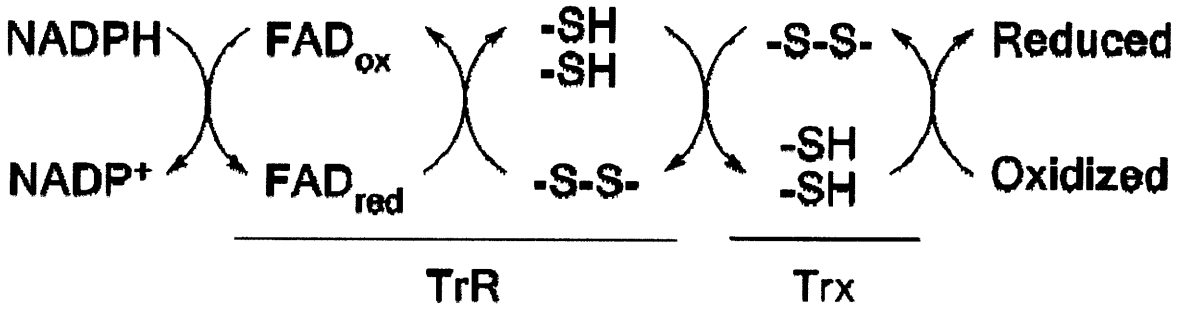


Figure 3.2. Alignment of selected TrxR sequences from PSI-BLAST (1). Sequences were aligned using the program ClustalW (2) and the output was rendered using the web interface ESPript 2.2 (3). Sequences are *E. coli* (gi:16128855), *V. parahaemolyticus* (gi:28898025), *V. cholerae* 0395 (gi:147674572), *S. boydii* BS512 (gi:82543374), *S. enterica* subsp. *enterica serovar Typhi* str. CT18 (gi:62179482), *Y. pestis* CO92 (gi:16121654), and *S. pneumoniae* R6 (gi:15903355). White letters in red blocks are strictly conserved residues. Red letters in black boxes are more than 50% conserved residues.

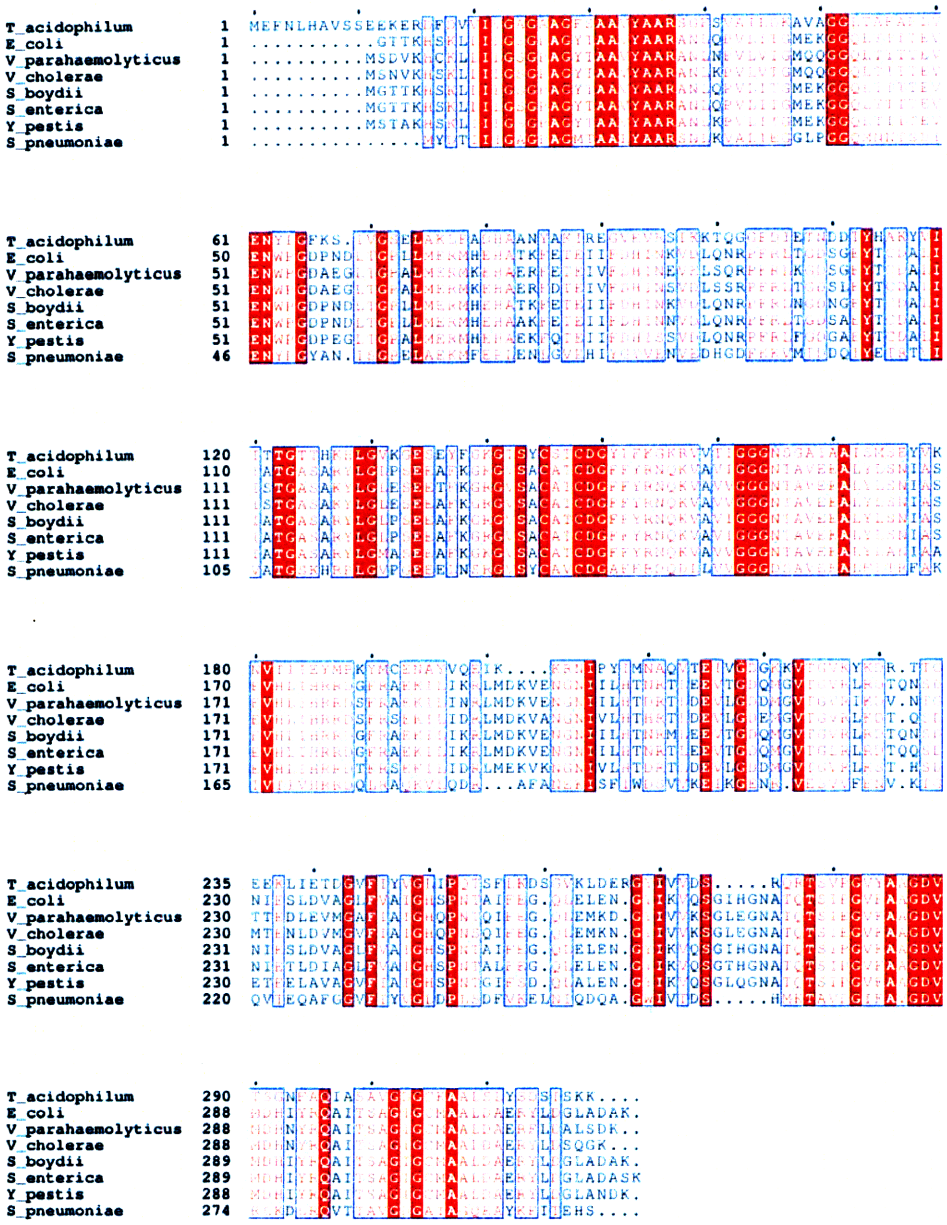


Figure 3.3. Alignment of selected Trx sequences from PSI-Blast (1). Sequences were aligned using the program ClustalW (2) and the output was rendered using the web interface ESPript 2.2 (3). Sequences are *E. coli* (gi:67005950), *V. parahaemolyticus* (gi:28899775), *V. cholerae* 0395 (gi:147675356), *S. boydii* BS512 (gi:82546135), *S. enterica* subsp. *enterica serovar Typhi* str. CT18 (gi:62182391), *Y. pestis* CO92 (gi:16124003), *S. pneumoniae* R6 (gi:15903644). White letters in red blocks are strictly conserved residues. Red letters in black boxes are more than 50% conserved residues.

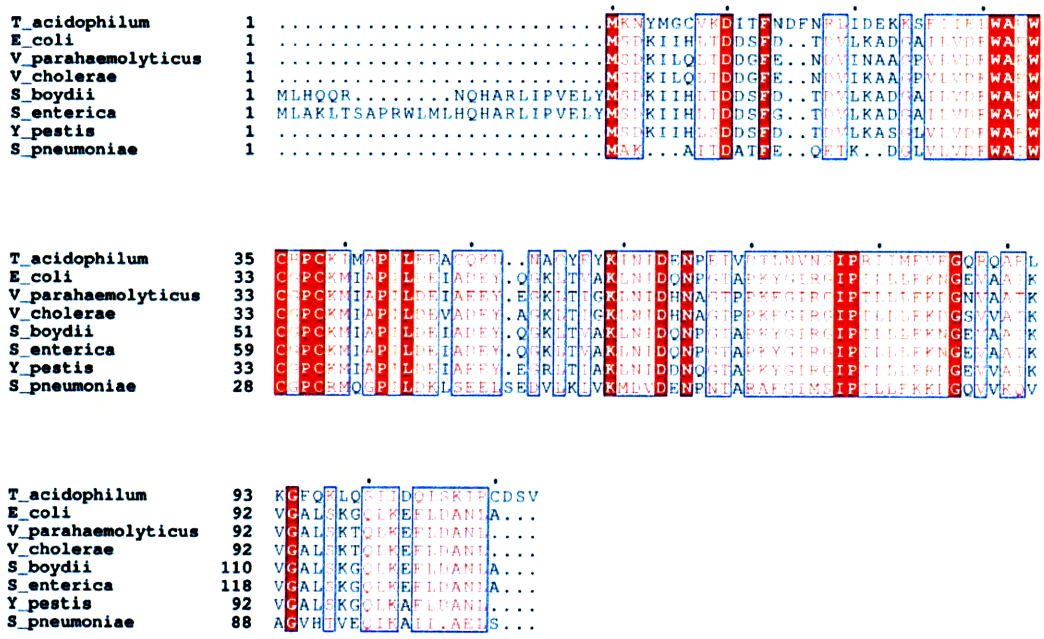


Figure 3.4. DNA and corresponding amino acid sequence of ORF Ta0984 from *T. acidophilum* encoding for a putative TrxR. DNA sequence was confirmed by the MIT Biopolymers Laboratory.

```
atggaatttaacctgcatgcagtatcatcagaggaaaaagaaagagattttgatgttgtc  60
 M  E  F  N  L  H  A  V  S  S  E  E  K  E  R  D  F  D  V  V
atcgtcggcgccggtgctgccggtttttcagcggcagtctatgctgcaaggtcaggattc 120
 I  V  G  A  G  A  A  G  F  S  A  A  V  Y  A  A  R  S  G  F
agcgttgcaattctcgacaaggccgtagctggaggccttactgcagaggcaccactggta 180
 S  V  A  I  L  D  K  A  V  A  G  G  L  T  A  E  A  P  L  V
gagaattatcttggtttcaagagcattgtcggatcagagcttgcgaaacttttcgccgac 240
 E  N  Y  L  G  F  K  S  I  V  G  S  E  L  A  K  L  F  A  D
catgcagcgaactatgccaagataagggagggcgtagaggtaagatccataaagaaaacg 300
 H  A  A  N  Y  A  K  I  R  E  G  V  E  V  R  S  I  K  K  T
caaggcggctttgacattgagacaaatgacgatacgtaccatgccaaatatgtgataata 360
 Q  G  G  F  D  I  E  T  N  D  D  T  Y  H  A  K  Y  V  I  I
acgactggaacaacgcataagcatttaggtgtcaaaggcgaatccgaatactttgggaag 420
 T  T  G  T  T  H  K  H  L  G  V  K  G  E  S  E  Y  F  G  K
ggtacctcttactgctctacctgcgatggatacctcttcaagggaaaacgtgtggtcacg 480
 G  T  S  Y  C  S  T  C  D  G  Y  L  F  K  G  K  R  V  V  T
ataggtggcggaaactcaggggccatagctgccatatcaatgagtgaatacgtcaaaaac 540
 I  G  G  G  N  S  G  A  I  A  A  I  S  M  S  E  Y  V  K  N
gtcaccataattgaatacatgcctaagtacatgtgcgaaaacgcctacgtacaggagatc 600
 V  T  I  I  E  Y  M  P  K  Y  M  C  E  N  A  Y  V  Q  E  I
aagaaaagaaacataccctacataatgaacgctcaggtaaccgagatcgttggggatgga 660
 K  K  R  N  I  P  Y  I  M  N  A  Q  V  T  E  I  V  G  D  G
aagaaggtcaccggagttaaatacaaagacaggacaacgggtgaagaaaaactcatagaa 720
 K  K  V  T  G  V  K  Y  K  D  R  T  T  G  E  E  K  L  I  E
actgacggcgtcttcatatacgtcggcctcataccgcagacttcatttctcaaggatagt 780
 T  D  G  V  F  I  Y  V  G  L  I  P  Q  T  S  F  L  K  D  S
ggcgtaaaactggatgagagagggtacatagtggtggattcaaggcagagaactagcgtt 840
 G  V  K  L  D  E  R  G  Y  I  V  V  D  S  R  Q  R  T  S  V
ccaggggtatatgcagcaggagacgttacatccggcaacttcgcacagattgcctctgcc 900
 P  G  V  Y  A  A  G  D  V  T  S  G  N  F  A  Q  I  A  S  A
gttggagatggctgcaaggcagcgctttccttgtattctgattctatatccaaaaaa     957
 V  G  D  G  C  K  A  A  L  S  L  Y  S  D  S  I  S  K  K
```

Figure 3.5. DNA and corresponding amino acid sequence of ORF Ta0866 from *T. acidophilum* encoding for a putative Trx. DNA sequence was confirmed by the MIT Biopolymers Laboratory.

```
atgaagaattatatgggctgcgttaaagatataacattcaatgatttcaatcgtctcata  60
 M  K  N  Y  M  G  C  V  K  D  I  T  F  N  D  F  N  R  L  I
gatgagaaaaaaagcttcataattgagttatgggctgattggtgccatccctgcaagatc 120
 D  E  K  K  S  F  I  I  E  L  W  A  D  W  C  H  P  C  K  I
atggcgccatacctggaggaggcgtgtcagaagttgaacgcctgctactttacaagata 180
 M  A  P  Y  L  E  E  A  C  Q  K  L  N  A  C  Y  F  Y  K  I
aatatcgatgagaatcctgagatcgttgacacgcttaatgtgaacagcataccaaggata 240
 N  I  D  E  N  P  E  I  V  D  T  L  N  V  N  S  I  P  R  I
ataatgtttgttgaaggccagcgccaggccgagcttaagggatttcagaagcttcaatca 300
 I  M  F  V  E  G  Q  R  Q  A  E  L  K  G  F  Q  K  L  Q  S
attattgaccagatttcaaaaattccctgcgattccgtc                      339
 I  I  D  Q  I  S  K  I  P  C  D  S  V
```

Figure 3.6. SDS-PAGE [4-20% (w/v)] gel of the isolated **(A)** *ta*TrxR and (**B**) *ta*Trx from the overexpression *E. coli* Rosetta™ (DE3) pLysS strain.

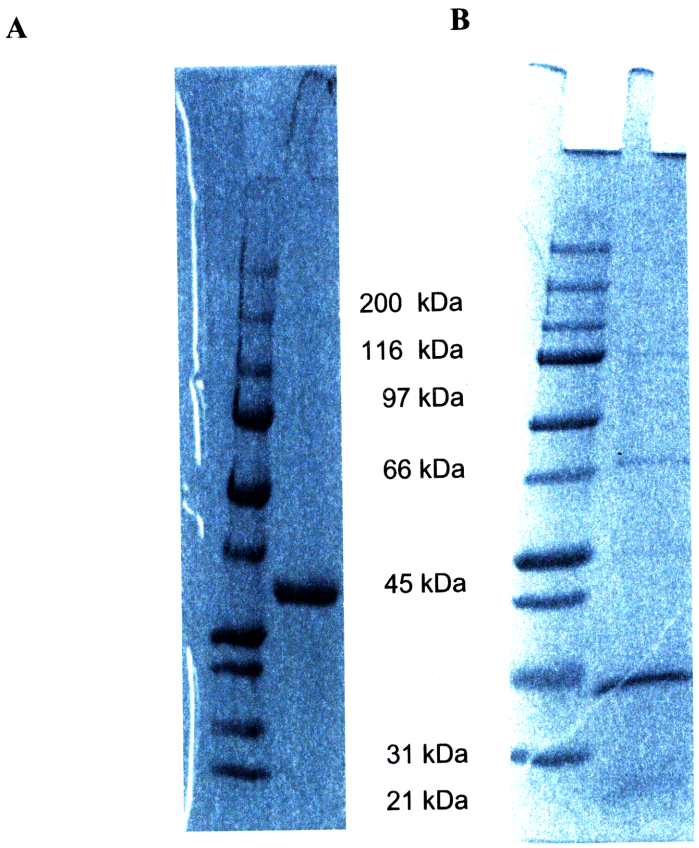


Figure 3.7. Spectral recordings of temperature dependent degradation of NADPH and NADH. **(A)** NADPH at 20° C (solid line), NADPH at 80° C (dashed line), and heat treated NADPH cooled down to 20° C (dotted line). **(B)** Spectra of temperature dependent degradation of NADH. NADH at 20° C (solid line), NADH at 80° C (dashed line), and heat treated NADH cooled down to 20° C (dotted line). **(C)** Time dependent degradation of NADPH at 20° C (square), 60° C (circle), and 80° C (triangle). **(D)** Time dependent degradation of NADH at 20° C (square), 60° C (circle), and 80° C (triangle).

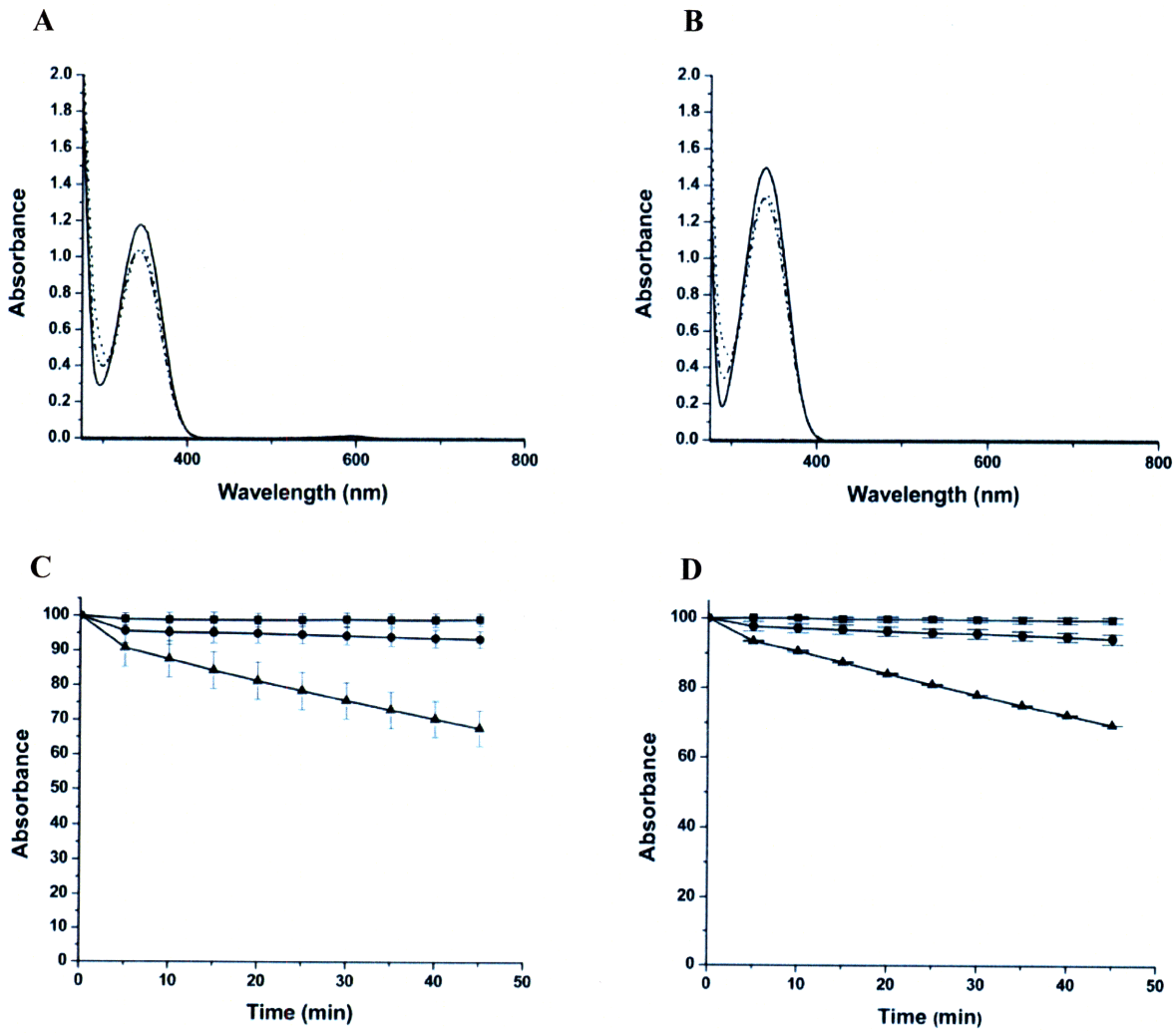


Figure 3.8. Anaerobic titration of *ta*TrxR and *ec*TrxR with NADPH and NADH. **(A)** Visible spectra of *ta*TrxR (43.3 µM) titrated with NADPH at 59 ºC. Solid line is without NADPH, dashed line is with 0.5 moles NADPH / mole FAD, dotted line is with 1.0 mole NADPH, dash dot line is with 2.0 mole NADPH, dash dot dot line is with 5 mole NADPH. **(B)** Visible spectra of *ta*TrxR titrated with NADH at 59 ºC. **(C)** Visible spectra of *ec*TrxR (40.5 µM) titrated with NADPH at 25 ºC. **(D)** Visible spectra of *ec*TrxR (40.5 µM) titrated with NADH at 25 ºC. Spectra (B), (C), and (D) were recorded as in (A). **(E)** Time dependent reduction of *ta*TrxR and NADPH at 59 ºC (closed triangles), *ta*TrxR and NADH at 59 ºC (closed circles), and *ec*TrxR at 25 ºC (closed squares).

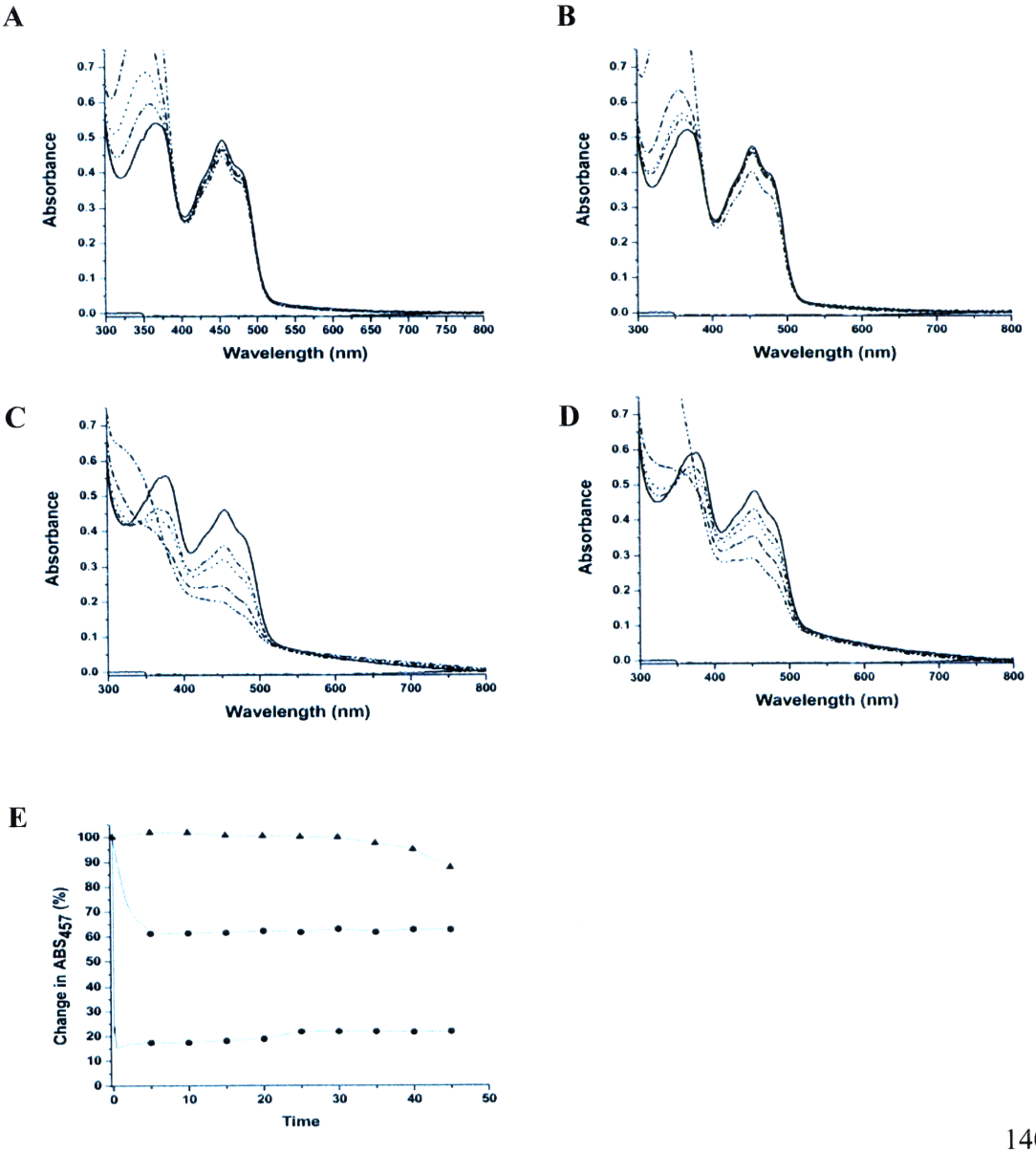


Figure 3.9. Reduction of insulin catalyzed by *ta*Trx. The DTT dependent reduction of bovine insulin was carried out as described in Methods. The increase in turbidity at 650 nm is plotted against the reaction time. Solid line is 70 μM *ta*Tr, dashed line is 66 μM *ta*Tr, and dotted line is 43 μM *ta*Tr.

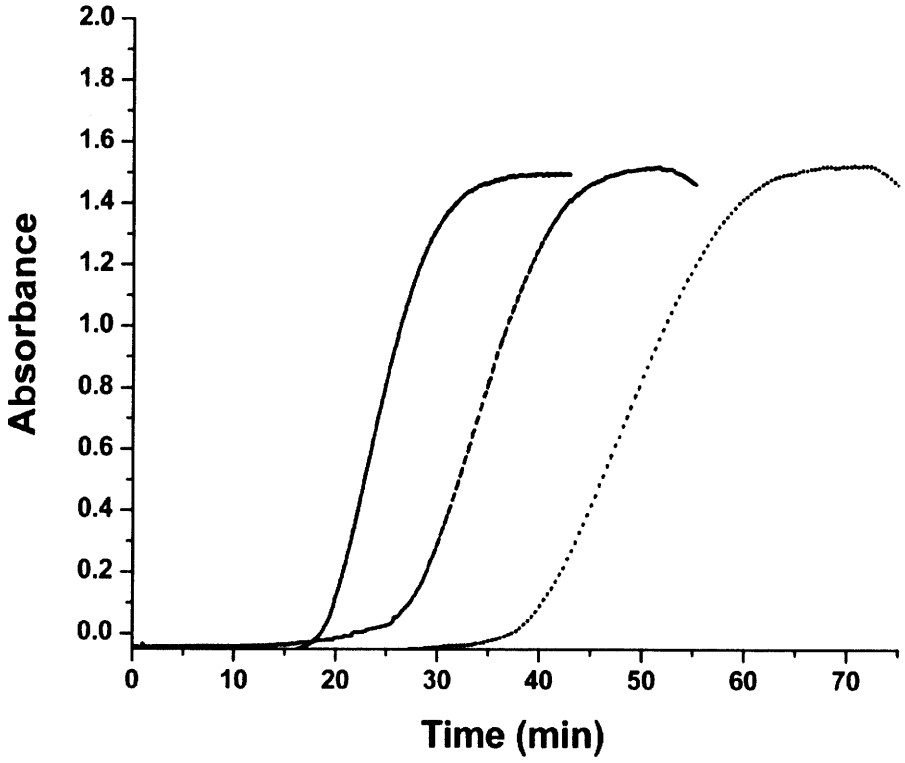


Figure 3.10. Tr activity in thioredoxin system using *ta*Trx and *ec*TrxR. The assays were performed using 3.8 μM *ec*TrxR and varying concentrations of *ta*Tr. Assays were performed as described in Methods.

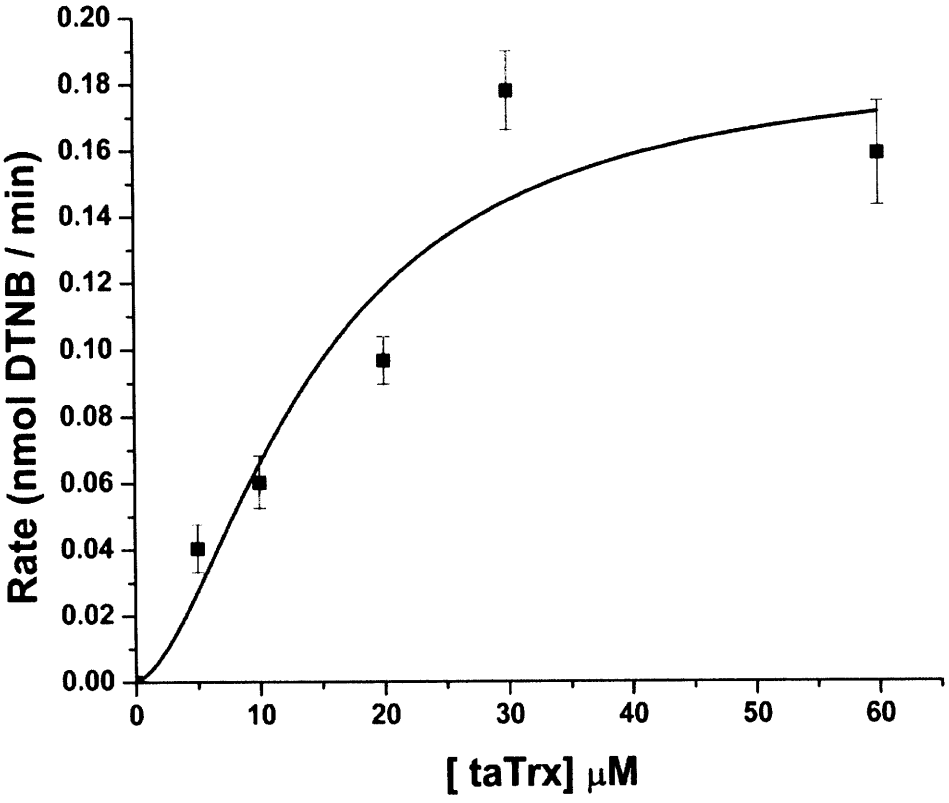


Figure 3.11. Demonstration of *ta*TrxR/Trx reactivity. Reduction of oxidized *ta*TrxR (4 μM) was achieved by use of a modified xanthine/xanthine oxidase system, as described in the Materials and Methods section. Oxidized *ta*TrxR (solid line) was reduced with xanthine as the source of electrons, over the course of 90 min, in order to build up a stable reduced form of the enzyme (dashed line). Upon introduction of 2 equivalents of the *ta*Trx, the sample was completely reoxidized within 1 min (dotted line).

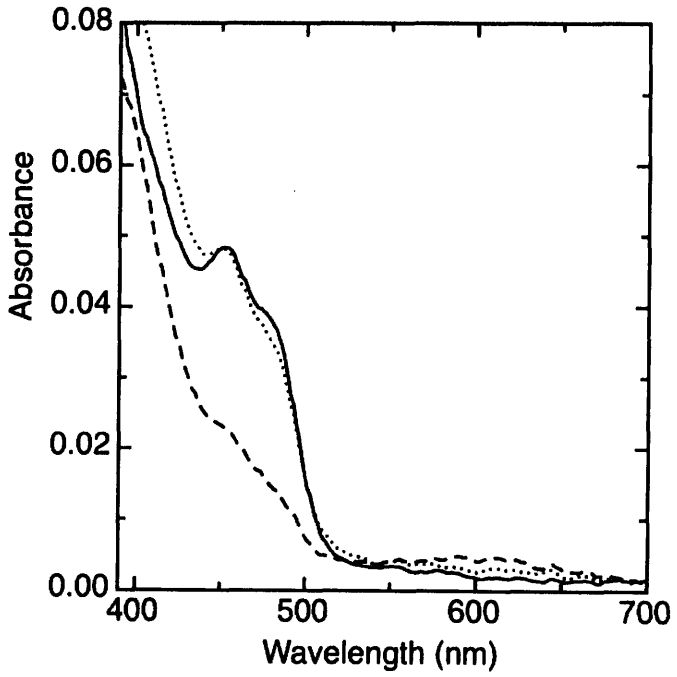


Figure 3.12. Ribbon diagram structure of *ta*TrxR dimer. (A) Monomers of *ta*TrxR are colored by chain. The flavin molecule (1 per monomer) is represented as space filled spheres and colored in CPK color scheme. Circle highlights C145/ C148 disulfide. (B) Superposition of the *ta*TrxR NADPH binding domain with *ec*TrxR NADPH binding domain. (C) Superposition of the *ta*TrxR FAD binding domain with *ec*TrxR FAD binding domain. (*ta*TrxR is colored in yellow, *ec*TrxR is colored in cyan)

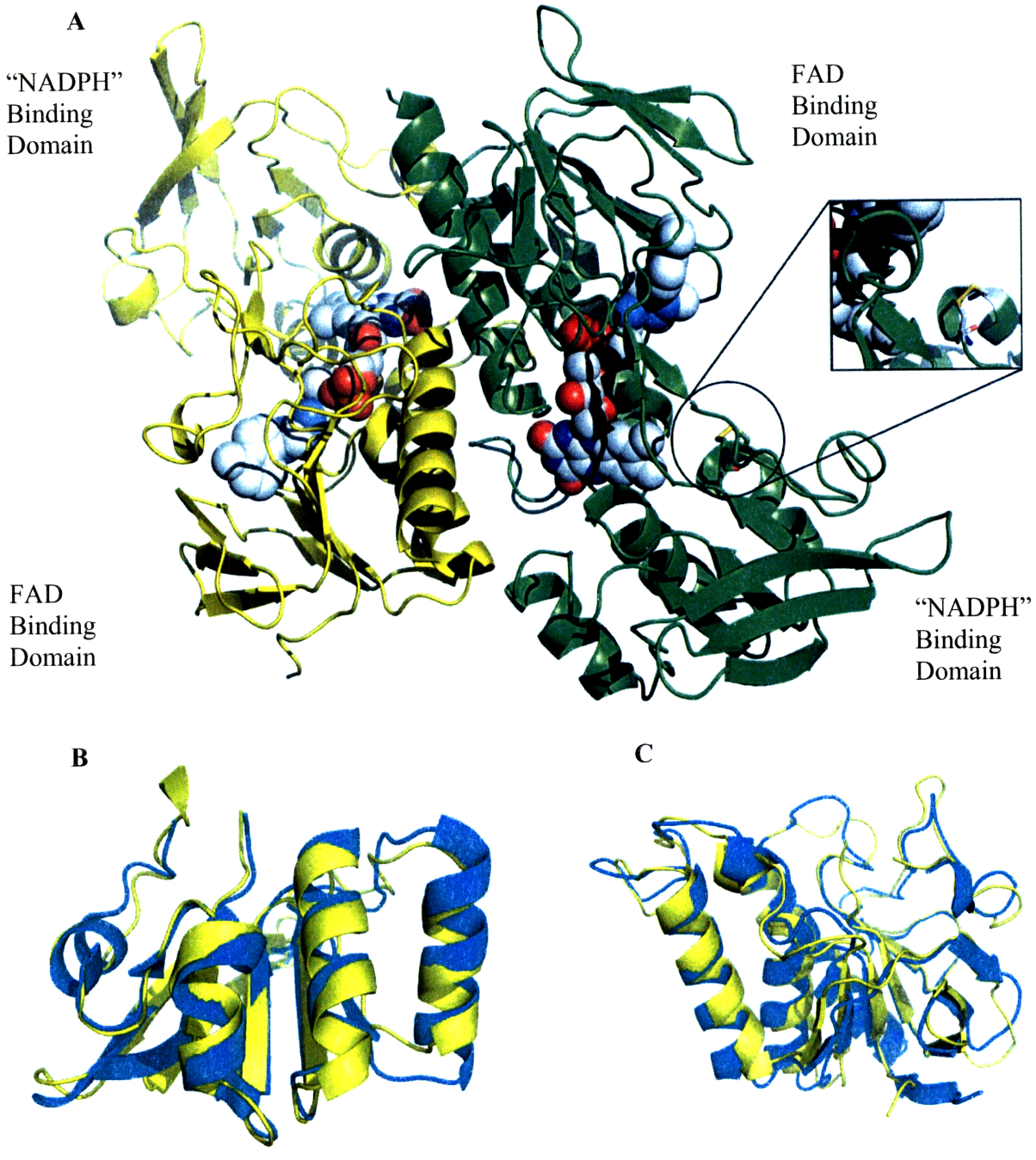


Figure 3.13. View of the FAD binding pocket of *ta*TrxR. The FAD molecule is shown in composite omit map density contoured at 1 σ. FAD atoms are colored in CPK format (C in light grey, N in blue, O in red, P in orange).

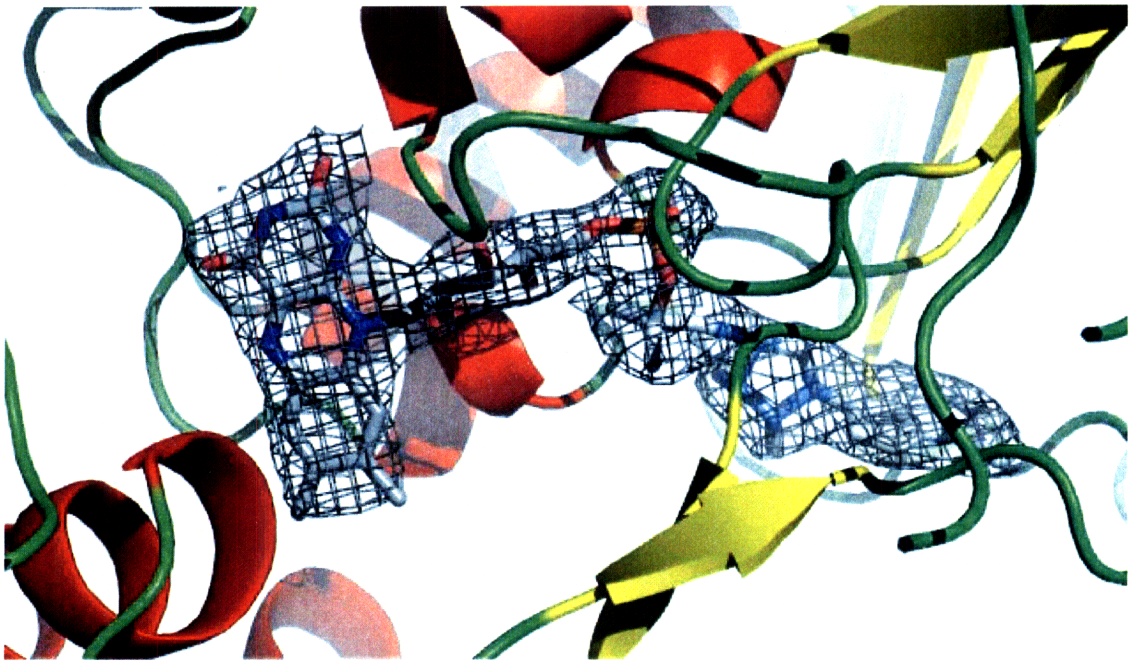


Figure 3.14. Schematic drawing of the FAD molecule in *ta*TrxR. Polar atoms of residues of molecule A that are within 4 Å of atoms of FAD are listed. Distances are in Å.

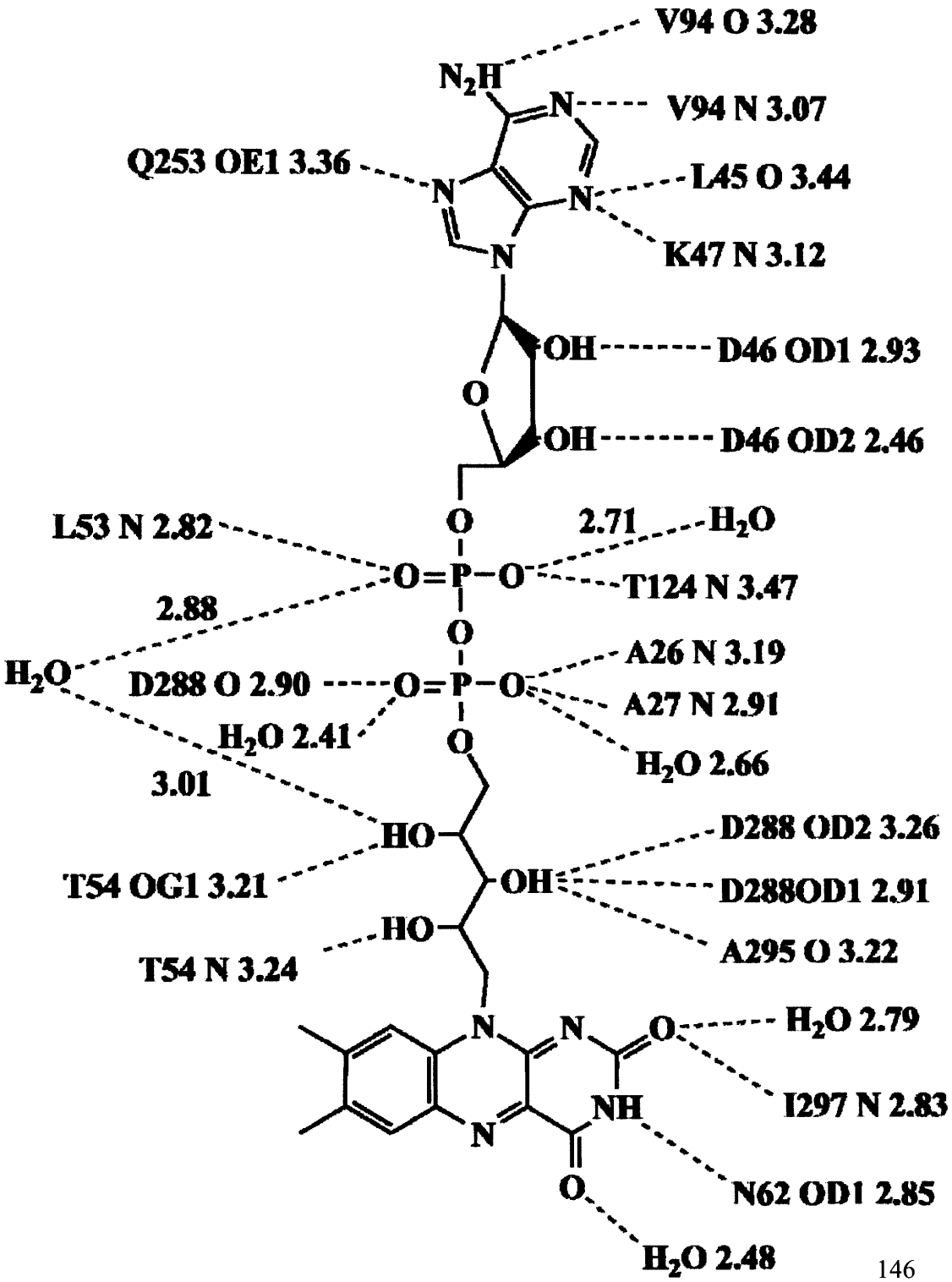


Figure 3.15. Ribbon representation of the *ec*TrxR NADPH binding domain (cyan) (PDB code 1TDF) and *ta*TrxR "NADPH" binding domain (yellow). (A) *ec*TrxR with NADPH binding amino acids, H175, R176, R177, and R181, represented in stick form. (b) *ta*TrxR structure with corresponding structurally aligned amino acids, E185, Y186, M187 and M191 represented in stick form with modeled NADPH. (C) Superposition of *ec*TrxR and *ta*TrxR NADPH binding domains with NADPH from *ec*TrxR structure (D) Sequence alignment of the NADPH binding domain regions.

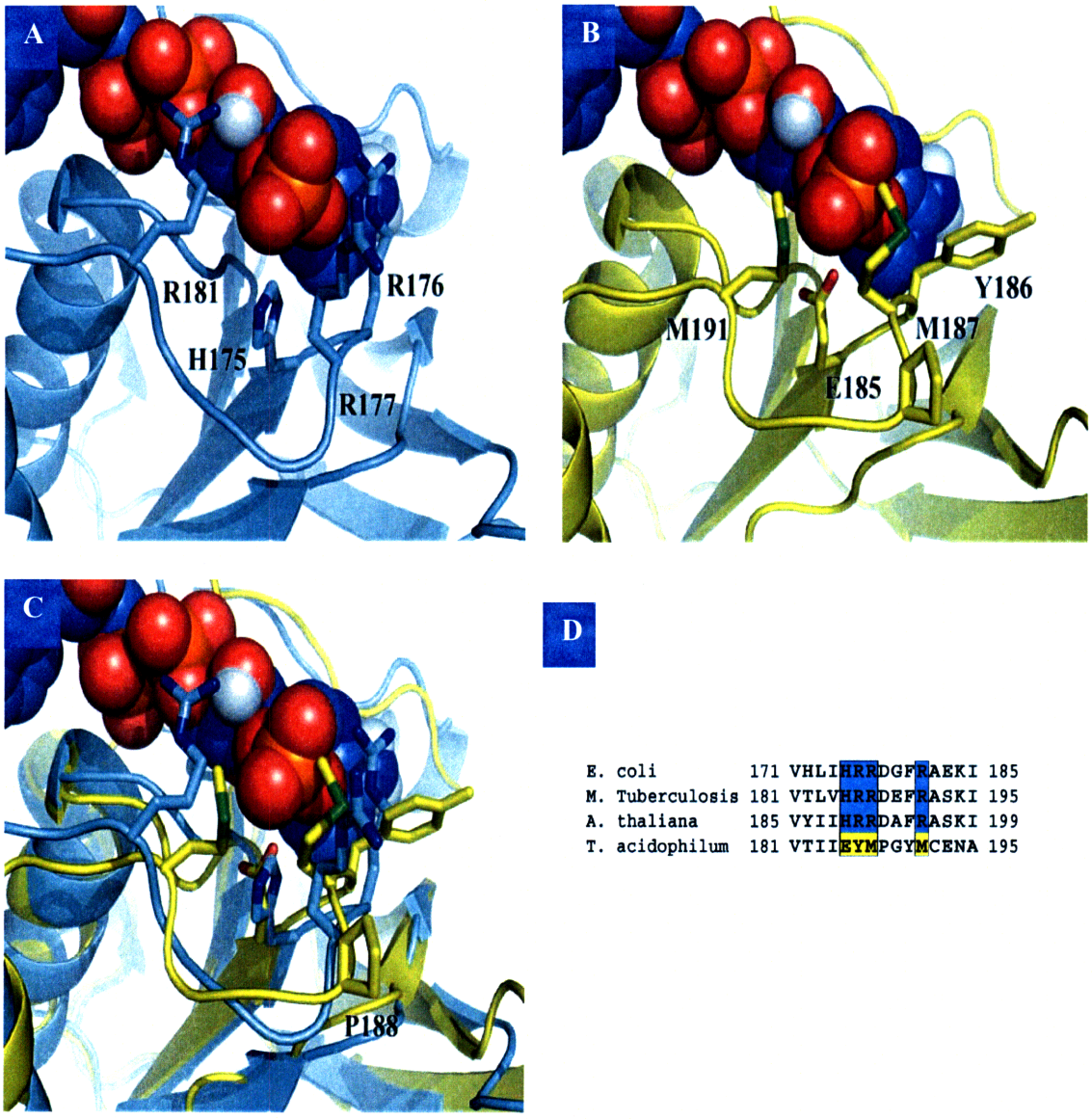


Figure 3.16. Space filling surface model of **(A)** *ec*TrxR and **(B)** *ta*TrxR NADPH binding domain. The positively charged residues are represented in blue and the negatively charged residues are represented in red. Neutral residues are represented as white. Surface calculations were made using the APBS module in Pymol (4).

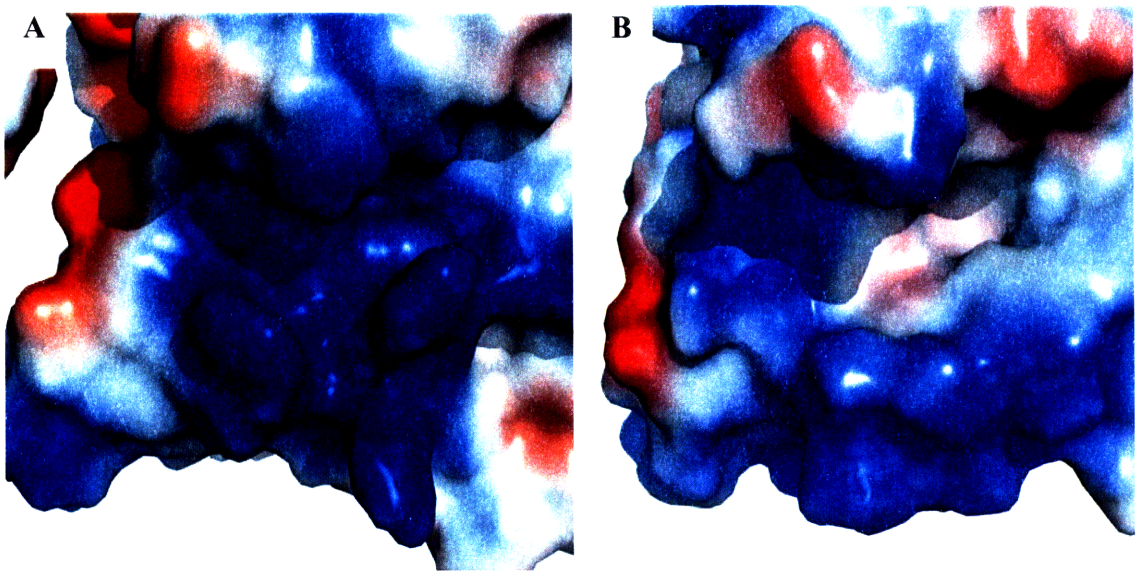


Figure 3.17. Structural superposition of TrxR NADPH binding domains. **(A)** The *ec*TrxR NADPH binding domains crystallized with NADPH analogue AADP (1TDF) is in teal and without NADPH (1TDE) is in cyan. **(B)** The *ec*TrxR (1TDF) co-crystallized with AADP, an NADPH analogue, is in teal, *A. thaliana* TrxR (1VDC) structure is in orange, *M. tuberculosis* TrxR (2A87) co-crystallized with NADP is in magenta, and *ta*TrxR is in yellow.

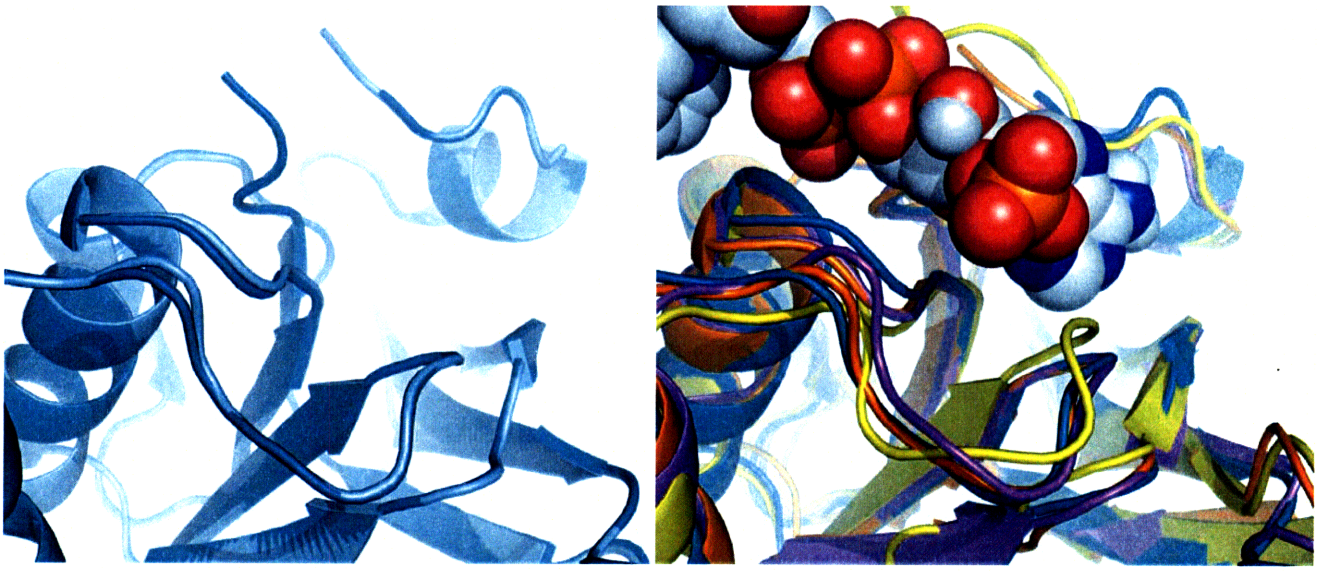
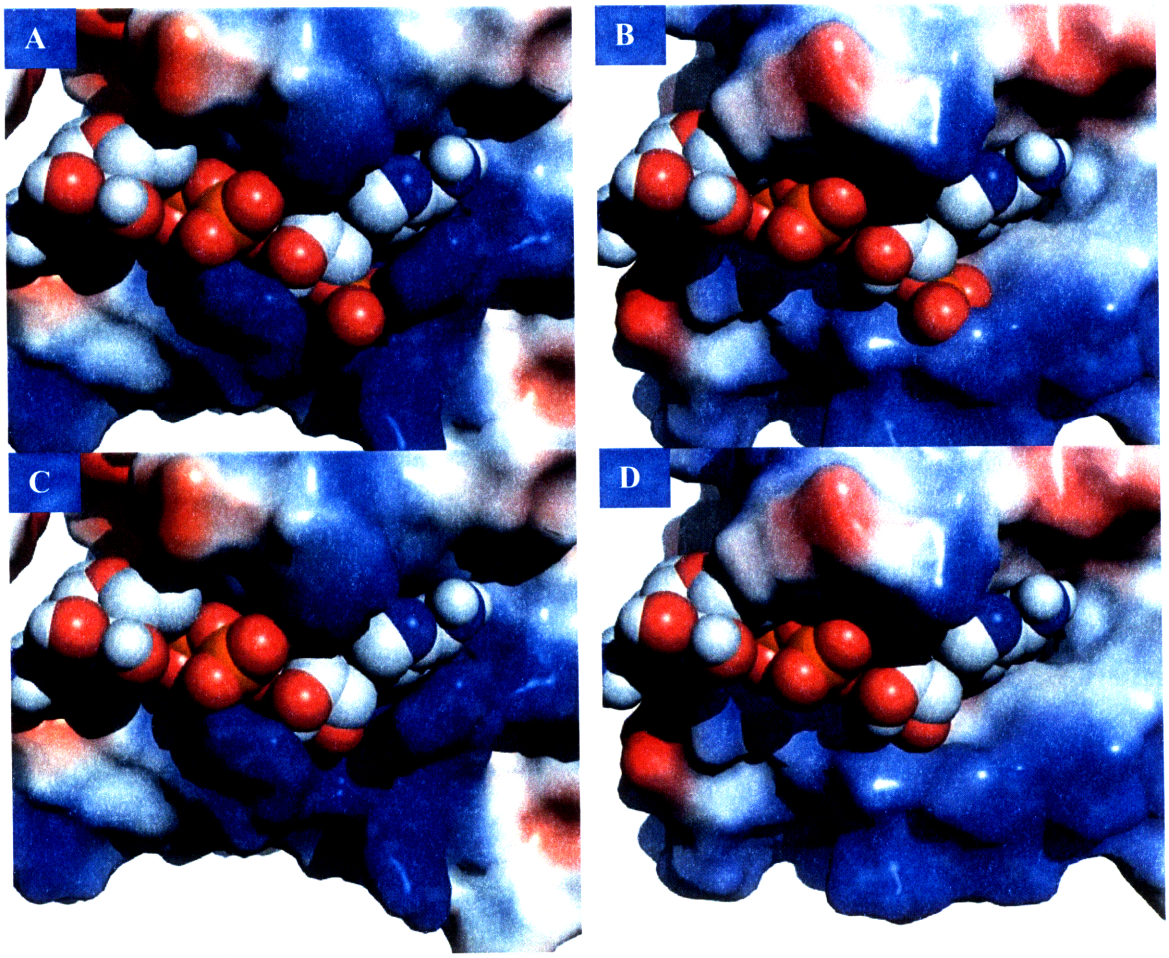


Figure 3.18. Surface representation of NADPH binding pocket with blue representing basic residues, red acidic residues, and white neutral residues. **(A)** *ec*TrxR structure with NADPH bound. (B) *ta*TrxR structure with modeled NADPH. **(C)** *ec*TrxR structure with NADH modeled in the NADPH binding site. **(D)** *ta*TrxR structure with NADH modeled in the NADPH binding site. **(E)** *ec*TrxR structure with coenzyme F420 modeled in the NADPH binding site [1].



REFERENCES

1. Jaeger, T., Flohe, L. (2006) The thiol-based redox networks of pathogens: unexploited targets in the search for new drugs. *Biofactors* 27, (1-4), 109-120.

2. Arner, E. S., Holmgren, A. (2000) Physiological functions of thioredoxin and thioredoxin reductase. *Eur J Biochem* 267, (20), 6102-6109.

3. Masutani, H., Ueda, S., Yodoi, J. (2005) The thioredoxin system in retroviral infection and apoptosis. *Cell Death Differ* 12 Suppl 1, 991-998.

4. Mustacich, D., Powis, G. (2000) Thioredoxin reductase. *Biochem J* 346 Pt 1, 1-8.

5. Stroher, E., Dietz, K. J. (2006) Concepts and approaches towards understanding the cellular redox proteome. *Plant Biol (Stuttg)* 8, (4), 407-418.

6. Toledano, M. B., Kumar, C., Le Moan, N., Spector, D., Tacnet, F. (2007) The system biology of thiol redox system in *Escherichia coli* and yeast: differential functions in oxidative stress, iron metabolism and DNA synthesis. *FEBS Lett* 581, (19), 3598-3607.

7. Herrick, J., Sclavi, B. (2007) Ribonucleotide reductase and the regulation of DNA replication: an old story and an ancient heritage. *Mol Microbiol* 63, (1), 22-34.

8. Nordlund, P., Reichard, P. (2006) Ribonucleotide reductases. *Annu Rev Biochem* 75, 681-706.

9. Shao, J., Zhou, B., Chu, B., Yen, Y. (2006) Ribonucleotide reductase inhibitors and future drug design. *Curr Cancer Drug Targets* 6, (5), 409-431.

10. Hashemy, S. I., Ungerstedt, J. S., Zahedi Avval, F., Holmgren, A. (2006) Motexafin gadolinium, a tumor-selective drug targeting thioredoxin reductase and ribonucleotide reductase. *J Biol Chem* 281, (16), 10691-10697.

11. Darland, G., Brock, T. D., Samsonoff, W., Conti, S. F. (1970) A thermophilic, acidophilic mycoplasma isolated from a coal refuse pile. *Science* 170, (965), 1416-1418.

12. Searcy, D. G. (1976) *Thermoplasma acidophilum*: intracellular pH and potassium concentration. *Biochim Biophys Acta* 451, (1), 278-286.

13. Hsung, J. C., Haug, A. (1975) Intracellular pH of *Thermoplasma acidophila*. *Biochim Biophys Acta* 389, (3), 477-482.

14. Ruepp, A., Graml, W., Santos-Martinez, M. L., Koretke, K. K., Volker, C., Mewes, H. W., Frishman, D., Stocker, S., Lupas, A. N., Baumeister, W. (2000) The

genome sequence of the thermoacidophilic scavenger *Thermoplasma acidophilum*. *Nature* 407, (6803), 508-513.

15. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25, (17), 3389-3402.

16. Thompson J.D., Higgins D.G., Gibson T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22, 4673-4680.

17. Gouet, P., Courcelle, E., Stuart, D. I., Metoz, F. (1999) ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics* 15, (4), 305-308.

18. Holmgren, A. (1979) Thioredoxin catalyzes the reduction of insulin disulfides by dithiothreitol and dihydrolipoamide. *J Biol Chem* 254, (19), 9627-9632.

19. Arner, E. S., Zhong, L., Holmgren, A. (1999) Preparation and assay of mammalian thioredoxin and thioredoxin reductase. *Methods Enzymol* 300, 226-239.

20. Massey, V., Brumby, P. E., Komai, H. (1969) Studies on milk xanthine oxidase. Some spectral and kinetic properties. *J Biol Chem* 244, (7), 1682-1691.

21. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* 276, 307-326.

22. Waksman, G., Krishna, T. S., Williams, C. H., Jr., Kuriyan, J. (1994) Crystal structure of *Escherichia coli* thioredoxin reductase refined at 2 Å resolution. Implications for a large conformational change during catalysis. *J Mol Biol* 236, (3), 800-816.

23. McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C., Read, R. J. (2005) Likelihood-enhanced fast translation functions. *Acta Crystallogr D* 61, (Pt 4), 458-464.

24. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., Warren, G. L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D* 54, (Pt 5), 905-921.

25. McRee, D. E. (1999) XtalView/Xfit--A versatile program for manipulating atomic coordinates and electron density. *J Struct Biol* 125, (2-3), 156-165.

26. Kleywegt, G. J., Jones, T. A. (1998) Databases in protein crystallography. *Acta Crystallogr D* 54, (Pt 6 Pt 1), 1119-1131.

27. Williams, C. H., Jr., Zanetti, G., Arscott, L. D., McAllister, J. K. (1967) Lipoamide dehydrogenase, glutathione reductase, thioredoxin reductase, and thioredoxin. *J Biol Chem* 242, (22), 5226-5231.

28. Thelander, L. (1967) Thioredoxin reductase. Characterization of a homogenous preparation from *Escherichia coli B*. *J Biol Chem* 242, (5), 852-859.

29. Williams, C. H., Arscott, L. D., Muller, S., Lennon, B. W., Ludwig, M. L., Wang, P. F., Veine, D. M., Becker, K., Schirmer, R. H. (2000) Thioredoxin reductase two modes of catalysis have evolved. *Eur J Biochem* 267, (20), 6110-6117.

30. Biaglow, J. E., Miller, R. A. (2005) The thioredoxin reductase/thioredoxin system: novel redox targets for cancer therapy. *Cancer Biol Ther* 4, (1), 6-13.

31. Williams, C. H., Jr. (1976); Flavin-Containing Dehydrogenases. In *The Enzymes*, (Boyer, P. D., Ed. Academic Press) Third Edition ed.; New York, Vol. 1, pp 89-173.

32. Lennon, B. W., Williams, C. H., Jr., Ludwig, M. L. (2000) Twists in catalysis: alternating conformation of *Escherichia coli* thioredoxin reducatase.*Science 289*, 1190-1194.

33. Wang, P. F., Veine, D. M., Ahn, S. H., Williams, C. H., Jr. (1996) A stable mixed disulfide between thioredoxin reductase and its substrate, thioredoxin: preparation and characterization. *Biochemistry* 35, (15), 4812-4819.

34. O'Donnell, M. E., Williams, C. H., Jr. (1983) Proton stoichiometry in the reduction of the FAD and disulfide of *Escherichia coli* thioredoxin reductase. Evidence for a base at the active site. *J Biol Chem* 258, (22), 13795-13805.

35. Copley, S. D., Dhillon, J. K. (2002) Lateral gene transfer and parallel evolution in the history of glutathione biosynthesis genes. *Genome Biol* 3, (5), 1-16.

36. Reher, M., Schönheit, P. (2006) Glyceraldehyde dehydrogenases from the thermoacidophilic euryarchaeota *Picrophilus torridus* and *Thermoplasma acidophilum*, key enzymes of the non-phosphorylative Entner-Doudoroff pathway, constitute a novel enzyme family within the aldehyde dehydrogenase superfamily. *FEBS Lett* 580, (5), 1198-1204.

37. Nishiya, Y., Tamura, N., Tamura, T. (2004) Analysis of bacterial glucose dehydrogenase homologs from thermoacidophilic archaeon *Thermoplasma acidophilum*: finding and characterization of aldohexose dehydrogenase. *Biosci Biotechnol Biochem* 68, (12), 2451-2456.

38. Smith, L. D., Budgen, N., Bungard, S. J., Danson, M. J., Hough, D. W. (1989) Purification and characterization of glucose dehydrogenase from the thermoacidophilic archaebacterium *Thermoplasma acidophilum*. *Biochem J* 261, (3), 973-977.

39. Fischer, M., Bacher, A. (2005) Biosynthesis of flavocoenzymes. *Nat Prod Rep* 22, (3), 324-350.

40. Graupner, M., Xu, H., White, R. H. (2002) The pyrimidine nucleotide reductase step in riboflavin and F(420) biosynthesis in archaea proceeds by the eukaryotic route to riboflavin. *J Bacteriol* 184, (7), 1952-1957.

41. Mack, M., Grill, S. (2006) Riboflavin analogs and inhibitors of riboflavin biosynthesis. *Appl Microbiol Biotechnol* 71, (3), 265-275.

42. Baker, N. A., Sept, D., Joseph, S., Holst, M. J., McCammon, J. A. (2001) Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc Natl Acad Sci U S A* 98, (18), 10037-10041.

Chapter 4:

Direct electrochemical characterization of four thioredoxins from *Archaeoglobus Fulgidus*

Genomic analysis, cloning, amplification, purification, and insulin assays were performed by Hector H Hernandez. Protein film voltammetry was performed by Sarah E. Chobot.

Reproduced in part from: Chobot, S. E., Hernandez, H. H., Drennan, C. L., Elliott, S. J.,(2007) Dirrect Electrochemical Characterization of Archaeal Thioredoxins. Angew Chem Int Ed Engl 46, 22 4145-4147

INTRODUCTION

The thioredoxin (Trx) superfamily of proteins contains small soluble proteins that function as 2 e^-/2 H^+ electron-transfer agents by virtue of a redox-active disulfide bond. Although the members of this superfamily are known to contain disulfide bonds that span a range of midpoint potential of at least 150 mV, (1) a detailed picture of the molecular determinants of the disulfide-bond potential has yet to be attained. Herein, we demonstrate that this goal is feasible through the application of protein-film voltammetry (PFV), (2) an electrochemical technique that we use to directly observe the reversible 2 e^- redox couple of thioredoxins.

Successful PFV yields a fast electrochemical connection between a submonolayer of protein analyte and an electrode. Previous electrochemical investigations of plant-type and *Escherichia coli* Trx disulfides did not yield reversible 2 e^- voltammetry: instead quasireversible 1 e^- cyclic voltammetry for the disulfide/disulfide radical potential and an second irreversible feature (corresponding to the reduction of the radical intermediate) was observed (3, 4). Martin and co-workers have developed a modified gold electrode to investigate His-tag-labeled *E. coli* Trx; they too were unable to directly observe a reversible 2 e^- potential (5) and could only establish reversible voltammetry for the 1 e^- disulfide/disulfide radical couple (5, 6). As the biologically significant reaction for thioredoxins involves cooperative 2e^- chemistry (7, 8), we have examined a series of members of the Trx superfamily to observe a reversible 2 e^-/2 H^+ reaction by using PFV. Herein, we report the successful extension of PFV to the direct measurement of a reversible 2 e^- disulfide/dithiol couple.

Members of the Trx superfamily contain a single disulfide bond within a -CXXC- motif, and the differences in redox properties of Trx proteins are attributed to variation of the -CXXC- motif (Figure 4.1) (1, 9). The difference in potential of the disulfide is correlated to function: *E. coli* Trx is reducing in the cellular milieu (8) with a reported potential from solution-based experiments of -270 mV (versus hydrogen) (10). More oxidizing members of the thioredoxin superfamily, such as *E. coli* DsbA, bear a disulfide bond with a reported reduction potential of -90(11) to - 122 mV (12). In comparison, the disulfide bond proximal to the Rieske [2Fe-2S] center of the *Thermus thermophilus* Rieske protein is electrochemically stable at sustained potentials as low as -0.85 V (13).

Trx acts as a reducing agent for a wide array of critical biological pathways (e.g., ribonucleotide biosynthesis, oxidative-stress defense, and transcription-factor activation) (14), whereas DsbA oxidizes target proteins, installing disulfides for proper protein folding and maturation (15). Structures of the oxidized, disulfide-containing loops of *E. coli* Trx and DsbA are shown in Figure 4.1 (16, 17).

Probing the influence of sequence and structure on the reduction potential of Trx proteins has been limited by the methods used for determining 2 e^- disulfide bond potentials. Such potentials are typically measured by coupled, solution-based processes, such as glutathione equilibria, which indirectly give a value for the disulfide potentials of interest, though such studies have yielded variable values of potential previously (11, 12, 18, 19).

MATERIALS AND METHODS

Materials. *Archaeoglobus fulgidus* organism was obtained from the American Tissue Culture Center. TOPO and TOPO Zero Blunt PCR Cloning kits were obtained from Invitrogen (Carlsbad, CA). NovaBlue Cells Competent Cells, Rosetta™ (DE3)pLysS Competent Cells, and Plasmid pET28-A were purchased from Novagen (Madison, WI, USA). *pfu*Turbo Polymerase and PCR Optimization Kit were obtained from Strategene (Cedar Creek, TX, USA). Primers were obtained from Integrated DNA Technologies (Coralville, IA, USA). Nickel nitrilotriacetic (Ni-NTA) acid-agarose gel was purchased from Qiagen (Valencia, CA, USA). Restriction enzymes were purchased from New England BioLabs (Beverly, MA, USA). Kanamycin, chloroamphenicol, TRIS-HCl, EDTA, and phenyl methyl sulfonyl fluoride (PMSF) were obtained from Sigma-Aldrich (St. Louis, MA, USA).

Genomic analysis of Archaeoglobus fulgidus genome. The genome of *Archaeoglobus fulgidus* (RefSeq NC_000917.1, GenBank AE000782.1) was analyzed using the web interface at the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov) for putative thioredoxin (Trx) open reading frames (ORFs). Four ORFs encoding a 119 amino acid protein, AF0711, a 93 amino acid protein, AF0769, a 134 amino acid protein, AF1284, and a 91 amino acid protein,

AF2144, were identified by sequence homology to known Trxs using PSI-BLAST (20). Protein sequences were aligned using the program ClustalW (21) and the output was rendered using the web interface ESPript 2.2 (Figure 4.2 – 4.8)(22).

Amplification of afuTrxs. The open reading frames (ORFs) coding for the putative *A. fulgidus* thioredoxins, (*afu*Trx) (AF0711, AF0769, AF1284, and AF2144) were retrieved from the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov). The Trx genes were amplified by the polymerase chain reaction (PCR) using genomic DNA of *A. fulgidus*. The oligonucleotides primers were designed based on the open reading frames coding for the proteins: Trx 1 (AF0711) - 5' - ATC GCA TAT GCC AAT GGT GAG GAA AG CTG CAT TTT ATG C-3' (N-terminus) and 5' - TAG GAT CCT CAT TTC ATC CCC TCT ATT GCT GTC CTA ACC-3' (C-terminus); Trx 2 (AF0769) - 5' - ATC GCA TAT GGT TTT CAC CAG CAA ATA CTGCCC CTA C - 3' (N-terminus) and 5' - TAG GAT CCT CAG TTT TTG TAA TCA GGC TTA AGA AAT GC - 3' (C-terminus); Trx 3 (AF1284) - 5' - ATC GCA TAT GGA CGA GTT AGA GCT AAT CAG ACA G - 3' (N-terminus) and 5' - TAG GAT CCT CAG AGA TTC CTC TGA ACC CAT CTC TTC AGC - 3' (C-terminus); and Trx 4 (AF2144) - 5' - GGC AGC CAT ATG GAG CGT CTT AAC AGT GAG AGG TTT AGG GAG G - 3' (N-terminus) and 5' - GGT CGG ATC CTC ATC ATC AGG CTC CAA GTG CTT TCT CTA TTT CCG - 3' (C-terminus).

PCR amplification of the four Trx genes was performed via 30 cycles using the following conditions: initial denaturing at 94 °C for 10 min, 94 °C for 30 s, 55 °C for 30 s, 68 °C for 30 s) using *pfu*Turbo DNA polymerase with genomic DNA as a template along with each respective primer. PCR fragments were gel purified and ligated into TOPOBlunt vector. NovaBlue competent cells were transformed with ligated construct and plated on kanamycin agar plates. Colonies were picked and the plasmids were purified and followed by enzymatic digestion using *Nde*I and *BamH*I, then gel purified. cDNA fragments were ligated into pET28a vector cut with the same restriction enzymes. NovaBlue competent cells were transformed with pET28a/Trx construct and plated on kanamycin agar plates. Colonies were grown and miniprepped as described above and cDNA was sent to the Massachusetts General Hospital Sequencing facility or to the MIT

Biopolymers Sequencing Facility. After confirmation of the correct sequences, the resulting constructs were transformed into Rosetta™ (DE3)pLysS cells for facile protein expression.

Purification of recombinant proteins. One liter TB media was inoculated with overnight-growth, grown at 37 °C to OD_{600}=0.8, induced with 0.8 mM isopropylthio-β-D-galactoside (IPTG), and incubated for an additional 3 hrs. Cell paste was obtained by centrifugation at 6,000 rpm for 10 min, resuspended in lysis buffer (10 mM TRIS-HCl, pH 8.0, 150 mM NaCl, 10 mM imidazole, 1 mM PMSF) and cell lysis was accomplished by ultrasonication. Cells were spun down at 18,000 rpm for 30 min. Supernatant was heated at 70°C for 45 min, and again centrifuged for 30 min. The resulting supernatant was loaded on a Ni-NTA column and eluted with elution buffer (10 mM TRIS-HCl, pH 8.0, 150 mM NaCl, 200 mM imidazole) and fraction samples run on SDS-PAGE 4-20% gel. Fractions containing proteins with comparable size to Trxs were concentrated using an Amicon stirred ultrafiltration cell. The molecular masses of the proteins with the (6) His-tag were estimated to be about 15.6 kDa for AF0711, 12.4 kDa for AF0769, 17.7 kDa for AF1284, and 14.3 kDa for AF2144. A SDS-PAGE is in agreement with the expected molecular weights (Figure 4.9).

Imidazole removed by dialysis into thrombin digest buffer (20 mM HEPES, pH 7.5, 20 mM NaCl, 0.5 mM (ethylenediamine)tetra acetic acid (EDTA), 1 mM 1,4-dithio-DL-threitol (DTT)) followed by His-Tag removal in thrombin digest buffer. Protein was loaded onto a SEC-30 high-prep size exclusion column from Amersham Biosciences (Piscataway, NJ, USA) column and eluted into storage buffer (10 mM HEPES, pH 7.0, 20 mM NaCl). Final protein fraction was concentrated using Centriprep YM-3filter from Amicon (Millipore, Bedford, MA, USA). Purity of protein was analyzed by 4-20% gradient SDS/PAGE. Protein concentration was determined from theoretical extinction coefficient (ProtParam, ExPaSy web server).

Thioredoxin Insulin Assay Activity Assays. Thioredoxin activity was determined by the insulin precipitation assay (23). The assay mixture was composed of 100 mM potassium phosphate, pH 7.0, 1 mM EDTA, and 0.13 mM bovine insulin in the presence and

absence of Trxs. The reaction was allowed to equilibrate for one min and then it was initiated by the addition of 1 mM 1,4-dithio-DL-threitol (DTT). Reaction was carried out on a Cary 300 spectrophotometer by monitoring increase in absorbance at 650 nm at 30 °C.

Protein Film Voltammetry. Protein-film voltammetry was carried out in a three-electrode configuration by using a standard calomel reference electrode and constructed PGE working electrodes. The electrochemical cell was water jacketed for temperature control and the reference electrode maintained in a separate compartment at a fixed (room) temperature. Electrodes were constructed of pyrolytic graphite (Advanced Ceramics), which was machined into cylinders, mounted onto a steel rod with silver epoxy, and then embedded in epoxy, yielding a graphite edge plane as the working surface. PGE electrodes were polished with an aqueous slurry of 1.0 μm alumina (Beuhler) and then sonicated prior to use.

Electroactive protein films were generated by rotating the working electrode at 200 rpm in a diluted protein solution (10 μM Trx in 20 mM 3-(*N*-morpholine) propanesulfonic acid (MOPS), 150 mM NaCl; pH 7.5) for 15 minutes while cycling the applied potential from 0.2 to -0.5 V. The protein-containing solution was then remove, the cell rinsed, and the working electrode then replaced in protein-free buffer solution containing 150 mM electrolyte. PFV was initially conducted at 10 °C to observe an electrochemical response at uniform conditions prior to adjusting the pH value or the temperature. Baseline subtraction was achieved by first measuring the baseline response of the electrode for a given set of buffer solution conditions prior to deposition of the protein film.

RESULTS

Sequence Alignments. The *afu*Trxs vary in length from 91 amino acids (*afu*Trx 4) to 134 amino acids (*afu*Trx 3). The shorter *afu*Trxs have sequence and length homology to glutaredoxins while the longer ones resemble thioredoxins (see Table 4.1, Figures 4.4). Analysis of the CxxC motifs for the cloned proteins reveals that each Trx has a unique CxxC sequence (Table 4.1). Homology between *E. coli* Trxs and *afu*Trxs ranges from

12% for *afu*Trx 2 to 40% for *afu*Trx 3. The *afu*Trx 3 resembles *E. coli* Trx1 and Trx 2 in that it contains the flanking tryptophan (WCxPC) (see Figure 4.2). *afu*Trx 1 and *E. coli* protein disulfide isomerase A, (dsbA) share a conserved CPHC motif (Figure 4.3), but little other similarity. The *afu*Trx 2 and *afu*Trx 4 CxxC motifs are identical to the *E. coli* glutaredoxin CPYC motif (see Figure 4.4). However, there is little other conservation of sequence between the *afu*Trxs and the *E. coli* glutaredoxins. All sequence alignments were performed using ClustalW alignment program from European Bioinformatics Institute. Alignment figures were generated using ESPript (22).

The purified *afu*Trxs yields proteins of an apparent MW of 15 kDa for AF0711, 13 kDa for AF0769, 18 kDa for AF1284, and 14 kDa for AF2144. A SDS-PAGE is in agreement with the expected molecular weights (Figure 4.9). Subsequent gel-filtration using a Sephacryl SEC-30 HR 16/60 column suggests that the *afu*Trxs are monomers in their native state.

All four *afu*Trxs reduced the insulin disulfide bonds. *afu*Trx 3 had the fastest rate, followed by *afu*Trx 2, than *afu*Trx 4, and than *afu*Trx 1 (Figures 4.10 - 4.13). Doubling the concentration of the protein decreased the time for reduction of insulin showing that the reduction of insulin was enzyme dependent. The controls with out protein or without insulin did not display a change in absorbance over time caused by turbidity of the assay.

Figure 4.14 shows the cyclic voltammetric response of four purported thioredoxins from the thermophilic archaeon *Archaeoglobus fulgidus*, at a pyrolytic graphite edge (PGE) electrode. Baseline subtraction of the non-faradaic component of the current reveals a single set of nernstian peaks in all cases. The peak height (I_p) corresponds linearly to the scan rate (v) indicating that the Trx-based signal is due to protein immobilized upon the PGE surface (data not shown).

The peaks resulting from oxidative and reductive scans appear highly symmetric, and all of the Trx proteins demonstrate a value of peak width at half height (δ) that indicates a cooperative redox process; that is, the number of electrons (n) is greater than 1(24). Figure 4.15A shows that at lower temperatures, measured values of δ are between those found for n=1 e^- and n=2 e^- process, though at 25 °C, δ approaches the anticipated value for an immobilized system undergoing a n=2 e^- reaction (equivalent to two highly rapid n=1 e^- steps) (24). Values of δ decrease as a function of temperature, suggesting a

dispersion of Trx conformations exists at the electrode surface, yielding artificial peak broadening at low temperatures and a nonzero peak separation. Upon heating, conformations that result in faster electrokinetics dominate and peak narrowing is observed. All experiments gave identical data regardless of the redox poise of the Trx. The determined values of the midpoint potential (E_m) and δ are given in Table 4.2 for a single set of conditions, demonstrating that E_m for the 2 e^- couple varies significantly between Trxs.

The pH dependence of the Trx 3 electrochemical response is shown in Figure 4.15B. In all cases, Trx films are stable upon the electrode surface from pH 4.5 to 9.5, though at pH values greater than 9.0, the intensity of the electrochemical signal becomes discernibly smaller (the depletion of peak height is fully reversible, indicating that protonation promotes conformations that yield reversible voltammetry). The slope of the pH dependence is -57 mV per pH unit, implying a 1 e^-/1 H^+ or 2 e^-/2 H^+ stoichiometry. As the observed peak widths indicate that n=2 e^-, the pH dependence indicates a 2 e^-/2 H^+ process is at work. Within the pH values studied, pK_a events are not observed, that is, E_m always appears to have a linear dependence upon the pH value. The absence of a pK_a value for the buried thiol of the reduced form may be due to the instability of the Trx-electrode interaction at pH values greater than 9.5. Protein films readily desorbed at highly basic pH values and recent studies of the *E. coli* and *Chlamydomonas reinhardtii* Trx, indicate that pK_a values for the Trx thiol greater than 10.0 are common for Trx proteins (25).

Trx 1 has a strongly oxidizing potential, in analogy with DsbA from *E. coli.* This variation can be generally understood by the hypothesis of Raines and co-workers that the presence of a protonatable residue in the -CXXC- motif, which contains the Trx disulfide bond, leads to a high potential (1). Indeed, the Trx 1 disulfide loop possesses the -CPHC- sequence found in DsbA, suggesting that functionally Trx1 is not a thioredoxin at all, but an oxidizing protein. The identity of the residues in the -CXXC- motif cannot be the sole determinant of the Trx reduction potential: Trx 3 and Trx 4 contain the same motif (-CMPC-) and yet vary in potential by 20 mV.

DISCUSSION

Although the four Trxs from *A. fulgidus* vary in sequence, they all contain the CxxC motif in the active site. Direct electrochemistry experiments reveal that the midpoint potential of the four *afu*Trxs range from -309 to -32 mV with *afu*Trx 1 being the most oxidizing member of the enzymes (Table 4.1). Analysis of the amino acid variation of the *afu*Trx 1 –CxxC- motif reveal an intervening The most oxidizing member of the Trx family is DsbA (Table 4.2) (11). Expanding the CxxC motif to include the neighboring amino acids shows a serine in N-terminal to the motif (SCPHCR) (Table 4.2). There are many hypotheses for the wide range of midpoint potentials of Trxs, and many mutagenic studies have tried to address this topic, (Reviewed in (26)), but model has emerged that explains the span of midpoint potentials. While one can speculate on the physiological role of afuTrx 2-4 in the cell, the role of *afu*Trx 1 is a mystery.

In summary, we have characterized four Trxs from *A. fulgidus*. The reduction potential of the Trx proteins in this archaeon represent the greatest range in a specific organism. Also, we have determined that PFV is a useful tool in the study of the archaeal thioredoxins as it does not depend upon other solution equilibria and it can relate quantitative information regarding redox cooperativity. Thus, employment of PFV will enable detailed studies of the relationship between sequence, structure, and redox chemistry of thioredoxins.

Table 4.1. Measured midpoint potentials and peak widths of Trx proteins; pH 7.0, 10 °C.

Sample	E_m [mV]	δ^I [mV]	-CXXC-
Trx 1	-32	62	CPHC
Trx 2	-301	65	CPYC
Trx 3	-287	56	CMPC
Trx 4	-309	64	CMPC

Table 4.2. Properties of *Escherichia coli* disulfide reductases and *Archaeoglobus fulgidus* Trxs. The table lists the length of the different proteins in the alignments and the amino acid composition of the CxxC motif plus and minus one amino acid.

Protein Length and xCxxCx motif composition

Protein	Number of AA	xCxxCx	Em (V)
Trx 1	109	WCGPCK	-0.270
Trx 2	130	WCGPCR	N/A
*afu*Trx 1	119	SCPHCR	-0.032
*afu*Trx 2	93	YCPYCR	-0.301
*afu*Trx 3	134	WCMPCK	-0.287
*afu*Trx 4	91	TCPYCP	-0.309
*ec*dsbA	108	FCPHCY	-0.122
*ec*dsbC	136	TCGYCH	-0.130
*ec*Glutaredoxin I	85	GCPYCV	-0.233
*ec*Glutaredoxin III	83	TCPYCH	-0.198

Figure 4.1. The –CXXC- motif of the thioredoxin superfamily is illustrated by (A) *E. coli* Trx (-CGPC-) and (B) DsbA (-CPHC-).

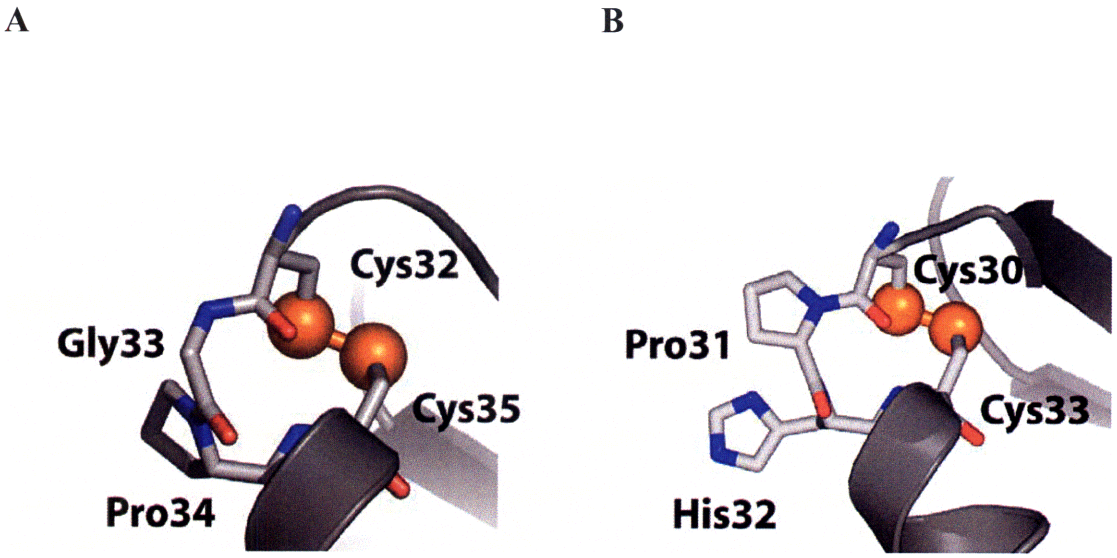


Figure 4.2 Alignment of *ec*Trx 1 and *ec*Trx 2 with *afu*Trx 1-4. The redox active CxxC motif is conserved in all sequences. Red boxes with white letters and blue boxes with red letters show identical and homologous residues, respectively. Last row displays consensus of sequences. Capital letters show identical residues and lower case letters show residues with greater than 50% occurrence.

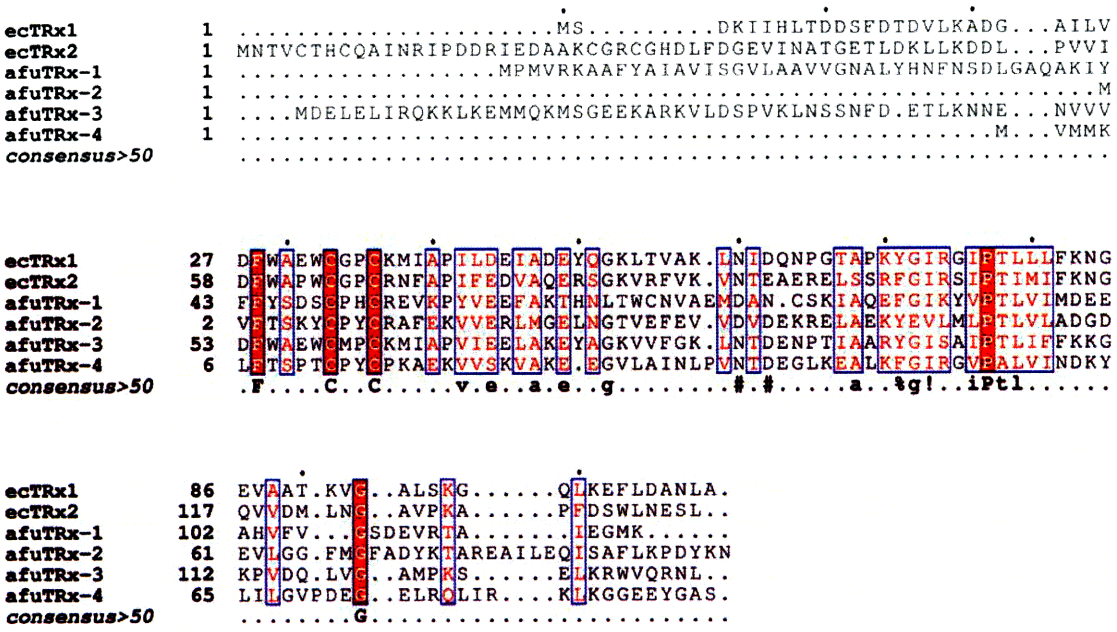


Figure 4.3 Alignment of *ec*dsbA and *ec*dsbC with *afu*Trx 1-4. The redox active CxxC motif is conserved in all sequences. Red boxes with white letters and blue boxes with red letters show identical and homologous residues, respectively. Last row displays consensus of sequences. Capital letters show identical residues and lower case letters show residues with greater than 50% occurrence.

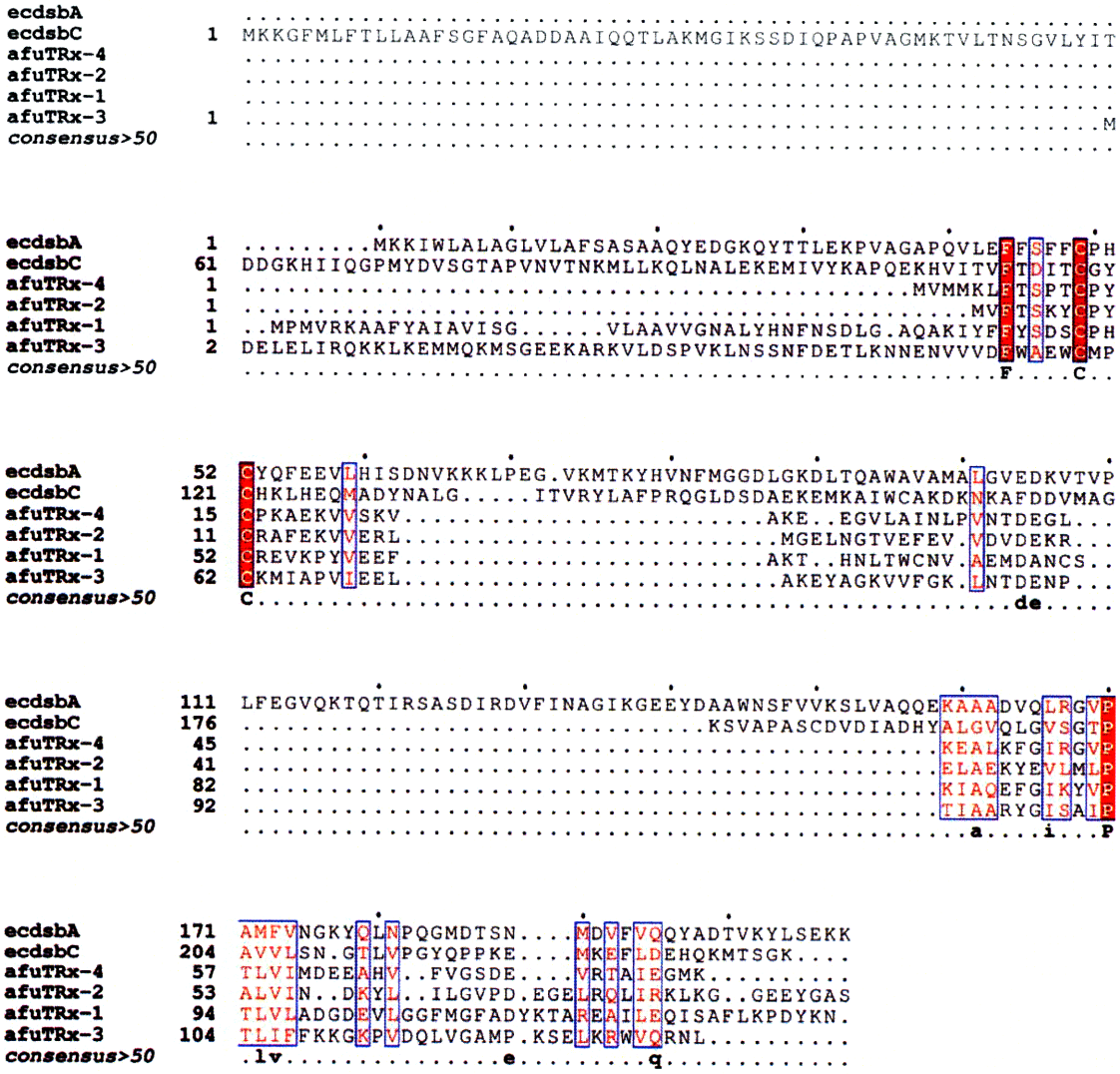


Figure 4.4 Alignment of *ec*Glutaredoxin 1 and *ec*Glutaredoxin 3 with *afu*Trx 1-4. The redox active CxxC motif is conserved in all sequences. Red boxes with white letters and blue boxes with red letters show identical and homologous residues, respectively. Last row displays consensus of sequences. Capital letters show identical residues and lower case letters show residues with greater than 50% occurrence.

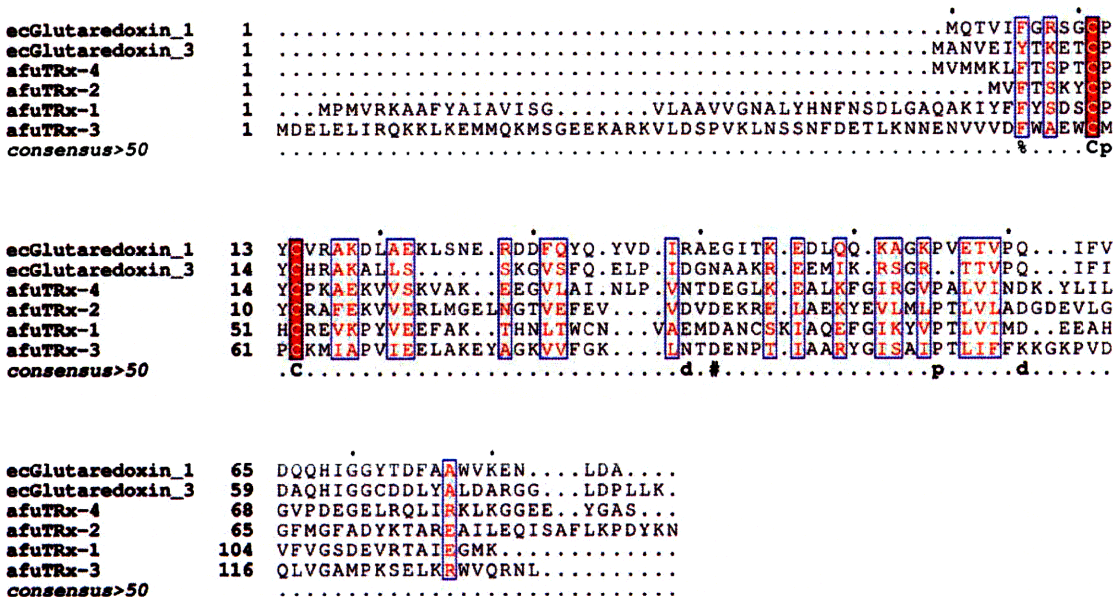


Figure 4.5 DNA and corresponding amino acid sequence of ORF AF0711 from *A. fulgidus* encoding for a putative Trx 1. DNA sequence was confirmed by the MIT Biopolymers Laboratory.

```
  1  atgccaatgg tgaggaaagc tgcattttat gcaattgcag tcatttcagg agttctcgct  60
       m  p  m     v  r  k     a  a  f  y     a  i  a     v  i  s     g  v  l  a

 61  gccgtagtgg ggaatgcctt gtaccacaat ttcaactcgg atttaggtgc tcaagctaaa 120
       a  v  v     g  n  a     l  y  h  n     f  n  s     d  l  g     a  q  a  k

121  atttacttct tctactccga ctcatgcccc cactgccgcg aagtgaaacc gtacgtggag 180
       i  y  f     f  y  s     d  s  c  p     h  c  r     e  v  k     p  y  v  e

181  gagtttgcaa aaacgcacaa cttaacctgg tgcaatgtcg cggaaatgga cgcaaactgc 240
       e  f  a     k  t  h     n  l  t  w     c  n  v     a  e  m     d  a  n  c

241  tccaaaatag ctcaggagtt tggaattaag tacgtgccta ctctcgttat aatggacgag 300
       s  k  i     a  q  e     f  g  i  k     y  v  p     t  l  v     i  m  d  e

301  gaggctcacg tttttgtggg ttctgatgag gttaggacag caatagaggg gatgaaatga 360
       e  a  h     v  f  v     g  s  d  e     v  r  t     a  i  e     g  m  k  -
```

Figure 4.6 DNA and corresponding amino acid sequence of ORF AF0769 from *A. fulgidus* encoding for a putative Trx 2. DNA sequence was confirmed by the MIT Biopolymers Laboratory.

```
  1  atggttttca ccagcaaata ctgcccctac tgcagggcct ttgagaaagt tgtcgagaga  60
       m  v  f     t  s  k     y  c  p  y     c  r  a     f  e  k     v  v  e  r

 61  ttgatggggg agcttaacgg cactgtggaa tttgaagtcg tggacgttga cgagaaaagg 120
       l  m  g     e  l  n     g  t  v  e     f  e  v     v  d  v     d  e  k  r

121  gagcttgctg agaagtacga ggttttgatg ctccccacct tggttcttgc agatggagat 180
       e  l  a     e  k  y     e  v  l  m     l  p  t     l  v  l     a  d  g  d

181  gaggttttgg gaggcttcat gggctttgcc gactacaaaa cagcccgaga agccatattg 240
       e  v  l     g  g  f     m  g  f  a     d  y  k     t  a  r     e  a  i  l

241  gagcagattt cagcatttct taagcctgat tacaaaaact ga                    282
       e  q  i     s  a  f     l  k  p  d     y  k  n     -
```

Figure 4.7 DNA and corresponding amino acid sequence of ORF AF1284 from *A. fulgidus* encoding for a putative Trx 3. DNA sequence was confirmed by the MIT Biopolymers Laboratory.

```
  1  atggacgagt tagagctaat cagacagaaa aaactcaagg agatgatgca gaaaatgagt 60
       m  d  e     l  e  l     i  r  q  k     k  l  k     e  m  m     q  k  m  s

 61  ggtgaggaga aggctagaaa ggttttggat tctccggtaa agctgaactc atcgaatttc 120
       g  e  e     k  a  r     k  v  l  d     s  p  v     k  l  n     s  s  n  f

121  gacgagacgc tgaaaaacaa cgagaacgtg gtagtggact tctgggcgga gtggtgcatg 180
       d  e  t     l  k  n     n  e  n  v     v  v  d     f  w  a     e  w  c  m

181  ccgtgcaaaa tgattgctcc ggttattgag gagctggcaa aggagtacgc gggaaaggtg 240
       p  c  k     m  i  a     p  v  i  e     e  l  a     k  e  y     a  g  k  v

241  gtgttcggaa agctgaacac agacgagaat ccaaccatcg cggcaaggta cggaatttca 300
       v  f  g     k  l  n     t  d  e  n     p  t  i     a  a  r     y  g  i  s

301  gccattccaa cgctgatatt tttcaaaaaa ggaaagccgg tggaccagct tgtgggagca 360
       a  i  p     t  l  i     f  f  k  k     g  k  p     v  d  q     l  v  g  a

361  atgccgaaaa gcgagctgaa gagatgggtt cagaggaatc tctaa                 405
       m  p  k     s  e  l     k  r  w  v     q  r  n     l  -
```

Figure 4.8 DNA and corresponding amino acid sequence of ORF AF2144 from *A. fulgidus* encoding for a putative Trx 4. DNA sequence was confirmed by the MIT Biopolymers Laboratory.

```
  1  atggagcgtc ttaacagtga gaggtttagg gaggtaatcc agagcgacaa gctggtggtg  60
       m  e  r  l  n  s  e  r  f  r  e  v  i  q  s  d  k  l  v  v

 61  gtggacttct acgccgactg gtgcatgccc tgcaggtaca tctcgccgat acttgagaag 120
       v  d  f  y  a  d  w  c  m  p  c  r  y  i  s  p  i  l  e  k

121  ctctcgaagg agtacaacgg cgaagtggag ttctacaagc tcaacgtcga tgaaaaccag 180
       l  s  k  e  y  n  g  e  v  e  f  y  k  l  n  v  d  e  n  q

181  gatgtggcct ttgaatacgg cattgcgagc atcccaacag tgctgttctt caggaacggc 240
       d  v  a  f  e  y  g  i  a  s  i  p  t  v  l  f  f  r  n  g

241  aaggttgttg gtggattcat tggagcgatg cctgaaagtg cggtgagggc ggaaatagag 300
       k  v  v  g  g  f  i  g  a  m  p  e  s  a  v  r  a  e  i  e

301  aaagcacttg gagcctga                                            318
       k  a  l  g  a  -
```

Figure 4.9 SDS-PAGE gel of *afu*Trxs. SDS-PAGE gel electrophoresis of *afu*Trxs. Locations of molecular weight standards are given by the positions of the molecular weight labels. Lanes of purified Trx proteins are labeled as per the thioredoxin.

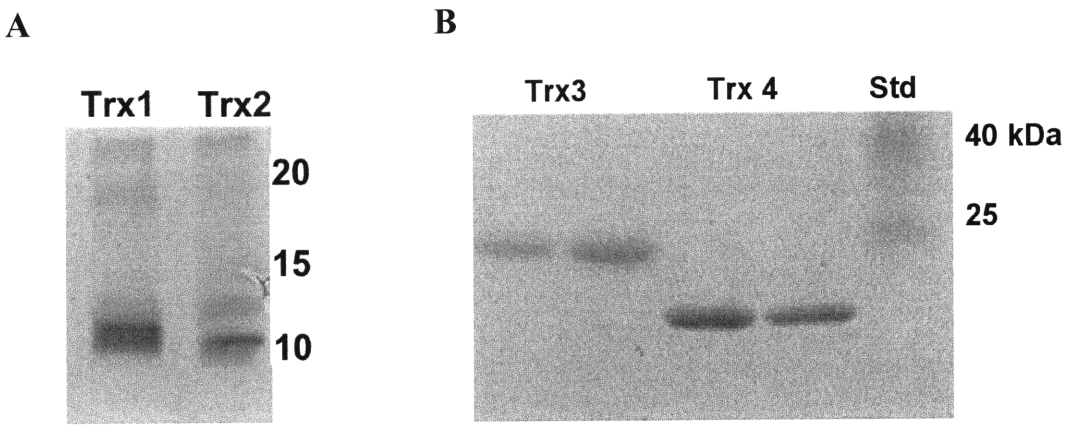


Figure 4.10 Reduction of insulin by *afu*Trx 1. Blue line is 6.7 µM *afu*Trx 1, red line is 3.3 µM *afu*Trx 1, green line is afuTrx 1 and DTT without insulin and black line is insulin and DTT with no enzyme.

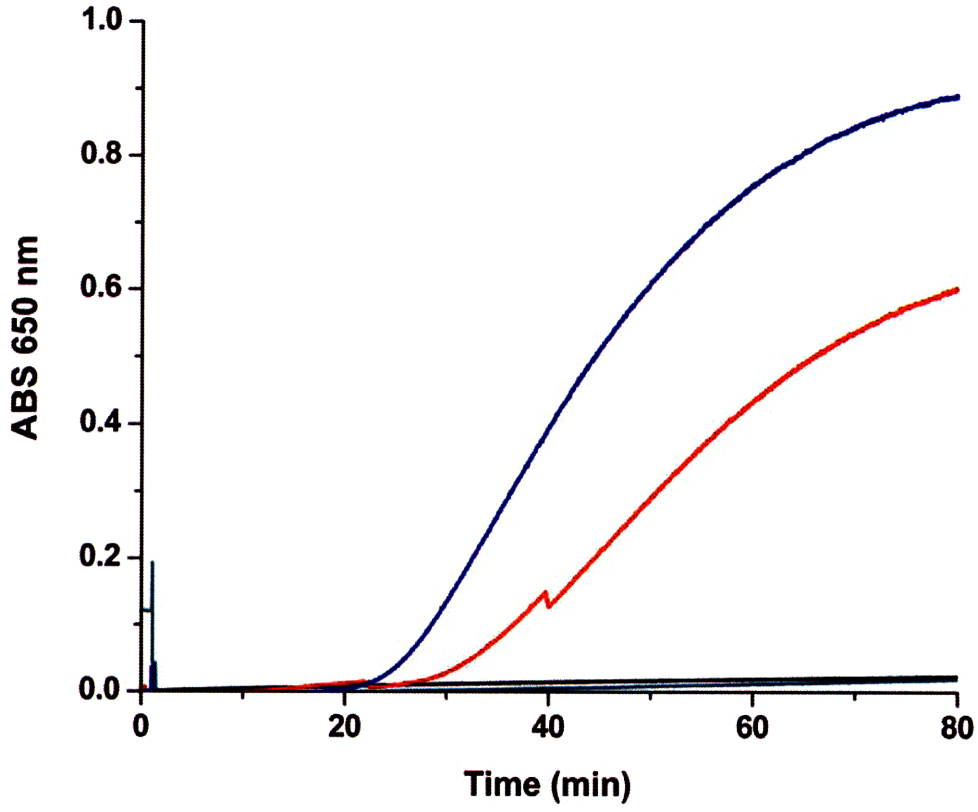


Figure 4.11 Reduction of insulin by *afu*Trx 2. Blue line represents 6.8 μM *afu*Trx2, red line represents 3.4 μM *afu*Trx2, and black line is insulin and DTT with no enzyme.

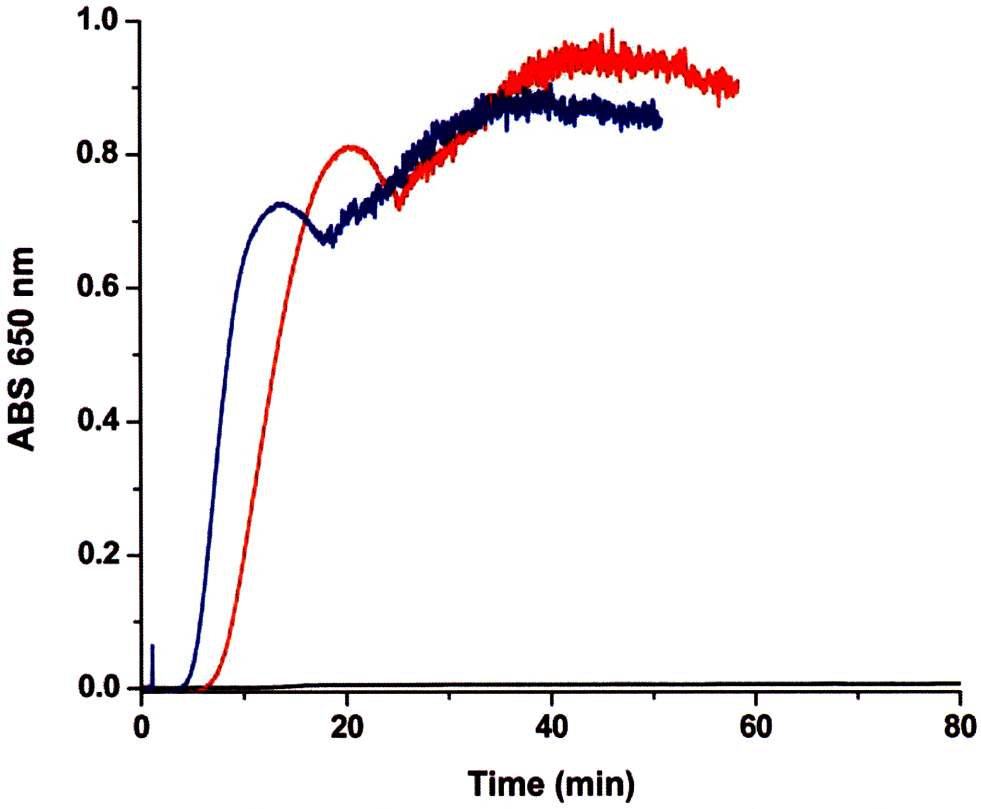


Figure 4.12 Reduction of insulin by *afu*Trx 3: Blue line is 7.8 µM *afu*Trx 3, red line is 3.9 µM *afu*Trx 3, and black line is insulin and DTT with no enzyme.

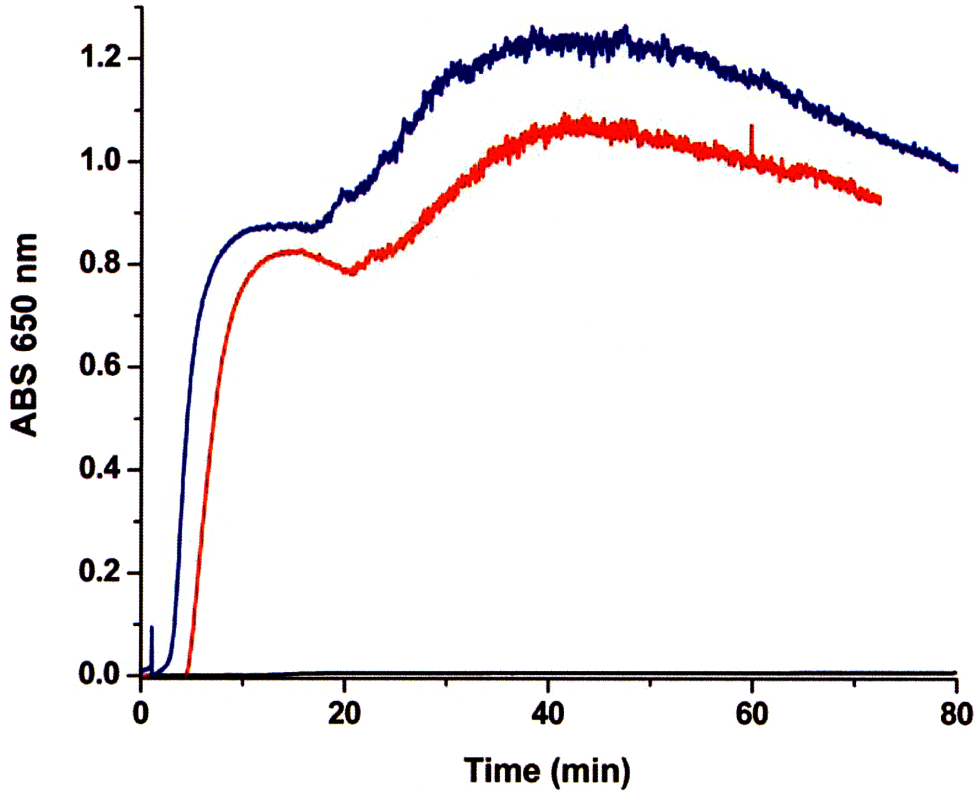


Figure 4.13 Reduction of insulin by *afu*Trx 4: Red line is 6.8 µM *afu*Trx 4, blue line is 3.4 µM *afu*Trx 4, and black line is insulin and DTT with no enzyme.

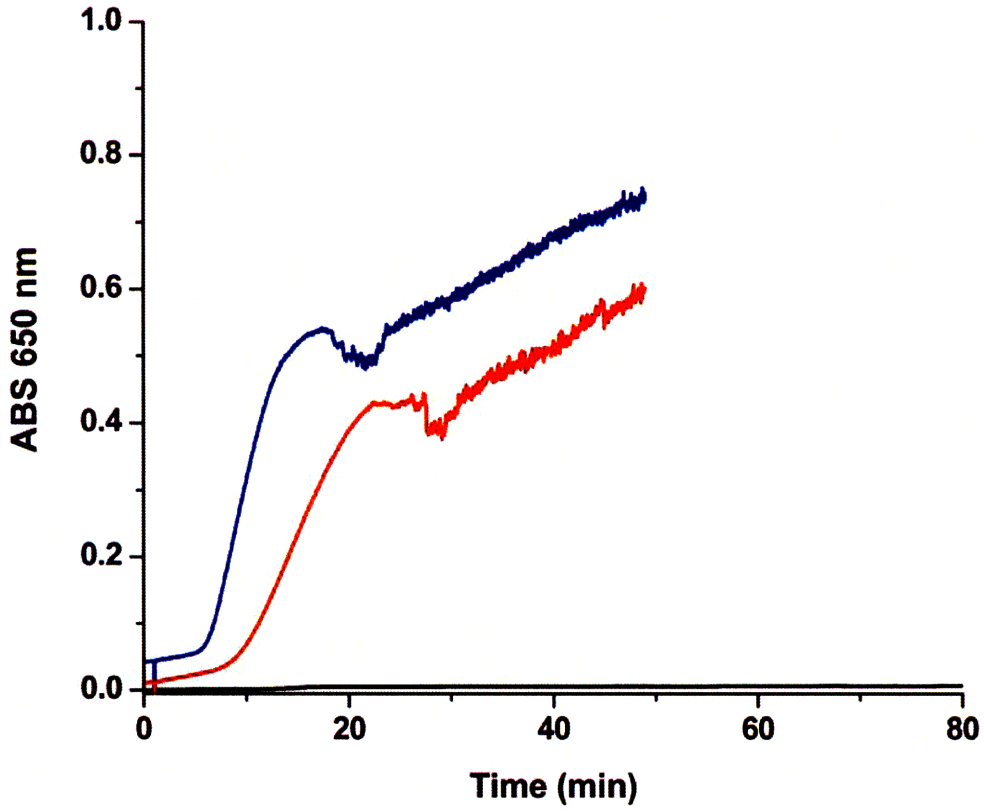


Figure 4.14 Voltammetric response of *A. fulgidus* Trx proteins as both raw and baseline-subtracted data at 10 °C. All data were acquired at a pyrolytic graphite edge electrode with pH 7.75 for Trx 1, Trx 2, and Trx 3, and with pH 4.5 for Trx 4. SHE=standard hydrogen electrode.

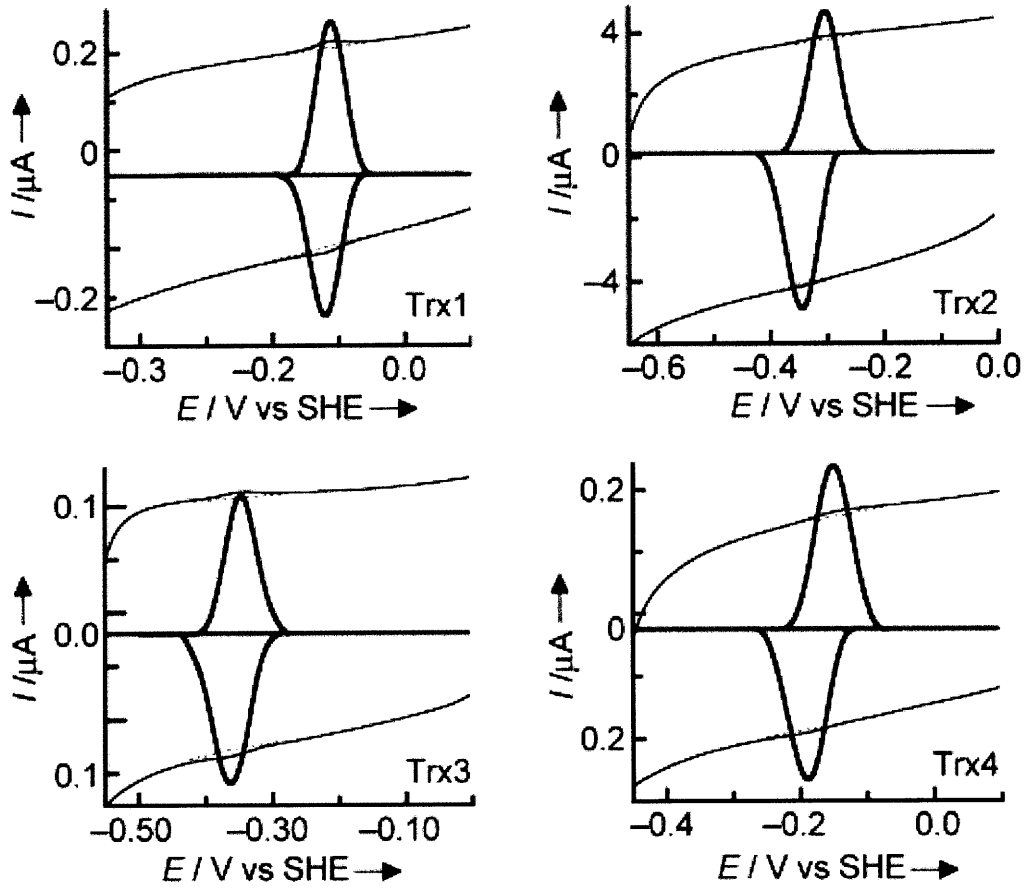
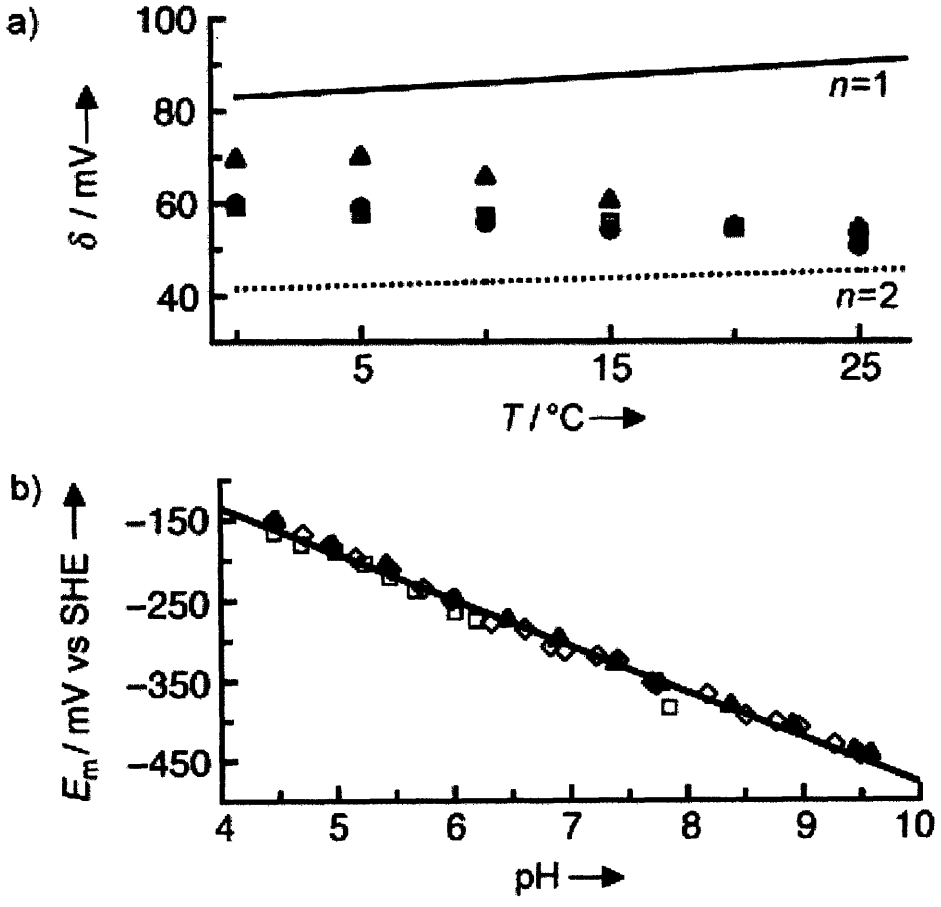


Figure 4.15 Temperature and pH dependence of the *afu*Trxs electrochemical response. a) Experimentally determined values of peak width for Trx 1 (□), Trx 2 (▲), and Trx3 (▪) (Trx 4 protein films were not sufficiently robust to heating) compared with calculated values for n=1 and n=2 electron reactions. b) The pH-dependent behavior of Trx 3 midpoint potentials for three distinct sets of experiments. The data are fit to a linear progression with a slope indicating a 1 e^-/1 H^+ process. The buffer solution was composed of 10 mM β-morpholinoethanesulfonic acid (MES), MOPS, 3-[tris(hydroxymethyl)methylamino]-1-propanesulfonic acid (TAPS), and 2-cyclohexylaminoethanesulfonic acid (CHES) with 150 mM NaCl, and the pH value was determined independently for each data point.



REFERENCES

1. Chivers, P. T., Prehoda, K. E., Raines, R. T.,(1997) The CXXC motif: a rheostat in the active site. *Biochemistry* 36, (14), 4061-4066.

2. Armstrong, F. A. H., Hendrik A.; Hirst, Judy (1997) Reactions of complex metalloproteins studied by protein-film voltammetry *Chemical Society Reviews* (26(3)), 169-179

3. Salamon, Z., Gleason, F. K., Tollin, G.,(1992) Direct electrochemistry of thioredoxins and glutathione at a lipid bilayer-modified electrode. *Arch Biochem Biophys* 299, (1), 193-198.

4. Salamon, Z., Tollin, G., Hirasawa, M., Gardet-Salvi, L., Stritt-Etter, A. L., Knaff, D. B., Schurmann, P.,(1995) The oxidation-reduction properties of spinach thioredoxins f and m and of ferredoxin:thioredoxin reductase. *Biochim Biophys Acta* 1230, (3), 114-118.

5. Johnson, D., Norman, S., Tuckey, R. C., Martin, L. L.,(2003) Electrochemical behaviour of human adrenodoxin on a pyrolytic graphite electrode. *Bioelectrochemistry* 59, (1-2), 41-47.

6. Johnson DL, M. L.,(2005) Controlling protein orientation at interfaces using histidine tags: an alternative to Ni/NTA. *J Am Chem Soc.* 127(7), (Feb 23), 2018-2019.

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

8. Holmgren, A.,(1985) Thioredoxin. *Annu Rev Biochem* 54, 237-271.

9. Grauschopf, U., Winther, J. R., Korber, P., Zander, T., Dallinger, P., Bardwell, J. C.,(1995) Why is DsbA such an oxidizing disulfide catalyst? *Cell* 83, (6), 947-955.

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

11. Wunderlich M, G. R.,(1993) Redox properties of protein disulfide isomerase (DsbA) from Escherichia coli. *Protein Sci.* , (2(5)), 717-726.

12. Huber-Wunderlich M, G. R.,(1998) A single dipeptide sequence modulates the redox properties of a whole enzyme family. *Fold Des*, (3(3)), 161-171.

13. Zu, Y., Shannon, R. J., Hirst, J.,(2003) Reversible, electrochemical interconversion of NADH and NAD+ by the catalytic (Ilambda) subcomplex of mitochondrial NADH:ubiquinone oxidoreductase (complex I). *J Am Chem Soc* 125, (20), 6020-6021.

14. Arner, E. S., Holmgren, A.,(2000) Physiological functions of thioredoxin and thioredoxin reductase. *Eur J Biochem* 267, (20), 6102-6109.

15. Hiniker, A., Bardwell, J. C.,(2003) Disulfide bond isomerization in prokaryotes. *Biochemistry* 42, (5), 1179-1185.

16. Katti, S. K., LeMaster, D. M., Eklund, H.,(1990) Crystal structure of thioredoxin from Escherichia coli at 1.68 A resolution. *J Mol Biol* 212, (1), 167-184.

17. Guddat, L. W., Bardwell, J. C., Glockshuber, R., Huber-Wunderlich, M., Zander, T., Martin, J. L.,(1997) Structural analysis of three His32 mutants of DsbA: support for an electrostatic role of His32 in DsbA stability. *Protein Sci* 6, (9), 1893-1900.

18. Hawkins, H. C., de Nardi, M., Freedman, R. B.,(1991) Redox properties and cross-linking of the dithiol/disulphide active sites of mammalian protein disulphide-isomerase. *Biochem J* 275 (Pt 2), 341-348.

19. Lundström, J., Holmgren, A.,(1993) Determination of the reduction-oxidation potential of the thioredoxin-like domains of protein disulfide-isomerase from the equilibrium with glutathione and thioredoxin. *Biochemistry*, (32(26)), 6649-6655.

20. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., Lipman, D. J.,(1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25, (17), 3389-3402.

21. Higgins D., T. J., Gibson T.Thompson J.D., Higgins D.G., Gibson T.J,(1994) CLUSTAL W: improving the sensitivity of progressivemultiple sequence alignment through sequence weighting,position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.*, (22), 4673-4680.

22. Gouet, P., Courcelle, E., Stuart, D. I., Metoz, F.,(1999) ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics* 15, (4), 305-308.

23. Holmgren, A.,(1979) Thioredoxin catalyzes the reduction of insulin disulfides by dithiothreitol and dihydrolipoamide. *J Biol Chem* 254, (19), 9627-9632.

24. Bard, A. J., Faulkner, L. R., 2001; *Electrochemical Methods: Fundamentals and Applications*. 2nd ed.; Wiley: New York.

25. Setterdahl, A. T., Chivers, P. T., Hirasawa, M., Lemaire, S. D., Keryer, E., Miginiac-Maslow, M., Kim, S. K., Mason, J., Jacquot, J. P., Longbine, C. C., de Lamotte-Guery, F., Knaff, D. B.,(2003) Effect of pH on the oxidation-reduction properties of thioredoxins. *Biochemistry* 42, (50), 14877-14884.

26. Carvalho, A. P., Fernandes, P. A., Ramos, M. J.,(2006) Similarities and differences in the thioredoxin superfamily. *Prog Biophys Mol Biol* 91, (3), 229-248.

This page intentionally left blank

Appendix 1:

Direct Electrochemical Analyses of a Thermophilic Thioredoxin Reductase: interplay between conformational change and redox chemistry

Reproduced in part from: Michael J. Hamill, Sarah E. Chobot, Hector H. Hernandez, Catherine L. Drennan, and Sean J. Elliott

INTRODUCTION

The ubiquitous protein Thioredoxin (Trx)1 functions as a reductant for many key cellular processes, including DNA synthesis, oxidative stress response and transcriptional activation (1). The reduced form of Trx is generated by thioredoxin reductase (TrR), and the pair maintains the reducing conditions of the cytoplasmic environment, as well as being responsible for the reduction of ribonucleotide reductase and peroxiredoxins, and the reductive activation of several transcription factors (2). The maintenance of redox environment is achieved through the coupling of reduction equivalents available from NAD(P)H to the TrR active site, which includes a flavin moiety, and a redox-active disulfide bond (3, 4, 5). TrRs are homodimeric enzymes that are differentiated by complexity and molecular weight: the prokaryotic enzymes (exemplified by the enzyme from *Escherichia coli*) are of relatively low *Mr* (~35,000 Da per monomer) when compared to their higher *Mr* mammalian homologs (~55,000 Da per monomer) which contain an additional redox active unit, such as a selenocysteine-cysteine pair (6). To date the *E. coli* TrR (*ec*TrR) has proven to be a robust model system for studies of low *Mr* TrR structural dynamics and catalytic chemistry. Preliminary structural analyses by Kuriyan revealed that NADPH binding occurs far from the flavin domain of the protein, while the redox active disulfide pair (Cys135 and Cys138) are on the *re* face of the flavin, in a buried state, forbidding direct interaction with Trx (*7*). Further analysis of this conformer, the flavin oxidizing (FO) state, by Waksman suggested that flavin reduction could be achieved by a 67° ball-in-socket rotation of the NADPH-binding and flavin domains (8). Additional kinetic and spectroscopic studies of wild-type and mutant forms of *ec*Trx have indicated that not only is a large conformation change required in the course of *ec*TrR catalysis for the generation of the flavin reducing (FR) form of the enzyme, but that the two forms, FO and FR are in equilibrium in solution (9, 10, 11,12). By crystallographically characterizing a site-directed mutant of TrR crosslinked to a Trx "doorstop", Ludwig and co-workers successfully provided the first structural snapshot of the FR conformation of *ec*Trx, as shown in Figure A1.1 (13, 14).

The picture of the *ec*TrR mechanism that emerges requires a ternary complex (15) that redox cycles between the two- and four-electron reduced states (16), where redox transformations must be coupled to potential conformational changes. However, it has proven challenging to develop a detailed picture of the impact of conformational dynamics upon the redox properties of the FAD cofactor and the Cys135-Cys138 disulfide. Previous potentiometric titrations suggest that the thermodynamic potentials of the *ec*TrR redox cofactors (FADH2/FAD and the Cys135-Cys138 disulfide) are very close, with *Em* varying by only 7 mV at pH 7.6 (16, 17). Specifically, the enzyme FAD/FADH2 potential has been calculated as -263 mV and -280 mV, for the active site disulfide and dithiol forms, respectively; while the enzyme active site disulfide/dithiol potential has been calculated as -270 mV and -287 mV for the flavin oxidized and flavin reduced forms, respectively (16, 17). However, such calculations have been previously based upon data that cannot discriminate between the FR and FO conformers, which rapidly interconvert at room temperature (18). Thus, we hypothesize that the former measurements represent potentials that are best considered as an average between the populations of each conformation.

Here we have looked to direct electrochemical methods in order to parse the relationship between conformational dynamics and redox chemistry. We have used protein film voltammetry (PFV) as a method to investigate the redox characteristics of the low *Mr* TrR enzymes from *E. coli,* as well as the enzyme for the archaeon *Thermoplasma acidophilum,* an enzyme that has been biochemically and structurally characterized in the accompanying report (19). PFV has the ability to interrogate redox equilibria that are coupled to additional chemical processes – particularly when chemical steps result in modulations in the midpoint potential of the system. For example, studies of proton-coupled steps in redox reactions of bacterial ferredoxin (20, 21) has resolved proton transfer reactions of a single sidechain transfer. No electrochemical studies of TrR (of any type) have been reported to date, and the reports of potentials have been limited to techniques that will average away any contribution due to distinct conformational changes, which are rapid on the timescale of potentiometry. However, typical electrochemical investigations of flavoenzymes are at times hindered by the stability of

the flavin-bound form of the protein: protein binding of the flavin cofactor can be compromised upon exposure to a typical graphite working electrodes frequently used in PFV experiments, as has been reported for flavodoxin and P450 BM3 (22, 23). Given this, we have opted to study both the low *M*r *ec*TrR and a novel TrR from the thermophilic organism *Thermoplasma acidophilum*, whose crystal structure and preliminary biochemical characterization are reported in the accompanying paper (19). Using this strategy, it is our hope that the greater protein stability of the thermophilic enzyme will yield not only better electrochemical results, but will enable a dissection of the contribution of FO and FR conformations to the control of flavin and disulfide reduction potentials. As described in the preceding paper, the *ta*TrR has recently been cloned, purified, and subjected to structural and biochemical analyses, demonstrating that this new TrR is an excellent homolog of the *ec*TrR, with the caveat that it cannot use NADPH as an electron donor (19).

MATERIALS AND METHODS

Protein purification. The *ta*TrR was purified as a His-tagged construct following cloning from genomic DNA from *T. acidophilum*. The details of protein purification and characterization are described in Chapter 3 of this thesis. The *ec*TrR was produced from the system described by Mulrooney (24), with the following modifications: cells were streaked on 2xYT agar plates with 100 μg/ml ampicillin, and 25 μg/ml kanamycin and incubated at 310 K overnight. One colony was used for a 600 ml culture that grew with gentle shaking overnight at 310 K in 2xYT media with 50 mM phosphate (pH 7), 20mM glucose, 100 μg/ml ampicillin, and 25 μg/ml kanamycin. The cell mixture was cooled on ice and centrifuged then the pellets were collected and frozen at 253 K overnight. Frozen cells were thawed and resuspended by adding 20 ml of 10mM phosphate (pH 7.6), 0.3 mM EDTA, and 0.5 mM phenylmethylsulfonyl fluoride. Cells were then sonicated 5 times for 10 seconds, allowing 20 seconds between sonications to cool on ice. After sonication, 100 mg streptomycin sulfate was added and the mixture was centrifuged, supernatant was saved, and ammonium sulfate was added to the final concentration of 0.56 g/ml. The precipitate was collected by centrifuging the mixture and the pellet was resuspended in 10 ml of 10mM phosphate (pH 7.6), 0.3 mM EDTA and dialyzed

overnight in 1 L 10 mM phosphate (pH 7.6), 0.3 mM EDTA, the solution was refreshed 3 times. The protein solution was then passed through a 45 μm syringe filter and then loaded onto a 3 cm x 5 cm 2`,5`-ADP Sepharose 4B column. Protein was eluted using an 120 ml 10 mM phosphate (pH 7.6), 0.3 mM EDTA buffer with a 0.15 mM – 1 M NaCl linear gradient. Fractions were combined based on 280 nm/ 456 nm ratio and then precipitated by adding ammonium sulfate to the final concentration of 0.56 g/ml. The protein precipitate was collected by centrifugation and the pellet was resuspended in 10 mM phosphate (pH 7.6), 0.3 mM EDTA, and 100 μM FAD, incubated for several hours at 4°C and then dialyzed overnight in 1 L 10mM phosphate (pH 7.6), 0.3 mM EDTA, the solution was refreshed 3 times. Protein was refrigerated for immediate use. The plasmid pTrR301 was the generous donation of Dr. Scott B. Mulrooney.

Protein electrochemistry.

Protein film voltammetry was conducted using a PGSTAT 12 electrochemical analyzer (EcoChemie), in a three electrode cell configuration, using a pyrolytic graphite edge (PGE) working electrode, a calomel reference electrode and a platinum wire counter electrode. All experiments were performed in a water-jacketed all-glass electrochemical cell that was thermostatted using a circulating water bath. The reference electrode was maintained at room temperature in the course of the experiments, and all potentials described here are reported versus the hydrogen potential, correcting for the discrepancy between SHE and the calomel electrode as a function of temperature. Electroactive films of *ta*TrR were generated by first polishing the PGE surface with and aqueous slurry of 1.0 μM alumina (Buhler), followed by sonication. The polished electrode was washed thoroughly with distilled water, and 1.0 μL of TrR (typically a 100 μM solution in pH 8.0 HEPES buffer with 150 mM NaCl) was deposited on the graphite surface. Excess protein was removed with washing, and the resulting protein film was placed immediately in the electrochemical cell compartment, in which 2.0 mL of a working cell buffer was purged under a continuous nitrogen blanket. Electroactive films of *ec*TrR were similarly obtained by first poising a freshly polished PGE at –510 mV in a protein free cell solution for five minutes. Then, 1.0 μL of *ec*TrR (260 μM) was directly to the electrode tip, and excess protein was removed by washing with a small amount of ice-cold buffer.

Electrochemistry was conducted as staircase cyclic voltammetry. The resulting voltammograms were subjected to baseline subtraction of a polynomial background, representing the non-faradaic (charging) component of the current. The entropic component of the redox reaction of TrR was determined using the slope of a plot of *Em* vs T (ΔSo rc = nF(d*Em*/dT)) for data collected using a cell in which the reference calomel electrode was maintained at a constant temperature (293 ± 0.5 K), while the working electrode-containing cell temperature was varied. The enthalpic component was similarly determined by a Gibbs-Helmholtz plot (*Em*/T vs 1/T).

Redox potentiometry.

Additional determinations of the thermodynamic potentials of *ta* and *ec*TrR were conducted using the xanthine/xanthine oxidase system described by Massey (25). Specifically, potentials were determined under anaerobic conditions in a glovebox using a S. I. Photonics 400 series spectrophotometer. The solution contained 100mM HEPES (pH 7.0), 1μM benzyl viologen, 1μM methyl viologen, 250 μM xanthine, xanthine oxidase, 18 μM enzyme, and 5μM Safranine O as a reference dye (-289 mV).

RESULTS

Protein electrochemistry.

The *ta*TrR was investigated at pyrolytic graphite edge electrodes using cyclic voltammetry, at 283 K, as shown in Figure 2 (solid line). The electrochemical response reveals a single pair of highly broadened voltammetric peaks, centered at a midpoint potential of -315 mV. On initial analysis, the data do not seem to be in accord with theoretical predictions of redox active molecules adsorbed at an electrode: ideal peak shapes would display zero separation of cathodic and anodic peaks, while the data shown in Figure 2A display a peak separation of nearly 110 mV. However, the relationship of cathodic and anodic peak heights (*ic* and *ia*) with respect to the scan rate (v) show a linear relationship, indicating an adsorbed species. Further, the ratio of *ic*/*ia* is invariant with respect to scan rate, indicative of a simple redox process with no further preceding, or following, chemical reactions (Figure A1.2A, inset). Like the unusual increase of peak

separation at 283 K, the peak-widths (δ) of both the cathodic and anodic peaks (Figure 2B) also deviate from a simple, ideal system. By theory, $\delta \approx 3.53(RT/nF)$, so at 283 K, a δ value of $86/n$ mV is predicted. Thus, even if the voltammogram represents the sum of two distinct, over-lapping electrochemical responses, each corresponding to an $n = 2$ redox reaction (for the FAD/FADH2 and active site disulfide/dithiol couples) the observed data appears to be unusually broad. Plotting the peak-width at half-height versus temperature (Figure A1.2B) shows that the peak-width decreases substantially as the temperature increases, becoming closer to the predicted value for a n=2 process at temperatures greater 313 K. Notably, theory states that for a routine redox process, the decrease should not occur (δ should increase as a function of temperature) indicating that other processes must be at work.

The observed voltammetric response is strongly affected by changes in temperature. The dependence of the voltammetric peak shape upon temperature is shown in Figure A1.3A. By warming the cell solution using externally controlled circulator, the peak separation, peak-width and midpoint potential of the *ta*TrR PFV response was altered. As mentioned above, the peak width decreases as a function of temperature (Figure A1.2B) and as Figure A1.3A indicates, no additional features are observed in the voltammetry. At low temperature, a broad envelop-like signal appears, while at elevated temperatures the voltammetric response sharpens considerably. In a similar way the peak separation also decreases as a function of temperature – which is also unexpected on the basis of electrochemical theory for a simple reversible redox equilibrium. Figure 3B shows that as the temperature increases to the upward limit of the stability of the electroactive protein film, 333 K, the peak separation decreases, though the ideal value of 0 mV is never attained. While perfect reversibility is not observed at higher temperatures, the peak separation (typically ~30-50 mV) is close to values observed for other systems analyzed with PFV. As discussed below, the non-zero peak separation may due to heterogeneity upon the electrode, including distinct global conformations of *ta*TrR at the electrode surface, but also the FR and FO conformations unique to low *MR* TrR enzymes. The intensity, or area, of the electrochemical response is also strongly affected by temperature – as Figure A1.3A indicates, the area of both the oxidative and the reductive

signals is reduced as the temperature is elevated. In the higher temperature regime (*e.g.*, 313 K), the *ta*TrR electrochemistry still corresponds to an immobilized species, as determined by a linear relation of *ia* or *ic* with *v* (data not shown). In the minutes required to attain and stabilize the protein at a specific temperature, desorption of the protein film did occur, though as all experiments were conducted in fresh, protein-free solutions, the desorption was not significant in the time-scale of a single cyclic voltammogram, collected at 150 mV/s. Figure A1.3A clearly indicates that cooled protein films initially yield pronounced surface coverage (heavy, solid line). Heating not only leads to peak narrowing and a decrease in peak separation, but when a protein film at elevated temperature is returned to the cool temperature (283 K), broadening and peak separation occurs again, and although the signal intensity is depleted (Figure A1.3A, dotted line) the peak width (Figure 2A, open square) and peak separation (Figure A1.3B, closed square) are nearly identical upon the return to 283 K, as compared to the initial experiments. We interpret the dramatic changes in the PFV response as a demonstration of the conformational dynamics of the FR and FO states as the temperature approaches the optimal growth temperature of *T. acidophilum*.

The redox thermodynamics for the overall *ta*TrR electrochemical response is illustrated in Figure A1.4, which depicts the measured midpoint potential of the overall voltammetric response (taking as *Em* the average of the cathodic and anodic peak) as function of temperature. The dependence of the observed midpoint potential displays a clear downward trend, as in the case of the redox thermodynamics of flavins measured in non-aqueous solvents (26), as well as in flavoproteins (23, 27). If the mid-point potentials are taken as a super-position of the two redox cofactors (*i.e.* the FAD/FADH2 and the disulfide/dithiol couples), and the behavior of the two couples is assumed to be similar, the slope of the linear fit shown in Figure A1.4A can be related to the entropy change for the redox reaction, $\Delta So\ rc/nF$, yielding a value of –78.1 J/K•mol. As the individual redox couples cannot be observed directly, this value can serve only as an estimate of $\Delta So\ rc$. A similar interpretation of a Gibbs-Helmholtz plot (Figure A1.4B) conveys $\Delta Ho\ rc$ for the process at the electrode, giving a value of –19.2 kJ/mol.

Possible Role of Adsorbed Flavin.

The propensity for several flavoproteins and enzymes to lose flavin cofactors at graphite electrodes forced us to consider the possibility that the apparent electrochemical reponse of *ta*TrR could be due to FAD adsorbed upon the electrode. To investigate this possibility, we have conducted several experiments: First, we have compared adsorbed FAD at a graphite electrode with the *ta*TrR response. A microliter of a buffered FAD solution (50 μM) was deposited on a graphite electrode, in a similar fashion to the *ta*TrR PFV. The resulting electrochemically active species displayed a midpoint potential ~50 mV more positive than the *ta*TrR response.

Second, in order to ensure the that our *ta*TrR data did not correspond to denatured protein in the presence of extra FAD in a non-native conformation, we have supplemented *ta*TrR solutions used to generate protein films with additional, 'free' flavin. Free FAD can be added to the electrochemical cell, with *ta*TrR already adsorbed onto the working electrode. In this experiment, the typical *ta*TrR PFV signal (Figure A1.5A, light solid line) in the absence of exogenous flavin, transforms into a voltammogram that is a convolution of overlapping 'free FAD' and '*ta*TrR-based' signals as a function of the added FAD (Figure A1.5A). Identical results are achieved with FMN or riboflavin addition, and it is clear that the FAD-based signal is discernible from the TrR-based signal, at all pH values in this study (Figure A1.5B). In the case of both FAD and *ta*TrR, a pH dependence of –60 mV/pH is observed (Figure A1.5B), indicating a 1:1 ratio of protons and electrons involved in the redox couples.

Third, the *E. coli* TrR (*ec*TrR) can be interrogated using PFV (Figure A1.6), and reveals a significant loss of 'free flavin' at the EPG electrode, indicating that a simple difference between the Ec and Ta enzymes is the enhanced stability of the TrR from the thermophile *T. acidophilum* in retaining flavin, when bound to the electrode. Similar to the *ta*TrR findings, the *ec*TrR shows a 'free flavin' based electrochemical signature that grows upon the addition of exogenous flavin into the cell solution (data not shown). In general, the *ec*TrR behaves very much like the *ta*TrR although the overall electroactive surface coverage (Γ) is significantly lower for the *ec*TrR as compared to the *ta*TrR. Yet, the average peak width of the *ec*TrR does decrease with increasing temperature (Figure

A1.6B), as observed for the *T. acidophilum* protein (Figure A1.2B); thus, we believe that the essential characteristic of redox function the TrR cofactors are nearly identical.

Finally, we have noted that electrodes modified with *ta*TrR behave very much like other protein film electrodes, in contrast with FAD-modified electrodes. Electroactive films generated from *ta*TrR solutions are easily removed by repolishing the electrode surface with an alumina slurry, and electrodes are themselves sensitive to changes in buffer conditions. Neither of these parameters is true for a graphite electrode modified with FAD, which is quite difficult to remove from the graphite surface, even with multiple cycles of mechanical abrasion and alumina polishing. Thus, due to these four lines of inquiry, we are convinced that the Ta TrR-based electrodes genuinely represent the redox chemistry of the *ta*TrR, and are not grossly compromised due to the loss of the flavin cofactor.

Verification of Potentials via Redox Potentiometry.

In order to verify that observed reduction potentials were analogous to measurements obtained by redox potentiometry for the TrR from *E. coli*, we measured the observable flavin potential for *ta*TrR using the xanthine/xanthine oxidase-coupled potentiometric assay described by Massey (25). Figure A1.7 depicts the result of this experiment, using the redox indicator safranine O (SafO). The solution-based method cleanly reduces the *ta*TrR flavin from the oxidized to reduced forms, without the generation of the flavin semiquinone, yielding a value of the flavin potential of -305 ±5 mV, at 303 K, agreeing with the measurement made with PFV (-322 mV for the entire *ta*TrR envelop, Figure A1.4). Control experiments were conducted using the *ec*TrR, which we analyzed using an identical protocol, resulting in a potential of -291 mV ± 5 mV, in a reasonable agreement with previous measurements of the *ec*TrR flavin potential, -252 mV, as determined by O'Donnell and Williams, using a potentiometric method involving NADPH- and Trx-based equilibria.

Variable Scan Rate Measurements.

In order to better describe the temperature-induced alteration of the *ta*TrR electrochemical response, we examined the impact of scan rates (and therefore, time scale) used in PFV. By collecting voltammetry data at faster and faster time-scales (increasing scan rate), we presumed that in the higher temperature regime, fast scan rates would allow us to detect the distinct FO and FR conformations on the electrochemical timescale. Thus, if conformational dynamics were a substantial contributor to the narrow voltammetric envelop observed at higher temperatures, we could reverse the narrowing effect by using fast scan rates. Figure A1.8 shows the impact upon the PFV experiment achieved by increasing the scan rate over two orders of magnitude, while the temperature is held at a constant value of 308 K. By normalizing the baseline-subtracted data for the scan rate (and small deviations overall area due to film-loss associated with the thermal desorption of the TrR from the electrode), it clear that as v is increased from 200 mV/s (Figure A1.8, dotted line) to 2 V/s (dashed line) to 20 V/s (solid), the data mimic the impact of lowering the temperatures. As shown in Figure A1.8, as v is increased by two orders of magnitude, the oxidative and reductive peaks shift from the mid-point position by over 110 mV, and the shape of the peak broadens by ~50%. At temperatures greater than 308 K identical behavior was observed when the scan rate is increased, though the electrochemical response is complicated by thermal desorption of the *ta*TrR which yields signals of smaller magnitudes (and for this reason, we show illustrative data at 308 K in Figure A1.8). The overall scan-rate dependent behavior is likely compounded with the natural increase of peak separation and peak-width at half-height that can otherwise be utilized to find interfacial electron transfer kinetics. Yet, qualitatively, we have found a striking similarity between experiments conducted using faster scan-rate yet higher temperatures, and those at slower scan-rates and lower temperatures. As discussed below, we take these findings as supportive of the model of TrR conformational dynamics upon the electrode.

DISCUSSION

We have shown that the direct electrochemistry can be used to interrogate the redox active cofactors of low Mr TrRs, through the study of the newly characterized enzyme from *T. acidophilum* as well as the TrR from *E. coli*. In the most general terms, the data

are in reasonable agreement with control experiments carried out with a xanthine/xanthine oxidase method for the determination of flavoprotein redox potentials. Other methods of spectrophotometrically titrating the reduction potential of the flavin would typically utilize the soluble redox agent responsible for supporting turnover *in vivo*, such as NADPH. As described in the preceding paper, however, the *ta*TrR cannot be reduced by NADPH, due to clear structural determinants that preclude binding of the co-substrate. At this time, the nature of the *in vivo* reduction system is unknown. Regardless, the use of the xanthine/xanthine oxidase method revealed a flavin potential for *ta*TrR of -305 mV at 303 K, consistent with the broad electrochemical response centered at -322 mV observed at the same temperature. We have ruled out the possibility that the PFV results are due to free flavin cofactor that has dissociated from the *ta*TrR during the course of our experiments, and we have found that there is a similarity between the *ec*TrR electrochemistry, and that observed for the *ta* enzyme. In both cases, the voltammetric response shows broad, yet symmetrical voltammetry, that likely contains information about both the flavin cofactor as well as the redox-active disulfide group associated with thioredoxin reductases. However, two distinct voltammetric features are not clearly resolved, regardless of the temperature, pH, or ionic strength (not shown), each of which could be a factor that would preferentially modulate one of the two cofactors' redox properties. Further, the redox thermodynamics of the voltammetric feature observed shows a single, linear dependence of *Em* as a function of temperature, indicating that PFV is reporting upon redox processes which are behaving similarly at the electrode. However, *ta*TrR reveals a very stable electrochemical response that displays a peak shape that is very sensitive to the temperature of the electrochemical cell. While requiring significantly lower temperatures, McEvoy and Armstrong have also noted that cryotemperature-induced broadening can occur for thermophilic proteins, such as the *ta* ferredoxin (28). Here we hypothesize that low temperature induced "broadening" is further enhanced by the dynamics of the TrR enzyme. In following points, we will attempt to resolve an electrochemical model of the data, which rationalizes how the distinct conformations associated with TrR activity (FR and FO) result in a modest modulation in the potentials for redox active cofactors associated with TrR, thus yielding the apparently broadened voltammograms.

General considerations of an electrochemical model.

As the FO and FR conformations can interconvert regardless of redox-state, a model describing the relationships between the conformational states and redox potentials must allow for each redox reaction and each interconversion between FO and FR, as shown in Scheme A1.A1. As depicted, either the flavin or the disulfide cofactor may undergo redox chemistry (indicated by either top or bottom face of the cube illustrated in Scheme A1.1) regardless of the conformer. The interconversion between the conformers is indicated as a transition between the top and bottom faces of the cube. Thus, if both FO and FR are populated with equal probability at the electrode, there are eight distinct microscopic potentials that should be associated with electrochemical steps. This scheme can be simplified to Scheme A1.1B, which considers that PFV will report upon thermodynamic potentials, resulting from just four features (*Em* for the FAD/FADH2 and the disulfide/dithiol couples, in both the FR and FO conformations). In the case of the *ta*TrR, the relative populations of the FR and FO conformers are not known; for our electrochemical analyses, this should make itself manifest as a specific electroactive coverage for both the FR (ΓFR) and the FO (ΓFO) conformers. In the following analysis, we will assume that the broad electrochemical envelope observed at low temperature is due to kinetic trapping of FR and FO states upon the electrode; thus, heating the *ta*TrR allows for interconversion between the states, and the resulting voltammetric features represent an averaging between specific redox couples (*e.g.*, the FAD/FADH2 couple) as experienced in the FO and FR conformers.

PFV results of the taTrR at 333 K.

As observed in the raw data (Figure A1.3A) conducting PFV at temperatures near the growth optimum of *T. acidophilum* results in highly symmetric, nearly reversible voltammograms. In our model this feature should contain information regarding the FAD/FADH2 and disulfide/dithiol redox couples of the *ta*TrR, averaged between the interconverting FO and FR conformations. Thus, we have initially attempted to describe this data by a set of two redox couples of equal electroactive area, as the stoichiometry of the flavin cofactor to the disulfide is 1:1. Figure A1.9 shows the result of fitting the

baseline-subtracted, oxidative scan, computing an electroactive coverage for the TrR, and determining a two-electron mid-point potential for both the flavin and the disulfide, each of which is described by the formalism derived by Laviron, given in Equation A1.1 (*29; 30*).

Eq. A1.1 .
$$i_{total} = \frac{F^2 \nu A \Gamma}{RT} \sum_k \left(n_s^k n_{app}^k \frac{exp[(n_{app}^k F/RT)(E - E^k)]}{(1 + exp[(n_{app}^k F/RT)(E - E^k)])^2} \right)$$

Where,

$$\delta = exp[f(E^0_{I/R} - E^0_{O/I})] \quad ,$$

$$\xi = exp[2f(E - E^0_{O/R})] \quad , \text{and}$$

$$f = F/RT \quad .$$

(Here, F is the Faraday, R the ideal gas constant, T the temperature, ν the scan rate, A the geometric area, and Γ the electroactive coverage.) In all cases, qualitatively identical results were observed for the reductive scan of each experiment. As Figure A1.9 illustrates, at 333 K the PFV response of *ta*TrR is modeled by just two n=2 e– features that are indicated by the dotted lines. In this way, the two voltammetric features at -295 mV and -330 mV are directly observed, and most likely represent the cooperative FAD/FADH2 and disulfide/dithiol couples. The two features are separated in potential by approximately 35 mV, somewhat larger than the difference of 7 mV suggested by Williams and co-workers for the *E. coli* enzyme (9; 16; 17).

Additionally, the higher potential feature is in good agreement with the value for the FAD cofactor determined by redox titration at 303 K, -305 mV. Figure A1.9A shows the fitting for a case where the cooperativity of each cofactor is perfect (by requiring a large separation in the crossed, formal one electron potentials described by Eq. A1.1). The result is only slightly improved by allowing for the possibility that the redox behavior is somewhat nonideal. Figure A1.9B shows that a discernible, but modest improvement in the fit (and less than 5 mV total separation in the two detectable potentials) is achieved by allowing for breakdown of perfect n=2 e– behavior,

represented here by slight broadening of the component features (Figure A1.9B, dotted lines). In the Laviron description of two-electron redox chemistry, a decrease in cooperativity corresponds to the two one-electron potentials becoming closer in value. The apparent loss of cooperativity could also be reflective of dispersion of the *ta*TrR molecules upon the electrode, that is still present at the elevated temperature. From this analysis, it is clear that the averaged potentials are close in absolute value, and cannot be more than ~35 mV separated from one another, or the electrochemical data would reveal a clear indication of two peaks in the raw data.

PFV results at 283 K.

A similar approach to the analysis of data collected at lower temperatures (here, 283 K) is shown in Figure A1.10. Attempting to describe the data in terms of two n = 2e– components is not at all possible, given the overall broadening observed in the data. However, allowing for the loss of cooperativity is illustrated in Figure A1.10A, which shows the fitting of the oxidative half-scan of data collected at 283 K, where the dotted lines correspond to two n = 2e– components that both display artificially broadened peak shapes. The breadth in the case of the fitting shown in Figure A1.10A arises from the loss of cooperativity for the redox-chemistry of both co-factors (i.e., for both features the two one-electron potentials are apparently no longer crossed.) However, there are no additional indicators from solution experiments that at 283 K, either the *ta*TrR flavin or the disulfide engage in one-electron chemistry. Indeed, the redox titrations of the flavin all show clean n = 2e– behavior at low temperatures (Figure A1.7). As such, we take the apparent loss of cooperativity to be synonymous with the effect dispersion generated by the "freezing out" of the FO and FR conformations. If we do attempt to account for the possibility of distinct conformations, by modeling the low-temperature response with four species, we can indeed obtain an excellent agreement with the data, shown in Figure A1.10B. In this model, four features are apparent, corresponding to both n = 2e– redox cofactors being present in the FO and FR conformations. The key result from this analysis is that we can estimate that the impact of FR and FO conformers is a modulation of 40-60 mV for both the FAD/FADH2 and disulfide/dithiol reduction potentials. However, we cannot assign specific redox couples to the voltammetric components of the

envelope. Instead, PFV elucidates the similarity in the formal potentials for the cofactors, and the dynamics of TrRs at an electrode interface.

While the development of a more precise mechanistic model will require greater insight into the assignment of the nernstian peaks observed in our fitting (potentially through the generation of a disulfide-free mutant of the *ta*TrR), an attractive hypothesis would allow for the flavin to possess a higher potential in the FR conformation, such that reduction of the flavin is efficient; and then in the FO conformation if the potential has been lowered by ~40-60 mV with respect to the disulfide, reduction of the disulfide will again be favored. Finally, when the FR conformer is adopted again, the resulting dithiol will be able to readily react with Trx in a thermodynamically favorable manner. Thus, the TrR scaffold may be able to achieve precise and unidirectional redox chemistry by tuning the potential of flavin and disulfide by discreet steps.

Concluding Remarks

Here we have interrogated low *Mr* thioredoxin reductases using direct electrochemistry for the first time. We investigated the redox properties of a classic TrR from *E. coli*, as well as newly described TrR from the archaeon *Thermoplasma acidophila* (19). Structural and biochemical studies have demonstrated that the *ta*TrR is very much like the better characterized bacterial homolog from *E. coli* except in its selection for its reductant. By using PFV we have been able to exploit the superlative thermostability of the *ta*TrR to our advantage, and demonstrate that protein dynamics between the FO and FR conformations contribute to the formal potentials of the flavin and disulfide cofactors.

Figure A1.1. The FO and FR conformations of *ec*TrR, superimposed in grey and wheat respectively. Orange spheres are used to represent the disulfide-forming cysteines in the FO conformation, and the cysteine/serine pair used in the crystallographic trapping of the *ec*TrR FR conformation. The image was constructed from the FO and FR pdb files 1TDF.pdb and 1F6M.pdb, respectively.

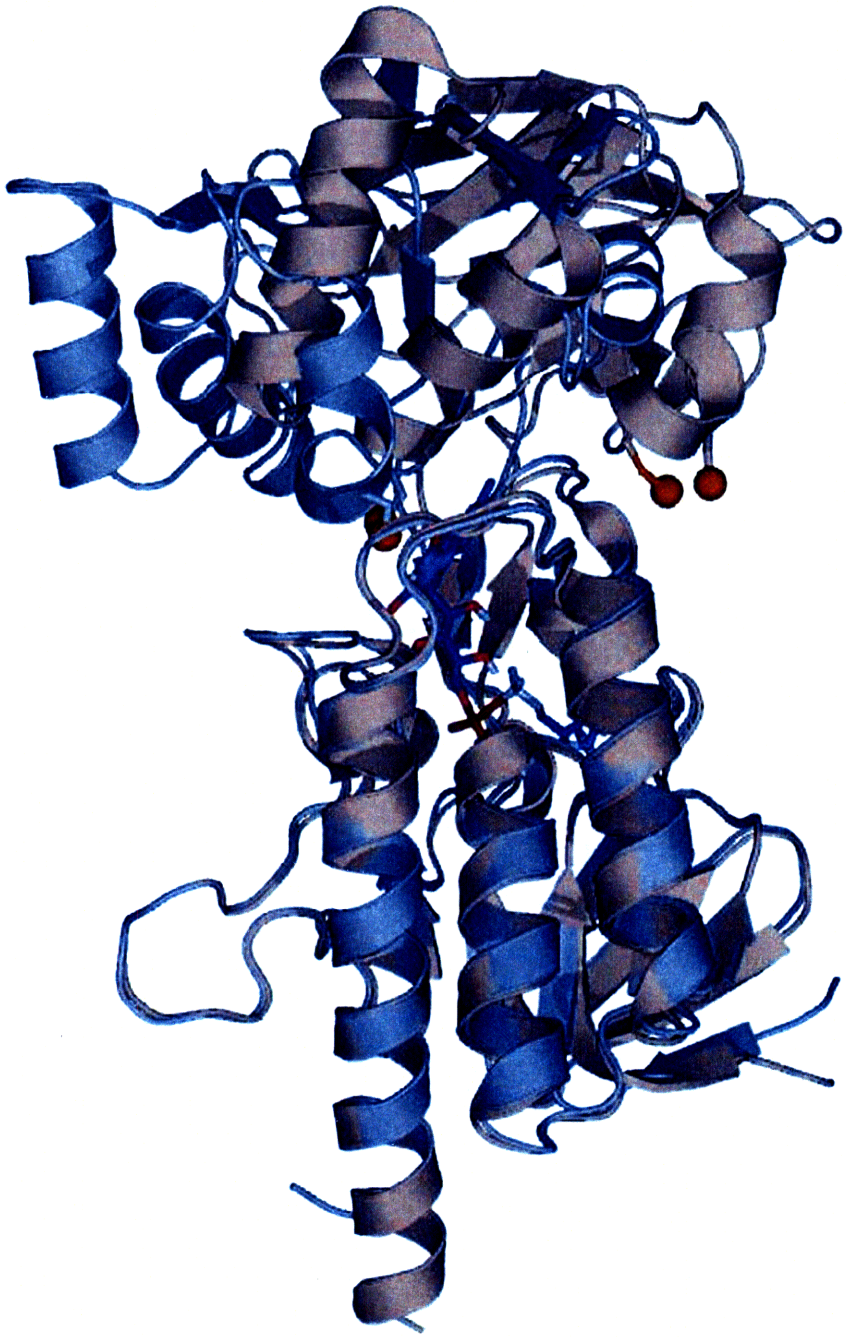


Figure A1.2. PFV response of *ta*TrR upon EPG electrodes. (A) Raw PFV data of *ta*TrR, at low temperature (solid line) superimposed upon the baseline response of the EPG electrode used for *ta*TrR immobilization (dashed line). Experimental conditions were as follows: scan rate, 100 mV/s; pH 7.0; 100 mM NaCl electrolyte; 283 K. (inset). Inset: dependence of peak height (*i*p) as a function of scan rate. (B) The average peak width at half height for oxidative and reductive scans, as a function of temperature (open squares), compared to computed values of peak-width for a n = 1e– (solid) or n = 2e– (dashed) features. The solid square represents the observed average peak width, after returning a heated *ta*TrR protein film to 293 K.

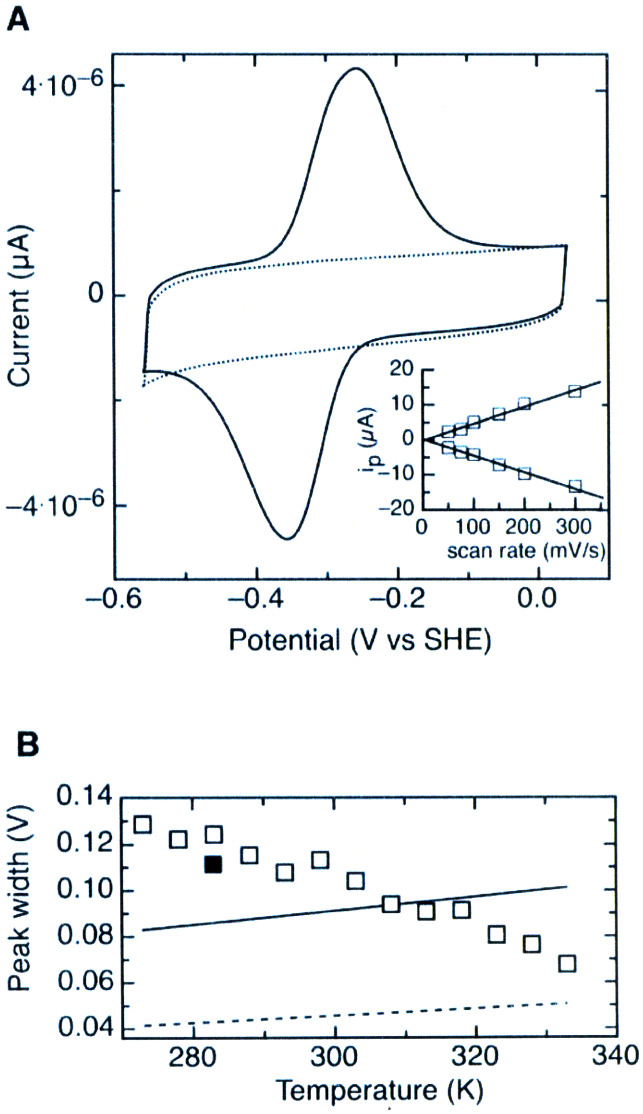


Figure A1.3. Variable temperature, PFV response of the *ta*TrR. (A) The *ta*TrR PFV response as a function of temperature. Protein films generated at 283 K (solid line, heavy) were then subsequently warmed to 293 (blue dashed), 303 (green dash), 313 (yellow) and then 323 K (red dash). Finally, the experimental temperature was lowered to 283 again (dotted line), indicating the loss of the protein surface coverage during the course of the experiment. All data were collected at pH 7.0, 100 mM NaCl electrolyte, and with a scan rate of 150 mV/s. (B) Peak separation as a function of temperature, for a single experiment, heating the protein film in increasing steps of five degrees; the open square represents the peak separation observed when returning the *ta*TrR film to 283 K.

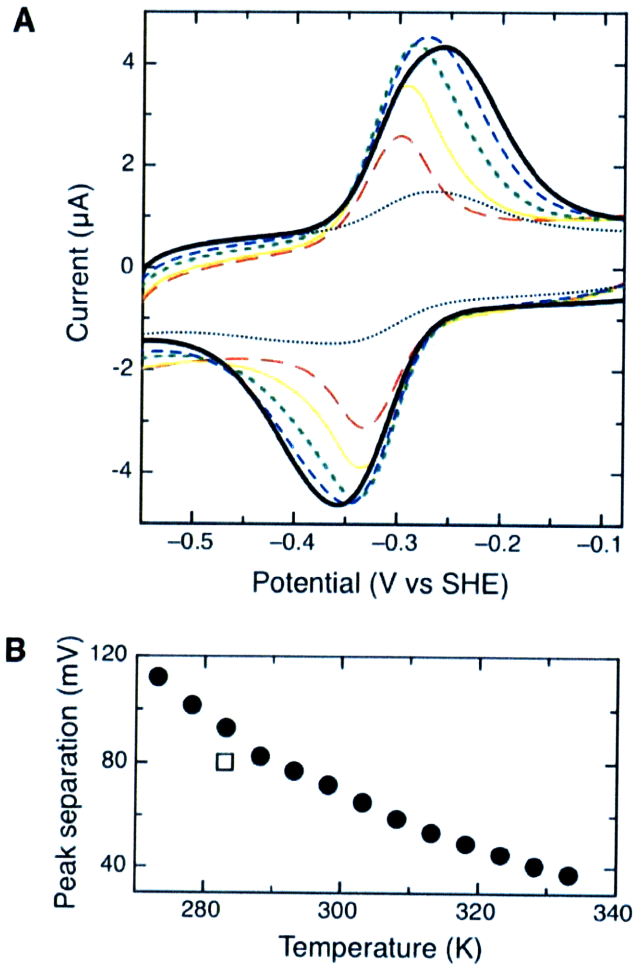


Figure A1.4. Redox thermodynamics of *ta*TrR protein films, modeling the data to two overlapping 2e– features. (A) ΔSo rc is determined from the slope of the *Em* vs T plot. Data are fit to a linear regression analysis, and the slope of which is interpreted as slope = ΔSo rc/*n*F, where *n* is assumed to be 2e–. (B) Gibbs-Helmholtz plot of the same data.

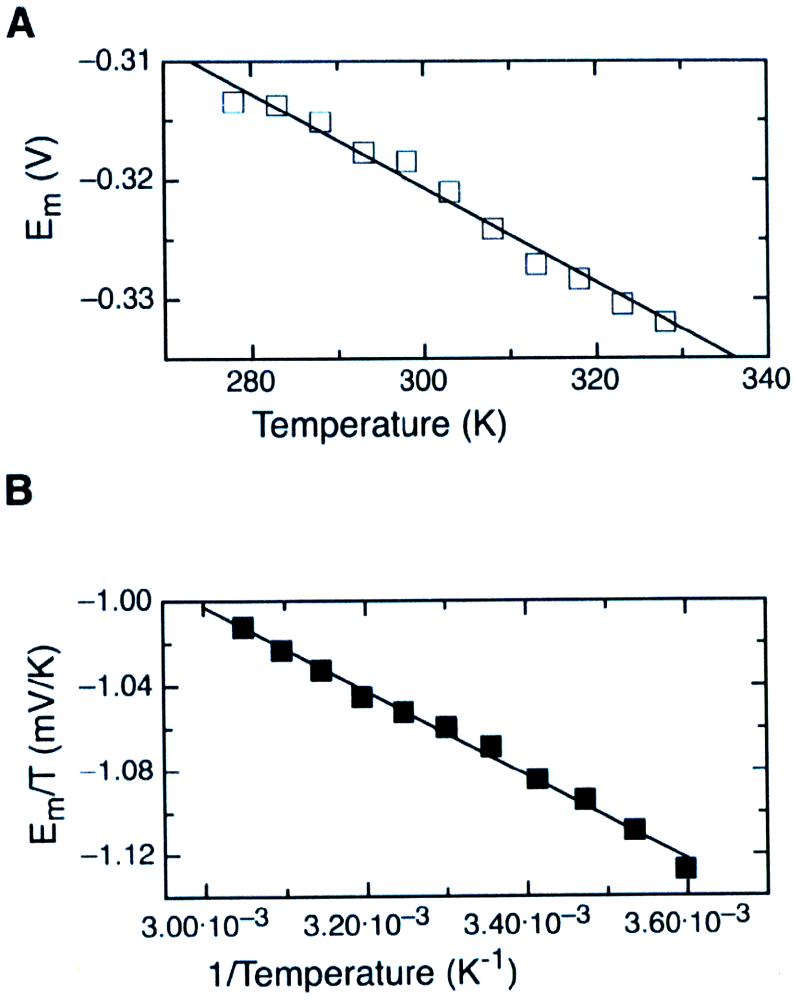


Figure A1.5. Demonstration that the PFV response of *ta*TrR is not due to flavin desorbed from the protein. (A) Addition of exogenous FAD to a *ta*TrR protein film. Build up of the voltammetric response of the FAD/FADH2 couple is observed by starting with a *ta*TrR-bound electrode (light solid line) and adding FAD to the electrochemical cell at final concentrations of 2, 4 and 10 μM (dotted, dashed, and heavy solid line). The build up of the free FAD species is observable at ~-220 mV. Other experimental features: pH 7.0, 10 oC, v = 150 mV/s. (B) Free FAD (•) and *ta*TrR (open square) midpoint potentials, as a function of pH.

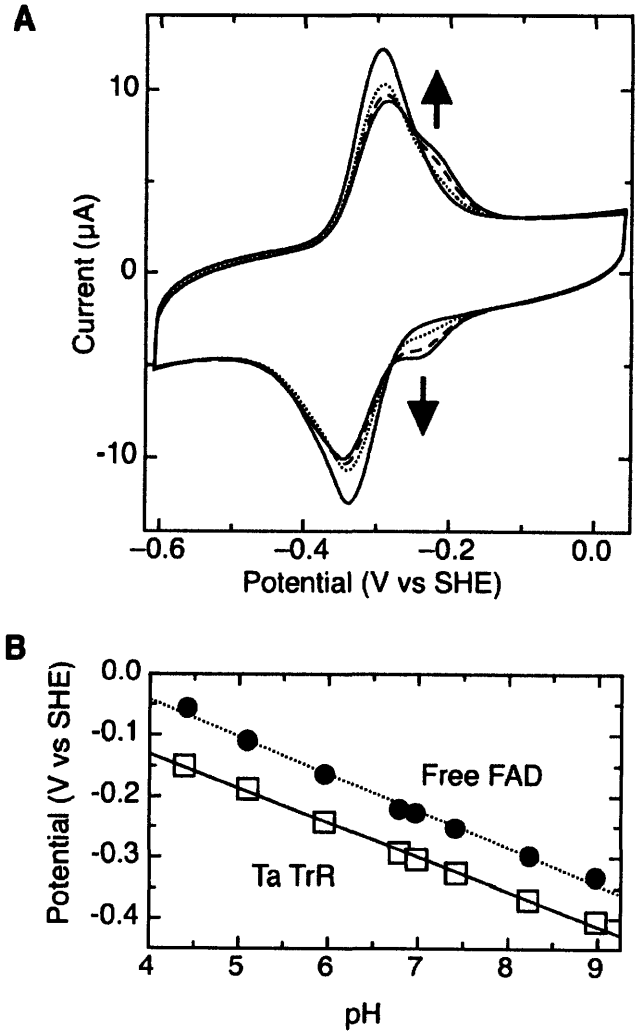


Figure A1.6. PFV response of *ec*TrR. (A) Raw PFV and baseline-subtracted voltammetric features for *ec*TrR; parameters for the experiment: pH 7.0, 10 oC, ν = 150 mV/s. (B) Average peak width at half height as a function of temperature, for the *ec*TrR component of the PFV response shown in (A).

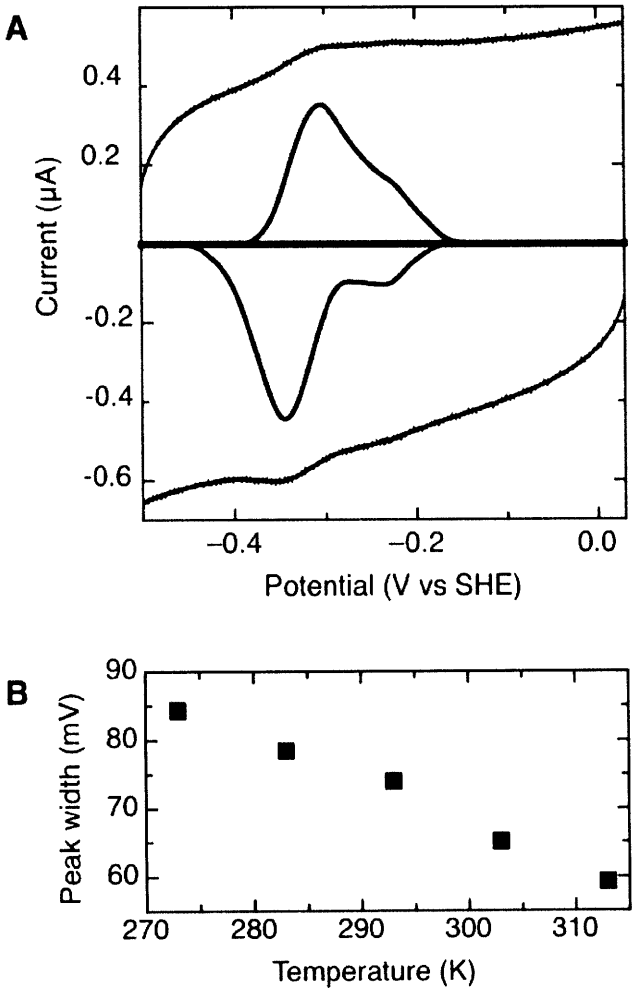


Figure A1.7. Redox titration of the *ta*TrR flavin using the xanthine/xanthine oxidase reduction methodology of Massey (*25*), showing the decrease in the visible absorption spectrum, as collected every ~5 minutes during the course of the reduction of the flavin cofactor. SafO was used as the reporter dye (maximum at 520 nm), as described in the materials and methods. Inset: Plot of log [Ox]/[Red] for SafO vs. log [Ox]/[Red] for TrR, used to find the midpoint potential with respect to the reference value of SafO (-289 mV).

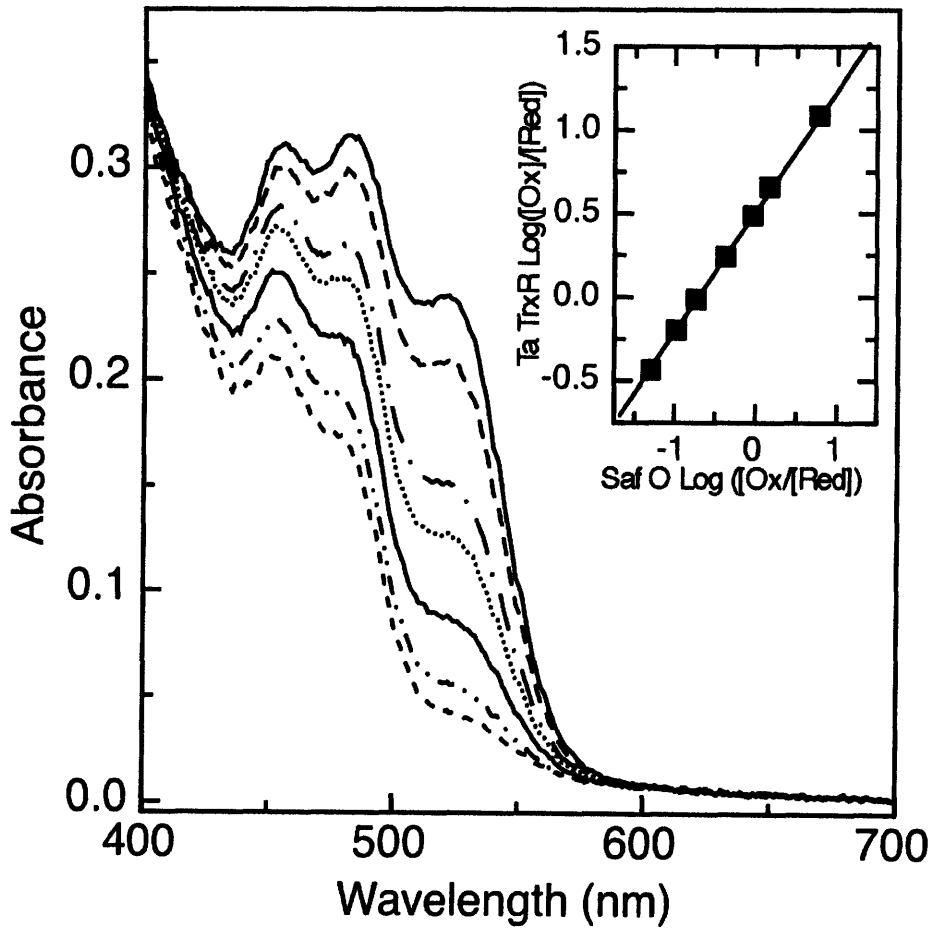


Figure A1.8. Impact of increasing scan-rate upon the *ta*TrR voltammetry at 308 K. Baseline substracted data are given to illustrate how at elevated temperatures, the scan rate of the experiment can be raised from 200 mV/s (dotted line) to 2 V/s (dashed) to 20 V/s (solid), which results in broadening due to the observation of FO and FR states.

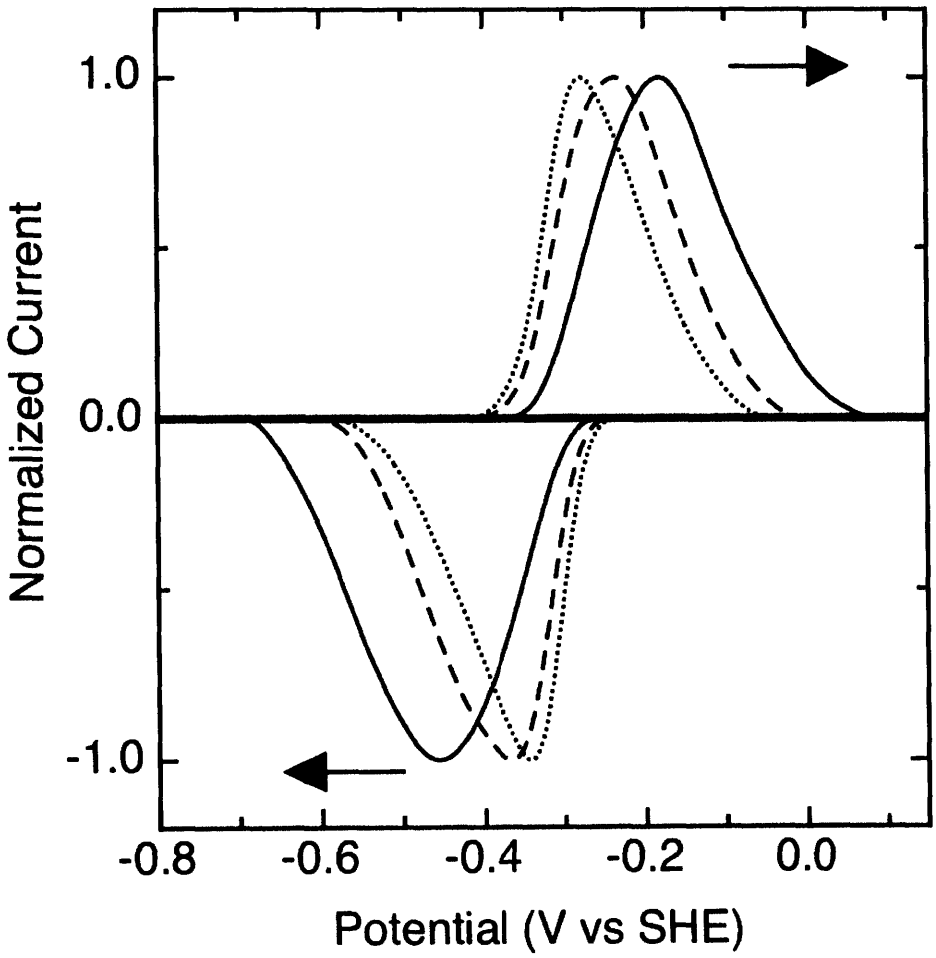


Figure A1.9. Fitting of the baseline-subtracted, oxidative half-scan of *ta*TrR to two distinct nernstian features, with an identical value of the electroactive coverage, at 333 K. For each feature the stoichiometric number of electrons was held to 2, while the separation between the *EI/R* and *EO/R* potentials was allowed to vary, as indicated by Equation 1. (A) The case where the two one-electron potentials are crossed, with the difference in potential being greater than 200 mV; (B) the difference in crossing of potentials is ~50 mV.

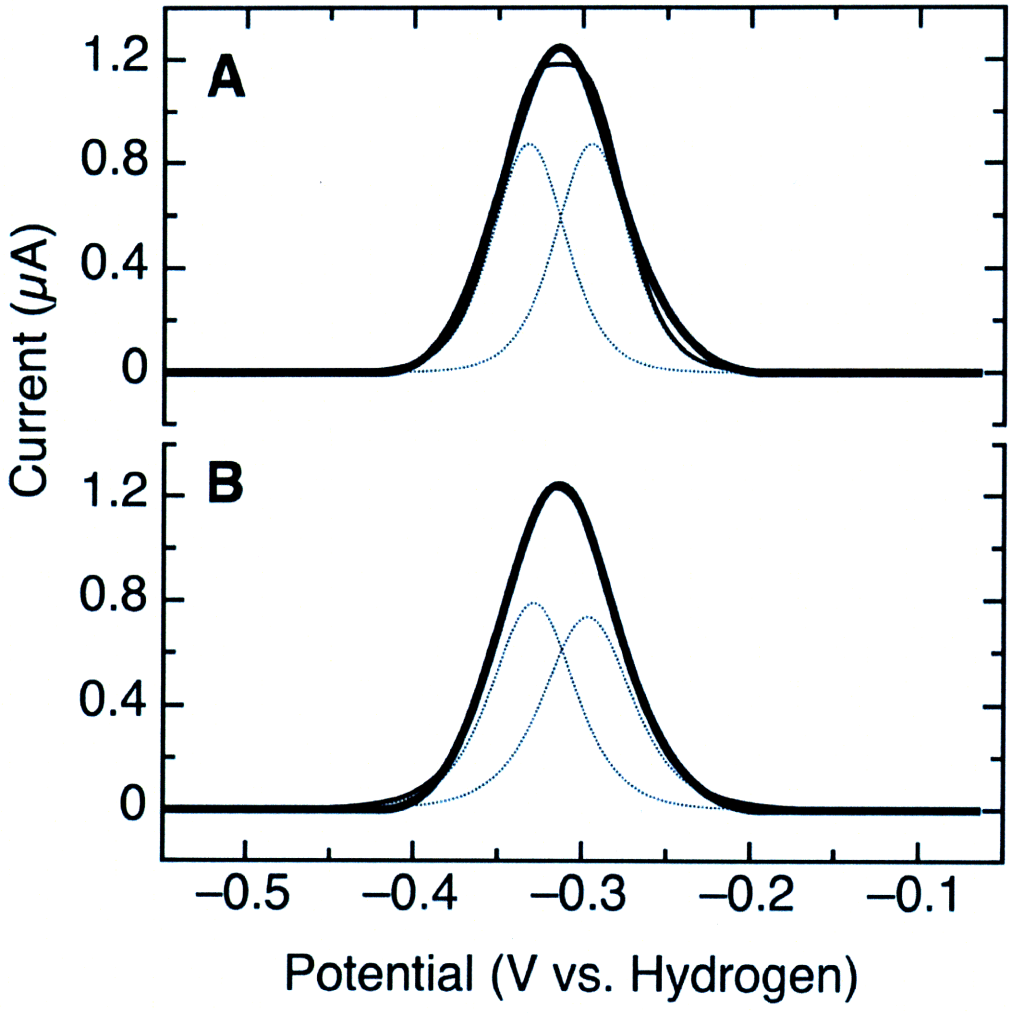
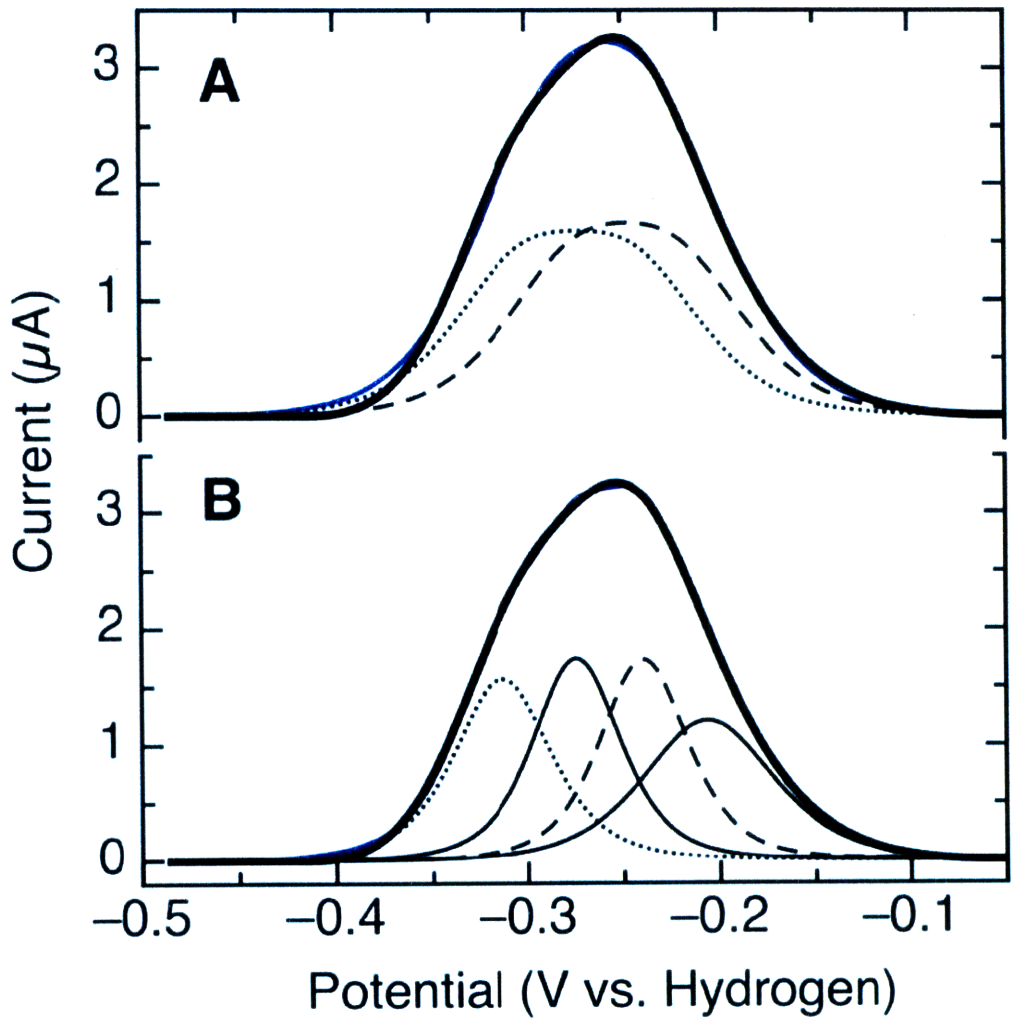
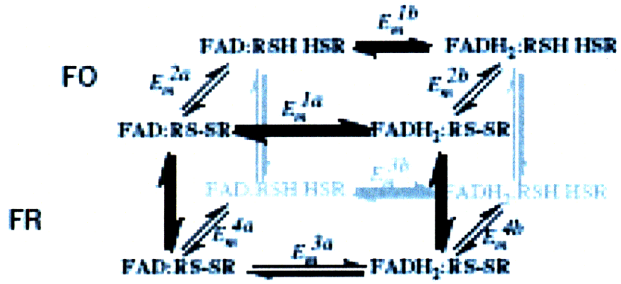
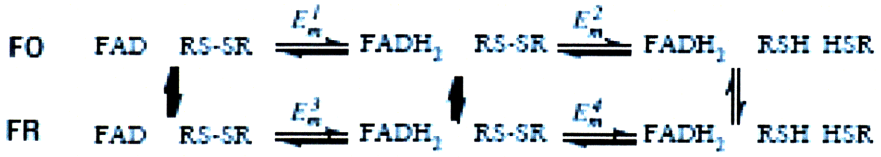


Figure A1.10. Fitting of the baseline-subtracted, oxidative half-scan of *ta*TrR to either two noncooperative (A) or four cooperative (B) n = 2e- nernstian features, assuming identical electroactive coverages, and a temperature of 283 K.



Scheme A1.1. (A) A model of microscopic potentials and the distinct FO and FR conformations for a low Mr TrR. (B) A simplified model in which individual microscopic potentials will not be addressable in the PFV experiment.





REFERNCES

1. Holmgren, A. (1985) Thioredoxin, *Ann. Rev. Biochem. 54*, 237-271.

2. Arner, E. S. J., and Holmgren, A. (2000) Physiological functions of thioredoxin and thioredoxin reductase, *Eur. J. Biochem. 267*, 6102-6109.

3. Moore, E. C., Reichard, P., and Thelander, L. (1964) Enzymatic synthesis of deoxyribonucleotiedes. V. Purification and properties of thioredoxin reductase from *Escherichia coli* B, *J. Biol. Chem. 239*, 3445-3452.

4. Russel, M., and Model, P. (1988) Sequence of thioredoxin reductase from *Escherichia coli* -Relationship to other flavoprotein disulfide oxidoreductases, *J. Biol. Chem. 263*, 9015-9019.

5. Zanetti, G., and Williams, C. H. (1967) Characterization of the active center of thioredoxin reductase, *J. Biol. Chem. 242*, 5232-5236.

6. Williams, C. H. (2000) Thioredoxin-thioredoxin reductase - a system that has come of age, *Eur. J. Biochem. 267*, 6101-6101.

7. Kuriyan, J., Wong, L., Russel, M., and Model, P. (1989) Crystallization And Preliminary-X-Ray Characterization Of Thioredoxin Reductase From *Escherichia coli*, *J. Biol. Chem. 264*, 12752- 12753.

8. Waksman, G., Krishna, T. S. R., Williams, C. H., and Kuriyan, J. (1994) Crystal-Structure Of Escherichia-Coli Thioredoxin Reductase Refined At 2-Angstrom Resolution - Implications For A Large Conformational Change During Catalysis, *J. Mol. Biol. 236*, 800-816.

9. Lennon, B. W., and Williams, C. H. (1997) Reductive half-reaction of thioredoxin reductase from *Escherichia coli*, *Biochemistry 36*, 9464-9477.

10. Mulrooney, S. B., and Williams, C. H. (1997) Evidence for two conformational states of thioredoxin reductase from *Escherichia coli*: Use of intrinsic and extrinsic quenchers of flavin fluorescence as probes to observe domain rotation, *Prot. Sci. 6*, 2188-2195.

11. Veine, D. M., Ohnishi, K., and Williams, C. H. (1998) Thioredoxin reductase from *Escherichia coli*: Evidence of restriction to a single conformation upon formation of a crosslink between engineered cysteines, *Prot. Sci. 7*, 369-375.

12. Wang, P. F., Veine, D. M., Ahn, S. H., and Williams, C. H. (1996) A stable mixed disulfide between thioredoxin reductase and its substrate, thioredoxin: Preparation and characterization, *Biochemistry 35*, 4812-4819.

13. Lennon, B. W., Williams, C. H., and Ludwig, M. L. (1999) Crystal structure of reduced thioredoxin reductase from *Escherichia coli*: Structural flexibility in the isoalloxazine ring of the flavin adenine dinucleotide cofactor, *Prot. Sci. 8*, 2366-2379.

14. Lennon, B. W., Williams, C. H., and Ludwig, M. L. (2000) Twists in catalysis: Alternating conformations of *Escherichia coli* thioredoxin reductase, *Science 289*, 1190-1194.

15. Lennon, B. W., and Williams, C. H. (1995) Effect Of Pyridine-Nucleotide On The Oxidative Half-Reaction Of *Escherichia coli* Thioredoxin Reductase, *Biochemistry 34*, 3670-3677.

16. Lennon, B. W., and Williams, C. H. (1996) Enzyme-monitored turnover of *Escherichia coli* thioredoxin reductase: Insights for catalysis, *Biochemistry 35*, 4704-4712.

17. O'Donnell, M. E., and Williams, C. H. (1983) Proton Stoichiometry In The Reduction Of The Fad And Disulfide Of *Escherichia coli* Thioredoxin Reductase - Evidence For A Base At The Active-Site, *J. Biol. Chem. 258*, 3795-3805.

26

18. Van den Berg, P. A. W., Mulrooney, S. B., Gobets, B., Van Stokkum, I. H. M., Van Hoek, A., Williams, C. H., and Visser, A. (2001) Exploring the conformational equilibrium of *E. coli* thioredoxin reductase: Characterization of two catalytically important states by ultrafast flavin fluorescence spectroscopy, *Prot. Sci. 10*, 2037-2049.

19. Hernandez, H. H., Jaquez, O. A., Hamill, M. J., Elliott, S. J., and Drennan, C. L. Thioredoxin reductase from *T. acidophilum:* a new twist on redox regulation, *submitted*

20. Chen, K. S., Bonagura, C. A., Tilley, G. J., McEvoy, J. P., Jung, Y. S., Armstrong, F. A., Stout, C. D., and Burgess, B. K. (2002) Crystal structures of ferredoxin variants exhibiting large changes in [Fe-S] reduction potential, *Nature Struct. Biol. 9*, 188-192.

21. Chen, K. S., Hirst, J., Camba, R., Bonagura, C. A., Stout, C. D., Burgess, B. K., and Armstrong, F. A. (2000) Atomically defined mechanism for proton transfer to a buried redox centre in a protein (vol 405, pg 814, 2000), *Nature 407*, 110-110.

22. Fleming, B. D., Tian, Y., Bell, S. G., Wong, L. L., Urlacher, V., and Hill, H. A. O. (2003) Redox properties of cytochrome P450(BM3) measured by direct methods, *Eur. J. Biochem. 270*, 4082-4088.

23. Heering, H. A., and Hagen, W. R. (1996) Complex electrochemistry of flavodoxin at carbon-based electrodes: Results from a combination of direct electron transfer, flavin-mediated electron transfer and comproportionation, *J. Electroanal. Chem. 404*, 249-260.

24. Mulrooney, S. B. (1997) Application of a single-plasmid vector for mutagenesis and high-level expression of thioredoxin reductase and its use to examine flavin cofactor incorporation, *Prot. Expr. Purif. 9*, 372-378.

25. Massey, V. (1991) in *Flavins and Flavoproteins 1990* (Curti, B., Ronchi, S. and Zanetti, G., Eds.) pp 59-66, Walter de Gruyter & Co, Berlin.

26. Cuello, A. O., McIntosh, C. M., and Rotello, V. M. (2000) Model systems for flavoenzyme activity. The role of N(3)-H hydrogen bonding in flavin redox processes, *J. Am. Chem. Soc. 122*, 3517-3521.

27. Lindgren, A., Larsson, T., Ruzgas, T., and Gorton, L. (2000) Direct electron transfer between the heme of cellobiose dehydrogenase and thiol modified gold electrodes, *J. Electroanal. Chem. 494*, 105-113.

28. McEvoy, J. P., and Armstrong, F. A. (1999) Protein film cryovoltammetry: demonstrations with a 7Fe ([3Fe-4S]+[4Fe-4S]) ferredoxin, *Chem. Comm.* 1635-1636.

29. Plichon, V., and Laviron, E. (1976) Theoretical study of a two-step reversible electrochemical reaction associated with irreversible chemical reactions in thin layer linear potential sweep voltammetry, *J. Electroanal. Chem. 71*, 143-156.

30. Savéant, J. M. (2006) Elements of molecular and biomolecular electrochemistry, John Wiley & Sons, Hoboken, N.J.

Appendix 2:

Substantial Energetic Improvement with Minimal Structural Perturbation in a High Affinity Mutant Antibody

K.S. Midelfort, H.H. Hernandez, S.M. Lippow, B. Tidor, C.L. Drennan, and K.D. Wittrup

- APPENDIX 2 – Substantial Improvement with Minimal Structural Perturbations in a High Affinity Mutant Antibody

INTRODUCTION

The mechanisms differentiating very high affinity protein interactions (sub-picomolar dissociation constants) from high affinity (nanomolar dissociation constants) have seldom been possible to study (1, 2), because few protein interaction families that span this range of affinity have been identified from either nature or engineering efforts. Antibodies provide a natural affinity series of related proteins but, *in vivo*, rarely surpass a 0.1 nanomolar dissociation constant (K_d) due to physiological limits on the selection (3). Further affinity improvements are often desired for pharmaceutical and biotechnology uses, and can be attained through directed evolution protein engineering techniques (4, 5, 6, 7). Computational methods have also provided some limited predictions of specific changes needed to improve high affinity interactions (8), although they have yet to realize substantial improvements experimentally. To date, the transitions from micromolar to nanomolar (9, 10, 11, 12, 13, 14) and from femtomolar back to picomolar (2, 15) affinity have been studied, providing little guidance into the mechanistic means to engineer an enhancement from nanomolar to femtomolar affinity.

Previous studies of antibody maturation from weak binding (micromolar) to high affinity binding (nanomolar) have shown that a number of factors are involved in improving affinity across this range. Several studies have noted that decreased loss of entropy on binding, with creation of a lock and key fit mechanism, contribute to affinity improvements (11, 14, 16, 17). In a structural, thermodynamic, and computational energetic analysis of an esterolytic antibody bound to a *p*-nitrophenyl phosphonate hapten, affinity maturation occurred through a reorganization of the binding site geometry so as to optimize gaining favorable electrostatic interactions with the hapten and losing those with solvent during the binding process (9, 18). In a study of an anti-testosterone antibody affinity maturation, improvement was derived from small structural changes providing more comprehensive packing around the antigen (12). A study of an anti-(4-hydroxy-3-nitrophenyl) acetate affinity matured antibody suggests that "relief of cramped contacts" in the wild-type provided the increased affinity in the mutant (19). In the first structural mechanistic study of affinity matured antibodies to a protein antigen,

the higher affinity was attributed to an increased burial of hydrophobic surface area on binding and improved shape complementarity between antigen and antibody (13). Across these affinity maturation studies, mutations were found in antigen contact and non-contact regions; however, it was not always clear whether all the cited mutations were contributing to the improvements. It is perhaps to be expected that insights derived from studies of micromolar to nanomolar affinity improvements may not translate directly to mechanisms involved in fine-tuning an interaction from nanomolar to femtomolar affinity levels.

The monoclonal antibody 4-4-20 directed against the hapten fluorescein was originally isolated to facilitate studies of immunological recognition (20). The 4-4-20 binding has been extensively studied as the whole IgG, the Fab, the Fv, and the scFv by thermodynamic, kinetic, structural, computational, spectroscopic, and mutational methods (21, 22, 23, 24, 25, 26, 27, 28). A recent comparative structural study compared the 4-4-20 Fab structure with an idiotypically related lower affinity Fab, 9-40 (29). This study revealed a reorganization of the binding site and a more closed structure with fewer water molecules at the binding pocket for 4-4-20, compared with 9-40, to account for the change in binding affinity.

Ten femtomolar affinity mutants of 4-4-20 were previously identified through directed evolution with random mutagenesis over the whole Fv followed by screening with yeast surface display (7). Clone 4M5.3 from that study was chosen for further study here. This clone contains 14 mutations, including all ten of the consensus mutations observed in the final directed evolution round clones.

Here, we compare 4M5.3 to 4-4-20 by kinetic, thermodynamic, structural, and theoretical analyses. No individual structural change was unambiguously associated with the −4.5 kcal/mol improvement in binding free energy. However, these results implicate a sum of many small structural changes in combination with specific mutations that improve electrostatics, which together result in a large binding free energy change.

MATERIALS AND METHODS

Over-expression and purification of scFv proteins

The 4-4-20 and 4M5.3 scFvs (7) were subcloned into the pRS316 backbone with the Gal1-10 promoter (40). A Flag tag (amino acid sequence Asp-Tyr-Lys-Asp-Asp-Asp-Asp-Lys) was inserted between the EagI and NheI sites N-terminal to the scFv sequence and the first expressed amino acid was changed to alanine by site-directed mutagenesis (Stratagene, La Jolla, CA). The sequence expressed was Ala-Ala-Arg-Pro-(Flag-tag)-(scFv). The scFvs were expressed solubly in *Saccharomyces cerevisiae* strain YVH10, which has an integrated PDI gene, using the growth conditions of SD (2% (w/v) dextrose, 0.67% (w/v) yeast nitrogen base, 50 mM sodium phosphate buffered to pH 6.6) containing 2× SCAA amino acid supplement (lacking Trp, Leu, and Ura, as described (40)) with one liter of culture in a 2.5 l fully baffled Tunair flask (Shelton Scientific, Shelton, CT) at 30 °C, 250 rpm shaker for 32–48 hours. For induction, the cells were centrifuged to remove the supernatant, and resuspended in YPG medium (1% (w/v) Bactoyeast extract, 2% (w/v) Bactopeptone, 2% (w/v) galactose); one liter of culture in a 2 l glass Erlenmeyer flask. The flasks were placed in a 20 °C, 250 rpm shaker for 48–60 hours. The supernatant was then separated by centrifugation and filter sterilized. The supernatant was concentrated to ~100 ml with a PREP/SCALE TFF 10k MWCO concentration unit (Millipore, Bedford, MA), and the buffer was exchanged to Tris-buffered saline (TBS: 50 mM phosphate buffer (pH 8), 150 mM NaCl), with 500–750 ml of TBS in four batches. The final volume was ~100 ml. The protein was purified with M2 anti-FLAG resin (Sigma, St. Louis, MO) by batch method. Final protein was identified as a single band by Coommassie blue staining of an SDS-PAGE gel and purified yields were ~5 mg and ~10 mg per one liter initial culture, for 4-4-20 and 4M5.3, respectively. Mass spectrometry was performed to confirm the desired sequence.

Equilibrium K_d of scFv

The concentrations of the scFv were determined by fluorescence quench titrations, as described using the quenching constant (Q) of 0.94 and 0.95 for 4-4-20 and 4M5.3, respectively (7). The soluble 4-4-20 equilibrium binding constant, K_d, was also determined by fluorescence quench titration, as described (7). The equilibrium K_d of

4M5.3 displayed on the surface of yeast was determined by a competition assay with soluble 4-4-20. The 4M5.3 was displayed on the surface of yeast as reported (7). Labeling reactions contained 2×10^6 4M5.3 displaying yeast ($\sim$0.2 nM 4M5.3 scFv), 1.7 nM fluorescein-biotin (Molecular Probes, Eugene, OR) and varying concentrations of soluble 4-4-20 (0.01–40 μM). The experiments were either mixed with the 4M5.3 displaying cells and fluorescein-biotin first, allowed to incubate for one hour at 25 °C, and then the 4-4-20 competitor was added, or the competitor and fluorescein-biotin were allowed to incubate together for one hour at 25 °C, and then the 4M5.3 displaying cells were added. All experiments contained a final concentration of 1×Pen-Strep. The tubes were then incubated at 25 °C for 15 days, with daily mixing of the samples. Controls were also performed as above with either no competitor protein or with 10 μM soluble 4M5.3 as the competitor. Fluorescein-biotin binding to the cells was checked by secondary labeling with streptavidin-Phycoerythrin (Pharmingen, BD Biosciences, San Diego, CA), with analysis by flow cytometry (XL cytometer, Beckman Coulter, Fullerton, CA).

Isothermal titration calorimetry

The enthalpy of binding was determined with both 4-4-20 and 4M5.3 scFv using a MicroCal isothermal titration calorimetry (ITC) instrument (VP-ITC, MicroCal Inc., Northampton, MA). The scFvs were at 1–20 μM concentrations in TBS in the cell, while the fluorescein-biotin was at 10–200 μM concentrations in TBS in the syringe being titrated. The final fully bound titration injections were used as the heats of dilution standards (41). The enthalpy of binding was determined from a fit to the resulting curves. The same enthalpy of binding results could also be obtained by subtracting the heats of dilution of the fully saturated titrations from the heats of binding of the unsaturated scFv. K_d values were not obtainable due to the steepness of the curves.

Differential scanning calorimetry

The melting temperatures of the scFvs in their bound and unbound forms in TBS were determined with a differential scanning calorimeter (VP-DSC, MicroCal Inc.,

Northampton, MA). Each 0.5 ml sample (~17 μM (~0.5 mg/ml) scFv or ~17 μM scFv+~30 μM fluorescein) was placed in the cell and the instrument was heated at 1 deg. C/minute from 20 °C to 90 °C. A reference trace of buffer without protein, collected as above, was subtracted from the protein sample traces.

Association rates

Association rates were determined by stopped-flow fluorescence measurements. Using a fluorimeter (Varian Eclipse Instrument, Varian Inc., Mulgrave, Victoria, Australia) connected to a manual stopped-flow instrument (RX-200 Rapid Kinetics Accessory from Applied Photophysics, Leatherhead, Surrey, United Kingdom), 10 nM fluorescein-biotin in TBS (pH 8), was mixed in a 1 : 1 (v/v) ratio at 25 °C with 10–60 nM scFv (final concentrations: 5 nM fluorescein biotin, 5–30 nM scFv), and data were collected on the millisecond time-scale for up to 240 seconds depending on the rate of the binding reaction. The quenching data were fit with a single association rate, using the dissociation rates determined previously, with the program Gepasi (42).

Crystallization and data collection

4M5.3 was concentrated to 5 mg/ml in a solution of 200 μM fluorescein (Aldrich, St. Louis, MO), 10 mM Tris–HCl (pH 8.0). Standard hanging-drop, vapor-diffusion methods were used to crystallize 4M5.3, at 4 °C in the dark, over a reservoir solution of 0.15 M sodium acetate (pH 4.6), 0.1 M ammonium acetate, 20% (w/v) PEG 4000, and 10% (v/v) glycerol (Hampton Scientific, Chalfont, PA). Crystals grown from these conditions appeared in four to seven days and all had an intense yellow color, presumably due to bound fluorescein. Prior to data collection, crystals were transferred to reservoir solution and flash-cooled in liquid nitrogen. The crystals belong to the space group $P2_12_12_1$ with unit cell dimensions a=55.78 Å, b=57.45 Å, and c=79.22 Å, $\alpha=\beta=\gamma=90°$, and contain two molecules of 4M5.3 per asymmetric unit. Data were collected to 1.5 Å resolution at the Stanford Synchrotron Radiation Laboratory (SSRL) beam line 9.11 and reduced using DENZO and SCALEPACK (43). See Table A2.1 for data collection statistics.

Structure determination and refinement

The structure of 4M5.3 was determined by molecular replacement with EPMR (44) using a polyalanine-substituted model derived from 4-4-20 ($IgG2a_{\kappa}$) with the fluorescein ligand bound (entry 1FLR in the Protein Data Bank) (24). Initial refinement using a simulated annealing method was carried out using CNS (45). Final refinement using positional and individual *B*-factor refinement was carried out using REFMAC (43). Anisotropic *B*-factor refinement was used in the final stages of refinement. Manual protein model rebuilding was done between rounds of refinement in XtalView (44, 45). Of 268 residues per monomer, residues 14–268 were observed in electron density maps, with one chain break per molecule (from 124–149, which corresponds to the linker region). Residues 15–124 correspond to V_L 1–111 and residues 151–269 correspond to V_H 1–117. For the structures, most (99%) chain angles fall into allowed regions of the Ramachandran plot with only one residue (L57) disallowed. This agrees with previous structures in which the corresponding residue (L56) is disallowed. The average *B*-factor for the structure is 14.26 $Å^2$. See Table A2.1 for refinement statistics. Numbering used here will correlate with the 4-4-20 Fab (1FLR) structure numbering. Shape complementarity and solvent-accessible surface area calculations were performed with the Sc and Areaimol programs in CCP4, respectively (31, 46). The solvent probe radius for Areaimol was 1.4 Å.

Computational structure preparations

The 1.85 Å 4-4-20 Fab structure (1FLR) (24) and the solved 1.5 Å structure of 4M5.3 (this work) were used as the basis for molecular modeling and continuum electrostatic calculations and theoretical analysis. The two structures were prepared in parallel to minimize differences that might bias comparative analysis. Only residues with corresponding crystallographic data in both structures were used, prompting the removal of the 4-4-20 C_L and C_{H1} domains (residues after L112 and after H117) and a few N and C-terminal residues corresponding to the flexible linker (4-4-20 V_L112 and 4M5.3 V_L0, V_H-2, V_H-1, V_H0), all of which are solvent-exposed and at least 15 Å from the fluorescein-binding pocket. Six water molecules were retained in each structure (4-4-20:

606, 608, 615, 618, 676, 689; 4M5.3: 6, 3, 1, 11, 26, 47); these six solvent molecules make corresponding interactions in the two structures. All other water molecules were removed from the structure files and were modeled implicitly. Side-chain titration states, single conformations for all multiple-occupancy residues, and the crystallographic carbon/nitrogen/oxygen uncertainties in asparagine, glutamine, and histidine side-chains were resolved based on examination of side-chain local environments. The fluorescein was modeled in its net charge −2 state reflecting the pH of 8 in the binding experiments, and this choice is supported by the local environments of the fluorescein protonatable sites in each structure.

Molecular mechanics was used to prepare the protein structures. The HBUILD facility (47) in the program CHARMM (48) was used to build all hydrogen atoms onto each structure. The CHARMm22 all-atom parameter set (49) with the CHARMM-adapted TIP3P water model was used for molecular mechanics calculations, assigning appropriate general atom types for the fluorescein. Fluorescein partial atomic charges were obtained as described (50) by first using the program GAUSSIAN98† with restricted Hartree–Fock and the 6-31G* basis set to optimize the geometry of fluorescein starting from that in the 4-4-20 structure, and subsequently fitting the electrostatic potential using the restrained-fitting methods (RESP) described (51).

A second set of 4-4-20 and 4M5.3 structures for comparison were prepared by subjecting each to a harmonically restrained minimization to convergence with 10 kcal/mol per $Å^2$ force constants on every atom in the system, no non-bonded cut-offs, and a $4r$ distance-dependent dielectric constant. The root-mean-square deviation for structures before and after minimization were 0.093 Å (0.094 Å) and 0.053 Å (0.053 Å) for 4-4-20 and 4M5.3, respectively, taken over all atoms (taken over side-chain atoms only).

Electrostatic calculations

Electrostatic contributions of individual functional groups to the binding free energy were computed using a rigid binding model with continuum solvent following

previous work (52). A locally modified version of the DELPHI program (53, 54, 55) was used to solve the linearized Poisson–Boltzmann equation with finite-difference methods. PARSE parameters (56) were used for atomic radii and partial atomic charges of protein. Fluorescein partial atomic charges were obtained from the RESP fitting procedure described above, and radii were assigned based on the PARSE convention. PARSE parameters do not include aliphatic hydrogen atoms and thus the appropriate hydrogen atoms were removed from the system. Fluorescein does not contain any aliphatic hydrogen atoms. The dielectric constant was assigned to 4 for protein, ligand, and explicit water, and 80 for the implicit solvent regions. A salt concentration of 0.145 M was used with a 2.0 Å Stern layer and a molecular surface generated with a 1.4 Å probe sphere. Each molecule was oriented to minimize the volume of the bounding cube. A focusing procedure was used that includes a low grid spacing using 23% fill and Debye–Hückel boundary conditions followed sequentially by higher resolution calculations first at 92% and then at 184% fill centered on the specific functional group of interest. Ten translations relative to the grid were performed and averages were used. The standard error of the mean was of the order of 0.001 kcal/mol, much less than the 0.1–1.0 kcal/mol range of differences identified between 4-4-20 and 4M5.3 components. A 129×129×129 grid was used, resulting in final grid spacings of 4.49 and 4.40 grid units per Å for 4-4-20 and 4M5.3, respectively. Convergence of free energies with respect to grid resolution (data not shown) indicates that a difference of 0.09 grid unit per Å results in about 0.01 kcal/mol change in net binding free energy. Therefore, the comparison between 4-4-20 and 4M5.3 electrostatic calculations on slightly different grids remains valid. For binding calculations, all explicit water molecules remained associated with the antibody in the unbound state. Overall electrostatic contributions to binding were dissected into component contributions from individual chemical groups and physical sources based on the work of Hendsch & Tidor (53).

RESULTS

Kinetics and thermodynamics

The fluorescein-biotin equilibrium binding constant for 4-4-20 in both the soluble form and displayed on the surface of yeast has been determined previously, with a K_d of 0.7 nM at 25 °C for both (7). The 4M5.3 binding constant with fluorescein-biotin was previously determined to be 270 fM through the ratio of dissociation and association kinetic rates (7). To test for the presence of kinetic intermediates that would cause the ratio of the measured association and dissociation rates to differ from the equilibrium binding constant, the 4M5.3 equilibrium binding constant was determined through equilibrium competition titration. A competition assay was developed with 4M5.3 displayed on the surface of yeast, with varying concentrations of soluble 4-4-20 scFv used to compete for binding to fluorescein-biotin. By this method, the equilibrium binding constant was found to be 0.27(±0.05) pM at 25 °C (Figure A2.1), in agreement with the previous kinetic ratio-determined K_d of 0.27 pM (7). Thus, the equivalence $K_d=k_{off}/k_{on}$ is consistent with a two-state transition from the unbound to the bound states, and is not consistent with the presence of any substantially populated kinetic intermediates. The ΔG of binding for 4-4-20 and 4M5.3 calculated from these K_d values is −12.4(±0.1) kcal/mol and −16.9(±0.1)kcal/mol, respectively, for a $\Delta\Delta G$=−4.5(±0.1)kcal/mol. Multi-step binding mechanisms such as an encounter complex intermediate are not consistent with the observed equivalence between K_d and k_{off}/k_{on}.

In order to investigate the fast association rates found previously by manual mixing fluorescence experiments for the presence of encounter complexes, conformational selection, or induced fit, (7) stopped-flow fluorescence binding kinetics were used to study the association rates of the 4-4-20 and 4M5.3 scFvs with antigen in the 0.5 to 240 second range. Non-linear regression of the fluorescein-biotin quenching curves was consistent with single, concentration-independent association rate constants for fluorescein-biotin binding by 4-4-20 or 4M5.3 (Figure A2.2). The association rates for 4-4-20 and 4M5.3 were 3.3(±0.3)×10^7 M^{-1} s^{-1} and 3.8(±0.3)×10^6 M^{-1} s^{-1} for 4-4-20 and 4M5.3, respectively. These rates were quantitatively consistent with those previously reported for soluble 4-4-20 and 4M5.3 binding (7), and are also in agreement with the association kinetics for 4-4-20 determined on the surface of yeast (data not shown). The kinetic traces shown in Figure A2.2 are well fit by a single rate constant, and therefore

provide no evidence for substantial population of an encounter complex, or the existence of induced fit or conformational selection. This is consistent with the equivalence between the equilibrium affinity constant and the ratio of kinetic rate constants.

In order to parse the binding free energy into enthalpic and entropic components, isothermal titration calorimetry (ITC) was used to measure binding enthalpy. The change of enthalpy on binding fluorescein-biotin ($\Delta H_{binding}$) for 4-4-20 and 4M5.3 were found to be −10.2(±0.3) kcal/mol and −14.2(±0.1) kcal/mol at 25 °C, respectively (Figure A2.3). The binding interactions for both 4-4-20 and 4M5.3 were enthalpically and entropically favorable at 25 °C. The $\Delta\Delta H_{binding}$ between 4-4-20 and 4M5.3 was −4.0(±0.3) kcal/mol of the total −4.5(±0.1) kcal/mol of the $\Delta\Delta G_{binding}$. The binding free energy change is, therefore, primarily enthalpic with only ~10% attributable to the entropic component of binding ($\Delta(-T\Delta S)$). This largely enthalpic change in 4M5.3 fluorescein-binding energetics could suggest improved 4M5.3/fluorescein contacts *via* electrostatics, hydrogen bonds, and/or van der Waals interactions.

Differential scanning calorimetry was performed on the 4-4-20 and 4M5.3 scFvs with and without ligand. The higher affinity 4M5.3 scFv had a slightly lower melting temperature of 49.2(±0.1) °C without ligand present, compared to a 53.0(±0.1) °C melting temperature for the 4-4-20 without ligand (data not shown). The fluorescein-bound scFvs both possessed higher melting temperatures but retained the pattern of the high affinity clone possessing the lower melting temperature; 64.8(±0.4) °C for the 4M5.3 bound complex compared to 65.8(±0.3) °C for the 4-4-20 bound complex melting temperature. It would be formally possible to convert these changes in T_m to ΔG of binding; however, this would require reversibility of the traces and the evaluation of ΔC_p for the binding reaction as well as for at least one of the unfolding reactions, since the melting temperatures of the complexes also occur at different temperatures. Determining ΔC_p by performing ITC at various temperatures is a fairly arduous procedure complicated somewhat by the narrow temperature range of folding stability for these scFvs.

ScFv structural comparison

- APPENDIX 2 – Substantial Improvement with Minimal Structural Perturbations in a High Affinity Mutant Antibody

The 4M5.3 scFv was crystallized in complex with fluorescein and the structure was solved by molecular replacement using the previously solved 1.85 Å resolution Fv portion of the 4-4-20 Fab fluorescein-bound structure (24). The final model had an R value of 16.48 and an R_{free} of 19.29 solved to 1.5 Å resolution (Table A2.1). The 4M5.3 structure contained 231 amino acid residues, including all of the light and heavy chain residues, but excluding most of the linker between the light and heavy chains, and the Flag epitope tag peptide at the N-terminal end. When the structures of the 4-4-20 Fv domains and the 4M5.3 scFv structure were aligned by their backbone atoms, the RMS difference was 0.60 Å, and with alignment by only the C^{α} atoms, the RMSD was 0.57 Å. No large differences were seen in alignment of 4-4-20 Fab and 4M5.3 scFv structures (Figure A2.4A); however, on the side opposite the binding site there is evidence of a heavy chain loop (residues H40–H44) packing differently in 4M5.3 than in 4-4-20. This is most likely due to the removal of the constant portion of the Fab in the 4M5.3 structure. The variable light (heavy) chain domains aligned with a backbone RMSD of 0.46(0.92) Å, which indicates greater rearrangement of variable heavy chain structure (the heavy chain contains a greater share of the mutations).

The 14 mutations in 4M5.3 are distributed throughout the scFv structure with eight mutations located in the complementarity determining regions (CDRs) and six in the framework regions (Figure A2.4B). Two of the mutations are in the light chain while 12 are in the heavy chain. Three of the mutations (AspH31His, SerH101Ala, and TyrH102Ser) are within 5 Å (first contact shell) of fluorescein, four mutations (AspH1Gly, SerH30Gly, IleH51Phe, and AspH106Glu) are in the second contact shell, and an additional three mutations (PheL60Val, AlaH24Thr, and TrpH108Leu) are in the third contact shell surrounding the fluorescein. The final four mutations are well outside of these ligand contact shells (SerL81Asn, ArgH16Gly, ProH17Ala, and MetH93Thr) and are all surface-exposed residues.

The interaction of fluorescein in the binding pocket of 4-4-20 has been previously observed in two crystal structures of the 4-4-20 Fab complex (4FAB and 1FLR) (24, 30). These show the binding pocket to be a relatively deep slot formed by a network of tryptophan and tyrosine side-chains, with the fluorescein participating in a network of

electrostatic interactions. The planar xanthonyl ring of the hapten is accommodated at the bottom of the slot while the phenylcarboxyl group is exposed to the solvent.

The binding site remains very similar in the 4M5.3 structure compared to the 4-4-20 Fab (Figure A2.5). The 4-4-20 fluorescein-binding slot is surrounded by four tyrosine residues (L37, H56, H102 and H103) and two tryptophan residues (L101 and H33) (Figure A2.5A). Of these aromatic residues, only H102 was mutated in 4M5.3 (to a serine). There is also an extensive hydrogen-bonding network in the binding pocket. The five hydrogen bonds formed with fluorescein in the 4-4-20 Fab remain in the 4M5.3 structure (Figure A2.5B). Four of these polar interactions become slightly closer, but one is lengthened in the 4M5.3 structure (Figure A2.5C). The very small differences in hydrogen-bonding distances (one or two tenths of an ångström unit) could very well be insignificant, given the different sources and space groups of the two structures.

From the structural alignment there appears to be a slight movement of the 4M5.3 residues toward the ligand. In order to quantify this, we tabulated the differences between the distance from the closest atom of each CDR residue to the closest atom of fluorescein in 4M5.3 compared to 4-4-20 (Figure A2.6). By this measure the majority of CDR residues are closer to the fluorescein ligand in the 4M5.3 structure. However, most of these changes are small, and individual refinement parameter differences or crystallization conditions between the two structures may account for some of the differences observed. All 11 residues that are greater than 0.5 Å closer to the fluorescein (four in V_L and seven in V_H) are at positions that are not within 5 Å contact. Three of the four residues that moved greater than 1 Å (V_L61Ser, V_H65Ser, and V_H67Lys) are surface-exposed side-chains. The fourth residue is a site of mutation, PheV_L60Val, and the new valine beta-branch adds a methyl group for the closer approach to the fluorescein, although the distance only reaches about 12 Å. Six of the CDR residues in 4M5.3 contained mutations from the 4-4-20 structure and of these locations only the V_L60 position moved more than 0.5 Å closer to the ligand. Furthermore, the mutated residues did not always move toward the ligand. Notably, the mutated residue SerV_H102Tyr moved further from the ligand in 4M5.3 by almost 0.5 Å. Nevertheless, the general trend

of residues within 5 Å contact of the fluorescein is a slight movement of 4M5.3 toward the fluorescein.

Another measure of the close packing can be obtained from the relative change in solvent-accessible surface area upon binding. The measure of the change in solvent-accessible surface area (ΔSAS) between the structures with and without fluorescein (rigid binding) yields 24 $Å^2$ less solvent-accessible surface area lost in the 4M5.3 structure compared with the 4-4-20 structure (Table A2.2). The change is slightly more localized to the V_L domain interface with the fluorescein, with a contribution of 15 $Å^2$, compared with a contribution of 9 $Å^2$ for the V_H domain interface with fluorescein. This reduced loss of solvent-accessible surface area for 4M5.3 is due, in net terms, to polar residues (25 $Å^2$), with 1 $Å^2$ gained from non-polar residues. Essentially, the 4-4-20 complex buries more polar residue surface area than the higher affinity 4M5.3 complex, which suggests that some form of electrostatic undercompensation may be operating in the 4-4-20 complex.

Shape complementarity (*Sc*) is a measure of the geometric shape matching between two molecular surfaces (31). Values range from 0 to 1.0, where a value of 1.0 describes an interface with perfect geometric fit. The shape complementarity was calculated for the crystal structures of both 4-4-20 and 4M5.3 with fluorescein (Table A2.2) and indicates negligible change in the overall complementarity between the scFvs and fluorescein. The shape complementarity of V_L with fluorescein remains the same between 4-4-20 and 4M5.3 while the shape complementarity of V_H with fluorescein is slightly less in 4M5.3. However, when the shape complementarity measure is applied to the interface between the V_H and V_L domains, there is a larger difference between 4-4-20 and 4M5.3. The 4M5.3 has higher shape complementarity between the two domains than 4-4-20.

Electrostatic contributions to binding

The role of improved electrostatic interactions was investigated through continuum electrostatic binding calculations on the two prepared, unminimized crystal

structures. The overall electrostatic contribution to binding computed by this model is summarized in Table A2.3a. An overall improvement of −3.5 kcal/mol for binding 4M5.3 was computed due to improved interactions with fluorescein (−5.7 kcal/mol) despite increased desolvation penalty for the fluorescein (1.8 kcal/mol). Interestingly, the antibody desolvation cost was very similar for both complexes and in neither case was the full cost of desolvating the binding partners recovered in intermolecular interactions. These energetic calculations of desolvation are in contrast to the simple solvent-accessible surface area measurements. To further understand these results, the −3.5 kcal/mol total electrostatic affinity difference was dissected into individual "components". A separate component for each backbone amino (C^{α}–N–H), each backbone carbonyl (C═O), and each side-chain group beyond C^{α} was computed, and each component was divided into desolvation (changed interactions with solvent due to binding) and interaction components (new interactions made intermolecularly with the binding partner in the complex). The components provide additive contributions to the computed difference in binding affinity. This type of dissection is possible in a strictly additive manner because of the superposition properties of the linearized Poisson–Boltzmann equation. Table A3.3B shows the difference of component analyses between the two crystal structures; a negative (positive) value corresponds to a contribution towards improvement (reduction) for 4M5.3 binding relative to 4-4-20. The scFv components are ranked by the metric *SumAbs*, given by equation (1), which gives a measure of the sum of the unsigned difference of the contributions for each component:

$$SumAbs_i = \left|Desolv_i^{4M5.3} - Desolv_i^{4\text{-}4\text{-}20}\right| + \sum_{j=1}^{n} \left|Inter_{i-j}^{4M5.3} - Inter_{i-j}^{4\text{-}4\text{-}20}\right|$$

$Desolv_i^X$ represents the desolvation penalty for component i in complex X, and $Inter_{i-j}^X$ represents the interaction free energy between the atoms in component i and the atoms in component j in complex X. Three of the top four components and four of the top 20 components, ranked by this difference metric, are 4M5.3 side-chain mutations. The

direct role of the 14 mutations together was calculated by summing together the components involving the 14 mutated side-chains, which revealed a total of −1.1 kcal/mol (Table A2.3C). Four side-chains dominate this value. The remaining −2.4 kcal/mol in computed electrostatic improvement is the result of secondary effects resulting from the mutated residues (e.g. altering the binding conformation of fluorescein in the site). Three components dominate this −2.4 kcal/mol, His L31, Arg L39, and Ser L96, and were previously identified in Figure A2.5C, which identified hydrogen bond distance differences.

Because the results of this analysis produced significant free energy differences that were the sum of quite small values from individual changed interactions, further calculations were performed to investigate sensitivities to precise crystal structure atomic locations. The electrostatic component analyses were repeated on 4-4-20 and 4M5.3 crystal structures each subjected to a constrained minimization (see Materials and Methods). Table A2.4 summarizes the effects of minimization on electrostatic calculations. The net improvement decreased from −3.5 kcal/mol down to −1.3 kcal/mol. Analysis of the effects of the minimization on the component analyses reveals that changes in three interaction differences account for most of the net change. His L31, Arg L39, and Ser L96, the same three residues identified above as having shortened hydrogen bond lengths and a large role in the 4M5.3 improvement, are no longer calculated to be as different between 4-4-20 and 4M5.3 when using the minimized structures. The small changes in the hydrogen bond distances due to the minimization result in relatively large changes in the interaction differences. On the other hand, the contribution of the 14 mutated side-chain components after minimization is predominantly unchanged, still totaling −1.1 kcal/mol. Thus, the calculations suggest that the subtle structural differences between the crystal structures ("secondary effects") appear to coalesce during minimization and may not be real, but the direct computed effects appear robust to small structural changes.

Four residues of the 14 mutations account for almost all of the electrostatic differences in Table 3C: H31 Asp/His, H101 Ser/Ala, H102 Tyr/Ser, and H106 Asp/Glu. Examination of the 4-4-20 and 4M5.3 structures together with the electrostatic results

reveals several mechanistic hypotheses for 4M5.3 improved binding affinity. These mechanisms are outlined in Discussion. Here, we simply refer to Figure A2.7, which illustrates the desolvation and interaction differences of these four residues in a spatial diagram, and Figure A2.8, which shows structural details of these residue-interactions in the 4-4-20 and 4M5.3 complexes.

Discussion

The entropic and enthalpic components of protein–ligand binding free energy include: (1) covalent, van der Waals, hydrogen-bonding, and electrostatic interactions within and between protein and ligand; (2) solvation effects, including desolvation costs and the hydrophobic effect; and (3) overall translational and rotational, as well as internal (often represented as vibrational) degrees of freedom within the protein and ligand. It remains unclear how best to create substantial enhancements in binding affinity through structural changes that have multiple effects on free energetic contributions to binding. Here, we have analyzed the difference between a wild-type and mutant (evolved) complex to understand how affinity enhancements are arrived at in evolved systems. The results shown here exemplify the complex situation of a large number of small structural changes resulting in a large energetic effect.

The 4M5.3 affinity matured scFv has an 1800-fold higher affinity than its wild-type 4-4-20 parent with a difference in the change in free energy of binding ($\Delta\Delta G_{binding}$) of −4.5 kcal/mol between the 4-4-20 and 4M5.3 scFvs. The primary contribution to the $\Delta\Delta G_{binding}$ is enthalpic, −4.0 kcal/mol of the −4.5 kcal/mol. The small entropic contribution to the change in free energy of binding, combined with the lack of an observation of two-step association kinetics, suggests that the single-step binding model has not changed in the affinity maturation and remains a lock-and-key fit. This is in contrast to several previously studied lower affinity antibody maturations, where improvement was observed to be derived primarily from entropic factors and an induced fit in the wild-type to a lock and key model in the affinity matured mechanism was implicated from studies of bound and unbound structures (14, 16, 17).

Enthalpic and entropic changes in binding free energy often compensate for each other and lessen the possible overall gain in affinity in improved binding interactions (32, 33). In 4M5.3, both the enthalpic and entropic changes were favorable, with the enthalpic contribution ($\Delta\Delta H$) providing the majority of the change. A major component of entropic change is often contributed by solvent molecules fixed at, or liberated from, non-polar surfaces. Five buried water molecules are present near the fluorescein-binding pocket between the V_L and V_H domains and one additional water molecule is partly buried in the surface-exposed portion of the binding pocket in 4-4-20 and 4M5.3. Each of these six water molecules is in a corresponding location and each makes similar interactions in 4-4-20 and 4M5.3. The minimal change in entropic contribution to $\Delta\Delta G_{bind}$ in the mutant structure is supported by the lack of visible changes in the bound water locations.

The enthalpic change in binding for 4M5.3 over 4-4-20 is consistent with improved van der Waals interactions, hydrogen bonds, and electrostatic interactions between 4M5.3 and fluorescein. However, the crystallographic snapshot shows no large spatial changes between the 4-4-20 and 4M5.3 complexes. The majority of the CDR residues are closer to fluorescein in 4M5.3; however, the changes are all very small, with most below 0.5 Å. Individual changes of this magnitude may not be statistically significant; however, the overall trend shows that 50/63 of the differences shown in Figure 6 are negative. This fraction is different from 50% with confidence $p<3\times10^{-6}$, considering each difference as an independent binomial trial. The sum of these small changes may provide a significant addition to the enthalpic component through small favorable changes.

Thermal *B* values are sometimes interpreted as providing information on relative mobility between related crystal structures. In the 4M5.3 structure the apparent motions are of more similar magnitude across V_H and V_L than in 4-4-20, where the *B* factors vary substantially between the two domains. In the 4-4-20 structure the V_H domain has 38% higher average *B*-values compared to the V_L domain, while in the 4M5.3 structure, the V_H domain has average *B*-values only 13% higher than the V_L domain (data not shown). It is interesting that most of the mutations, 12 of the 14, were in the V_H domain, which is

the domain that has the higher *B*-values in the 4-4-20 structure, and that after these mutations, both domains have similar *B*-value averages. This result could be simply due to the differences in crystal lattice packing between the 4-4-20 and 4M5.3 crystal space groups (*P*1 *versus* $P2_12_12_1$, respectively); however, a different, consistent theory is that rigidification of V_H due to the mutations in 4M5.3 allows greater time-averaged contacts with fluorescein, thereby decreasing the mobility in the 4M5.3 V_H compared to 4-4-20.

Shape complementarity is considered important in high affinity binding. However, in this study, and in agreement with a recent study of protein–antibody interactions, shape complementarity did not significantly improve between the wild-type and higher affinity antibody–ligand binding interaction (13). The shape complementarity did improve between the Fv domains (V_H and V_L domains) in our study. The mechanism of antigen binding affinity modified by V_L and V_H domain interactions has been proposed to occur by changing the relative positions of the CDR loops (34). Although in this study the shape complementarity of the V_L and V_H improved in 4M5.3, no noticeable shift of the domains was observed.

The higher shape complementarity of the Fv domains did not produce an increased melting temperature in 4M5.3. The wild-type 4-4-20 was more thermally stable than the high affinity 4M5.3 in both the bound and unbound states. The melting temperature data are consistent with 4M5.3 destabilizing the unbound antibody at a cost of only slightly destabilizing the complex. It remains interesting to determine the effect of the four mutations that fall at the V_L–V_H interface in this study and their importance in both the V_L–V_H shape complementarity, stability, and antigen affinity contributions.

The effect of non-covalent interactions on binding was examined by comparing solvent-accessible surface (*SAS*) area in the presence and absence of fluorescein in the 4-4-20 and 4M5.3 complexes. This shows that the ΔSAS area has decreased slightly in 4M5.3, relative to 4-4-20, suggestive of a better fit. However, this better fit was not illustrated through the shape complementarity parameter. The change also did not correspond to a change in the calculated desolvation penalty between 4-4-20 and 4M5.3, suggesting that the small decrease in polar amino acid residue surface area buried is not

playing a significant role in the improved affinity of 4M5.3. Aromatic–aromatic interactions are known to contribute to specific, energetically favorable interactions (35). Since several tyrosine residues and tryptophan residues surround the aromatic rings of fluorescein, these interactions were considered as possible components of the 4M5.3 improved binding interactions. An ideal aromatic–aromatic interaction has been defined as occurring in pairs of phenyl ring centroid separations greater than 4.5 Å and less than 7 Å with dihedral angles of 30° to 90°. They are suggested to provide between −0.6 kcal/mol and −1.3 kcal/mol. Again, however, only small changes are seen in the aromatic ring positions surrounding the fluorescein in 4M5.3, and it is unclear whether any would be more favorable than in 4-4-20. Additionally, amino–aromatic interactions are possible with the aromatic rings of the fluorescein and the amino groups of nearby residues (36). This interaction could be possible with amino groups of lysine, arginine, glutamine, or histidine when within 6 Å of the fluorescein ring centroid, making van der Waals contact with the δ(−) π-electrons while avoiding the δ(+) ring edge. However, the nearby histidine and arginine residues are in hydrogen-bonding interactions with the fluorescein and do not appear to be involved in these possible amino-aromatic interactions.

Electrostatic components of the binding free energy were calculated using crystal structure snapshots and the framework of a rigid binding model with continuum solvent. Calculations on the two crystal structures revealed −3.5 kcal/mol in favor of 4M5.3, though subsequent analysis revealed that hydrogen bonds between fluorescein and the conserved residues Arg L39, Ser L96, and His L31 account for much of the −3.5 kcal/mol improvement and are highly sensitive to precise atomic locations. At the same time, however, four of the mutated residues dominated −1.1 kcal/mol of the binding affinity improvement, regardless of whether the structures used for analysis had been subjected to the constrained minimization. It is possible, but still uncertain, that mutations in 4M5.3 are responsible for small structural changes in conserved residues that improve binding to fluorescein. On the other hand, the four residues identified for a calculated −1.1 kcal/mol improvement reveal several specific mechanisms for mutated residues in the affinity maturation.

A diagram of the calculated differences between 4-4-20 and 4M5.3 in electrostatic components of binding is presented in Figure A2.7. Besides the three interaction differences from the three sensitive residues discussed above, seven out of the nine significant energetic changes (values in red) are associated with one or more of the four mutated residues H31, H101, H102, or H106. H31 and H102 are not coupled to other mutations while H101 and H106 are coupled by the significant change in indirect interaction between them, an initial suggestion for cooperativity. Improvements are found in all three terms of the electrostatic binding free energy: direct interactions (e.g. a hydrogen bond between the scFv and fluorescein), indirect interactions (e.g. a hydrogen bond within the scFv that is buried upon binding and hence strengthened due to loss of solvent-screening), and desolvation (i.e. the penalty paid by a polar group for leaving its unbound, aqueous environment and entering a buried environment upon binding). Binding affinity improvements can be the result of either the introduction/strengthening of a favorable interaction, the removal/weakening of an unfavorable interaction, or the reduction of a desolvation penalty.

Figure A2.8A and A2.8B show the local environment of the mutated sites H101, H106, and H108. H108 was not indicated by the electrostatic analysis for binding free energy, although the position is part of an scFv intramolecular hydrogen-bonding network that includes H101 and H106. Integrating the quantitative calculations in Table 3c with the structures shows the following mechanisms for enhanced affinity. The H101 serine to alanine mutation eliminates the unsatisfied hydroxyl oxygen, removing a small desolvation penalty, yet loses the favorable indirect interaction from the H106 aspartate hydrogen bond (mutated to glutamate), for a net favorable effect. This effect is captured in the context of the other 13 mutations, but this analysis does not preclude the possibility of it also being present in a single mutant (although some interaction with the mutation at H106 is likely). The H106 aspartate to glutamate mutation reduces the residue desolvation penalty by burying the negative charge further into the protein, but more notably facilitates the serine to alanine mutation by maintaining a clear intramolecular hydrogen-bonding network. Either the H101 serine or the H106 aspartate mutation alone might disrupt antibody stability (work in progress, S.M.L. & B.T.). Likewise, the H108

tryptophan to leucine mutation in the bordering region with two bound water molecules may be coupled to the H101 and H106 mutations *via* hydrogen bonds.

Figure A2.8c and A2.8d shows the local environments of H31 and H102. The mutation at H102 from tyrosine to serine creates a new intramolecular hydrogen bond to the backbone carbonyl group at residue H31. This hydrogen bond is buried and strengthened upon binding and appears as the −0.73 kcal/mol term in Figure A2.7. Though the hydrogen bond is to residue H31, it is to the backbone and therefore the independence from the side-chain of H31 makes it a distinct mechanism from the actual mutation at H31. The aspartate to histidine mutation at H31 removes a negative charge, which although solvent-exposed, was within 7 Å of the −2 charged fluorescein. Therefore, this mutation is calculated to improve the direct interaction by eliminating a long-range electrostatic repulsion (−0.3 kcal/mol).

The electrostatic calculations indicate four of the 14 mutations for a role in the affinity improvement. Mechanisms discovered include the removal of an unsatisfied hydrogen-bonding group (H101 Ser/Ala), reduction of desolvation penalty (H106 Asp/Glu), creation of an intramolecular hydrogen bond for an indirect effect (H102 Tyr/Ser), and removal of long-range charge repulsion (H31 Asp/His). Other hypothesized mechanisms include the maintenance of a well-ordered intramolecular scFv hydrogen-bonding network.

The 4-4-20 and 4M5.3 structures used for computational analysis were prepared to minimize bias from inherent crystal structure differences; however, a few differences remain between the structures used for computational analysis and the scFvs used in the binding experiments. The N-terminal residue of the heavy chain in 4-4-20 is a glutamate in all experiments but the crystal structure contains an aspartate at this position (24). This discrepancy is most likely minimal, since both side-chains are charged −1, solvent-exposed, and about 20 Å from the fluorescein. Second, the fluorescein used in all binding experiments was biotinylated at the 5-carbon position of fluorescein, which points out of the binding site, whereas both crystal structures were obtained with neat fluorescein. A concern is whether the 4M5.3 mutations enhance binding affinity through interaction

with atoms present only in the biotinylation. Two of the mutated residues, H31 Asp/His and H102 Tyr/Ser are within at least 5 Å of the beginning of the biotinylation, a thiourea group with hydrogen-bonding capabilities. Experiments to indicate the extent of the importance of the linker in the binding interaction have not yet been successful.

The results summarized here do not definitively elucidate the mechanism for the full 1800-fold affinity improvement of 4M5.3. The largest change in the enthalpic contribution to the binding in 4M5.3 is attributed to improved electrostatics. A lack of visible changes in the structural comparison with a large change of affinity is not unprecedented. In a structural study of the four binding-site tryptophan residues in the streptavidin–biotin complex, where the affinity for biotin dropped up to 10^7-fold with a Trp to Ala mutation at any of these sites, the crystal structures showed no major changes in the binding mode for biotin compared to the wild-type streptavidin (1). In a study of the Fv fragment of the D1.3 anti-hen egg-white lysozyme antibody, a single change in the V_L92 site from tryptophan to phenylalanine, aspartate, histidine, serine, or valine created a drop in affinity of ten- to 100-fold with no significant shifts in the positions of the protein atoms at the mutation site (37, 38). Although it remains difficult to attach structural parameterization to the binding enthalpy, recent work by Luque & Freire (39) suggested that at least the following three terms are needed to account for the structural parameterization of binding enthalpy of small ligands: (1) the interactions between the ligand, antibody, and solvent; (2) changes in conformations within the antibody or ligand upon binding; and/or (3) protonation/deprotonation events in binding. Most of the effort here has been placed on studying the intrinsic enthalpy from the binding between the ligand and antibody. Preliminary calorimetry data suggests that the protonation effects on binding are the same between 4-4-20 and 4M5.3 (data not shown). Changes in conformations within the 4-4-20 and 4M53 on binding have thus far been difficult to study but may hold more information about this affinity maturation mechanism.

It would also be of interest to observe the unbound structures of both scFvs. However, perhaps notably, we were not able to crystallize either the 4-4-20 or 4M5.3 scFv in the absence of ligand. In a study of the 4-4-20 and 4M5.3 scFvs on a size-exclusion gel-filtration column (Superdex 200), 4-4-20 scFv eluted as expected at the

predicted molecular mass of ~30 kDa (data not shown), while the 4M5.3 scFv eluted much later with a standard of 17 kDa. However, in dynamic light-scattering studies, both unbound scFv proteins had a radius of 2.55 nm, indicating the same globular size (data not shown). The results from size-exclusion chromatography nevertheless indicate that the 4M5.3 unbound structure is interacting differently with the column material than the unbound 4-4-20, although not in direct correlation with the protein radius by crystallography or light-scattering.

The studies presented here add insight into high affinity binding mechanisms. By studying improvements to a protein that began with high affinity, we are addressing the mechanisms that are most important in fine-tuning the subtle aspects of affinity. In studying a protein that was developed by a random mutagenesis approach, we may be exposing binding mechanisms important to affinity that may rarely be identified by site-directed loss-of-function approaches or from studying only a bound and unbound form of a single protein. Since the binding affinity improvement of 4M5.3 does not appear to be attributable to any single large change in structure, hydrogen bonds, or salt bridges, this 1800-fold improved interaction would be difficult to predict computationally at this time. This interaction is an example of a binding improvement formed through the interaction of a variety of interactions and the sum of many small changes. Future experimental studies to dissect the energetic contributions of the individual mutations and specific combinations should prove interesting in the understanding of which mutations are contributing to this very high affinity 4M5.3 mutant antibody.

Protein Data Bank accession code

The atomic coordinates for the complex of fluorescein with the 4M5.3 scFv have been deposited with the Protein Data Bank under accession code 1X9Q.

Acknowledgements

K.M. thanks Balaji Rao for extremely helpful discussions about the competition assays for K_d determination on the surface of yeast. Funding was provided from NCI CA96504 (to K.D.W.). S.M.L. was funded by an NSF Graduate Fellowship. Funding was

provided from the Alfred P. Sloan Foundation (to C.L.D.). H.H.H. was funded by an NIH biotechnology training grant (T32-GM08334). Portions of this research were carried out at the Stanford Synchrotron Radiation Laboratory, a national user facility operated by Stanford University on behalf of the US Department of Energy, Office of Basic Energy Sciences. The SSRL Structural Molecular Biology Program is supported by the Department of Energy, Office of Biological and Environmental Research, and by the National Institutes of Health, National Center for Research Resources, Biomedical Technology Program, and the National Institute of General Medical Sciences.

- APPENDIX 2 – Substantial Improvement with Minimal Structural Perturbations in a High Affinity Mutant Antibody

Table A2.1: Crystallographic data statistics

A. Data collection statistics	
Space group	$P2_12_12_1$
Unit cell (Å)	a=55.78 b=57.45 c=79.22
Wavelength (Å)	0.9840
Resolution (Å)	50–1.50
R_{sym}[lk and lk] (%)	3.0 (38.2)
Total observations	257,289
Unique reflections	40,648
Avg. redundancy	7.0
$\langle I/\sigma \rangle$[b]	56.8 (3.46)
Completeness[b] (%)	98.8 (97.4)
B. Refinement statistics	
Resolution (Å)	28.75–1.50
Reflections (working/test)	40,648/4049
R_{cryst}/R_{free}[c] (%)	16.48/19.29
Residues	231
Total protein atoms	2113
Total ligand atoms	81
Total solvent atoms	233
Disordered side-chains	11
Alternate conformations	21
RMS deviation	
Bond lengths (Å)	0.013
Bond angles (deg.)	1.448
Ramachandran analysis (%)	
Most-favored	176 (92.1)
Allowed	14 (7.3)
Disallowed	1 (0.6)

- APPENDIX 2 – Substantial Improvement with Minimal Structural Perturbations in a High Affinity Mutant Antibody

[a] $R_{sym} = (\sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|) / \sum_{hkl} \sum_i I_i(hkl)$ for n independent reflections and i observations of a given reflection. $\langle I(hkl) \rangle$, average intensity of ith observation.

[b] Numbers for the highest resolution shell are shown in parentheses.

[c] $R_{cryst} = \sum_h ||F_o(h)| - |F_c(h)|| / \sum_h |F_o(h)|$, where F_o and F_c are the observed and calculated structure factors, respectively. R_{free} was calculated using the same equation, and included a test set of reflections that were not used during refinement.

Table A2.2: Characteristics of antibody–fluorescein complexes

A. Solvent-accessible surface area (SAS) (Å^2)			
	4-4-20	**4M5.3**	**Δ**
Δ*SAS* (scFv&Flu–scFv)	−284	−260	24
Δ*SAS* (V_L&Flu–V_L)	−109	−94	15
Δ*SAS* (V_H&Flu–V_H)	−175	−166	9
Δ*SAS* polar residues	−217	−192	25
Δ*SAS* apolar residues	−67	−68	−1
Δ*SAS* V_L, polar residues	−101	−90	11
Δ*SAS* V_L, apolar residues	−8	−4	4
Δ*SAS* V_H, polar residues	−116	−102	14
Δ*SAS* V_H, apolar residues	−59	−64	−5
B. Shape complementarity	4-4-20	4M5.3	
(V_L & V_H) *versus* Flu	0.812	0.808	
V_L *versus* V_H	0.640	0.710	
V_L *versus* Flu	0.820	0.815	
V_H *versus* Flu	0.792	0.756	

Shape complementarity (*Sc*) is a measure of the geometric shape matching between two molecular surfaces.[31] Values range from 0 to 1.0, where a value of 1.0 describes an interface with perfect geometric fit. The shape complementarity was calculated for the crystal structures of both 4-4-20 and 4M5.3 with fluorescein (Table 2) and indicates negligible change in the overall complementarity between the scFvs and fluorescein. The shape complementarity of V_L with fluorescein remains the same between 4-4-20 and 4M5.3 while the shape complementarity of V_H with fluorescein is slightly less in 4M5.3. However, when the shape complementarity measure is applied to the interface between the V_H and V_L domains, there is a larger difference between 4-4-20 and 4M5.3. The 4M5.3 has higher shape complementarity between the two domains than 4-4-20.

Table A2.3: Comparison of computed electrostatic contributions to binding in 4-4-20 and 4M5.3

A. Computed total electrostatic contributions to binding, $\langle C^{\alpha}-N-H \rangle$, in kcal/mol

	scFv desolvation	Fluorescein desolvation	Interaction	Net binding
4-4-20	14.36	14.80	−19.97	9.20
4M5.3	14.68	16.64	−25.62	5.70
4M5.3−4-4-20	0.32	1.84	−5.65	−3.50

B. Electrostatic components of the difference, 4M5.3−4-4-20, in kcal/mol

Component	Desolv.[a]	Indirect[b]	Direct[c]	Contrib.[d]	Mut.[e]	*SumAbs*[f]
Fluorescein	1.84	N.A.	−5.65	−0.98	−3.81	
L 39 Arg	0.01	1.24	−1.45	−0.09	−0.20	6.57
H 106 Asp/Glu[g]	−0.33	0.11	−0.13	−0.34	−0.35	4.50
H 101 Ser/Ala[g]	−0.11	0.38	−0.79	−0.31	−0.52	4.01
H 102 Tyr/Ser[g]	0.68	−0.51	−0.46	0.19	−0.29	3.59
H 105 carbonyl	−0.32	0.28	−0.16	−0.26	−0.20	3.13
W 615 water	0.04	0.42	−0.85	−0.18	−0.40	2.53
H 31 carbonyl	0.21	−0.95	0.24	−0.15	−0.50	2.36
H 102 amino	−0.21	0.13	0.15	−0.07	0.07	2.18
H 101 carbonyl	−0.04	−0.17	0.14	−0.05	−0.07	2.15
L 31 His	−0.15	−0.30	−0.47	−0.53	−0.91	1.85
H 74 Arg	0.05	0.21	−0.18	0.07	0.08	1.68
H 102 carbonyl	−0.07	0.15	0.10	0.06	0.19	1.64
L 96 Ser	−0.02	0.15	−0.76	−0.32	−0.63	1.64
H 104 amino	−0.04	0.01	0.28	0.11	0.26	1.55
H 53 Asn	0.27	−0.46	0.11	0.10	−0.07	1.52
H 52 Arg	−0.01	0.15	−0.12	0.00	0.01	1.46
H 103 carbonyl	−0.01	−0.00	−0.23	−0.12	−0.24	1.31
H 31 Asp/His[g]	−0.01	−0.04	−0.29	−0.18	−0.34	1.20

A. Computed total electrostatic contributions to binding, $\{C^{\alpha}-N-H\}$, in kcal/mol						
	scFv desolvation	**Fluorescein desolvation**	**Interaction**	**Net binding**		
H 104 carbonyl	−0.06	0.17	−0.27	−0.11	−0.16	1.19
H 103 amino	−0.04	0.09	−0.04	−0.01	0.02	1.17

All values are the differences between analyses on 4M5.3 and 4-4-20. Negative (positive) values indicate an improvement (reduction) for 4M5.3.

[a] Desolvation penalty.

[b] Sum of indirect interactions with all other scFv components.

[c] Direct interaction between scFv component(s) and fluorescein.

[d] Contribution=desolvation+(1/2)indirect+(1/2)direct.

[e] Mutation=desolvation+indirect+direct.

[f] *SumAbs*, sum of the absolute value of desolvation and all interaction terms.

[g] Site of 4M5.3 mutation.

Table A2.4: Effect of crystal structure minimizations on electrostatic analyses

A. Computed net electrostatics of binding,

$$SumAbs_i = |Desolv_i^{4M5.3} - Desolv_i^{4-4-20}| + \sum_{j \neq i} |Inter_{i,j}^{4M5.3} - Inter_{i,j}^{4-4-20}|$$

, using minimized structures (kcal/mol)

	scFv desolvation	Fluorescein desolvation	Interaction	Net binding
4-4-20	14.39	14.28	−21.57	7.11
4M5.3	14.66	16.62	−25.45	5.84
4M5.3-4-4-20	0.27	2.34	−3.87	−1.27

B. The largest component–component interaction changes as a result of minimization

Specific interaction	Change in strength (kcal/mol)
L 39 Arg–fluorescein	+0.575
L 96 Ser–fluorescein	+0.548
L 31 His–fluorescein	+0.531

C. Net electrostatic impact of the mutated side-chain components after minimization

	Desolv.	Ind. In	Ind. out	Direct	Total
Total	0.29	−0.29	0.27	−1.42	−1.14

Table A2.5:

C. Electrostatic impact of the mutated side-chain components					
Component	**Desolv.[a]**	**Ind. in[b]**	**Ind. out[c]**	**Direct[d]**	**Total**
H 1 Glu/Gly	0.00	−0.01	0.03	−0.06	−0.03
H 31 Asp/His	−0.01	0.03	−0.09	−0.29	−0.37
H 51 Ile/Phe	0.00	0.00	0.02	−0.03	−0.02
H 101 Ser/Ala	−0.11	−0.20	0.78	−0.79	−0.31
H 102 Tyr/Ser	0.68	0.03	−0.57	−0.46	−0.32
H 106 Asp/Glu	−0.33	−0.20	0.50	−0.13	−0.15
H 108 Trp/Leu	0.00	0.03	−0.13	0.18	0.08
TOTAL	0.23	−0.32	0.54	−1.58	−1.13

For the mutations of L60 Phe/Val, L81 Ser/Asn, H16 Arg/Gly, H17 Pro/Ala, H24 Ala/Thr, H30 Ser/Gly, and H93 Met/Thr, each of the individual five electrostatic values (in columns in part A) are ≤0.01 in magnitude, and these rows are left out of the Table. All values are the differences between analyses on 4M5.3 and 4-4-20. Negative (positive) values indicate an improvement (reduction) for 4M5.3.

[a] Desolvation penalty.

[b] Half sum of indirect interactions with the 13 other mutated side-chain components.

[c] Sum of indirect interactions with all but the 13 other mutated side-chain components.

[d] Direct interaction with fluorescein.

- APPENDIX 2 – Substantial Improvement with Minimal Structural Perturbations in a High Affinity Mutant Antibody

Figure A2.1: Equilibrium binding constant for 4M5.3 by competition with 4-4-20. Four replicates are shown: two starting from pre-incubation of 4M5.3 with fluorescein-biotin (× and •), and two starting from pre-incubation with the 4-4-20 competitor with fluorescein-biotin (▫ and •), as noted in **Materials and Methods**.

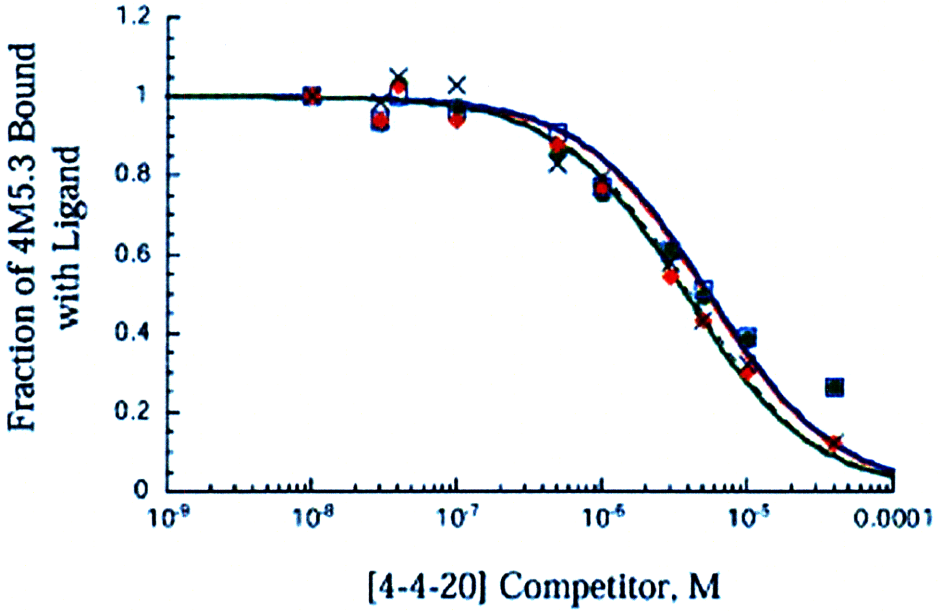
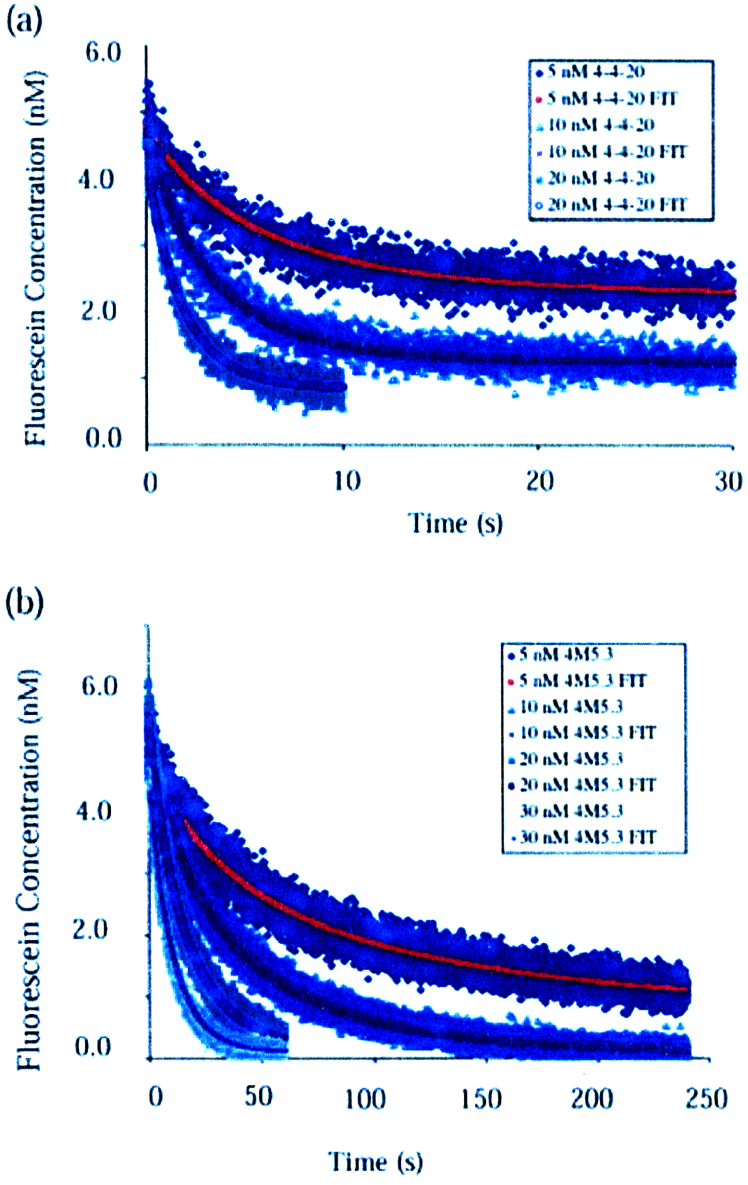
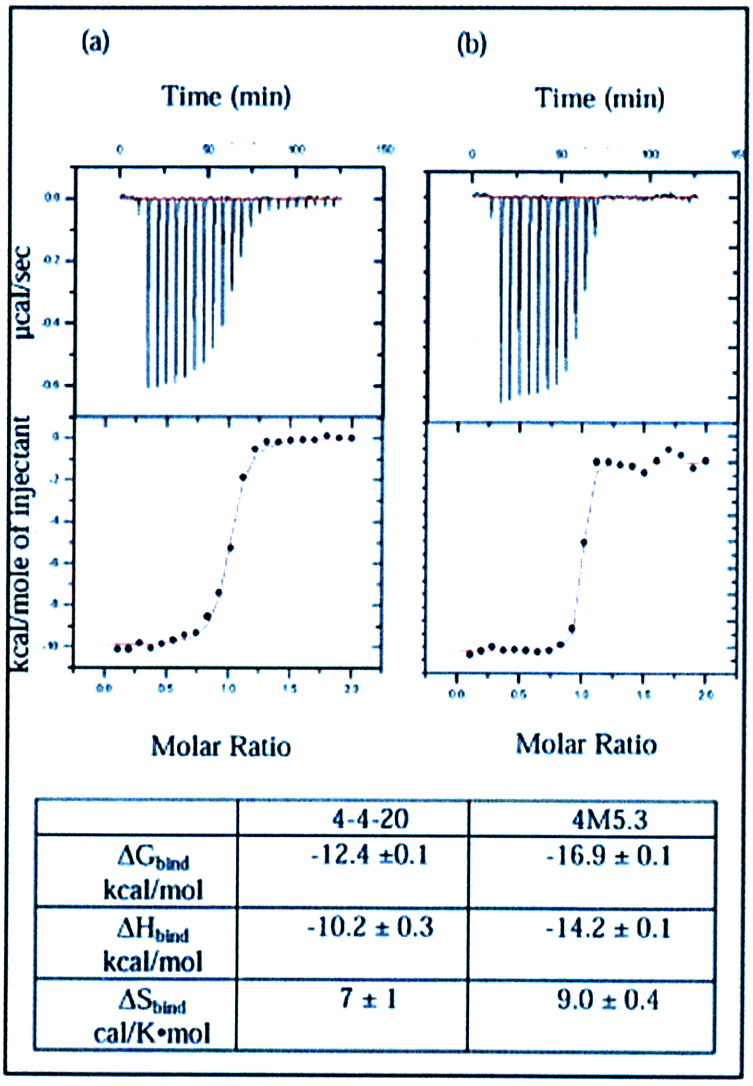


Figure A2.2: Stopped-flow association rates. Fluorescein-biotin (5 nM) was mixed with the specified concentrations of scFv. Curves were fit using k_{off} determined previously and solving for k_{on}, using the Gepasi software; see Materials and Methods. (a) 4-4-20, (b) 4M5.3.



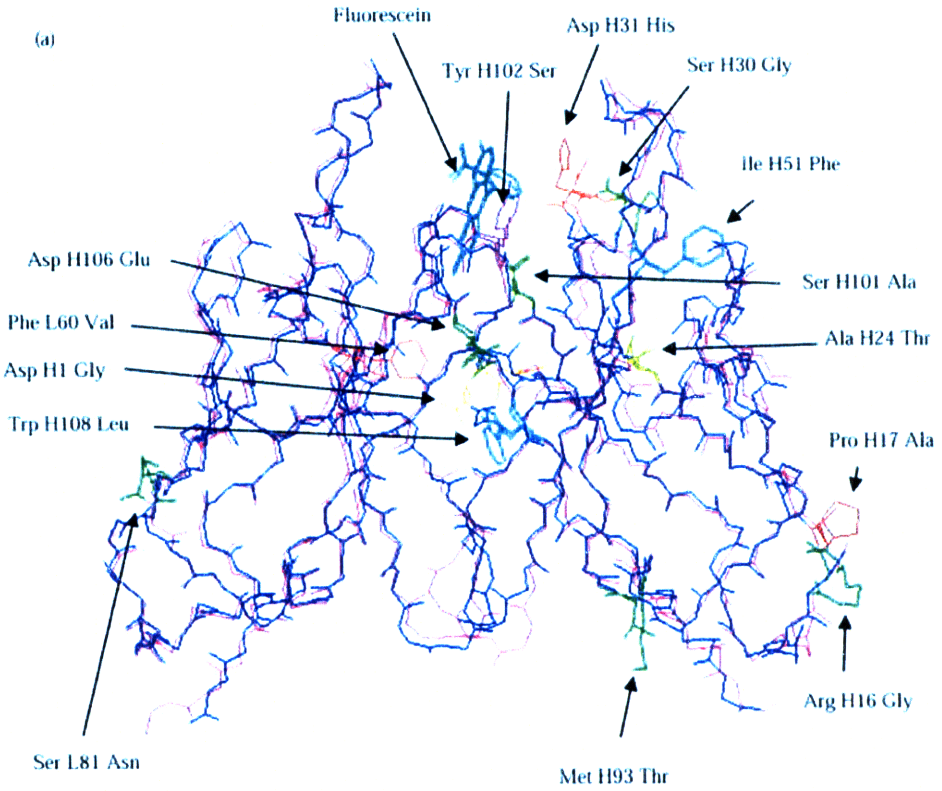
Conc. of scFv	4-4-20 k_{on} ($M^{-1} s^{-1}$)	4M5.3 k_{on} ($M^{-1} s^{-1}$)
5nM	3.7×10^7	4.1×10^6
10nM	3.1×10^7	3.5×10^6
20nM	3.1×10^7	3.8×10^6
30nM	N/D	3.7×10^6

Figure A2.3: Isothermal titration calorimetry of 4-4-20 and 4M5.3. Curves are fit to determine ΔH; ΔG is determined from previously determined K_d; ΔS is calculated from ΔG and ΔH. (a) 4-4-20, (b) 4M5.3.



	4-4-20	4M5.3
ΔG_{bind} kcal/mol	-12.4 ±0.1	-16.9 ± 0.1
ΔH_{bind} kcal/mol	-10.2 ± 0.3	-14.2 ± 0.1
ΔS_{bind} cal/K•mol	7 ± 1	9.0 ± 0.4

Figure A2.4: Comparison of 4-4-20 and 4M5.3 Fv domain structures. (a) Alignment of 4M5.3 and 4-4-20 crystal structures by protein backbone atoms. 4-4-20 is pink and 4M5.3 is blue. (b) The sequence of the 4-4-20 V_L and V_H domains. The 14 genetic changes between 4-4-20 and 4M5.3 and their structural locations are noted.



(b)

```
                                    --------CDR1--------                    --CDR2
4-4-20: VL  1 DVVMTQTPLS LPVSLGDQAS ISCRSSQSLV HSNGNTYLRW YLQKPGQSPK VLIYKVSNRF
4M5.3:        ---------- ---------- ---------- ---------- ---------- ---------V

                -                           --CDR3----
           61 SGVPDRFSGS GSGTDFTLKI SRVEAEDLGV YFCSQSTHVP WTFGGGTKLE IK (linker)
              ---------- ---------- N--------- ---------- ---------- --

                                    CDR1                  - -----CDR2-
4-4-20: VH  1 DVKLDETGGG LVQPGRPMKL SCVASGFTFS DYWMNWVRQS PEKGLEWVAQ IRNKPYNYET
4M5.3:        G--------- -----GA--- ---T-----G H--------- ---------- F---------

              --------                                -CDR3--
           61 YYSDSVKGRF TISRDDSKSS VYLQMNNLRV EDMGIYYCTG SYYGMDYWGQ GTSVTVS
              ---------- ---------- ---------- --T------- AS---E-L-- -------
```

Figure A2.5: Comparison of 4-4-20 and 4M5.3 Fv domain crystal structures. Alignment of 4M5.3 and 4-4-20 structures by protein backbone atoms. 4-4-20 is pink and 4M5.3 is blue. (a) The hydrophobic fluorescein-binding pocket is lined with tyrosine and tryptophan residues. (b) Hydrogen bonds with fluorescein. (c) Difference in distance (Å) of the five hydrogen bonds between scFv and fluorescein.

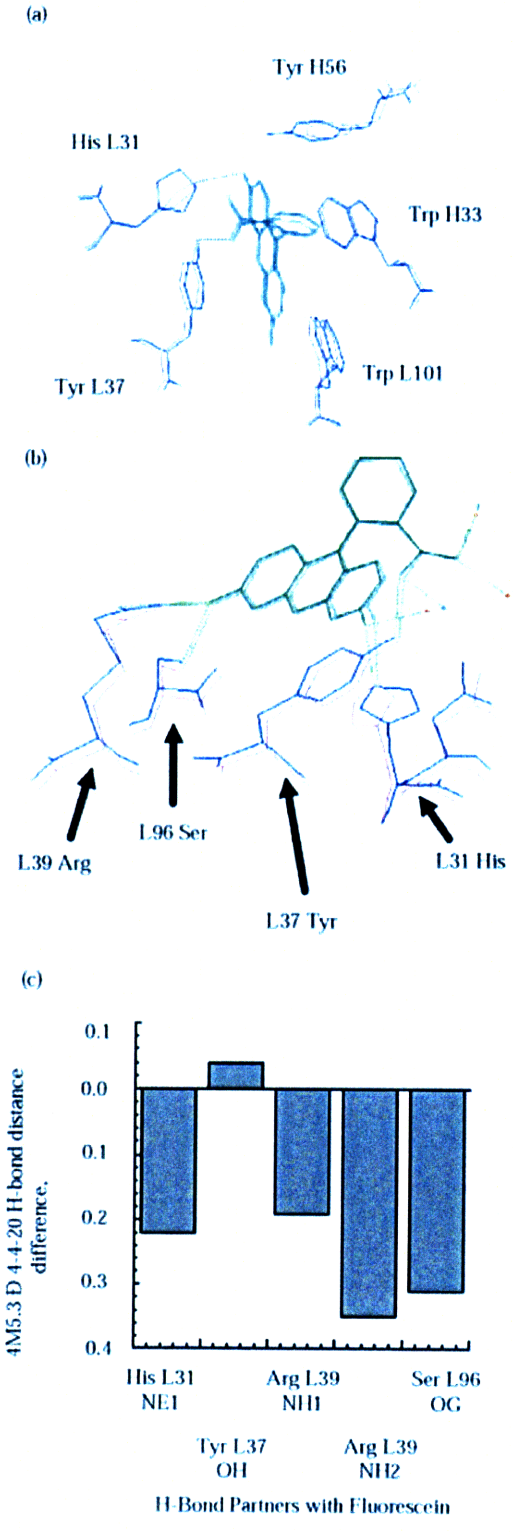


Figure A2.6: Difference in distance of closest atom of each CDR residue to fluorescein for the 4-4-20 and 4M5.3 structures. (a) Light chain CDR residues, (b) heavy chain CDR residues. * Indicates a mutated position; c indicates residue within contact (≤5 Å) of the fluorescein in the 4-4-20 and 4M5.3 structures; χ indicates a residue within contact (≤5 Å) of the fluorescein in only the 4M5.3 structure.

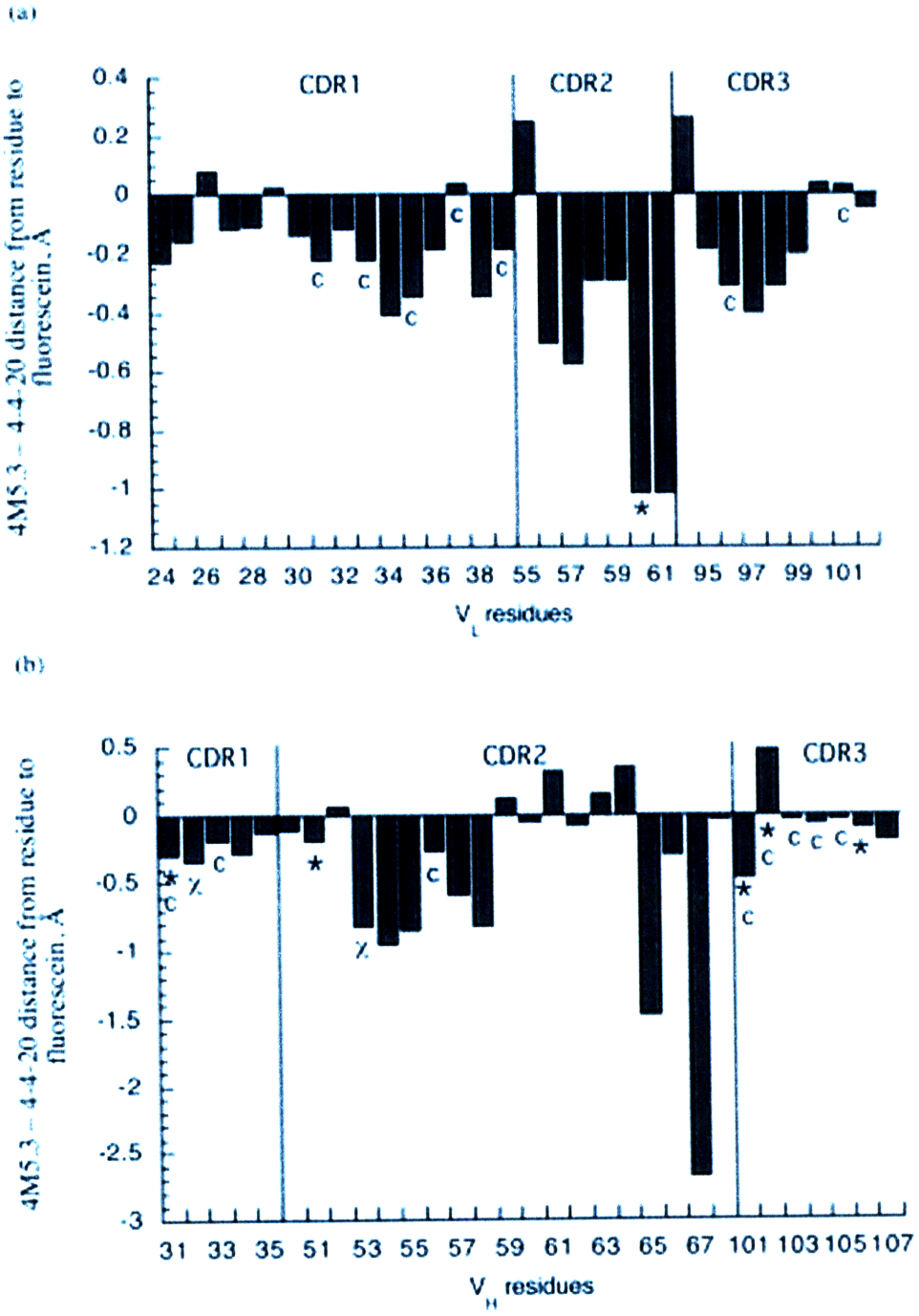


Figure A2.7: Diagram of all major (>0.2 kcal/mol) electrostatic interaction and desolvation penalty differences (4M5.3−4-4-20) among components in **Table 3**B. Arrows connecting to fluorescein correspond to changes in direct interactions, whereas arrows connecting scFv components correspond to changes in indirect interactions, and changes in desolvation penalties are shown in parentheses. The three interaction differences marked with * are those which dramatically decrease (become more positive) with minimization. All other values <−0.25 kcal/mol are shown in red. Mutated side-chains (four) are labeled in blue.

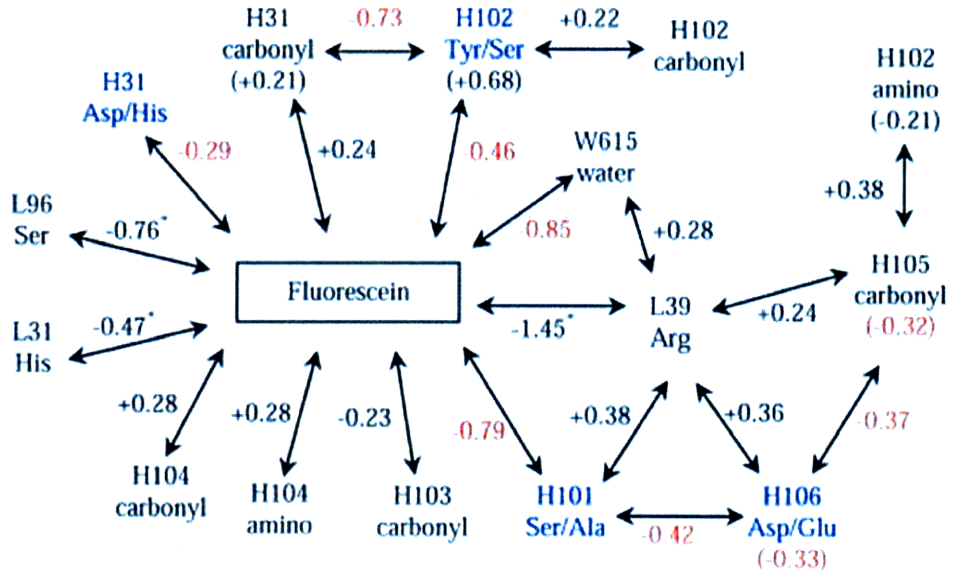
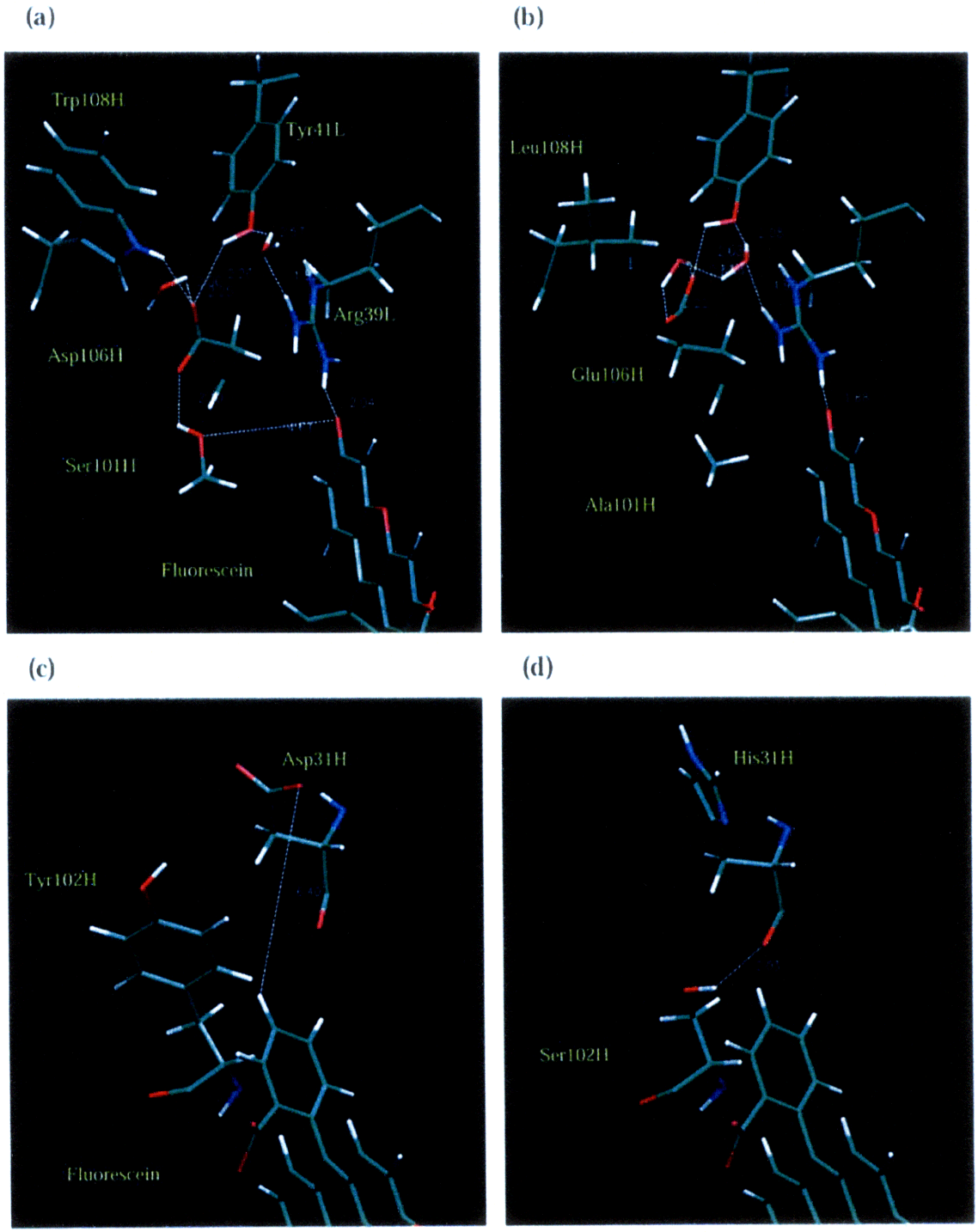


Figure A2.8: Two regions of the binding site that include the four residues identified by the electrostatic analysis. First, the region of H101 and H106 together with neighboring hydrogen-bonding residues L39, L41, and H108 are shown in (a) and (b) for 4-4-20 and 4M5.3, respectively. Second, the region of H31 and H102 is similarly shown in (c) and (d).



REFERENCES

1 S. Freitag, I. Le Trong, A. Chilkoti, L.A. Klumb, P.S. Stayton and R.E. Stenkamp, Structural studies of binding site tryptophan mutants in the high-affinity streptavidin–biotin complex, *J. Mol. Biol.* 279 (1998), pp. 211–221.

2 S. Freitag, I. Le Trong, L.A. Klumb, V. Chu, A. Chilkoti, P.S. Stayton and R.E. Stenkamp, X-ray crystallographic studies of streptavidin mutants binding to biotin, *Biomol. Eng.* 16 (1999), pp. 13–19.

3 J. Foote and H.N. Eisen, Kinetic and affinity limits on antibodies produced during immune responses, *Proc. Natl Acad. Sci. USA* 92 (1995), pp. 1254–1256.

4 W.P. Yang, K. Green, S. Pinz-Sweeney, A.T. Briones, D.R. Burton and C.F. Barbas III, CDR walking mutagenesis for the affinity maturation of a potent human anti-HIV-1 antibody into the picomolar range, *J. Mol. Biol.* 254 (1995), pp. 392–403.

5 R. Schier, A. McCall, G.P. Adams, K.W. Marshall, H. Merritt and M. Yim *et al.*, Isolation of picomolar affinity anti-c-erbB-2 single-chain Fv by molecular evolution of the complementarity determining regions in the center of the antibody binding site, *J. Mol. Biol.* 263 (1996), pp. 551–567.

6 Y. Chen, C. Wiesmann, G. Fuh, B. Li, H.W. Christinger and P. McKay *et al.*, Selection and analysis of an optimized anti-VEGF antibody: crystal structure of an affinity-matured Fab in complex with antigen, *J. Mol. Biol.* 293 (1999), pp. 865–881.

7 E.T. Boder, K.S. Midelfort and K.D. Wittrup, Directed evolution of antibody fragments with monovalent femtomolar antigen-binding affinity, *Proc. Natl Acad. Sci. USA* 97 (2000), pp. 10701–10705.

8 R.W. Dixon, R.J. Radmer, B. Kuhn, P.A. Kollman, J. Yang and C. Raposo *et al.*, Theoretical and experimental studies of biotin analogues that bind almost as tightly to streptavidin as biotin, *J. Org. Chem.* 67 (2002), pp. 1827–1837.

9 G.J. Wedemayer, P.A. Patten, L.H. Wang, P.G. Schultz and R.C. Stevens, Structural insights into the evolution of an antibody combining site, *Science* 276 (1997), pp. 1665–1669.

10 H. Wu, G. Beuerlein, Y. Nie, H. Smith, B.A. Lee and M. Hensler *et al.*, Stepwise *in vitro* affinity maturation of Vitaxin, an alphav beta3-specific humanized mAb, *Proc. Natl Acad. Sci. USA* 95 (1998), pp. 6037–6042.

11 J. Yin, E.C. Mundorff, P.L. Yang, K.U. Wendt, D. Hanway, R.C. Stevens and P.G. Schultz, A comparative analysis of the immunological evolution of antibody 28B4, *Biochemistry* 40 (2001), pp. 10764–10773.

12 J. Valjakka, A. Hemminki, S. Niemi, H. Soderlund, K. Takkinen and J. Rouvinen, Crystal structure of an *in vitro* affinity- and specificity-matured anti-testosterone Fab in complex with testosterone. Improved affinity results from small structural changes within the variable domains, *J. Biol. Chem.* 277 (2002), pp. 44021–44027.

13 Y. Li, H. Li, F. Yang, S.J. Smith-Gill and R.A. Mariuzza, X-ray snapshots of the maturation of an antibody response to a protein antigen, *Nature Struct. Biol.* 10 (2003), pp. 482–488.

14 T. Sagawa, M. Oda, M. Ishimura, K. Furukawa and T. Azuma, Thermodynamic and kinetic aspects of antibody evolution during the immune response to hapten, *Mol. Immunol.* 39 (2003), pp. 801–808.

15 P.S. Stayton, S. Freitag, L.A. Klumb, A. Chilkoti, V. Chu and J.E. Penzotti *et al.*, Streptavidin-biotin binding energetics, *Biomol. Eng.* 16 (1999), pp. 39–44.

16 V. Manivel, N.C. Sahoo, D.M. Salunke and K.V. Rao, Maturation of an antibody response is governed by modulations in flexibility of the antigen-combining site, *Immunity* 13 (2000), pp. 611–620.

17 J. Yin, A.E.T. Beuscher, S.E. Andryski, R.C. Stevens and P.G. Schultz, Structural plasticity and the evolution of antibody affinity and specificity, *J. Mol. Biol.* 330 (2003), pp. 651–656.

18 L.T. Chong, Y. Duan, L. Wang, I. Massova and P.A. Kollman, Molecular dynamics and free-energy calculations applied to affinity maturation in antibody 48G7, *Proc. Natl Acad. Sci. USA* 96 (1999), pp. 14330–14335.

19 W. Dall'Acqua, E.R. Goldman, W. Lin, C. Teng, D. Tsuchiya and H. Li *et al.*, A mutational analysis of binding interactions in an antigen–antibody protein–protein complex, *Biochemistry* 37 (1998), pp. 7981–7991.

20 D.M. Kranz and E.W. Voss Jr, Partial elucidation of an anti-hapten repertoire in BALB/c mice: comparative characterization of several monoclonal anti-fluorescyl antibodies, *Mol. Immunol.* 18 (1981), pp. 889–898.

21 D.M. Kranz, J.N. Herron and E.W. Voss Jr, Mechanisms of ligand binding by monoclonal anti-fluorescyl antibodies, *J. Biol. Chem.* 257 (1982), pp. 6987–6995.

22 J.N. Herron, D.M. Kranz, D.M. Jameson and E.W. Voss Jr, Thermodynamic properties of ligand binding by monoclonal anti-fluorescyl antibodies, *Biochemistry* 25 (1986), pp. 4602–4609.

23 J.N. Herron, A.H. Terry, S. Johnston, X.M. He, L.W. Guddat, E.W. Voss Jr and A.B. Edmundson, High resolution structures of the 4-4-20 Fab-fluorescein complex in two solvent systems: effects of solvent on structure and antigen-binding affinity, *Biophys. J.* 67 (1994), pp. 2167–2183.

24 M. Whitlow, A.J. Howard, J.F. Wood, E.W. Voss Jr and K.D. Hardman, 1.85 Å structure of anti-fluorescein 4-4-20 Fab, *Protein Eng.* 8 (1995), pp. 749–761.

25 K. Lim and J.N. Herron, Molecular dynamics of the anti-fluorescein 4-4-20 antigen-binding fragment. >1. Computer simulations, *Biochemistry* 34 (1995), pp. 6962–6974.

26 K. Lim, D.M. Jameson, C.A. Gentry and J.N. Herron, Molecular dynamics of the anti-fluorescein 4-4-20 antigen-binding fragment. 2. Time-resolved fluorescence spectroscopy, *Biochemistry* 34 (1995), pp. 6975–6984.

27 L.K. Denzin, G.A. Gulliver and E.W. Voss Jr, Mutational analysis of active site contact residues in anti-fluorescein monoclonal antibody 4-4-20, *Mol. Immunol.* 30 (1993), pp. 1331–1345.

28 R. Jimenez, G. Salazar, K.K. Baldridge and F.E. Romesberg, Flexibility and molecular recognition in the immune system, *Proc. Natl Acad. Sci. USA* 100 (2003), pp. 92–97.

29 S. Terzyan, P.A. Ramsland, E.W. Voss Jr, J.N. Herron and A.B. Edmundson, Three-dimensional structures of idiotypically related Fabs with intermediate and high affinity for fluorescein, *J. Mol. Biol.* 339 (2004), pp. 1141–1151.

30 J.N. Herron, X.M. He, M.L. Mason, E.W. Voss Jr and A.B. Edmundson, Three-dimensional structure of a fluorescein-Fab complex crystallized in 2-methyl-2,4-pentanediol, *Proteins: Struct. Funct. Genet.* 5 (1989), pp. 271–280.

31 M.C. Lawrence and P.M. Colman, Shape complementarity at protein/protein interfaces, *J. Mol. Biol.* 234 (1993), pp. 946–950.

32 F. Dullweber, M.T. Stubbs, D. Musil, J. Sturzebecher and G. Klebe, Factorising ligand affinity: a combined thermodynamic and crystallographic study of trypsin and thrombin inhibition, *J. Mol. Biol.* 313 (2001), pp. 593–614.

33 K. Sharp, Entropy–enthalpy compensation: fact or artifact?, *Protein Sci.* 10 (2001), pp. 661–667.

34 E. Vargas-Madrazo and E. Paz-Garcia, An improved model of association for VH-VL immunoglobulin domains: asymmetries between VH and VL in the packing of some interface residues, *J. Mol. Recognit.* 16 (2003), pp. 113–120.

35 S.K. Burley and G.A. Petsko, Aromatic–aromatic interaction: a mechanism of protein structure stabilization, *Science* 229 (1985), pp. 23–28.

36 S.K. Burley and G.A. Petsko, Amino–aromatic interactions in proteins, *FEBS Letters* 203 (1986), pp. 139–143.

37 X. Ysern, B.A. Fields, T.N. Bhat, F.A. Goldbaum, W. Dall'Acqua and F.P. Schwarz *et al.*, Solvent rearrangement in an antigen–antibody interface introduced by site-directed mutagenesis of the antibody combining site, *J. Mol. Biol.* 238 (1994), pp. 496–500.

38 E.J. Sundberg, M. Urrutia, B.C. Braden, J. Isern, D. Tsuchiya and B.A. Fields *et al.*, Estimation of the hydrophobic effect in an antigen–antibody protein–protein interface, *Biochemistry* 39 (2000), pp. 15375–15387.

39 I. Luque and E. Freire, Structural parameterization of the binding enthalpy of small ligands, *Proteins: Struct. Funct. Genet.* 49 (2002), pp. 181–190.

40 E.V. Shusta, R.T. Raines, A. Pluckthun and K.D. Wittrup, Increasing the secretory capacity of *Saccharomyces cerevisiae* for production of single-chain antibody fragments, *Nature Biotechnol.* 16 (1998), pp. 773–777.

41 B.M. Baker and K.P. Murphy, Dissecting the energetics of a protein–protein interaction: the binding of ovomucoid third domain to elastase, *J. Mol. Biol.* 268 (1997), pp. 557–569.

42 P. Mendes, Biochemistry by numbers: simulation of biochemical pathways with Gepasi 3, *Trends Biochem. Sci.* 22 (1997), pp. 361–363.

43 Z. Otwinowski and W. Minor, Processing of X-ray difraction data collected in oscillation mode, *Methods Enzymol.* 276 (1997), pp. 307–326.

44 C.R. Kissinger, D.K. Gehlhaar and D.B. Fogel, Rapid automated molecular replacement by evolutionary search, *Acta Crystallog. sect. D* 55 (1999), pp. 484–491.

45 A.T. Brunger, P.D. Adams, G.M. Clore, W.L. DeLano, P. Gros and R.W. Grosse-Kunstleve *et al.*, Crystallography & NMR system: a new software suite for macromolecular structure determination, *Acta Crystallog. sect. D* 54 (1998), pp. 905–921.

46 Collaborative Computational Project Number 4, CCP4 site: programs for crystallography, *Acta Crystallog. sect. D* 50 (1994), pp. 760–763.

47 A.T. Brunger and M. Karplus, Polar hydrogen positions in proteins: empirical energy placement and neutron diffraction comparison, *Proteins: Struct. Funct. Genet.* 4 (1988), pp. 148–156.

48 B.R. Brooks, R.E. Bruccolerire, B.D. Olafson, D.J. States, S. Swaminathan and M. Karplus, CHARMM – a program for macromolecular energy, minimazation, and dynamics calculations, *J. Comput. Chem.* 4 (1983), pp. 187–217.

49 F.A. Momany and R. Rone, Validation of the general-purpose QUANTA(R)3.2/CHARMM(R) force-field, *J. Comput. Chem.* 13 (1992), pp. 888–900.

50 D.F. Green and B. Tidor, Evaluation of *ab initio* charge determination methods for use in continuum solvation calculations, *J. Phys. Chem. B* 107 (2003), pp. 10261–10273.

51 C.I. Bayly, P. Cieplak, W.D. Cornell and P.A. Kollman, A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges – the RESP model, *J. Phys. Chem.* 97 (1993), pp. 10269–10280.

52 Z.S. Hendsch and B. Tidor, Electrostatic interactions in the GCN4 leucine zipper: substantial contributions arise from intramolecular interactions enhanced on binding, *Protein Sci.* 8 (1999), pp. 1381–1392.

53 M.K. Gilson and B. Honig, Calculation of the total electrostatic energy of a macromolecular system: solvation energies, binding energies, and conformational analysis, *Proteins: Struct. Funct. Genet.* 4 (1988), pp. 7–18.

54 M.K. Gilson, K.A. Sharp and B.H. Honig, Calculating the electrostatic potential of molecules in solution-method and error assessment, *J. Comput. Chem.* 9 (1988), pp. 327–335.

55 K.A. Sharp and B. Honig, Electrostatic interactions in macromolecules: theory and applications, *Annu. Rev. Biophys. Biophys. Chem.* 19 (1990), pp. 301–332.

56 D. Sitkoff, K.A. Sharp and B. Honig, Accurate calculation of hydration free-energies using macroscopic solvent models, *J. Phys. Chem.* 98 (1994), pp. 1978–1988.

APT1• 52 PARK STREET • SOMERVILLE, MA 02143
PHONE 857.998.1829 • FAX 617.258.7847 • E-MAIL HECTORH@MIT.EDU

HECTOR H. HERNANDEZ

EDUCATION

2000 - 2007 Massachusetts Institute of Technology Cambridge, MA
Ph.D. Biological Chemistry (on September 2007 degree list)

1998 - 2000 University of South Florida Tampa, FL
B.A. Chemistry

1996 - 1998 Valencia Community College Orlando, FL
A.A.

RESEARCH EXPERIENCE

2000 - present Massachusetts Institute of Technology Cambridge, MA
Graduate Research Assistant, Department of Chemistry, Professor Catherine L. Drennan, Advisor

- My research focuses on the elucidation of the structure of the enzymes thioredoxin (Trx) and thioredoxin reductase (TrR). These enzymes are implicated in control of redox regulation in the cell. They also mediate the re-reduction of disulfide bonds in cellular biochemical processes. As such, they are of medical and bioengineering importance. The understanding of the structural basis for redox control may lead to the development of unique therapies and to the engineering of tightly regulated biological redox systems.

2003 Summer New England Biolabs Beverly, MA
Research Assistant

- Using the *E. coli* selenocysteine insertion loop (SECIS), I investigated the insertion of selenocysteine by replacing the active site serine of β-lactamase. An overexpression vector was engineered to express the selenocysteine modified β-lactamase. This research was aimed at developing a system that can be used to express non-bacterial selenocysteine proteins in a bacterial expression system.

1998 - 2000 University of South Florida Tampa, FL
Research Assistant, Department of Biochemistry, University of South Florida College of Medicine, Professor Michael J. Barber, Mentor

- Spinach nitrate reductase catalyzes the rate-limiting step in the pathway of inorganic nitrate assimilation. By isolating, purifying, and characterizing the heme subunit of spinach nitrate reductase (SNR), my research facilitated further studies of amino acids involved in electron pathways as well as amino acids involved in sub-domain interactions. I designed and expressed a gene that was codon optimized for bacterial expression systems allowing the over-expression of protein needed for functional studies of the SNR heme electron transfer unit.

PUBLICATIONS

- Hamill MJ, Chobot SE, **Hernandez HH**, Drennan CL, Elliott SJ. Direct electrochemical analyses of a thermophilic thioredoxin reductase: interplay between conformational change and redox chemistry. (Submitted to *Biochemistry*)
- **Hernandez HH**, Jaquez OJ, Hamill MJ, Elliott SJ, Drennan CL. Thioredoxin Reductase from T. Acidophilum: a new twist on redox regulation. (Submitted to *Biochemistry*)
- Chobot SE, **Hernandez HH**, Drennan CL, Elliott SJ. Direct Electrochemical Characterization of Archaeal Thioredoxins. *Angew Chem Int Ed Engl.* 2007; 46(22) :4145-7
- Midelfort KS, **Hernandez HH**, Lippow SM, Tidor B, Drennan CL, Wittrup KD. Substantial energetic improvement with minimal structural perturbation in a high affinity mutant antibody. *J Mol Biol.* 2004 Oct 22;343(3):685-701.
- Barber MJ, Desai SK, Marohnic CC, **Hernandez HH**, Pollock VV. Synthesis and bacterial expression of a gene encoding the heme domain of assimilatory nitrate reductase. *Arch Biochem Biophys.* 2002 Jun 1;402(1):38-50.

ORAL PRESENTATIONS

- The Wednesday Seminars at Parson's, MIT, Cambridge, MA, 2008
- Massachusetts Institute of Technology Structural Group, MIT, Cambridge, MA 2007
- Society for the Advancement of Chicanos and Native Americans in Science, Tampa, FL, 2006
- Institute Leadership Dialogue, The Renewal of Research Universities: Building Capacities for Technological Innovation, MIT, Cambridge, MA 2006
- Boston DNA Repair and Mutagenesis Club, MIT, Cambridge, MA 2005
- 94th Annual Meeting of the American Society for Biochemistry and Molecular Biology Annual Meeting, San Diego, CA 2003

ABSTRACTS

- **Hernandez HH**, Elliott SJ, Drennan CL, Biochemical and Structural Characterization of thioredoxins from *Archaeoglobus fulgidus*, Annual Summer Symposium in Molecular Biology; Frontiers in Metallobiochemistry, 2006
- **Hernandez HH**, Brown II BA, Drennan CL, Cloning and expression of ribonucleotide reductase from *Archaeoglobus fulgidus*, ASBMB National Conference, 2003
- **Hernandez HH**, Brown II BA, Drennan CL, Cloning and expression of Ribonucleotide Reductase from Archaea Organisms, NIH Biotechnology Training Grant Retreat, 2002
- **Hernandez HH**, Brown II BA, Drennan CL, Cloning and expression of thioredoxin and thioredoxin reductase from *Thermoplasma acidophilum*, ACS National Conference, 2002
- **Hernandez HH**, Barber MJ, Expression and characterization of the heme domain of spinach nitrate reductase, Biophysics Conference, 2000
- **Hernandez HH**, Barber MJ, Modeling the cytochrome c reductase domain of spinach nitrate reductase, ASBMB, National Conference, 1999

FELLOWSHIPS

- NIH/NIGMS Ruth L. Kirschstein Pre-Doctoral Fellowship, MIT, 2004-2006

- NIH Biotechnology Training Program, MIT, 2001-2004
- NIH Bridges to the Baccalaureate Program, Valencia Community College, 1998
- Hispanic Scholarship Fund Undergraduate Award, Valencia Community College, 1996-1998

TEACHING AND MENTORING EXPERIENCE

2007 Massachusetts Institute of Technology Cambridge, MA

Teaching Assistant - 5.78 Biophysical Chemistry Techniques

- Taught and led project-based crystallography course for graduate students and advanced undergraduate students. The class focused on X-ray crystallography as a technique used in determining the 3-D structure of macromolecules. Discussion of crystallographic theory was complemented with exercises such as crystallization, data processing, and model building.

2000 - 2006 Massachusetts Institute of Technology Cambridge, MA

Undergraduate Research Opportunities Program (UROP) Research Mentor

- Trained and mentored MIT undergraduate students as an introduction to laboratory techniques including molecular biology, biochemistry concepts, and crystallography techniques. Guided students as they developed and completed a semester long projects. I have mentored five undergraduates over the course of four years as a researcher. My UROP students have gone on to post-graduate study at MIT, Cornel, and UCLA.

MIT Summer Research Summer Program (MSRP) Research Mentor

- Trained and mentored visiting students in molecular biology, introductory biochemistry concepts, and crystallography techniques. Guided students as they developed and completed a summer research project. I have mentored four undergraduates through the MSRP program. Subsequently two have applied and enrolled in graduate programs in chemistry: one at MIT and the other at Cornell University.

2004 - 2005 Massachusetts Institute of Technology Cambridge, MA

Biochemistry Instructor MITES Program

- To obtain independent, formal teaching experience, I designed, developed, and taught a general biochemistry course for economically disadvantaged high school students. Mentored students outside the classroom to facilitate their comprehension of biochemistry concepts. Lectures covered general titration and buffer equilibrium, amino and protein structures, carbohydrates, lipids, biological chemical reactions, enzyme kinetics, metabolic pathways, biochemical protocols, and DNA structure and function.

2000 - 2001 Massachusetts Institute of Technology Cambridge, MA

Teaching Assistant – 5.22 Biotechnology and Engineering

- Taught and mentored a case-study format course that examined the pharmaceutical industry. Students analyzed existing drugs and developed new uses and derivatives. Students were exposed to the process of obtaining patents, testing, evaluating, producing, and marketing a new drug. Emphasis was placed on teaching scientific and engineering principles, as well as the responsibilities scientists have for the consequences of their technologies.

Teaching Assistant – 5.07 Biochemistry

- Taught and reinforced detailed material covered in lectures regarding the structure of proteins, principles of catalysis, chemistry of organic and inorganic cofactors, and basic principles of metabolic and regulatory pathways within the cell.

SERVICE

2005 - 2007 Massachusetts Institute of Technology Cambridge, MA

Graduate Student Representative to the Faculty Policy Committee

- The Faculty Policy Committee sets the faculty agenda for MIT, and coordinates the work of all Faculty Committees. I provide a graduate student perspective on the programs and policies developed and implemented in Faculty Committees and by the Faculty Policy Committee.

2005 - 2006 Massachusetts Institute of Technology Cambridge, MA

Chair, Graduate Student Council (GSC) Task Force on Diversity

- The GSC Task Force on Diversity is charged with formulating the official Graduate Student Council statement on Diversity. This report provided a diversity policy framework that includes goals, program structure and outline, implementation timeline, and standards for evaluation and effectiveness.

2004 - 2005 Massachusetts Institute of Technology Cambridge, MA

Student Advisory Group to the Corporation Committee on the Presidency

- Facilitated open student discussions and feedback regarding their views on the challenges, opportunities, qualities, priorities, and nominations for the sixteenth president of MIT. Represented the student advisory group to the MIT Corporation.

Graduate Student Council Vice President

- Ensured and promoted MIT graduate student representation on Institute committees and governing boards. Provided graduated students with input and access to MIT decision-making processes, and interacted with faculty and administrators in advocating for student academic and extracurricular interests. Recruited graduate students to the position of departmental representatives to the Graduate Student Council, thus insuring proper representation of graduate student issues to Graduate Student Council.

Student Advisory Board to President Susan Hockfield

- Co-Vice Chair of joint graduate and undergraduate student group which elicited opinions from MIT students and then met for four lunches with President Hockfield to provide her with an overview of graduate and undergraduate student perspectives of academic, extracurricular, and social life at MIT. Was instrumental in the development of presentations and documents given to President Susan Hockfield.

AWARDS

- Recipient of the Undergraduate Research Opportunity Program Graduate Student Mentor Award, MIT, Spring 2005
- Recipient of the Dr. Martin Luther King Leadership Award, MIT, Spring 2003
- Excellence in Recitation Teaching by a Graduate Student Award, MIT, 2001
- ASBMB Undergraduate Poster Competition Winner, San Francisco, California,

1999

- Hispanic Scholarship Fund and Department of Energy Scholarship, 1998 -1999

PROFESSIONAL AFFILIATIONS

- American Association for the Advancement of Science (AAAS)
- American Chemical Society (ACS)
- American Crystallographic Association (ACA)
- American Society for Biochemistry and Molecular Biology (ASBMB)
- ASLO Minorities in the Aquatic Sciences Program
- Society for the Advancement of Chicanos and Native Americans in Science (SACNAS)

REFERENCES

Professor Catherine L. Drennan
Department of Chemistry
Massachusetts Institute of Technology
77 Massachusetts Avenue, Bldg 16-573
Cambridge, MA 02139
(617) 253-5622
cdrennan@mit.edu

Professor John M. Essigmann
William R. and Betsy P. Leitch Professor of Chemistry and Biological Engineering
Margaret MacVicar Faculty Fellow
Department of Chemistry
Massachusetts Institute of Technology
77 Massachusetts Avenue, Bldg 56-669
Cambridge, MA 02139
(617) 253-6227
jessig@mit.edu

Professor Bruce Tidor
Professor of Biological Engineering and Computer Science
CSBi Co-Director
Massachusetts Institute of Technology
77 Massachusetts Avenue, Bldg 32-212
Cambridge, MA 02139
(617) 253-7258
tidor@mit.edu

Professor Collin M. Stultz
W. M. Keck Assistant Professor of Biomedical Engineering
Electrical Engineering and Computer Science, Health Sciences
Massachusetts Institute of Technology

77 Massachusetts Avenue, Bldg 36-796
Cambridge, MA 02139