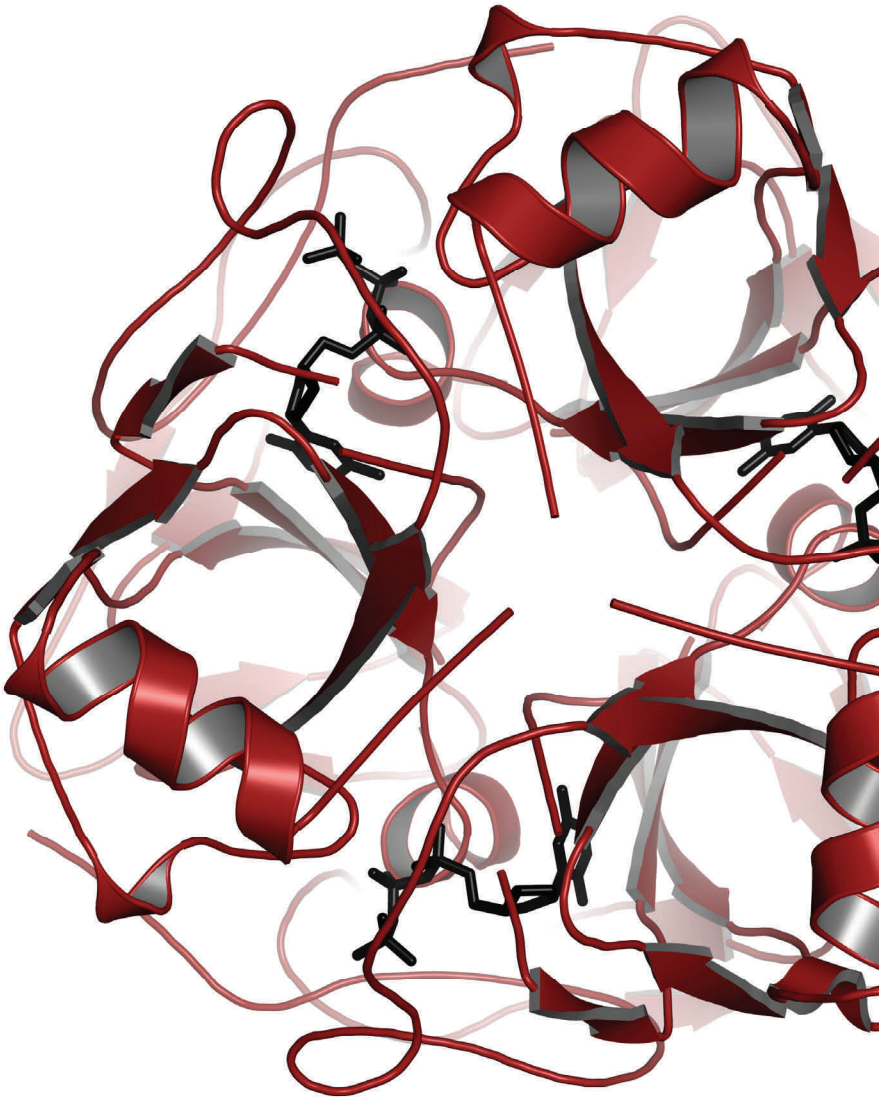
Structure and Characterization of Proteins and Enzymes Involved in Nucleotide Metabolism and Iron-Sulfur Proteins

Løvgreen, Monika Nøhr

Publication date:
2011

Document Version
Publisher's PDF, also known as Version of record

Structure and Charazterization of Proteins and Enzymes Involved in the Nucleotide Metabolism and Iron-Sulfur Proteins



Monika Nøhr Løvgreen

PhD Thesis, May 2011

Department of Chemistry
Technical University of Denmark

Structure and Characterization of Proteins and Enzymes involved in Nucleotide Metabolism and Iron-Sulfur Proteins

PhD Thesis

Monika Nøhr Løvgreen

May 2011

Department of Chemistry

Technical University of Denmark

Publications

Løvgreen MN, Harris P and Willemoës M. dTTP inhibition of the bifunctional dCTP deaminase-dUTPase from *Mycobacterium tuberculosis* is pH dependent: kinetic analyses and crystal structure of A115V variant. *Manuscript in preparation based on Chapter 9.*

Løvgreen MN, Martic M, Windahl MS, Christensen HEM and Harris P. Crystal structures of the all cysteinyl coordinated D14C variant of *Pyrococcus furiosus* ferredoxin: [4Fe-4S] ↔ [3Fe-4S] cluster conversion. *J. Biol. Inorg. Chem.* **(2011)** 16, 763-775 – available in Appendix G.

Løvgreen MN, Løvgreen M, Christensen HEM and Harris P. *MyCrystals* – a simple visual data management program for laboratory-scale crystallization experiments. *J. Appl. Cryst.* **(2009)** 42, 741-742 – available in Appendix G.

Conference contributions

Løvgreen MN, oral presentation, CoLuAa XIX crystallographers meeting. Formation of a disulfide bonded dimer of the all cysteinyl coordinated D14C variant of [3Fe-4S] *Pyrococcus furiosus* ferredoxin. Carlsberg Laboratory, 4-5 November **2010**.

Løvgreen MN, oral presentation, CoLuAa XVIII crystallographers meeting. Studies of WT and variants of the bifunctional dCTP deaminase-dUTPase from *Mycobacterium tuberculosis.* Novo Nordisk Bagsværd, 4-5 November **2009**.

Johannessen MN (Løvgreen MN), Nielsen MS, Harris P, Ooi BL and Christensen HEM. Abstract and poster, CoLuAa XVI crystallographers meeting. The crystal structure of the all cysteinyl coordinated D14C variant of [Fe_4S_4] *Pyrococcus fusiosus* ferredoxin. University of Aarhus, 8-9 November **2007**.

Front page illustration: Trimer of A115V Mycobacterium tuberculosis with dTTP bound. The figure was made using PyMol (by Delano WL, Schrödinger).

PREFACE AND ACKNOWLEDGEMENTS

This dissertation is submitted to the Technical University of Denmark in partial fulfillment of the requirements for the degree of PhD. The work was carried out at the Department of Chemistry, Technical University of Denmark, with Associate Professor Pernille Harris as main supervisor and Associate Professor Hans E. M. Christensen and the late Associate Professor Bee Lean Ooi as co-supervisors. Results are also included from my external stay at the Department of Biology, University of Copenhagen, which was made possible by Associate Professor Martin Willemoës. The PhD scholarship was funded by the Technical University of Denmark.

First of all, I would like to thank my supervisors for their help and guidance. A special thanks to Pernille Harris for her invaluable help, support and advice, she has been a great inspiration throughout this project. I am very grateful to Martin Willemoës for the opportunity to work at his laboratory and for including me in multidisciplinary studies of dCTP deaminase-dUTPase from *Mycobacterium tuberculosis* and especially for answering countless questions.

I would like to thank Torben Kibøl for his help in the laboratory, Bent Sigurskjold for running DSC, Hanne Nørgaard for running EPR, Maja Martic for her continued work on the ferredoxin variant and valuable scientific discussions. I thank ESRF and MAX-lab for allocating beam-time and Valerie Pye, Christian B. Oehlenschlæger and Pernille Harris are acknowledged for data collection at ESRF and MAX-lab.

I thank the members of X-ray and Protein Crystallography (DTU Chemistry) research for always being friendly and helpful and for providing a pleasant atmosphere. I have enjoyed Friday morning breakfasts and in particular your attentive nature when saving me a piece of cake. I also thank the members of the research groups Metalloprotein Chemisty and Engineering (DTU Chemistry) and Biomolecular Science (University of Copenhagen, Dept. of Biology) for providing pleasant working environments.

I would like to thank Anders Raffalt for his valuable help proofreading and joyful company at lunch. Also, I have appreciated the enjoyable company of my office-buddy Christian G. Frankær and the beneficial scientific discussions we have had. Finally, I would like to thank my family and friends for their love and support; a very special thanks to my incredibly patient and loving husband Mikkel Nøhr Løvgreen.

Monika Nøhr Løvgreen, Lyngby, May 24th 2011

ABSTRACT

This thesis presents work on two projects: the all cysteinyl coordinated D14C variant of the hyperthermostable *Pyrococcus furiosus* (*Pf*) ferredoxin (Fd) as well as wild-type (WT) and two variants of the bifunctional dCTP deaminase-dUTPase (DCD-DUT) from *Mycobacterium tuberculosis* (*Mt*). Furthermore, a program named *MyCrystals* has been developed to keep track of crystallization trials and results. The program combines pictures with crystallization conditions and is able to sort the pictures based on selected conditions. *MyCrystals* was used extensively throughout this work and allows for an overview of the crystallization results through the use of databases.

Changing the cluster coordinating aspartate to cysteine in *Pf* Fd proved to impair the ease with which the $[Fe_4S_4]$ cluster converted to the $[Fe_3S_4]$ cluster. A disulfide bonded dimer was observed at pH 8.0, whereas only the monomer was present at pH 5.8. The crystal structure of D14C $[Fe_3S_4]$ is the first structure with a $[Fe_3S_4]$ cluster, in which a cysteine from a full cysteine binding motif is unprotected and facing away from the cluster. The structure is in close resemblance with the WT $[Fe_3S_4]$ structure. Crystal packing in both D14C and WT $[Fe_3S_4]$ Fd shows extended β-sheet dimers. These dimers were not observed in solution and were likely a result of the high protein concentration in the crystals.

WT, A115V and A115G *Mt* DCD-DUT were successfully purified, and the crystal structure of the A115V variant with dTTP bound was solved. The variants were created to investigate the importance of steric hindrance on a water molecule suggested to play a key role in dephosphorylation. However, this water molecule was present in the structure of A115V:dTTP and the variant did not dephosphorylate dTTP. The dTTP pyrimidine moiety in the WT and A115V structures is rotated compared with the pyrimidine moiety of dUMPNPP in the structure of *Mehtanocaldococcus jannaschii* (*Mj*) DCD-DUT. This causes changes in the hydrogen bonding pattern of conserved residues in the active site and may give rise to less stabilization of the negative charge formed on the oxygen bridging the α-β-phosphorous of the nucleotide in the course of the dephosphorylation reaction. The flexible region consisting of residues 110-118 in the structure of A115V:dTTP is in close resemblance with the active conformation seen in *Mj* DCD-DUT:dUMPNPP. A115V *Mt* DCD-DUT was unable to bind dTTP in the inactive conformation due to steric hindrance caused by the introduced valine side chain. In contrast to the A115V:dTTP structure, an increased number of hydrogen bonds in the WT:dTTP structure favors dTTP binding in the inactive conformation. The reduced number of hydrogen bonds in A115V:dTTP may explain its reduced thermal stability compared with WT:dTTP.

dTTP inhibition of WT *Mt* DCD-DUT at pH 6.8 was confirmed, whereas the WT enzyme proved insensitive to dTTP at pH 8.0. The protonation state of the conserved His112 in the flexible loop is likely to play an important role herein. His112 is completely deprotonated at pH 8.0, where it is stabilized in the active conformation. The active conformation for the WT enzyme is likely to

be the same as seen in *Mj* DCD-DUT:dUMPNPP, and this conformation does not allow dTTP binding because of steric hindrance. Hyperbolic dCTP and dUTP saturation curves support that the WT enzyme was present solely in the active conformation at pH 8.0.

The A115V variant was inhibited by dTTP at both pH 6.8 and 8.0, albeit with a stronger inhibition at pH 8.0. The flexible 110-118 loop in the A115V:dTTP structure is more loose compared with *Mj* DCD-DUT:dUMPNPP. This wider loop could allow dTTP binding even with protonated His112. dTTP binds more easily at pH 8.0 because completely deprotonated His112 takes up less space near the nucleotide binding site.

The A115G variant showed an opposite pH effect of dTTP inhibition compared with the WT enzyme. A115G was very sensitive to dTTP at pH 8.0, while no substantial inhibition was observed at pH 6.8. Changing alanine to glycine, which has additional allowed backbone conformations, is likely to introduce additional flexibility to the already flexible 110-118 loop. Hence, dTTP binding at pH 8.0 could simply be explained by the less restrained structure. In contrast, the lack of inhibition at pH 6.8 for the A115G variant could be caused by the protonated histidine inducing stability by binding to Gly115 O and thereby structurally arranging the very flexible loop. At pH 8.0, the A115G variant shows sigmoidal dCTP saturation and hyperbolic dUTP saturation. This distinction between dCTP and dUTP as substrates could be related to Ala115 being involved in dCTP deamination in the WT enzyme. The flexible Gly115 backbone of the A115G variant may require adjustment for the deamination to take place, whereas the conformation of this residue is indifferent for dephosphorylation.

DANSK RESUMÉ

Denne afhandling omfatter arbejde på to projekter: D14C varianten af *Pyrococcus furiosus* (*Pf*) ferredoxin med fuldt cystein bindingsmønster samt vildtype (WT) og to varianter af den bifunktionelle dCTP deaminase-dUTPase (DCD-DUT) fra *Mycobacterium tuberculosis* (*Mt*). Endvidere blev programmet *MyCrystals* udviklet for at holde rede på krystallisationsforsøg og resultater. Programmet kombinerer billeder med krystallisationsbetingelser og er i stand til at sortere billederne ved valg af givne betingelser. *MyCrystals* blev brugt i vid udstrækning i dette arbejde og programmet giver et overblik over krystallisationsresultaterne ved hjælp af databaser.

Udskiftning af en klyngekoordinerende aspartat til cystein i *Pf* ferredoxin gjorde det sværere at omdanne klyngen fra [Fe_4S_4] til [Fe_3S_4]. En dimer bundet sammen af en disulfidbro blev observeret ved pH 8.0, mens kun monomeren fandtes ved pH 5.8. Krystalstrukturen af D14C [Fe_3S_4] er den første struktur med en [Fe_3S_4] klynge, hvori en fri cystein fra et fuldt cystein bindingsmønster peger væk fra klyngen. Strukturen er meget lig WT [Fe_3S_4] strukturen. Krystalpakningen af både D14C og WT [Fe_3S_4] Fd er som β-sheet dimerer. Disse dimerer blev ikke observeret i opløsning og var sandsynligvis et resultat af den høje proteinkoncentration i krystallerne.

WT, A115V og A115G *Mt* DCD-DUT blev oprenset og krystalstrukturen af A115V med dTTP bundet blev løst. Varianterne blev lavet for at undersøge betydningen af et vandmolekyle, der måske havde en vigtig rolle i defosforyleringsreaktionen. Dette vandmolekyle var dog til stede i A115V:dTTP strukturen og varianten var ikke i stand til at defosforylere dTTP. Pyrimidindelen af dTTP i WT og A115V strukturerne er roteret i forhold til pyrimidindelen af dUMPNPP i *Methanocaldococcus* jannaschii (*Mj*) DCD-DUT. Denne drejning forårsager ændringer i hydrogenbindingsmønsteret for konserverede rester i det aktive center og kan muligvis medføre mindre stabilisering af den negative ladning, som dannes på oxygenet, der forbinder α-β-fosfaterne i nukleotidet, i løbet af defosforyleringsreaktionen. Den fleksible region, der udgøres af aminosyreresterne 110-118, i strukturen af A115V *Mt* DCD-DUT med dTTP bundet er meget lig den aktive konformation af *Mj* DCD-DUT:UMPNPP. A115V varianten var ikke i stand til at binde dTTP i den inaktive konformation på grund af sterisk hindring fra valinsidekæden. Et højere antal hydrogenbindinger i WT:dTTP strukturen favoriserer dTTP binding i den inaktive konformation i modsætning til A115V:dTTP strukturen. Det reducerede antal hydrogenbindinger i A115V:dTTP kan forklare dennes lavere termostabilitet sammenlignet med WT:dTTP.

dTTP inhibering af WT *Mt* DCD-DUT ved pH 6.8 blev påvist, hvorimod WT enzymet var upåvirket af dTTP ved pH 8.0. Protoneringen af den konserverede His112 i det fleksible loop har sandsynligvis en vigtig rolle heri. His112 er fuldstændigt deprotoneret ved pH 8.0, hvor den er stabiliseret i den aktive konformation. Den aktive konformation for WT enzymet er sandsynligvis magen til konformationen af *Mj* DCD-DUT:dUMPNPP, og denne konformation tillader ikke dTTP

binding på grund af sterisk hindring. Hyperbolske dCTP og dUTP mætningskurver underbygger at WT enzymet udelukkende findes i den aktive konformation ved pH 8.0.

A115V varianten var inhiberet af dTTP både ved pH 6.8 og ved 8.0, om end stærkest inhiberet ved pH 8.0. Det fleksible 110-118 loop er løsere i A115V:dTTP strukturen end i *Mj* DCD-DUT:dUMPNPP. Dette bredere loop kan tillade dTTP binding selvom His112 er protoneret. dTTP har lettere ved at binde ved pH 8.0, idet en fuldstændig deprotoneret His112 optager mindre plads nær nukleotidbindingssitet.

A115G varianten viste en modsat pH effekt af dTTP inhibering i forhold til WT enzymet. A115G var meget sensitiv overfor dTTP ved pH 8.0 og overvejende upåvirket ved pH 6.8. Udskiftning af alanin til glycin, der har flere tilladte hovedkædekonformationer, vil sandsynligvis tilføje yderligere fleksibilitet til det i forvejen fleksible 110-118 loop. En simpel forklaring på dTTP binding ved pH 8.0 er, at der er mere plads i strukturen. Den modsatte mangel på dTTP inhibering ved pH 6.8 for A115G er muligvis forårsaget af den protonerede histidin, der kan give mere stabilitet i strukturen ved at binde til Gly115 O og derved skabe en strukturel begrænsning i det meget fleksible loop. Ved pH 8.0 udviser A115G varianten sigmoid dCTP mætning og hyperbolsk dUTP mætning. Denne skelnen mellem dCTP og dUTP som substrater kan skyldes, at Ala115 i WT enzymet er involveret i dCTP deaminering. Det er muligvis nødvendigt strukturelt at tilpasse positionen af det fleksible Gly115 backbone i A115G varianten for at deaminering kan finde sted, hvorimod konformationen af denne rest ikke er vigtig for defosforylering.

List of Abbreviations

A115G	A115G variant of *Mycobacterium tuberculosis* dCTP deaminase-dUTPase
A115V	A115V variant of *Mycobacterium tuberculosis* dCTP deaminase-dUTPase
A media	2 % $(NH_4)_2SO_4$, 7.5 % $Na_2HPO_4 \cdot 2H_2O$, 3 % KH_2PO_4, 3 % NaCl
ASU	Asymmetric unit
BisTris	Bis-(2-hydroxyethyl)imino-tris(hydroxymethyl)methane
CV	Column volume
DCD	dCTP deaminase
DCD-DUT	dCTP deaminase-dUTPase
dCTP	deoxycytidine triphosphate
Dd	*Dictyostelium discoideum*
DE52	Diethylaminoethyl
Dg	*Desulfovibrio gigas*
dUMP	deoxyuridine monophosphate
dUMPCPP	2'-deoxy-uridine-5'-[(α,β)-methyleno]triphosphate
dUMPNPP	2'-deoxy-uridine-5'-[(α,β)-imido]triphosphate
dTTP	deoxythymidine triphosphate
DUT	dUTPase
dUTP	deoxyuridine triphosphate
EDTA	2,2',2'',2'''-(Ethane-1,2-diyldinitrilo)tetraacetic acid
Ec	*Escherichia coli*
EPR	Electron paramagnetic resonance
ESRF	European Synchrotron Radiation Facility (Grenoble, France)
Fd	Ferredoxin
FeS	Iron-Sulfur
HEPES	N-(2-hydroxyethyl)-piperazine-N'-2-ethanesulfonic acid
HPLC	High-pressure liquid chromatography
IPTG	Isopropyl β-D-1-thiogalactopyranoside
LB media	1 % tryptone, 0.5 % yeast extract, 0.5 % NaCl
MAX-lab	National electron accelerator laboratory for synchrotron radiation research, nuclear physics and accelerator physics (Lund University, Sweden)
MIB	Buffer system containing malonic acid, imadizole and boric acid
Milli-q	18.2 MΩ cm water, Milli-Q Synthesis system by Millipore
MPD	2-Methyl-2,4-pentanediol
Mj	*Methanocaldococcus jannaschii*
MPEG	Polyethylene glycol monomethyl ether
Mt	*Mycobacterium tuberculosis*
NCS	Non-crystallographic symmetry
NDK	Nucleoside diphosphate kinase

NMR	Nuclear magnetic resonance
NMWL	Normal molecular weight limit
OD	Optical density; wavelength given as subscript
Pdb	Protein Data Bank: www.pdb.org
PEG	Polyethylene glycol
Pf	*Pyrococcus furiosus*
QTOF	Quadrupole Time of Flight
RMS	Root mean square
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
Ss	*Sulfolobus solfataricus*
TB/carb	Terrific Broth with 0.5 % (w/w) glycerol and 50 µg/mL carbenicillin
TLS	Translation Libration Screw-motion
Tris	2-amino-2-(hydroxymethyl)-1,3-propandiol
UV-vis	Ultraviolet-visible
WT	Wild-type

CONTENTS

Chapter One

1 Introduction and Outline

Proteins participate in virtually every process within cells and are an essential part of life, as they are required for structure, function and regulatory purposes. Enzymes are proteins that catalyze biochemical reactions. They are also of vital importance, as they catalyze the majority of the reactions involved in metabolism and are also involved in transcription, DNA-replication and -repair. The three-dimensional structure of proteins and enzymes is absolutely essential for the function, and structure determination is thus a valuable method for discovering shared features among members of molecular families. Detailed three-dimensional protein structures can be determined using NMR spectroscopy and X-ray crystallography, the latter of which was used throughout this work.

Successful protein crystallization is required in order to solve the crystal structure by X-ray diffraction. This can be an immense task involving numerous crystallization conditions, and meticulous bookkeeping of the results is important in order to optimize the conditions. Chapter 2 presents a database program named *MyCrystals* that was developed for managing and storage of such information. The unique feature of *MyCrystals* is the option to sort crystallization conditions coupled with a picture. *MyCrystals* is available online and is also enclosed in Appendix A along with databases containing crystallization conditions and pictures of protein crystallization results throughout this work.

Metalloproteins are crucial for many essential processes and exist extensively throughout nature. Ferredoxin (Fd) from *Pyrococcus furiosus* (*Pf*) is hyperthermostable and contains a single [Fe_4S_4] cluster with incomplete cysteinyl coordination [1,2]. Studies of the D14C variant ferredoxin, in which the cluster coordinating aspartate has been changed to cysteine, may contribute to the understanding of cysteine ligation and its effects on protein function and stability. The D14C variant of *Pf* Fd is introduced in Chapter 3 along with a brief introduction to iron-sulfur proteins. Conversion of [Fe_4S_4] to [Fe_3S_4] is a preceding step in the formation of heterometallic clusters, in which an externally imposed metal ion is part of the cluster [MFe_3S_4]. Formation of novel metalloproteins presents an opportunity to engineer new or improve existing functional properties with relevance for a wide range of applications within the pharmaceutical- and biotechnological industries [3]. Structural studies of metalloproteins are important because the structure-function relationship is vital for protein design. Expression and purification of D14C [Fe_4S_4] *Pf* Fd as well as cluster conversion to [Fe_3S_4] and subsequent purification are described in Chapter 4. Crystallization and crystal structures of D14C *Pf* Fd are described in Chapter 5.

Building blocks for DNA and RNA are in continuous demand, and (deoxy)ribonucleotide synthesis is thus vital for cell growth. During evolution, the base pairing in DNA changed from including the uracil base (found in RNA) to include the thymine base. Studying the synthesis of deoxythymidine triphosphate (dTTP) may provide clues as to how this evolution transpired. The bifunctional enzyme dCTP deaminase-dUTPase (DCD-DUT) is involved in the biosynthesis of dTTP and is described in Chapter 6. Kinetic and inhibition studies can be used to characterize enzymes in order to gain an understanding of the function. In this work, the bifunctional DCD-DUT from *Mycobacterium tuberculosis* (*Mt*) has been studied. The infectious bacterial disease tuberculosis caused by *Mt* killed an estimated 1.7 million people in 2009 [4], and enzymes involved in deoxyribonucleotide synthesis provide potential drug targets. Structures and inhibition mechanisms of these enzymes are important in the design of specific inhibitors that leave the human host unaffected. Wild-type (WT) and two variants of the bifunctional *Mt* DCD-DUT have been studied in this work. An introduction to the variants and the applied equations for kinetic analyses is given in Chapter 7. Purification of the enzymes is described in Chapter 8, followed by structural and kinetic studies of the enzymes in Chapter 9.

Numerous figures of protein structures are shown throughout this thesis, both in chapters outlining the literature and in experimental chapters describing the obtained results. All these figures have been made using PyMol [5], occasionally in combination with other programs, which will be stated explicitly in the figure legends.

Chapter Two

2 *MyCrystals* – A Simple Visual Data Management Program for Laboratory Scale Crystallization Experiments

2.1 Introduction

Macromolecular structure determination by X-ray diffraction requires successful crystallization. This can be a difficult task where numerous crystallization experiments lead to a large amount of data. Crystallization projects can be worked on by several people over longer periods of time and keeping track of, as well as sharing, results can be made much easier by the use of a database. A program named *MyCrystals* has been developed with the main purpose of managing and storing pictures and crystallization conditions, as well as sorting them to give an overview of the results. A program description is available in this chapter and a descriptive note of the program has been published [6].

There are many other available programs and applications to store crystallization data. Crystallization robots usually include commercial software that is able to handle large amount of data. Some of the available free programs are *XtalBase* [7], *XAct* [8], *Xtrack* [9], *CLIMS* [10] and *LISA* [11]. Images combined with scores present a simple way to assess crystallization results and also simplifies the optimization procedure because of the transparency of the results. *XtalBase* [7] is a comprehensive system that amongst many other features can manage and display pictures of the results, and *CLIMS* [10] is able to store visual observations of the drops. Laboratory Information Management Systems like *PiMS* [12] are designed to manage various kinds of data such as target, construct and experiment data through a web interface.

MyCrystals is simple to use, easy to navigate, and requires minimal time investment by the researcher to utilize the software. The unique feature of *MyCrystals* is its ability to display and sort crystallization pictures. The pictures are stored locally or on external media, and the program creates a database with file paths to identify the pictures. *MyCrystals* is also able to sort pictures by the entered data to view the effect of e.g. changing the pH or precipitant. This makes it easier to establish the best conditions and to assess how they can be optimized. *MyCrystals* can also be used as a notebook program to keep track of attempted crystallization conditions without necessarily using the picture viewing part. Furthermore, *MyCrystals* may be useful in other experimental optimization procedures, such as purification, where a lot of slightly different

procedures are usually tested. These may be stored in the database and linked to an image file containing the resulting chromatogram.

2.2 Program Description and Use

MyCrystals is primarily designed for viewing and sorting pictures in a simple and intuitive user interface. The *File* menu is used for creating new databases and loading or saving existing databases. Creating a new database in the file menu will auto-generate a standard set of conditions. The user interface of *MyCrystals* is divided into three simple main tabs: *Edit Data*, *Filter Data* and *Data Table*. Data entry and edit are carried out in the *Edit Data* tab as described in section 2.2.2. Picture display and selection based on conditions are carried out in *Filter Data* as described in section 2.2.3. *Data Table* lists all entered data in a table to give an overview of the contents of the database.

2.2.1 Data Structure

The data consist of two parts. The first part is an XML-database that contains all entered conditions including the appropriate image path for each set of conditions (data entries). The second part is the images, which are placed in a directory on the hard drive. *MyCrystals* recognizes graphics file formats JPEG, GIF, PNG and TIFF.

The full path to each image consists of two paths, a *Project Directory* and a relative *Image Path*. The *Image Path* is unique for each data entry while the *Project Directory* is shared. This arrangement facilitates the moving of all images to another location, as only the *Project Directory* needs to be changed.

2.2.2 Edit Data

In the *Edit Data* tab, the database can be edited by adding new data entries or editing already existing data entries. The tab consists of three parts: *Project Directory*, *Column Properties* and *Data Entry*.

The *Project Directory* is the path to a directory where all the images used in the database are placed. The rest of the path is stored as a property (*Image Path*) for each data entry. The *Column Properties* makes it possible to change the data columns of the database. It is possible to delete or add columns, change the names and types, and switch the positions of two columns. An intuitive user interface renders this possible, see Figure 2.1.

The *Data Entry* area is used to add, change or remove data entries in the database. The *Selected Data Entry* operates on an existing data entry. To add a new data entry into the database, use the *Add Data Entry* button, which creates a new data entry with a consecutive index number.

The data entry fields are empty when the first data are typed in, and after *Add Data Entry* has been used, the entry fields will contain the information of the latest data entry. Drop down menus of previously entered data are available when entering or correcting database entries. The *Selected Data Entry* can advantageously be used when entering data, as an entire set of previously entered conditions can be loaded by giving the index number. Minor changes can be applied to the set of conditions and *Add Data Entry* will generate a new data entry in the database.

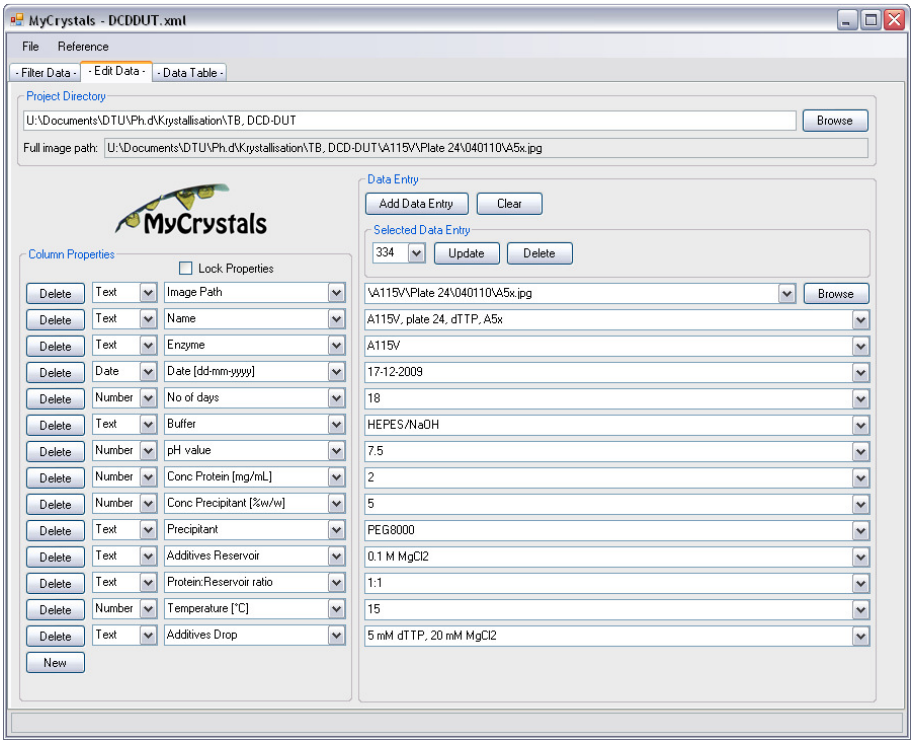


Figure 2.1: A screenshot showing how the Edit Data tab of *MyCrystals* can appear. Checking the *Lock Properties* option will simplify the *Column Properties* by only showing the name.

2.2.3 Filter Data

The *Filter Data* tab is used for viewing and sorting pictures entered into a database. The sorting options are available in a drop down menu and consist of all entered data. First, a condition is selected and then a filter option is applied. Text strings can be sorted as absolutes, while numerical strings also can be sorted using >, <, ≤, ≥, ≠. It is possible to apply 15 filters simultaneously. A list of entries that match the criteria will be generated and displayed in the tab, see Figure 2.2.

The information shown for each data entry in the list can be customized using the field *Filter text*. Writing '*[1]*' will list the data entered in the first field in the *Data Entry* area of the *Edit Data* tab, '*[2]*' will list the data in the second field and so forth. Selecting one of the listed entries will display the matching picture to the right of the tab as shown in Figure 2.2.

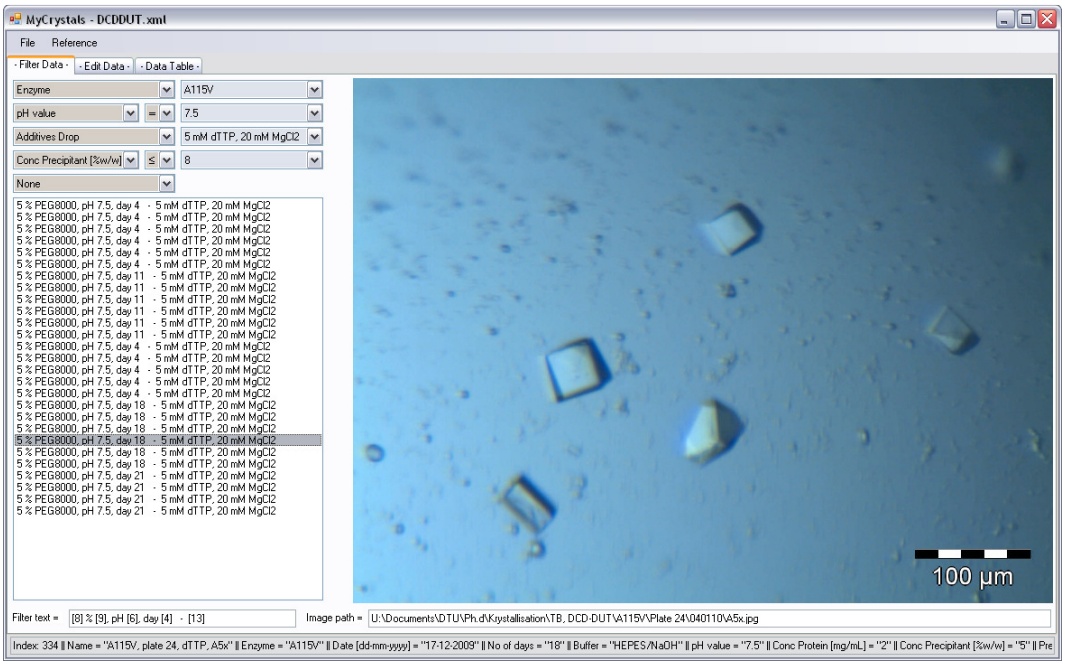


Figure 2.2: A screenshot showing how the Filter Data tab of *MyCrystals* can appear. Data with the following conditions have been sorted: *Enzyme:* A115V, *pH* = 7.5, *Additives Drop:* 5 mM dTTP, 20 mM $MgCl_2$ and *Precipitant concentration* ≤ 8 (w/w%). In the displayed image: 5 % PEG8000. The scale-bar is part of the depicted image.

2.3 Platforms and Availability

The program is written in C# as a Windows form in Microsoft Visual Studio. The program consists of a single executable file, and no installation is needed. *MyCrystals* runs under Windows and has no special hardware requirements. *MyCrystals* is available free of charge for download from http://xray.kemi.dtu.dk/mycrystals/ and is enclosed on DVD in Appendix A. Furthermore, the DVD includes three databases and matching pictures of crystallization results carried out as described in Chapters 5 and 9, and in Appendix F.

2.4 Conclusion

The unique main feature of the program *MyCrystals*, presented here, is the displaying and sorting of crystallization pictures based on the entered crystallization conditions. This enables the user to view the effect of changing conditions, which helps to identify parameters for optimization.

Chapter Three

3 Iron-Sulfur Proteins in Brief

3.1 Introduction to Iron-Sulfur Proteins

Iron-sulfur (FeS) proteins are present in all kinds of organisms and cellular compartments. They are predominantly involved in electron transfer, but also in substrate binding, catalysis, regulatory and sensing functions [13-15]. FeS proteins are proposed to be ancient structures that may have played a role during the emergence of life on this planet [16,17]. The formation of FeS clusters in a living cell is catalyzed by various complex biosynthetic machineries and a range of diseases are associated with defects in FeS proteins in humans [15].

The predominant product in autoassembly involving iron, sulfide and thiolates is the $[Fe_4S_4]$ cluster [18]. The stability of FeS clusters in proteins depends on the protein surroundings. The preferred organic ligands of FeS are cysteines, but also histidine, glutamine or arginine can be used as ligands. FeS proteins are in general present in a single protein domain and have a reasonably rigid ligand framework. Exposed FeS clusters are uncommon and very unstable. Examples of large and specialized FeS clusters are the $[Fe_8S_7]$ (P cluster) and $[MoFe_7S_9]$ (FeMoco) clusters of nitrogenase, the Fe_2-$[Fe_4S_4]$ active site of [FeFe] hydrogenases and the Ni_2-$[Fe_4S_4]$ (A cluster) and $NiFe_4S_5$ (C cluster) clusters of CO dehydrogenase/acetylcoenzyme A synthase [13,19]. Figure 3.1 shows simple and more complex FeS cluster structures.

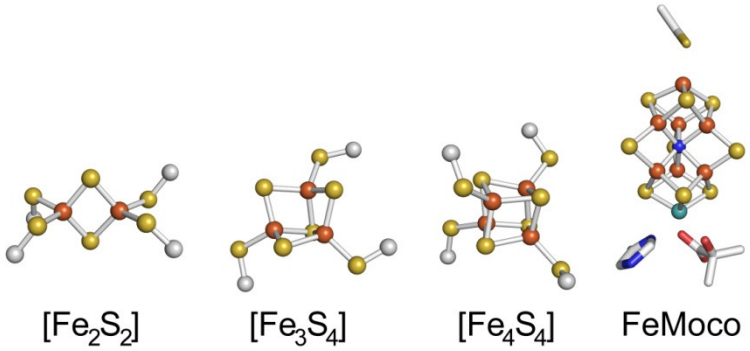


Figure 3.1: Four examples of iron-sulfur clusters: $[Fe_2S_2]$ (pdb 1RFK). $[Fe_3S_4]$ (pdb 1SJ1). $[Fe_4S_4]$ (pdb 2Z8Q). $[MoFe_7S_9]$ (pdb 1M1N).

The redox potentials of FeS proteins have a very large span of about 1 V. The $[Fe_4S_4]$ ferredoxin cluster changes oxidation state from +2 to +1 during reduction, while the $[Fe_3S_4]$ cluster changes from +1 to 0 during reduction [14]. High-potential iron-sulfur proteins (HiPIPs) are small globular proteins (6-10 kDa) that act as electron donors. They contain a $[Fe_4S_4]^{2+/3+}$ cluster bound to four conserved cysteine residues and are distinguished by high reduction potentials [13,20].

The electrochemical properties of ferredoxins are sensitive to changes in the cluster environment. More positive reduction potentials are expected for less buried clusters because the reduced cluster including ligands is more charged than the oxidized cluster, $[Fe_4S_4(Cys^-)_4]^{3-/2-}$. Increased solvent accessibility of the cluster thus increases the reduction potential [21-23]. More and stronger hydrogen bonds to cluster coordinating cysteine Sγ increase the reduction potential, most likely by stabilizing the reduced, more charged cluster by attenuating the charge density [22,24]. Other effects discussed to influence the reduction potential include ligand conformation, hydration effects and volume of the cavity hosting the cluster [22,25,26].

3.2 $[Fe_4S_4] \leftrightarrow [Fe_3S_4]$ Cluster Interconversions

Aconitase is a widely used example of incomplete cysteinyl coordination in which the cluster conversion plays a crucial role for the enzyme activity. The inactive enzyme contains a $[Fe_3S_4]^+$ cluster coordinated to the protein by three cysteines. The enzyme is active when a fourth iron coordinated by hydroxyl is incorporated to complete the cubane $[Fe_4S_4]^{2+}$ cluster [27-29].

The $[Fe_4S_4]$ cluster is a common cluster in bacterial ferredoxins and is bound to the polypeptide chain with a typical binding motif: -Cys-X_2-Cys-X_2-Cys-X_n-Cys-. Common single cluster ferredoxins have two additional cysteines forming an intramolecular disulfide bond, which replace a second cluster in more ancient ferredoxins [30]. Incomplete cysteine coordination of a $[Fe_4S_4]$ cluster is seen in *Pf* ferredoxin and in one of two $[Fe_4S_4]$ clusters in *Desulfovibrio africanus* (*Da*) ferredoxin III, which both readily convert to $[Fe_3S_4]$ clusters [2,31]. Site-directed change of a cluster coordination aspartate to cysteine has been reported to impair the ease with which $[Fe_4S_4]$ converts to $[Fe_3S_4]$ [31].

Formation of a $[Fe_3S_4]$ cluster is a preceding step to the formation of heterometallic clusters in which a fourth exogenous metal (M) is incorporated into the cluster $[MFe_3S_4]$. Protein templates used for this type of cluster redesign include ferredoxins from *Pf, Da* and *Desulfovibrio gigas* [32,33]. $[MFe_3S_4]$ *Pf* Fd has been studied with M = Cr^{2+}, Mn^{2+}, Zn^{2+}, Ni^{2+}, Co^{2+}, Cd^{2+}, Cu^+ and Tl^+ [34-38]. Design or modification of metal binding sites in proteins opens for the possibility to engineer new or improve existing functional properties. This can contribute to the understanding of structure-function relationships and result in new enzymes for a wide range of applications [3].

3.3 *Pyrococcus furiosus* Ferredoxin and the D14C variant

Ferredoxin from the strict anaerobic hyperthermophilic archaeon *Pyrococcus furiosus* functions as an electron carrier [1]. The 7.5 kDa ferredoxin consists of 66 amino acids and includes a single $[Fe_4S_4]$ cluster coordinated to the protein by three cysteine residues (residues 11, 17, 56) and one aspartate (residue 14) [2,39]. The iron coordinated to Asp14 is lost during a reversible cluster conversion to $[Fe_3S_4]$. Reassembly of the $[Fe_4S_4]$ cluster occurs during incubation with Fe^{2+} under reducing conditions [2]. Two additional cysteine residues (21, 48) form a redox active

disulfide bond [40]. The disulfide bond exists in a double conformation with a left-handed spiral conformation destabilized compared to the right-handed spiral conformation [39,41,42].

The crystal structure of $[Fe_3S_4]$ *Pf* Fd was solved by Nielsen *et al.* [41,43] (pdb 1SJ1). In the crystal structure, the ferredoxin forms an extended β-sheet dimer. The dimer has been proposed to be present at low ionic strength, while monomers are present at high ionic strength [44,45]; a behavior, which has been reported for *Pyrococcus woesei* ferredoxin [46].

Structural, electronic and functional characteristics of incomplete cysteinyl coordination have been studied extensively in *Pf* Fd by replacing Asp14 with cysteine [47-52]. D14C $[Fe_4S_4]$ *Pf* Fd show the same extreme thermostability as WT *Pf* Fd with no observed denaturation after 12 h at 95 °C [1,50]. Differences between D14C and WT *Pf* Fd are seen as an increase in the absorbance ratio, A_{390}/A_{280} = 0.73 for D14C compared with 0.56 for WT (UV-vis spectra are shown in Figure 3.2) and an increase in the extinction coefficient, ε_{390} = 20.2 $mM^{-1}cm^{-1}$ for D14C compared with 17.0 $mM^{-1}cm^{-1}$ for WT [48,49]. Also, D14C $[Fe_4S_4]$ *Pf* Fd has a more negative redox potential (-427 mV) compared with WT $[Fe_4S_4]$ *Pf* Fd (-368 mV) [50]. The efficiency of electron acceptance is also affected by the mutation, as the D14C variant has a decreased V_{max} compared with WT *Pf* Fd [50]. A detailed description of expression, purification and properties of WT, WT-recombinant and several mutants of *Pf* Fd is given by Kim *et al.* [53].

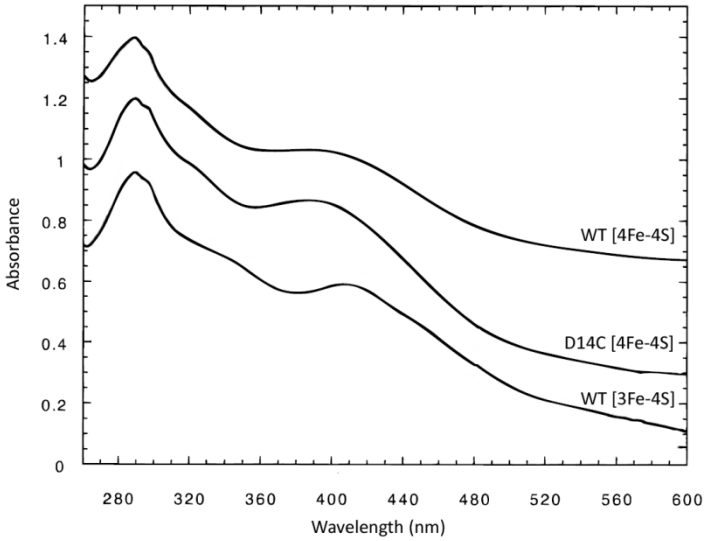


Figure 3.2: UV-vis spectra of *Pf* ferredoxins: WT and D14C [4Fe-4S] and WT [3Fe-4S]. Adapted with permission from [50]. Copyright 2011 American Chemical Society.

Chapter Four

4 Oxidation and Purification of the D14C variant of *Pyrococcus furiosus* Ferredoxin

4.1 Introduction

Oxidation of $[Fe_4S_4]$ ferredoxins to $[Fe_3S_4]$ ferredoxins is a preceding step to the formation of heterometallic clusters $[MFe_3S_4]$. D14C *Pf* Fd contains a $[Fe_4S_4]$ cluster and the difference in oxidation to $[Fe_3S_4]$ compared with the WT Fd may provide information on the effects of complete cysteinyl coordination. Aerobic purification of WT *Pf* Fd produces two peaks with 2 mS/cm elution difference during anion exchange on a Source 30Q column; first the Fd with a $[Fe_4S_4]$ cluster followed by the Fd with a $[Fe_3S_4]$ cluster [54]. Differences are seen in the characteristics of the UV-vis spectra of WT $[Fe_4S_4]$ and $[Fe_3S_4]$ *Pf* Fd shown in Figure 3.2. The absorbance maximum moves from 390 nm to 408 nm and a shoulder around 340 nm becomes more pronounced when the cluster in WT *Pf* Fd is converted from $[Fe_4S_4]$ to $[Fe_3S_4]$ [50].

Prior to this project[1], oxidation of purified D14C $[Fe_4S_4]$ *Pf* Fd was attempted, but proved unsuccessful (section 4.1.1) and several experiments were made to oxidize the Fd (section 4.2.3). Cultivation and purification of D14C *Pf* Fd are also described in this chapter.

All solutions were prepared with 18.2 MΩ cm water from a Milli-Q Synthesis system by Millipore (milli-q water). Protein purifications were carried out on an ÄKTA™ Purifier 100 HPLC system. Ultrafiltration was carried out at 4 °C in a stirred Amicon cell with a PLBC NMWL 3 kDa membrane.

4.1.1 Initial Oxidation Procedure and Results

The initial oxidation procedure presented in this section was carried out prior to this project, following the procedure described in reference [47] with minor modifications.

In experiment A, 15x molar excess of $K_3[Fe(CN)_6]_3$ was used. The solution with ferricyanide and 10 mg D14C *Pf* Fd in 20 mM Tris/HCl pH 8.0 was incubated at room temperature for 15 minutes with gentle stirring. The protein solution was exchanged into 20 mM Tris/HCl pH 8.0 by ultrafiltration and then loaded onto a 16 mm/10 cm Source 30Q column equilibrated with 20

[1] Monika Nøhr Johannessen (Løvgreen) Master's Thesis, Department of Chemistry, Technical University of Denmark, October 2007

mM Tris/HCl pH 8.0. The column was washed with 2 column volumes (CV) 20 mM Tris/HCl pH 8.0 and elution was carried out using a linear salt gradient of 0.15 M to 0.4 M NaCl in 20 mM Tris/HCl pH 8.0 over 17 CV. The chromatogram is shown in Figure 4.1 along with the UV-vis spectra of protein in the main *peak a* and D14C [Fe_4S_4] *Pf* Fd for comparison. Protein from the main *peak a* was collected and crystallized. The structure proved to contain a [Fe_4S_4] cluster.

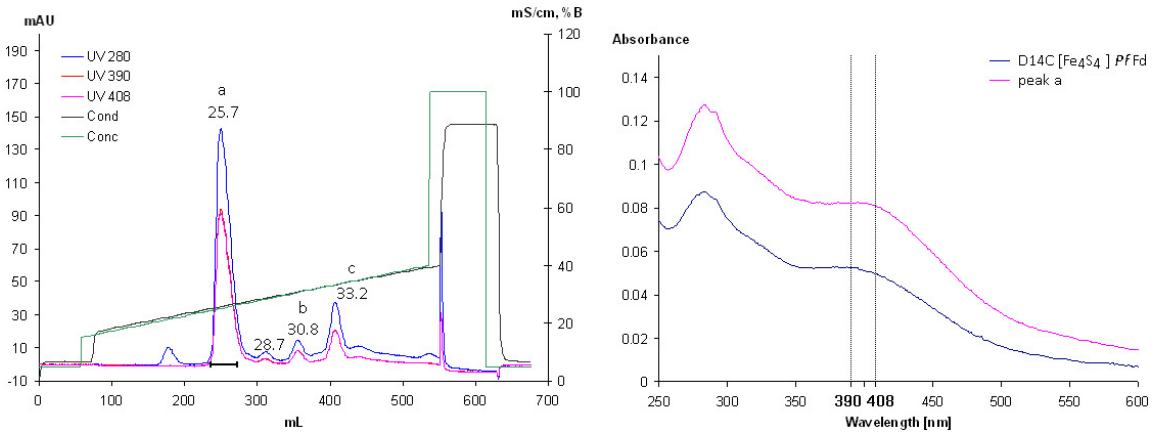


Figure 4.1: Left: chromatogram from purification of oxidized D14C *Pf* Fd. Collected fractions of *peak a* are marked with a black line. Right: UV-vis spectrum of peak a from the chromatogram and D14C [Fe_4S_4] *Pf* Fd.

4.2 Experimental

4.2.1 Expression of D14C *Pf* Fd

The expression strain was an *E. coli* strain over expressing the D14C variant of *Pf* Fd. Cultivation of the cells was carried out in Terrific Broth with carbenicillin as selection marker. Sterilization was applied to all instruments that came in contact with the cells to avoid contamination. The cultivation procedure is based on a previously developed method in the Metalloprotein Chemistry and Engineering Research Group for expressing native *Pf* Fd [54].

The bacteria cells expressing D14C *Pf* Fd from a glycerol stock were streaked on an agar plate containing Luria Broth, 100 µg/mL ampicillin. A single colony was selected and added to 50 mL Terrific Broth with 0.5 % (w/w) glycerol and 50 µg/mL carbenicillin (TB/carb) in a 300 mL triple baffled shake flask. Two 50 mL cultures were prepared per 6 x 650 mL large-scale cultures. The two cultures were incubated at 37 °C with shaking until the optical density at 600 nm (OD_{600}) was 0.6-1.0. The cultures were incubated at 4 °C overnight. Each culture was then transferred to a 50 mL sterile plastic tube and centrifuged at 4 °C. The clear supernatants were decanted off and discarded, and each of the two pellets was resuspended in 23 mL TB/carb. 6.5 mL culture was transferred into 650 mL TB/carb in a 2 L triple baffled shake flask to give a total of six 650 mL large-scale cultures. The cultures were incubated at 30 °C with shaking until OD_{600} was 2.4. IPTG was added to a final concentration of 0.1 mM and incubation continued at 30 °C for 16 hours.

Each of the 650 mL cells were transferred to a Beckmann centrifuge tube and centrifuged at 4 °C. The clear supernatants were decanted off and discarded. The cells in each centrifuge tube were resuspended in 25 mL ice-cold 20 mM Tris/HCl pH 8.0 and poured into 50 mL sterile plastic centrifuge tubes. The six 50 mL tubes were centrifuged at 4 °C. The clear supernatants were then decanted off and discarded. Cell pellets were stored at -80 °C.

4.2.2 Purification of D14C *Pf* Fd

Six tubes, each containing cells from 650 mL culture (section 4.2.1), were slowly defrosted on ice. Each tube was added approximately 30 mL 20 mM Tris/HCl pH 8.0 buffer and the cells resuspended. 80 µL 1 M sodium dithionite was then added to each tube. The cells were lysed by sonication three times for 40 seconds. The suspension was kept on ice during and in between sonications. After sonication, the suspensions were incubated for 10 minutes in a 70 °C water bath, and then centrifuged at 4 °C and the clear supernatants decanted off. The pooled supernatants were diluted four times with 20 mM Tris/HCl pH 8.0, 2 mM sodium dithionite.

The protein solution was loaded onto a 50 mm/7 cm Q Sepharose Fast Flow column equilibrated with 20 mM Tris/HCl pH 8.0. The column was washed with 2 CV 20 mM Tris/HCl pH 8.0 and 2 CV 0.15 M NaCl in 20 mM Tris/HCl pH 8.0. The protein was eluted isocratically using 0.4 M NaCl in 20 mM Tris/HCl pH 8.0. The volume was reduced by ultrafiltration. Less than 30 mg protein reduced to a maximum volume of 9 mL was then loaded onto a HiLoad 26 mm/60 cm Superdex 75 column equilibrated with 0.15 M NaCl in 20 mM Tris/HCl pH 8.0 and eluted using the same buffer. Fractions containing the protein were pooled and exchanged into a 20 mM Tris/HCl pH 8.0 buffer by ultrafiltration. A maximum of 15 mg protein was loaded onto a 16 mm/10 cm Source 30Q column equilibrated with 20 mM Tris/HCl pH 8.0. The column was washed with 2 CV 20 mM Tris/HCl pH 8.0 and elution was carried out using a linear salt gradient of 0.15 M to 0.4 M NaCl in 20 mM Tris/HCl pH 8.0 over 17 CV. The protein was exchanged into 20 mM Tris/HCl pH 8.0 by ultrafiltration. UV-vis spectrophotometry was used to determine the protein concentration and yield.

4.2.3 Oxidation and Purification

Different parameters were varied during the oxidation process. The experimental details are given in Appendix B and summarized in Table 4.1. In brief, purified D14C [Fe_4S_4] *Pf* Fd was mixed with ferricyanide and left with mild stirring for a variable amount of time and then exchanged into 20 mM Tris/HCl pH 8.0 and purified by anion exchange on a 16 mm/10 cm Source 30Q column equilibrated with the same buffer unless otherwise specified. In addition to these experiments (A to N), three other types of experiments (O, P and Q) were attempted. $CoCl_2$ was added in experiment O in order to investigate whether the [Fe_3S_4] cluster was unstable and would allow for a heterometal to replace the fourth iron of the cluster. In experiment P, it attempted to break the cluster by adding a large excess of dithionite as suggested by Moura *et*

al. [55]. In experiment Q, ferricyanide was substituted by $[Fe(phen)_3]^{3+}$ that has a potential of +1127 mV compared with +360 mV of ferricyanide and -426 mV D14C $[Fe_4S_4]$ *Pf* Fd [50,56,57].

Table 4.1: Overview of the experimental procedures for oxidation of D14C *Pf* Fd. $K_3[Fe(CN)_6]_3$ was present in 15x molar excess, except in experiments P and Q. Subsequent purification was performed by anion exchange at pH 8.0 unless otherwise stated.

ID	Buffer	pH	Time	Notes
A	20 mM Tris/HCl	8.0	15 min	-
B	20 mM Tris/HCl+0.3 M NaCl	8.0	15 min	-
C	100 mM Tris/HCl	7.8	15 min	-
D	100 mM Tris/HCl	7.8	45 min	-
E	100 mM Tris/HCl	7.8	3 hrs	-
F	100 mM BisTrisPropane/HCl	6.5	15 min	Anaerobic purification
G	20 mM Tris/HCl	8.0	45 min	1.5x molar excess EDTA
H	50 mM Na Acetate/Acetic Acid	4.5	45 min	1.5x molar excess EDTA
I	20 mM BisTris/HCl	5.8	3 hrs +overnight at 4 °C	1.5x molar excess EDTA
J	20 mM BisTris/HCl	5.8	overnight	1.5x molar excess EDTA
K	20 mM BisTris/HCl	5.8	3 hrs	1.5x molar excess EDTA Ion exchange at pH 5.8
L	20 mM BisTris/HCl	5.8	3 hrs	1.5x molar excess EDTA Ion exchange at pH 5.8
M	20 mM BisTris/HCl	5.8	overnight	1.5x molar excess EDTA Ion exchange at pH 5.8
N	100 mM Tris/HCl	7.8	15 min	Ion exchange pH 7.8
O	20 mM Tris/HCl	8.0	45 min	$CoCl_2$ present
P	0.8 M Tris/HCl	7.6	2.5 hrs	Excess dithionite
	0.1 M Tris/HCl+0.8 M NaCl	7.6	2.5 hrs	Excess dithionite
Q	20 mM Tris/HCl	8.0	20 min	Excess $[Fe(phen)_3]^{3+}$

4.2.4 Molecular Weight Determination and EPR

Three protein samples from experiment G were sent to molecular weight determination by mass spectrometry at Alphalyse A/S. The mass spectra were obtained using a Micromass QTOF mass spectrometer in positive ion mode and deconvoluted with MassLynx software, MaxEnt1 [58]. A detailed description of the selected samples from the purification is given with the results in section 4.3.3. The samples were in 5 mM Tris/HCl pH 8.0 with a concentration of 0.1 mg/mL. Two additional samples of were also sent: 0.6 mg/mL in 5 mM Tris/HCl pH 8.0 and 0.4 mg/mL in milli-q water.

Protein from experiment J (*EPR peak a sample*) and a mix of protein from experiments K and L (*EPR peaks c sample*) were concentrated to 500 µM in 20 mM Tris/HCl pH 8.0 by ultrafiltration. Oxidized EPR samples were prepared by transferring 350 µL of each sample directly to an EPR

sample tube. Reduced EPR samples were prepared in a glove-box by adding dithionite to a final concentration of 10 mM and then transferring 350 µL to EPR sample tubes, which were subsequently sealed with Suba-Seal silicon rubber septas. All samples were frozen in liquid nitrogen. EPR spectra were obtained by Hanne Nørgaard[2] under the following conditions: temperature 10 K, field modulation amplitude: 10 G, receiver gain 50, microwave frequency: 9.63GHz, microwave power: 14 mW and corrected by withdrawing the spectrum of an empty EPR tube.

4.2.5 Resource Anion Exchange

Some of the collected fractions of oxidized D14C *Pf* Fd after anion exchange (section 4.2.3, experiments A, J and L) were run on a 16 mm/3 cm ResourceQ anion exchange column to improve the resolution. Also, purified D14C [Fe_4S_4] *Pf* Fd was run twice. The oxidized samples were loaded onto a 16 mm/3 cm ResourceQ column equilibrated with 20 mM Tris/HCl pH 8.0. The column was washed with 2 CV 20 mM Tris/HCl pH 8.0 and elution was carried out using a linear salt gradient of 0.15 M to 0.4 M NaCl in 20 mM Tris/HCl pH 8.0 over 17 CV.

4.2.6 Gel Filtration

Selected fractions of oxidized D14C *Pf* Fd after anion exchange (section 4.2.3, experiments J, K and L) were run on a 10 mm/30 cm Superdex75 column to estimate if protein from different fractions differed in size. Purified D14C [Fe_4S_4] *Pf* Fd was run twice. The oxidized samples were loaded using a 250 µL superloop onto a 10 mm/30 cm Superdex75 column equilibrated with 0.15 M NaCl in 20 mM Tris/HCl pH 8.0 and eluted using the same buffer.

4.3 Results

4.3.1 Expression and Purification of D14C *Pf* Fd

The cells (section 4.2.1) had a light brown color. The yield was 5-10 mL cells per 650 mL culture.

The experimental purification procedure for D14C *Pf* Fd is given in section 4.2.2 and summarized in Figure 4.2. During the HPLC operated steps, protein elution was monitored with a UV-detector at three wavelengths: 280 nm, 390 nm and 408 nm.

Resuspension of the cells in buffer followed by cell lysis and centrifugation produced a brown supernatant. The initial anion exchange purification was not performed using HPLC and the brown colored protein was collected based on a visual estimation of its elution. Gel filtration of the protein solution on a HiLoad 26 mm/60 cm Superdex 75 column was carried out in several

[2] Metalloprotein Chemistry and Engineering, Department of Chemistry, Technical University of Denmark. Current address: Dept. of Chemistry, The Pennsylvania University, University Park, PA

runs and a typical chromatogram is shown in Figure 4.3. The marked fractions were collected and pooled for further purification. The chromatogram from typical anion exchange purification on the 16 mm/10 cm Source 30Q column is shown in Figure 4.3. The marked fractions were collected and pooled. The yield of D14C [Fe_4S_4] *Pf* Fd as determined by UV-vis spectrophotometry (Figure 4.3) was 14 mg/L cell culture.

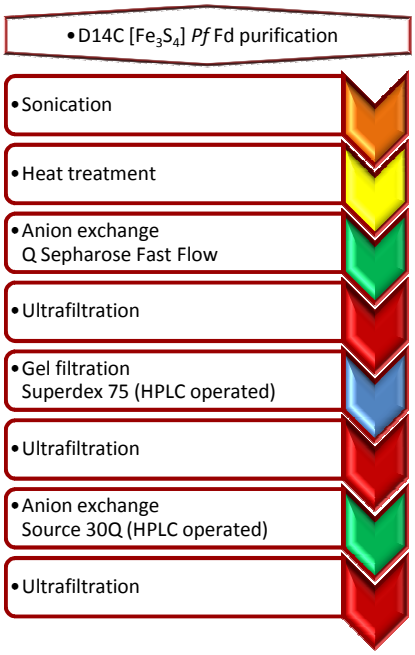


Figure 4.2: Steps involved in purification of D14C [Fe_4S_4] *Pf* Fd, see section 4.2.2 for details. Chromatograms were obtained for the HPLC operated steps.

4.3.2 Oxidation and Purification

Three experiments, O, P and Q (Table 4.1), had a different set-up compared with experiments A to N. The purpose of experiment O was to attempt to insert Co to form a heterometallic cluster simultaneously with the oxidation. The chromatogram (Appendix C, Figure C.2), however, did not differ from comparable experiments without Co (such as experiment D, Appendix C, Figure C.5) and this type of experiment was not pursued further. Experiment P was based on work by Moura *et al.* 1982 [55], which suggested that excess dithionite could degrade the cluster. The chromatograms and UV-vis spectra (Appendix C, Figure C.3) from these experiments were identical to pure D14C [Fe_4S_4] Fd and no cluster degradation was observed. Addition of $[Fe(phen)_3]^{3+}$ to oxidize the Fd (experiment Q) resulted in protein precipitation and only minor traces of protein were visible in the chromatogram (Appendix C, Figure C.4).

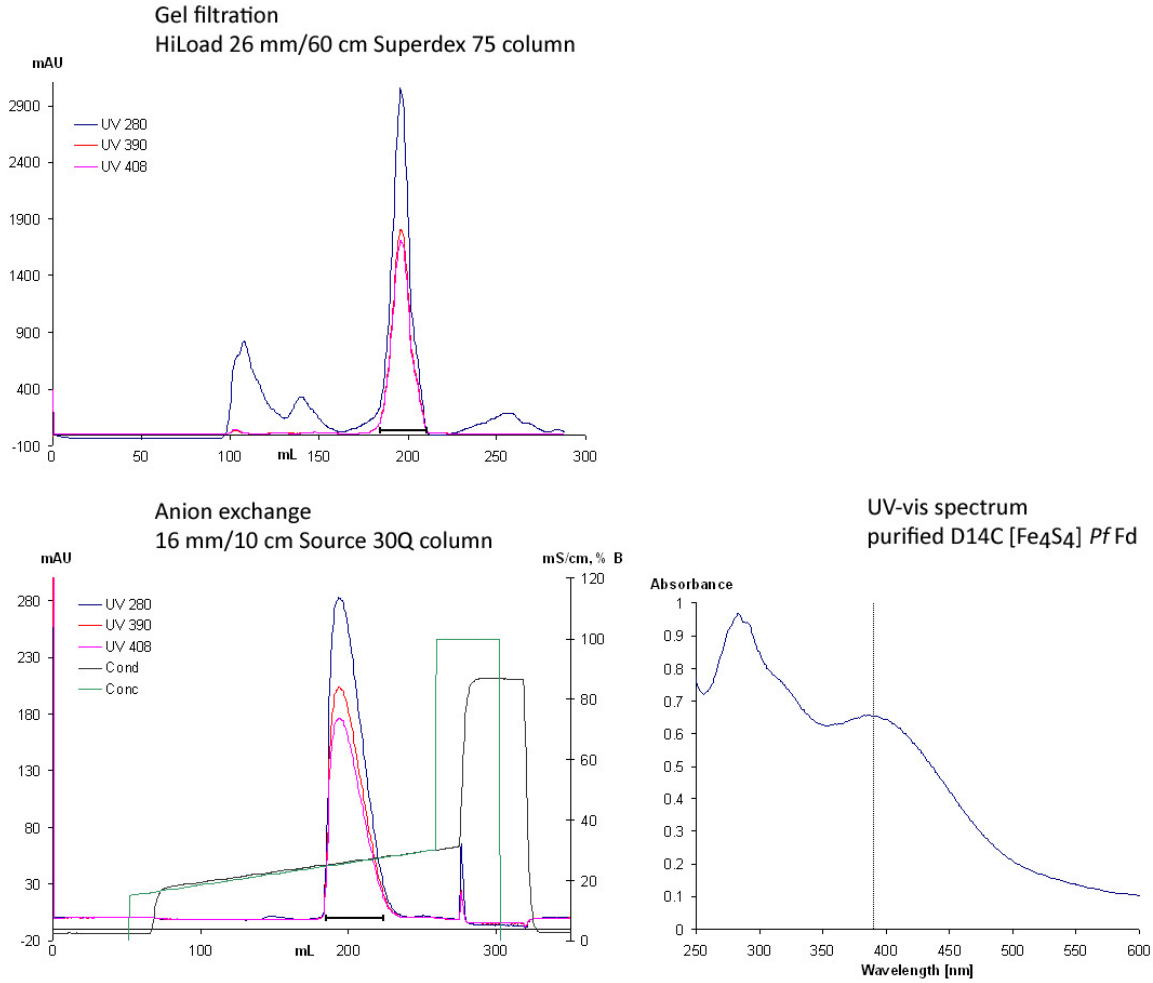


Figure 4.3: Chromatograms from purification of D14C *Pf* Fd on HiLoad 26 mm/60 cm Superdex 75 column and on 16 mm/10 cm Source 30Q column. The collected fractions are marked with a black line. To the right is shown a UV-vis spectrum of purified D14C [Fe_4S_4] *Pf* Fd, the vertical line marks 390 nm.

Addition of salt during the oxidation (experiment B) does not affect the chromatogram (Appendix C, Figure C.1) compared to Figure 4.1. Increasing the incubation time leads to an increase in the amount of protein in *peak c* and produces more irregular peaks. Chromatograms and UV-vis spectra from experiments C (15 min), D (45 min) and E (3 hours) are shown in Appendix C, Figure C.5. The UV-vis spectra show that after oxidation, the absorption maximum at 390 nm had moved towards 408 nm, while the curve was slightly flattened.

Oxidation at pH 6.5 followed by anaerobic purification at pH 8.0 (experiment F, chromatogram shown in Appendix C, Figure C.6) increases the amounts of protein in *peaks b* and *c*. Oxidation at pH 7.8 as described by Duderstadt *et al.* [47] followed by ion exchange at pH 7.8 (experiment N) produces *peak a* and a broad, irregular *peak c* (Appendix C, Figure C.7).

An increased amount of protein in *peaks b* and *c* is observed after EDTA addition (experiment G). Lowering the pH during the oxidation (experiment H) reduces *peak b*. The chromatograms and UV-vis spectra from experiments G and H are shown in Appendix C, Figure C.8. The UV-vis

spectra show the absorption maximum had a tendency to move towards 408 nm, but the curve was slightly flattened.

Overnight incubation (experiments I and J) produces *peak a* and an irregular *peak c*. Experiment M is identical to experiment J, but the purification was performed at pH 5.8. A single peak at 23.7 mS/cm is observed at pH 5.8. Chromatograms and UV-vis spectra are shown in Figure 4.4. The UV-vis spectra from experiments J and M show a more pronounced peak at 408 nm. In comparison, D14C [Fe_4S_4] Fd elutes at 25.6 mS/cm at pH 5.8.

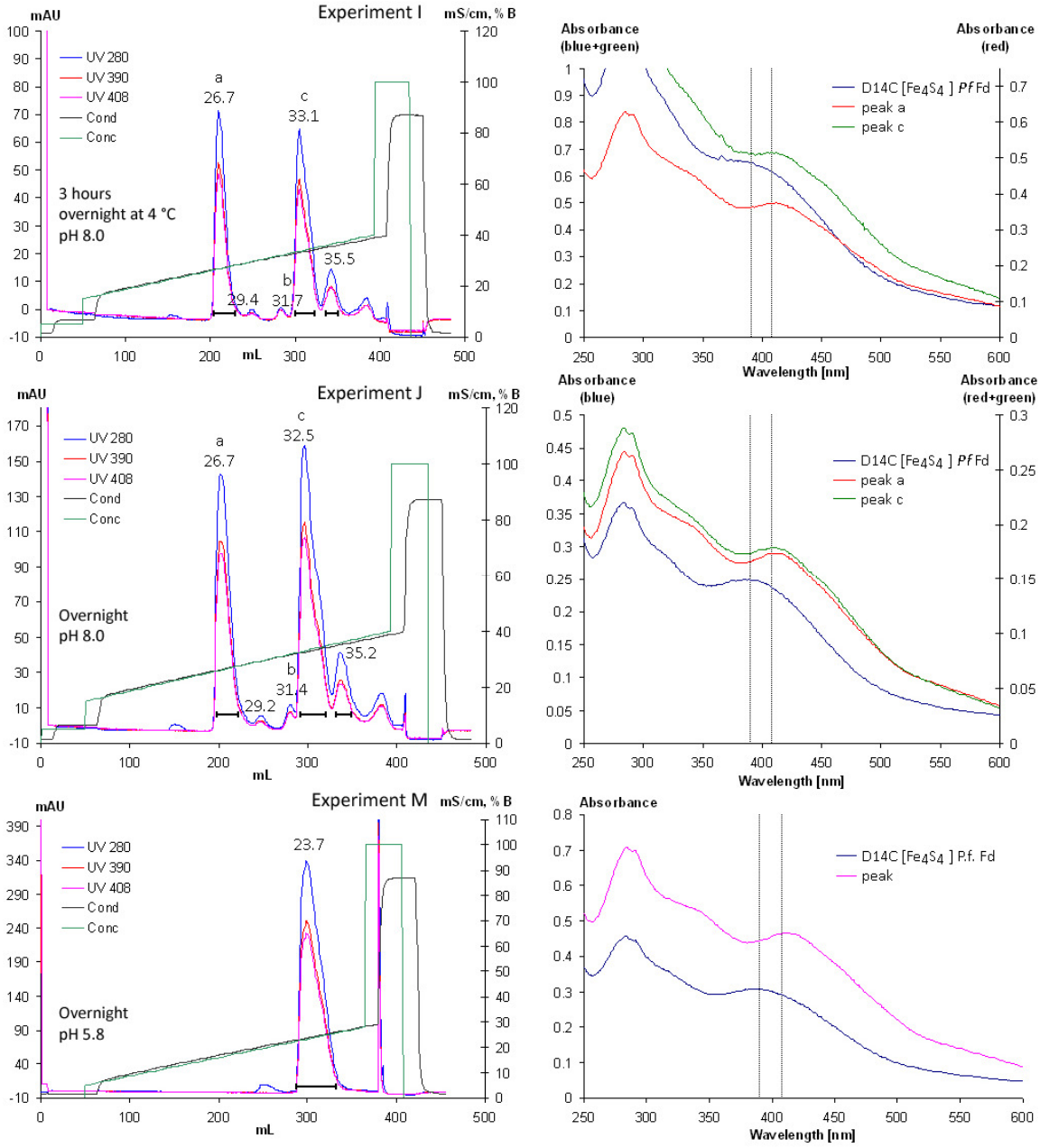


Figure 4.4: Oxidation of D14C *Pf* Fd; chromatograms and UV-vis spectra from experiment I, J and M. UV-vis spectra corresponding to the chromatograms are shown to the right.

The single peak at pH 5.8 from experiment M is also seen in experiments K at 25.1 mS/cm and 25.6 mS/cm in a re-run (Figure 4.5), as well as in experiment L at 24.8 mS/cm (Appendix C, Figure C.9). Exchanging the protein from this peak into 20 mM Tris/HCl pH 8.0 and running the purification at pH 8.0 produces even amounts of protein in *peak a* and *c*, see Figure 4.5. In contrast, performing the purification at pH 8.0 under anaerobic conditions produces more protein in *peak a* compared with *peak c*. Experiment L shows that a re-run of the protein from pH 5.8 anaerobically at pH 8.0 increases the amount of protein in *peak a* compared with *peak c* (Appendix C, Figure C.9 bottom chromatogram). Experiment L also shows that mixing protein purified at pH 5.8 with $[Fe_4S_4]$ protein increases the amount of protein in *peak a* compared with *peak c* (Appendix C, Figure C.9 second to bottom chromatogram). The UV-vis spectra show a tendency of the absorption maximum to move towards 408 nm after oxidation, except in the case where D14C $[Fe_4S_4]$ Fd was added (Figure C.9, UV-vis second to bottom red spectrum).

4.3.3 Molecular Weight Determination and EPR

Samples of protein in *peaks a, b* and *c* from experiment G (chromatogram shown in Appendix C Figure C.8) were sent to molecular weight determination by mass spectrometry at Alphalyse A/S. Two additional samples of protein from *peak c* were also sent. The molecular weights of the protein in *peaks a* and *b* were 7503 Da and 7425 Da, respectively. No mass was determined for an intact metalloprotein in *peak c* because the samples were very heterogenous. The molecular weight determination by Alphalyse A/S of protein in *peak b* does not correspond to a calculated mass of D14C *Pf* Fd with an intact cluster. The calculated mass of D14C $[Fe_4S_4]^{2+}$ *Pf* Fd with an intramolecular disulfide bond is 7503 Da, which corresponds very well with the molecular weight determination by Alphalyse A/S of protein in *peak a* from experiment G (spectra shown in Appendix C, Figure C.10).

A total of four samples were prepared for EPR analysis: *EPR peak a sample* contained protein from *peak a* (experiment J, Figure 4.4), while *EPR peaks c sample* contained protein from four *c peaks*: a mix from experiments K (re-run at pH 8.0 Figure 4.5 bottom chromatogram) and L (pH 8.0 re-runs, three bottom chromatograms in Figure C.9, Appendix C). Both oxidized and reduced samples of protein in *peak a* and *peaks c* were measured. The spin of $[Fe_4S_4]^{+,2+}$ is 0 and ½ for the oxidized and reduced cluster, respectively, and the spin of $[Fe_3S_4]^{0,+}$ is ½ and 2 for the oxidized and reduced cluster, respectively. Thus, only the reduced $[Fe_4S_4]^{+}$ and the oxidized $[Fe_3S_4]^{+}$ clusters produce an EPR signal. EPR spectra are shown in Figure 4.6. The signal for both *peaks c* samples and the oxidized *peak a* sample are at g = 2.01, which is characteristic for $[Fe_3S_4]^{+}$ clusters [2]. The observed signal in the reduced *peaks c* sample could be caused by a residual amount of oxidized protein in the sample or by the sample being re-oxidized. The reduced *peak a* sample show no signal, which supports the presence of a $[Fe_3S_4]^{0}$ cluster.

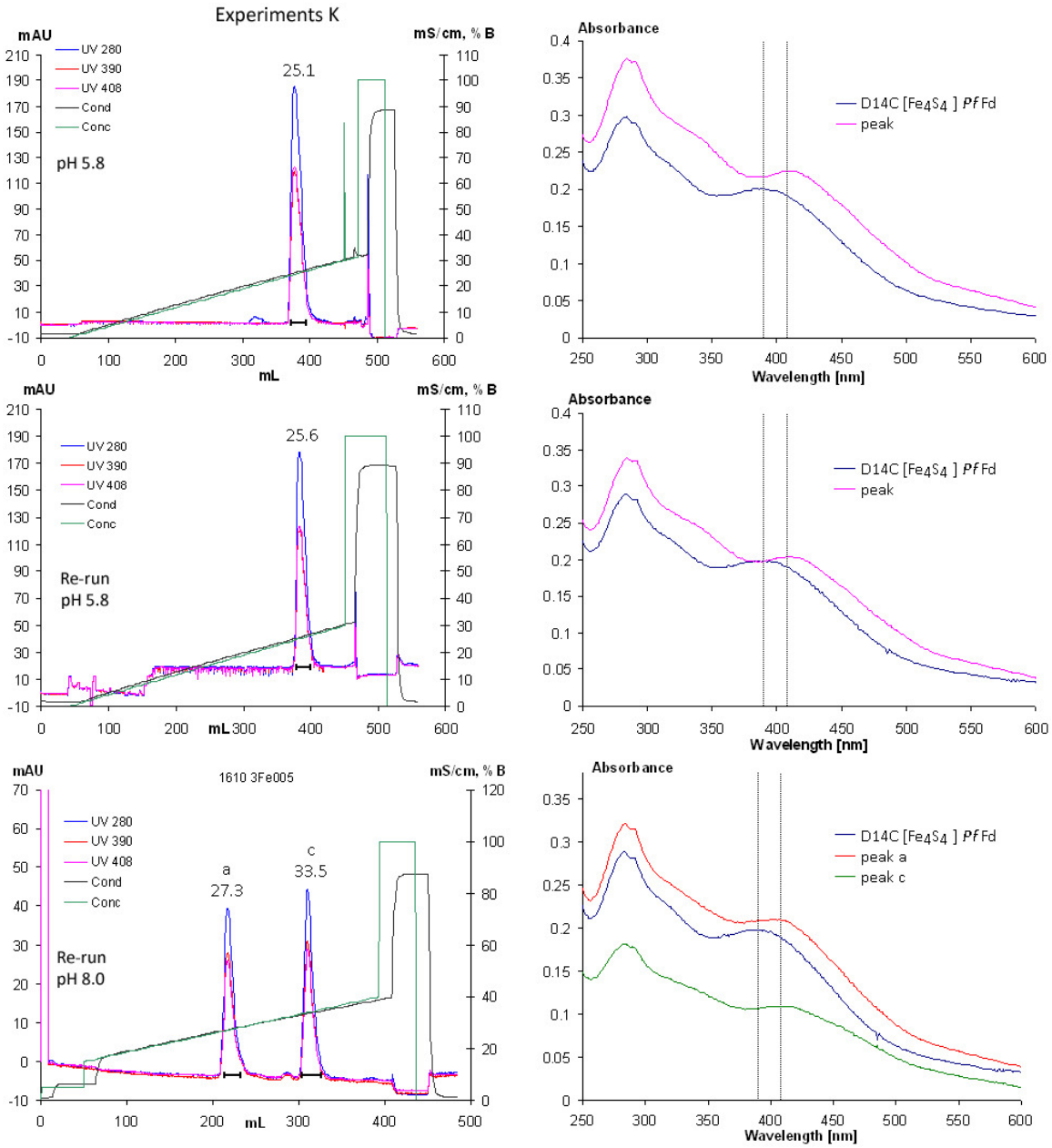


Figure 4.5: Oxidation of D14C *Pf* Fd; chromatograms and UV-vis spectra from experiments K. Top: chromatogram from first purification at pH 5.8. Mid: aliquot of collected fractions from the top chromatogram re-run at pH 5.8. Bottom: aliquot of collected fractions from top chromatogram re-run at pH 8.0.

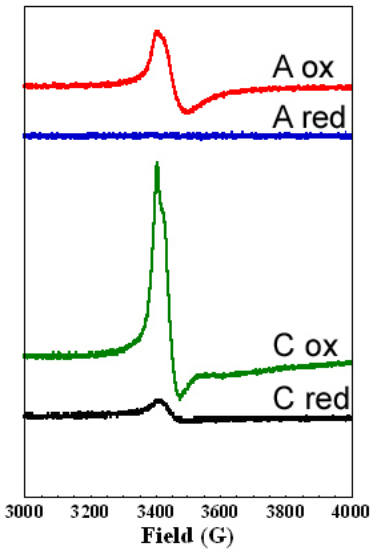


Figure 4.6: EPR spectra of *EPR peak a sample* and *EPR peaks c sample*; oxidized or reduced. Figure made by Hanne Nørgaard[3], reprinted with permission.

4.3.4 Resource Anion Exchange

Purified D14C [Fe_4S_4] *Pf* Fd elute at 26.7 mS/cm from the 16 mm/3 cm ResourceQ column (chromatograms not shown). In comparison, protein from *peak a* from experiment A (Figure 4.1) elute at 27.7 mS/cm (Figure 4.7 top left); a difference of 1 mS/cm. A general trend is observed in the Resource chromatograms in Figure 4.7: one peak around the elution of D14C [Fe_4S_4] Fd and a peak eluting at a later point. Conductivities suggest these peaks are equivalent to *peak a* and *c* observed in Source 30Q runs.

During experiment J, the oxidation process was left for incubation overnight followed by purification at pH 8.0 (Figure 4.4), *peak a* was collected and the protein re-run on the Resource column, top right chromatogram in Figure 4.7. It is seen, that the protein has rearranged and now elute significantly later. The oxidation process of experiment L had an incubation time of 3 hours and was purified at pH 5.8. An aliquot was purified at pH 8.0 and protein from *peak a* (Appendix C, Figure C.9 second to the top) was loaded onto the ResourceQ column and re-run, bottom left chromatogram in Figure 4.7, which shows two peaks equivalent to *peak a* and *c* in even amounts. Another aliquot from the 5.8 purification was mixed with purified D14C [Fe_4S_4] Fdx and purified at pH 8.0 (Appendix C, Figure C.9 second to the bottom) and protein from *peak a* was re-run on the Resource column, bottom right chromatogram in Figure 4.7, which shows a major *peak a* and also *peak c*.

[3] Metalloprotein Chemistry and Engineering, Department of Chemistry, Technical University of Denmark. Current address: Dept. of Chemistry, The Pennsylvania University, University Park, PA

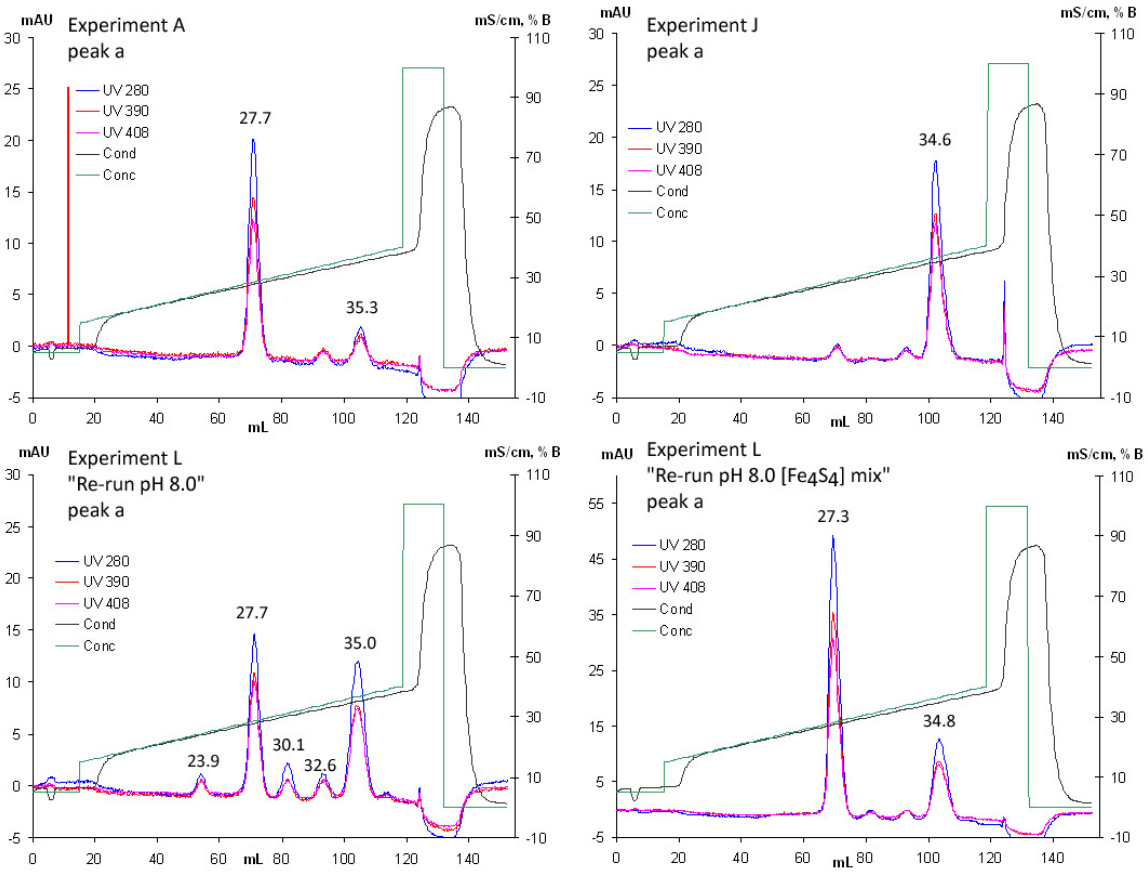


Figure 4.7: Chromatograms from re-runs on 16 mm/3 cm ResourceQ column. Top left: *peak a* from experiment A (Figure 4.1). Top right: *peak a* from experiment J (Figure 4.4). Bottom left: *peak a* from experiment L purified first at pH 5.8 and then at pH 8.0 (Appendix C, Figure C.9 second to the top). Bottom right: *peak a* from experiment L purified first at pH 5.8 and then mixed with D14C [Fe_4S_4] Fd and re-run at pH 8.0 (Appendix C, Figure C.9 second to the bottom).

4.3.5 Gel Filtration

Protein from *peaks a* and *c* were run on a 10 mm/30 cm Superdex 75 gel filtration column in order to estimate if the proteins differ in size. Purified D14C [Fe_4S_4] *Pf* Fd was run for comparison (top left chromatogram in Figure 4.8), it elutes at 12.4 mL. Protein from *peak a* (experiment J) elutes at the same volume (bottom left chromatogram in Figure 4.8). Protein from four *c peaks* was mixed: experiments K (re-run at pH 8.0, Figure 4.5 bottom chromatogram) and L (pH 8.0 re-runs, three bottom chromatograms in Appendix C, Figure C.9), the chromatogram from gel filtration is the top right part of Figure 4.8. Protein in these *peaks c* elute at 11.0 mL, which is earlier than protein from *peak a* and D14C [Fe_4S_4] Fd, indicating a larger protein size. The bottom right part of Figure 4.8 is protein from a peak that eluted later than *peak c* in experiment J (Figure 4.4) and it contains proteins with different sizes.

The ratios of the absorbance at 390 nm and 408 nm are given for each peak in Figure 4.8. A ratio slightly greater than one is observed for D14C [Fe_4S_4] Fd, which is consistent with the UV-vis

spectrum (Figure 4.1). Ratios slightly below one is observed for protein in *peak a* and *c*, which is consistent with the tendency of the absorption maximum to move towards 408 nm upon oxidation.

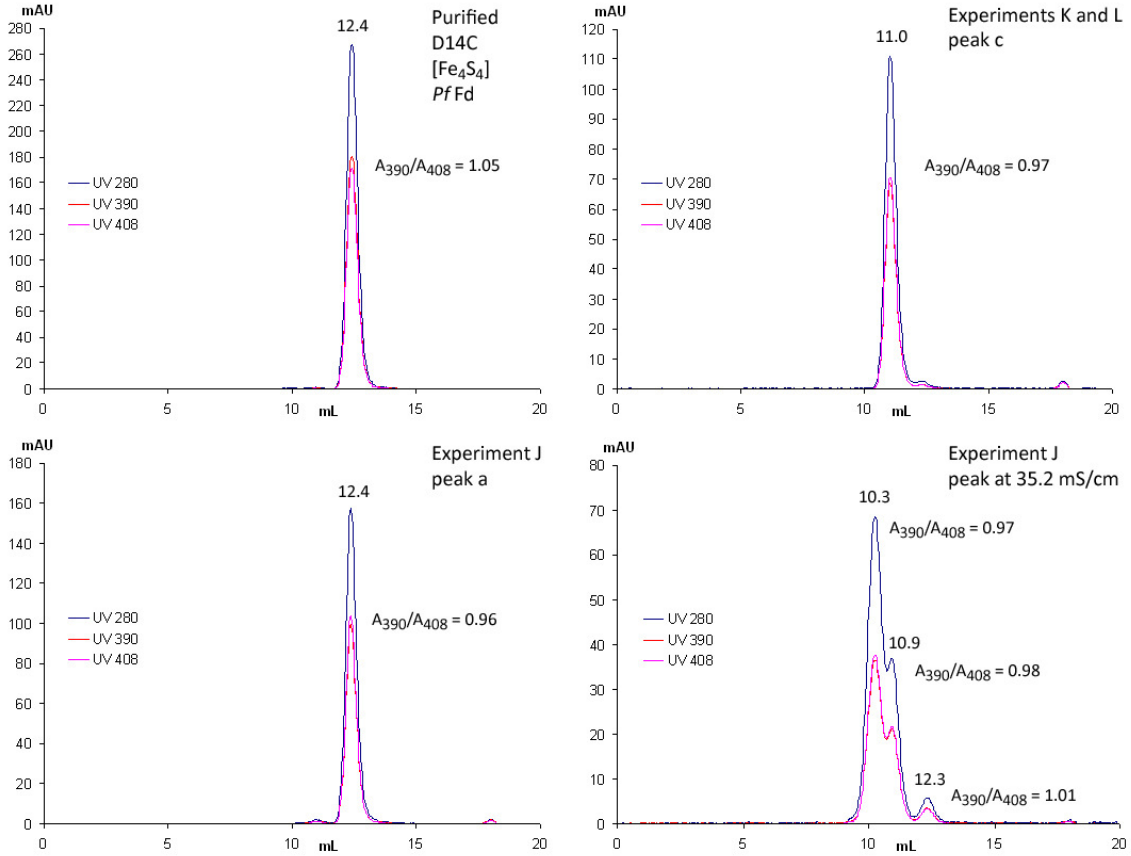


Figure 4.8: Chromatograms from gel filtration on a 10 mm/30 cm Superdex 75 column. Elution volumes in mL and the ratio A_{390}/A_{408} are given for each peak. Top left: elution of D14C $[Fe_4S_4]$ Fd. Top right: a mix of *peaks c* from experiments K and L (Figure 4.5 bottom chromatogram and Appendix C, Figure C.9 three bottom chromatograms). Bottom left: *peak a* from experiment J (Figure 4.4). Bottom right: late elution peak at 35.2 mS/cm from experiment J (Figure 4.4).

4.4 Discussion

The objective of the experiments was to obtain D14C $[Fe_3S_4]$ *Pf* Fd. The first attempt (section 4.1.1) proved to contain D14C $[Fe_4S_4]$ Fd as determined by X-ray diffraction. Purification of WT *Pf* Fd on a Source 30Q column produces two Fd peaks with an elution difference of 2 mS/cm; $[Fe_4S_4]$ Fd elutes first followed by $[Fe_3S_4]$ Fd [54]. D14C $[Fe_4S_4]$ *Pf* Fd eluted at 25-26 mS/cm and the collected fractions of *peak a*, experiment A (Figure 4.1), eluted at 25.6 mS/cm. During the oxidation experiments, section 4.2.3, it was attempted to produce D14C $[Fe_3S_4]$ *Pf* Fd with a UV-vis spectrum with distinct $[Fe_3S_4]$ cluster containing characteristics.

EPR supports that protein in *peak a* from experiment J contained a $[Fe_3S_4]$ cluster. This is in agreement with the UV-vis spectrum (Figure 4.4), which shows a significant peak at 408 nm. A more flattened UV-vis spectrum around wavelengths 390-408 nm is observed in experiment G (Appendix C, Figure C.8). Mass spectrometric analysis by Alphalyse supports the presence of D14C $[Fe_4S_4]$ *Pf* Fd in experiment G *peak a*. The differences between preparations in experiment G and J is pH and incubation time; experiment G was carried out at pH 8.0 with 45 minutes incubation, while experiment J was carried out at pH 5.8 with overnight incubation. Low pH is thus preferable during oxidation.

Purification at pH 5.8 produced a single peak in experiment M (Figure 4.4), experiments K (Figure 4.5) and experiment L (Appendix C, Figure C.9). Elution of protein in these peaks were at the same conductivity as elution of purified D14C $[Fe_4S_4]$ *Pf* Fd. UV-vis spectra at pH 5.8 strongly resembled a protein with a $[Fe_3S_4]$ cluster with an absorption maximum at 408 nm and a shoulder around 340 nm. Collecting the protein at pH 5.8 and running it on a Source 30Q at pH 8.0 produced two peaks in even amounts (*peaks a* and *c*, Figure 4.5). Protein from *peak a* rearranged to produce protein in *peak c*, since collecting protein from *peak a* and running it on a ResourceQ column produced both peaks (Figure 4.7 bottom left). Experiments L (Appendix C, Figure C.9) showed that anaerobic purification at pH 8.0 of an aliquot of protein oxidized and purified at pH 5.8 increased the amount of protein in *peak a* compared with *peak c*. Mixing D14C $[Fe_4S_4]$ Fd with an aliquot oxidized and purified at pH 5.8 and purifying it at pH 8.0 also increased the amount of protein in *peak a*. In summary, $[Fe_4S_4]$ and oxidized Fd (assumed $[Fe_3S_4]$) elute at the same conductivity at pH 5.8 in a single peak, and at pH 8.0 in *peak a* with an additional *peak c*. Anaerobic conditions favor protein elution in *peak a*, and the conversion of protein from *peak a* to *peak c* (Figure 4.7) appears to be an oxidative process.

Oxidizing D14C $[Fe_4S_4]$ Fd to D14C $[Fe_3S_4]$ *Pf* Fd produces a free cysteine. In WT $[Fe_3S_4]$ *Pf* Fd, the free residue is Asp14 [41], which corresponds to Cys14 in D14C Fd. The pK_a of Cys14 was estimated using PROPKA [59,60] with structure coordinates of WT $[Fe_3S_4]$ *Pf* Fd [41], where Asp14 was mutated to Cys using WinCoot [61,62]. pK_a was estimated to 8.8, which is an increase compared with the normal Cys pK_a of 8.3. At pH 5.8, the free cysteine is fully protonated and a single peak is observed during anion exchange. Protonation of the free cysteine can explain why the $[Fe_4S_4]$ Fd and $[Fe_3S_4]$ Fd elute at the same conductivity. The charges including ligands are identical: $[Fe_4S_4(Cys\text{-}S^-)_4]^{2-}$ and $[Fe_3S_4(Cys\text{-}S^-)_3(Cys\text{-}SH)]^{2-}$. At pH 8.0, D14C $[Fe_3S_4]$ Fd is only partly protonated and protein in *peak a* can thus be Fd with protonated Cys (elution at the same

conductivity as $[Fe_4S_4]$ Fd), while protein in *peak c* can be related to the deprotonated Cys portion of $[Fe_3S_4]$ Fd. This explains why re-running protein in *peak a* on a ResourceQ column produces two peaks: *a* and *c* (Figure 4.7). Gel filtration experiments support that protein in *peak c* is larger than protein in *peak a* (Figure 4.8). Protein in *peak c* can thus contain dimers, where a disulfide bond has formed between two deprotonated monomers. Such a dimer will be very flexible, as the two monomers have many possible orientations. This is consistent with *peak c* often appearing very irregular. The pH dependent equilibrium is illustrated in Figure 4.9.

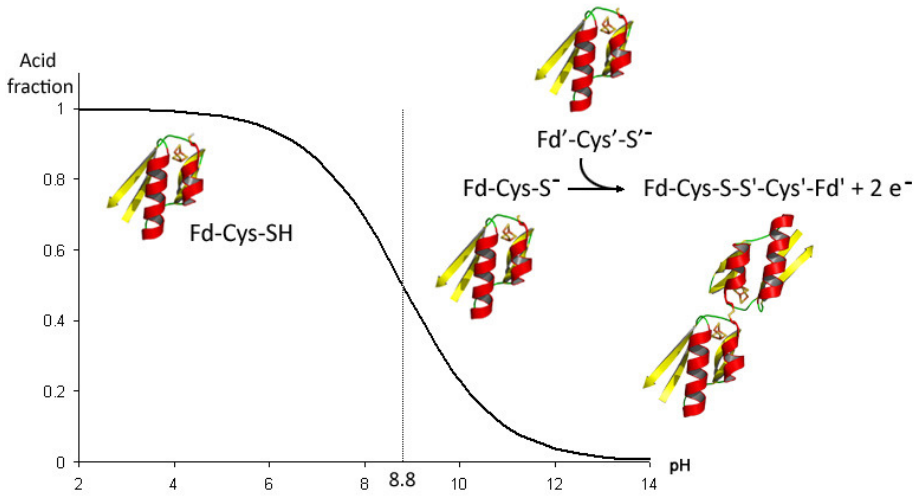


Figure 4.9: Bjerrum diagram illustrating the pH dependent equilibrium (pK_a=8.8) of D14C $[Fe_3S_4]$ *Pf* Fd between protonated and deprotonated monomers and formation of a disulfide bonded dimer. Two ferredoxin monomers were graphically connected to illustrate the dimer. The curve is an illustration, it does not reflect stability of the Fd in the given pH range.

Mass spectrometric analysis by Maja Martic [63] verified the presence of D14C $[Fe_3S_4]^+$ *Pf* Fd as the protein in the single peak from purification at pH 5.8 from an experiment prepared in the same way as experiment M (Figure 4.4). Re-running the protein purified at pH 5.8 at pH 8.0 produced two peaks (Figure 4.5) and mass spectrometric analysis on protein from a repetition of this experiment verified that the first peak contains D14C $[Fe_3S_4]^+$ *Pf* Fd, while the second peak contains a disulfide bonded dimer: D14C $[Fe_3S_4$ (Cys-S)-(S'-Cys') $Fe'_3S'_4]^{2+}$ *Pf* Fd. The apparent mass of the monomer was 7447.6 Da compared to a calculated mass of 7447.9 Da of D14C $[Fe_3S_4]^+$ *Pf* Fd with one intramolecular disulfide bond. The apparent mass of the disulfide bonded dimer was 14893.1 Da compared to a calculated mass of 14893.7 Da. For experimental details, see the article by Løvgreen *et al.* (2011) [64] (Appendix G).

4.5 Conclusion

Cells expressing D14C [Fe_4S_4] *Pf* Fd were successfully cultivated and the Fd was purified. The yield of pure D14C [Fe_4S_4] *Pf* Fd was 14 mg/L cell culture.

Purified D14C [Fe_4S_4] *Pf* Fd was oxidized with several variations in the experimental conditions. It proved difficult to separate D14C [Fe_4S_4] Fd from the oxidized D14C [Fe_3S_4] Fd as they have the same charge at low pH: $[Fe_4S_4(Cys\text{-}S^-)_4]^{2-}$ and $[Fe_3S_4(Cys\text{-}S^-)_3(Cys\text{-}SH)]^{2-}$. The pK_a value of the protonated cysteine, Cys14, is 8.8 and purifying the oxidized Fd at pH 5.8 is advantageous because the free cysteine is fully protonated. At pH 8.0, the free cysteine is partly deprotonated and able to form a disulfide bonded dimer.

In summary, in order to obtain D14C [Fe_3S_4] *Pf* Fd, it is advisable to oxidize purified D14C [Fe_4S_4] *Pf* Fd at pH 5.8 and leave it overnight with ferricyanide at room temperature with gentle stirring. Purifying and storing the protein at pH 5.8 is preferred in order to avoid formation of a disulfide bonded dimer. D14C [Fe_3S_4] *Pf* Fd was formed by incubation of purified D14C [Fe_4S_4] *Pf* Fd overnight with 15x molar excess ferricyanide in 20 mM BisTris/HCl pH 5.8 and purification by anion exchange on a Source 30Q column at pH 5.8.

Chapter Five

5 Crystal Structures of D14C *Pyrococcus furiosus* Ferredoxin

5.1 Introduction

Crystallization of D14C [Fe_3S_4] *Pf* Fd was very difficult due to solution inhomogeneity and results are described in section 5.3.1. The structure of D14C [Fe_4S_4] *Pf* Fd was solved prior to this project[4]. A structural description and discussions regarding the Fd electrochemical properties are included in results and discussion (section 5.3) and can be found in the article by Løvgreen *et al.* (2011) [64] (Appendix G).

There are currently two reported crystal structures of [Fe_3S_4] ferredoxins with complete cysteine binding motifs (-Cys-X_2-Cys-X_2-Cys-X_n-Cys-): *Desulfovibrio gigas* (pdb 1FXD [65]) and *Bacillus thermoproteolyticus* (pdb 1WFT [66]). However, these structures both have additional chemical groups bound to protect the free cysteine, and the structure of D14C [Fe_3S_4] *Pf* Fd reported here is the first crystal structure of a ferredoxin in which a cysteine from a complete cysteine binding motif is unprotected and facing away from the cluster.

5.2 Experimental

5.2.1 Crystallization and X-Ray Data Collection

The starting point of the crystallization trials of oxidized D14C *Pf* Fd was similar to the crystallization conditions of WT [Fe_3S_4] *Pf* Fd [43]. Several factors were varied during crystallization: protein concentration, buffer and pH, precipitant type and concentration, additives and protein:reservoir volume ratio. The crystal used for data collection was produced with protein collected from *peak a* in experiment J (Figure 4.4, page 18). The UV-vis spectrum is characteristic for a [Fe_3S_4] cluster containing Fd and EPR measurements strongly support the presence of a [Fe_3S_4] cluster (section 4.3.3). A total of 12 trays with up to 24 different reservoirs per tray and up to 3 drops per reservoir were set up. A complete set of all tested conditions is not provided here. Instead, a variety of crystallization conditions are available for picture entries in the enclosed database (Appendix A, Ferredoxin.xml, see Chapter 2 for a program description).

[4] Monika Nøhr Johannessen (Løvgreen) Master's Thesis, Department of Chemistry, Technical University of Denmark, October 2007

Crystals to be tested for diffraction were mounted in a loop and cryo-cooled directly in liquid nitrogen. Around 50 crystals were tested at MAX-lab in Lund, Sweden, at a number of visits, and data were collected to 2.8 Å at ESRF in Grenoble, France. Data collection statistics are shown in Table 5.1. Data collection and processing were performed by Pernille Harris with MOSFLM [67] and space group determination and averaging with Pointless and SCALA [68]. Indexing was only successful once and could not be reproduced.

Table 5.1: Data collection statistics[a] for D14C [Fe_3S_4] *Pf* Fd.

Beamline	ESRF, ID14-3
Detector	ADSC Quantum Q315r
Wavelength (Å)	0.9765
Temperature (K)	100
Crystal space group	$P2_12_12_1$
Unit cell parameters	
a (Å)	47.4
b (Å)	49.8
c (Å)	51.2
Resolution (Å)	35.7-2.8 (2.95-2.8)
No. of reflections	17,078 (2504)
No. of unique reflections	3,254 (456)
Redundancy	5.2 (5.5)
Mosaicity (°)	Fixed 0.8
Completeness (%)	99.9 (99.9)
$I/\sigma(I)$	3.6 (1.9)
R_{merge}	0.252 (0.755)

[a] Values for the outermost resolution shell are given in parenthesis.
$R_{merge} = \Sigma_i \left| I_i - \langle I_i \rangle \right| / \Sigma_i I_i$

5.2.2 Structure Determination and Refinement

The CCP4 [69] suite was used for structure determination and refinement. MOLREP [70] was used to solve the structure with molecular replacement. The search model was the peptide chain of molecule A D14C [Fe_4S_4] *Pf* Fd (pdb ID: 2Z8Q). Two molecules were found in the asymmetric unit. Molecular replacement was also carried out with the search model of molecule B D14C [Fe_4S_4] *Pf* Fd. The *R* factor and score after molecular replacement with molecule B were 0.622 and 0.455, respectively, compared to 0.592 and 0.534 for molecule A, respectively. An increase in R_{free} was observed during refinement when using molecule B compared with molecule A for molecular replacement. Molecule A of D14C [Fe_4S_4] *Pf* Fd is thus a considerably better model for molecular replacement than molecule B. Refinement was performed with REFMAC5 [71]. Inspection of the $2F_{obs} - F_{calc}$ and the $F_{obs} - F_{calc}$ sigma-A weighted difference maps was done using winCoot [61,62]. Electron density appeared clearly from the [Fe_3S_4] cluster in each of the two molecules. Four Co atoms were added. Refinement statistics are given in Table 5.2. The structure is available in the protein data bank with pdb entry code 1PNI.

Table 5.2: Refinement statistics.

R	0.279
R_{free}	0.318
No. of atoms in model	
Protein	988
FeS clusters	14
Cobalt atoms	4
Mean temperature (B) factors (Å^2)	
Protein chain A	39
Protein chain B	39
FeS cluster chain A	34
FeS cluster chain B	28
Ramachandran plot [72]	
No. of residues in favored region	116
No. of residues in allowed region	12
No. of residues in outlier region	0

5.3 Results

5.3.1 Crystallization and X-Ray Data Collection

Crystals were difficult to grow and drops were generally inhomogeneous with a mixture of several combinations of phase separation, precipitation, protein films and crystalline clusters with very high mosaicity. Figure 5.1 shows some typical crystal clusters and phase separations of D14C [Fe_3S_4] *Pf* Fd. At pH below 7, where a homogeneous solution of monomers is expected, grey precipitation generally occurred indicating that the ferredoxin had lost its iron-sulfur cluster. An elaborate summary of crystallization results can be found in the enclosed database (Appendix A, Ferredoxin.xml, see Chapter 2 for a program description) that includes 461 picture entries with matching crystallization conditions.

Diffraction patterns revealed several crystals with poor diffraction to around 7 Å as well as several salt crystals when ammonium dihydrogen phosphate was used as an additive. Data were collected to 2.8 Å on ESRF in Grenoble on one of the crystals shown in Figure 5.2. Crystallization conditions were 35 % PEG1500 in 100 mM Tris/HCl pH 8.5. The drop contained 2 µL protein with a concentration of 6.4 mg/mL, 2 µL reservoir solution and 0.3 µL 100 mM $[Co(NH_3)_6]Cl_3$.

5.3.2 Structure of D14C [Fe_4S_4] *Pf* Fd

This section briefly outlines significant differences between two molecules, type A and B, in the asymmetric unit of D14C [Fe_4S_4] *Pf* Fd (pdb 2Z8Q). A more detailed description of the structure is part of the article by Løvgreen *et al.* (2011) [64] (Appendix G).

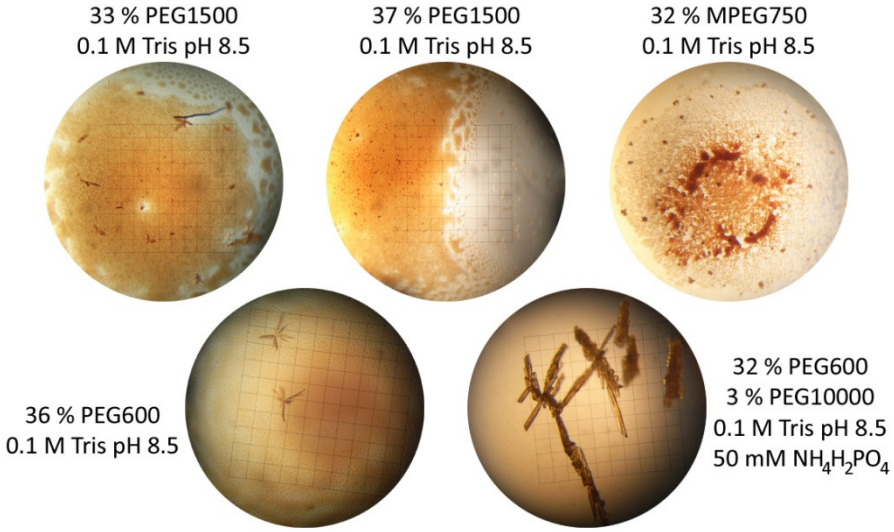


Figure 5.1: Crystal clusters of D14C [Fe_3S_4] *Pf* Fd. Grid spacing is 0.125 mm on the top three pictures and 0.089 mm on the bottom two pictures.

Figure 5.2: Crystal clusters of D14C [Fe_3S_4] *Pf* Fd. The small brown spheres are phase separations and the light brown background is a film that crackles when touched. Grid spacing is 0.125 mm.

Differences between molecules A and B of D14C [Fe_4S_4] *Pf* Fd can be related to the crystal packing and the conformation of the intramolecular disulfide bond. Molecule A packs as an extended β-sheet dimer with an adjacent A molecule, see Figure 5.3A, while molecule B packs as monomers in a less rigid position, see Figure 5.3B. Also, molecule B shows significantly larger B factors (18Å^2) compared with molecule A (7.5 Å^2), which may be related to the much more flexible position of the molecules in the crystal packing.

D14C [Fe_4S_4] *Pf* Fd has an intramolecular disulfide bond connecting Cys21 and Cys48. This disulfide bond is found in a double conformation in molecule A and in a single conformation in molecule B. WT [Fe_3S_4] *Pf* Fd also displays a double conformation of the disulfide bond [41], albeit with a shift in occupancy compared with molecule A of D14C [Fe_4S_4] *Pf* Fd. The single conformation of the disulfide bond in molecule B is possible due to an approximately 1 Å shift of

α-helix 2 accompanied by a displacement of the backbone at Asp42, Glu43 and Glu44, see Figure 5.3C.

In general, molecule A shows greater resemblance to WT [Fe_3S_4] *Pf* Fd than to molecule B from the same asymmetric unit when comparing molecule packing, hydrogen bonds and intramolecular disulfide bond conformation. Superposing the A and B molecules of D14C [Fe_4S_4] *Pf* Fd gives an RMS deviation of 0.8 Å, while superposing molecule A and B onto WT [Fe_3S_4] *Pf* Fd gives RMS deviations of 0.3 Å and 0.8 Å, respectively. RMS deviations were calculated using the program *superpose* [73].

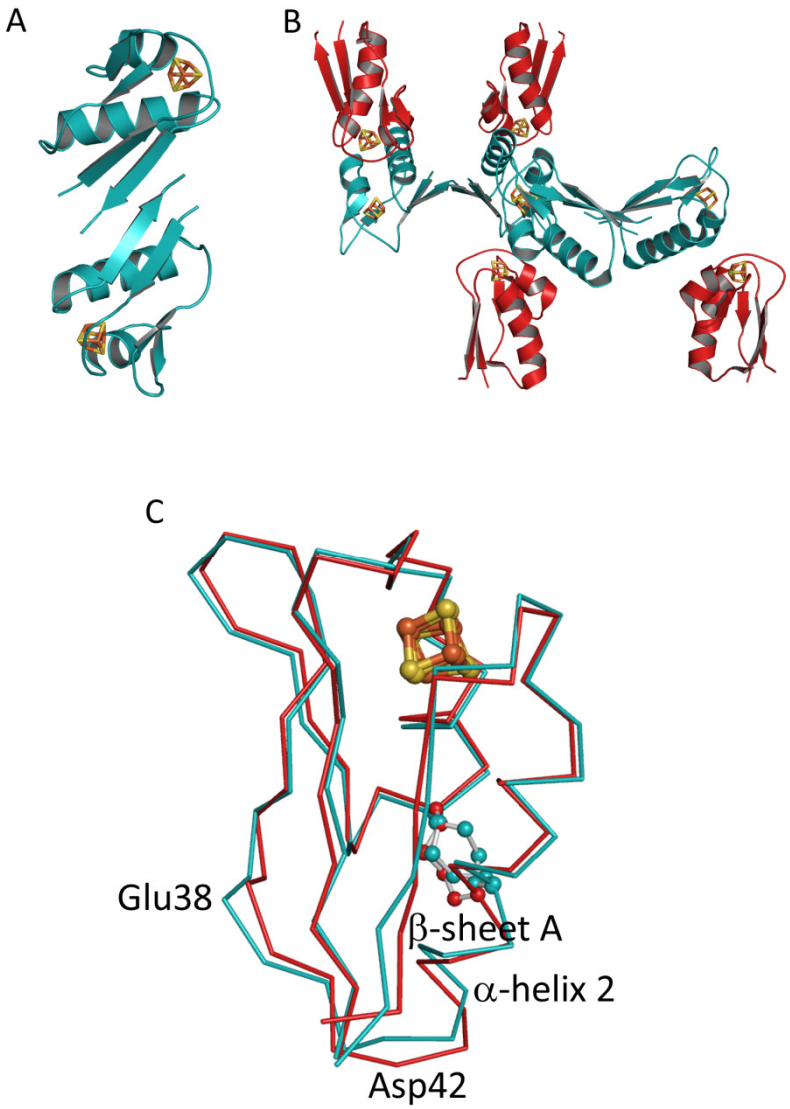


Figure 5.3: D14C [Fe_4S_4] *Pf* Fd molecule A shown in cyan, molecule B shown in red. A: crystal packing of two molecules A as an extended β-sheet dimer. B: crystal packing of molecules A and B. C: overlay of molecule A and B of D14C [Fe_4S_4] *Pf* Fd, disulfide bonds and clusters are shown in ball and stick.

5.3.3 Structure-reduction potential relationship

A sequence alignment of *Pf* Fd with related ferredoxins is given in Figure 5.4. Related ferredoxins were selected based on similarity and availability of crystal structures with an intact [Fe_4S_4] cluster. The cluster environment show a high degree of structural integrity, see Figure 5.5, even though there are different residues between the cysteines in the binding motif -Cys-X_2-Cys-X_2-Cys-X_2-Cys-X_n-Cys-.

Sγ of the second cysteine in the coordination motif (Cys14 in *Pf* Fd) is the least buried atom of the [Fe_4S_4] cluster including ligands [22] and solvent accessibilities of these atoms are given in Table 5.3 for [Fe_4S_4] ferredoxins from D14C *Pf*, *Thermotoga maritima*, *Desulfovibrio africanus* I and *Bacillus thermoproteolyticus* along with reduction potentials of the ferredoxins. In agreement with the literature [21-23], a correlation is seen where less buried [Fe_4S_4] clusters have higher reduction potentials. Table 5.3 also lists the N-H···S bonds of Sγ of cluster coordinating cysteines, which has been discussed in the literature to have an influence on the reduction potential of [Fe_4S_4] clusters [22,24]. The ferredoxins have equal numbers of bonds with comparable distances. A shift is seen in the reduction potential of D14C versus WT *Pf* Fd as the D14C Fd has a more negative reduction potential in its [Fe_4S_4] form and a more positive reduction potential in its [Fe_3S_4] form when it is compared with WT *Pf* Fd, see Table 5.3. Changing a cluster coordinating aspartate to cysteine in ferredoxin III from *D. africanus* also results in a more negative reduction potential in the [Fe_4S_4] form and a more positive reduction potential in the [Fe_3S_4] form [31]. For the [Fe_4S_4] cluster containing ferredoxins this can be explained by aspartate being a better electron withdrawing group than cysteine [50]. For the [Fe_3S_4] cluster containing ferredoxins it can be explained by the pK_a difference between aspartate and cysteine; aspartate is completely deprotonated while cysteine is partly protonated and the increased charge density of the WT ferredoxin with a free aspartate results in a lower reduction potential in the [Fe_3S_4] form compared with the D14C ferredoxin with a free (partly protonated) cysteine near the cluster.

```
D14C P.furiosus  MAWKVSVDQDTCIGCAICASLCPDVFEMNDEGKAQP----KVEVIEDE-ELYNCAKEAME
T.maritima       --MKVRVDADACIGCGVCENLCPDVFQLGDDG--------KAKVLQPE-TDLPCAKDAAD
D.africanus I    MARKFYVDQDECIACESCVEIAPGAFAMDPEIE-------KAYVKDVEGASQEEVEEAMD
B.thermoprot.    -PKYTIVDKETCIACGACGAAAPDIYDYDEDGIAYVTLDDNQGIVEVPDILIDDMMDAFE

D14C P.furiosus  ACPVSAITIEEA----------
T. maritima      SCPTGAISVEE----------
D. africanus I   TCPVQCIHWEDE----------
B.thermoprot.    GCPTDSIKVADEPFDGDPNKFE
```

Figure 5.4: Sequence alignment of ferredoxins from D14C *Pf*, *T. maritima*, *D. africanus* I and *B. thermoproteolyticus*. All cysteines are highlighted in yellow, cluster coordinating cysteines are shown in black and cysteines forming an intramolecular disulfide bond are shown in green. Conserved residues are shown in red. Sequence alignment was made using *ClustalW* [74] with sequence information from *UniProt* [75,76].

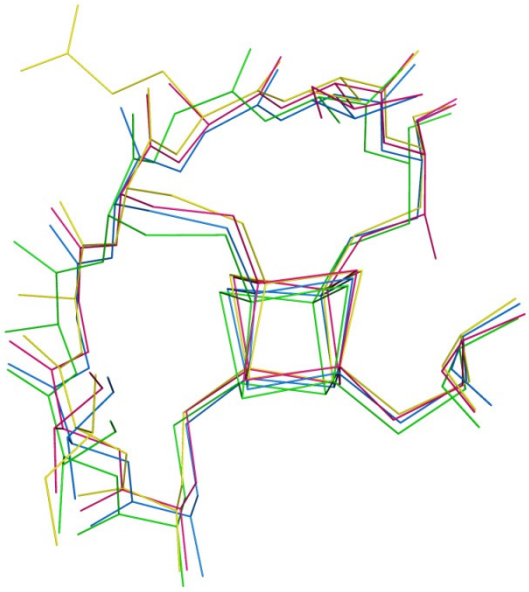


Figure 5.5: Structural overlay of ferredoxins with a [Fe_4S_4] cluster. D14C *Pf* molecule A (pink, pdb 2Z8Q), *D. africanus* I (yellow, pdb 1FXR), *T. maritima* (blue, pdb 1VJW) and *B. thermoproteolyticus* (green, pdb 1IQZ). The cluster and cysteines from the -Cys-X_2-Cys-X_2-Cys- segment are shown.

Table 5.3 Solvent accessibility of Sγ of the second cluster coordinating cysteine, number of N-H···S bonds to Sγ of cluster coordinating cysteines and their average distances (Å) and reduction potentials of ferredoxins.

	PDB code	Reduction potential	Solvent accessibility of second Cys Sγ	No. of N-H···S (average distance Å)
[Fe_4S_4] ferredoxin:				
P. furiosus	*	-368 mV [50]		
D14C *P. furiosus*	2Z8Q [64]	-427 mV [50]	A/B: 0/0	A/B: 10/10 (3.40/3.44)
T. maritima	1VJW [77]	-388 mV [78]	0.6	10 (3.47)
D. africanus I	1FXR [79]	-385 mV [80]	0.1	10 (3.58)
B. thermoproteolyticus	1IQZ [81]	-280 mV** [82]	3.9	10 (3.49)
[Fe_3S_4] ferredoxin:				
P. furiosus	1SJ1 [41]	-203 [50]		
D14C *P. furiosus*	3PNI [64]	-155 mV [63]		

* No structure is available for *Pf* [Fe_4S_4] Fd

** Reduction potential is for the very closely related *Bacillus stearothermophilus* ferredoxin

5.3.4 Structure of D14C [Fe_3S_4] *Pf* Fd

The resolution of the solved structure of D14C [Fe_3S_4] *Pf* Fd is inadequate to make detailed comparisons. However, the overall structure of D14C [Fe_3S_4] *Pf* Fd is very similar to both WT [Fe_3S_4] *Pf* Fd and molecule A of D14C [Fe_4S_4] *Pf* Fd, with RMS deviations of 0.4 Å and 0.5 Å, respectively. Less similarity is seen with molecule B of D14C [Fe_4S_4] *Pf* Fd, which shows an RMS deviation of 0.8 Å. RMS deviations were calculated using the program *superpose* [73].

Electron density clearly shows the presence of a [Fe_3S_4] cluster, in agreement with UV-vis (Figure 4.4, page 18) and EPR (section 4.3.3). Figure 5.6 shows two views of the cluster environment of D14C [Fe_3S_4] *Pf* Fd with a $2F_{obs} - F_{calc}$ σ = 1.3 electron density map.

Crystal packing is seen as extended β-sheet dimers in the structures of D14C [Fe_3S_4] *Pf* Fd, WT [Fe_3S_4] *Pf* Fd and in molecule A of D14C [Fe_4S_4] *Pf* Fd. The overall crystal packing is very similar for D14C [Fe_3S_4] *Pf* Fd and WT [Fe_3S_4] *Pf* Fd, see Figure 5.7, even though the space groups differ. Dimerization of WT *Pf* Fd has been proposed to occur depending on ionic strength [44,45]. Although extended β-sheet dimers are observed in the crystal structures, these dimers have not been observed in solution (Chapter 4) and could be an artifact of the high protein concentration in the crystal.

Crystals of D14C [Fe_3S_4] *Pf* Fd were grown at pH 8.5 (section 5.3.1) and although both monomers and disulfide bonded dimers should be present at this pH (Chapter 4), no disulfide bonded dimers are seen in the crystal structure. Crystallization of monomers is likely to be favored because of the large flexibility expected around the intermolecular disulfide bond in the dimer.

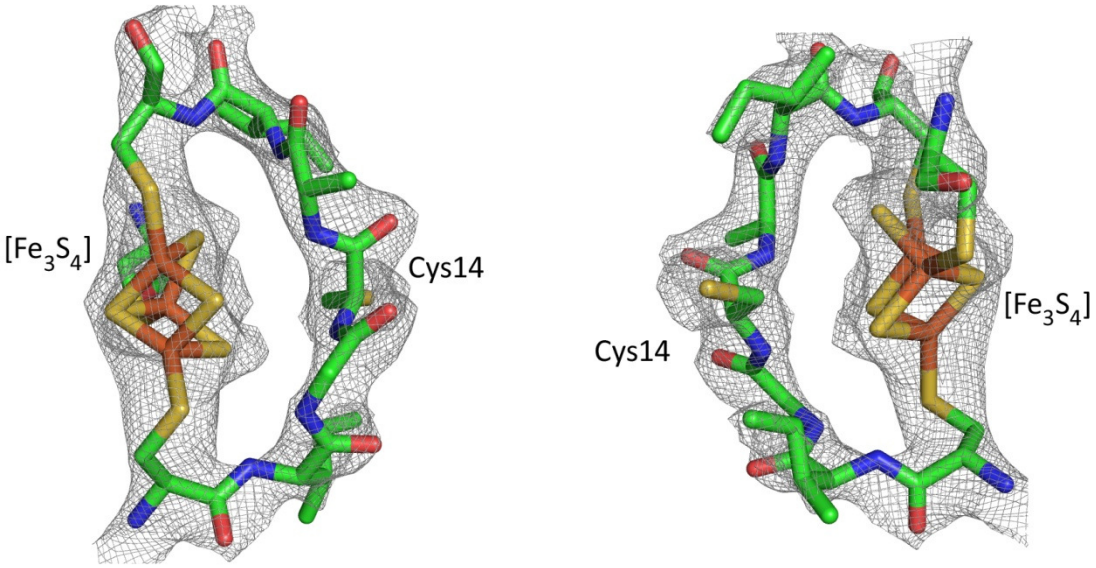


Figure 5.6: Two views of D14C [Fe_3S_4] *Pf* Fd with $2F_{obs} - F_{calc}$ σ = 1.3 electron density map.

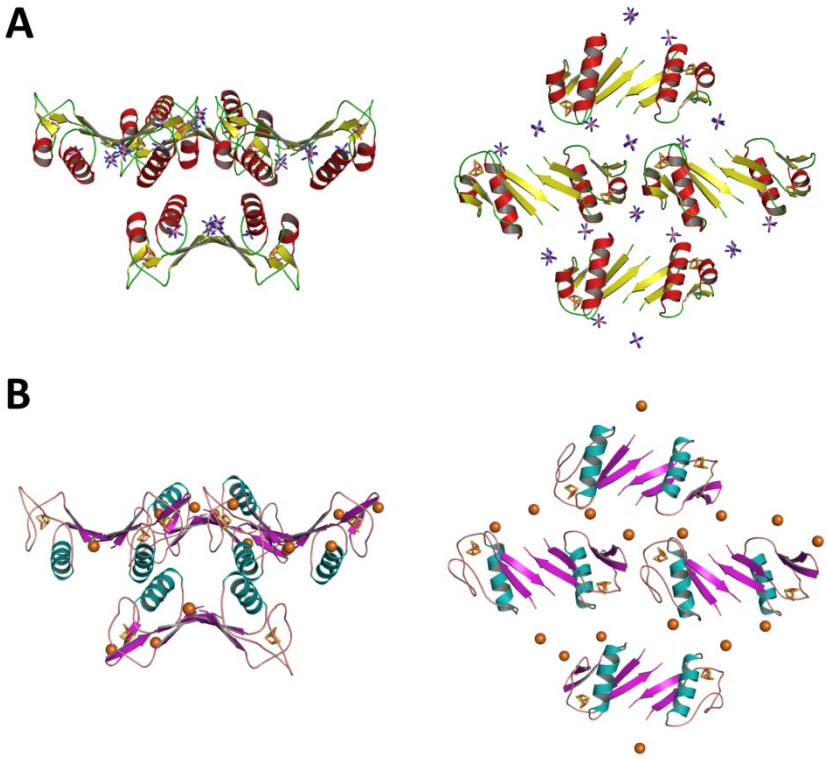


Figure 5.7: A: Two views of WT [Fe_3S_4] *Pf* Fd crystal packing including hexaaminecobalt(III) complexes. B: two views of D14C [Fe_3S_4] *Pf* Fd crystal packing including cobalt atoms.

5.4 Conclusion

Crystallization of D14C [Fe_3S_4] *Pf* Fd was very difficult and numerous crystallization conditions were tested. Small crystals or crystalline clusters only formed at pH 8.5 where the protein solution is very inhomogeneous (Chapter 4). Data were collected to 2.8 Å resolution and the crystal structure of D14C [Fe_3S_4] *Pf* Fd was solved. Close resemblance is seen between monomers of WT [Fe_3S_4], D14C [Fe_3S_4] and D14C [Fe_4S_4] molecule A, whereas D14C [Fe_4S_4] molecule B show some differences that can be related to the differences seen in the intramolecular disulfide bond conformation and crystal packing. Crystal packing as extended β-sheet dimers is seen for WT [Fe_3S_4], D14C [Fe_3S_4] and for adjacent A molecules of D14C [Fe_4S_4] *Pf* Fd. In contrast, D14C [Fe_4S_4] molecule B packs as monomers next to the extended β-sheet dimers formed between adjacent A molecules. The conformations of the highly conserved cysteines in the [Fe_4S_4] binding motif (-Cys-X_2-Cys-X_2-Cys-X_n-Cys-) are in close resemblance.

The crystal structure of D14C [Fe_3S_4] *Pf* Fd is the first structure in which a [Fe_3S_4] containing ferredoxin with a complete cysteine binding motif has an unprotected cysteine facing away from the cluster.

Chapter Six

6 Selected Enzymes Involved in Deoxyribonucleotide Synthesis

6.1 Introduction to deoxyribonucleotide synthesis

Nucleotides are essential for cell growth because they are necessary for replication and transcription of the genome. The building blocks for RNA are ribonucleotides, while the building blocks for DNA are deoxyribonucleotides. The building blocks consist of a purine or pyrimidine base linked to a sugar (RNA: β-D-ribose, DNA: β-D-deoxyribose) and a phosphate group. Nucleotides are able to interconvert between mono-, di- and triphosphates. In RNA, the base can be adenine, guanine, uracil or cytosine, whereas in DNA, thymine replaces uracil. This replacement in DNA prevents mutagenesis as a result of spontaneous deamination of cytosine to uracil, which pairs with adenine instead of guanine. A DNA repair system recognizes uracil to be foreign in DNA [83].

Deoxyribonucleotides are all, except dTTP, synthesized from their corresponding ribonucleotide, a reaction catalyzed by ribonucleotide reductase. Instead, dTTP is synthesized from dUMP, see Figure 6.1 for a schematic overview. dUMP is first methylated to form dTMP, which is then phosphorylated forming dTDP and phosphorylated again forming dTTP. dUMP is thus an important precursor in the formation of dTTP to be used for DNA synthesis. Different pathways exist for synthesis of dUMP: it can originate from a uracil base or it can be formed by deamination of a cytosine base as illustrated in Figure 6.1. The majority of dUMP is formed through deamination of the cytosine base either by deamination of dCTP followed by dephosphorylation or by deamination of dCMP [84].

Recently, a bifunctional enzyme was discovered that catalyzes both the deamination and dephosphorylation of dCTP thereby forming dUMP directly; Figure 6.2 shows the reaction. The bifunctional enzyme dCTP deaminase-dUTPase (DCD-DUT) has been found in the bacterium *Mycobacterium tuberculosis* (*Mt*) [85] and in the archaea *Methanocaldococcus jannaschii* (*Mj*) [86,87]. dTTP inhibition is observed for enzymes with dCTP deaminase activity as illustrated in Figure 6.1 and Figure 6.2. This chapter presents structures, inhibition, catalytic mechanisms and details on important residues for activity of the bifunctional enzymes as well as the monofunctional enzymes dCTP deaminase (DCD) and dUTPase (DUT). For keeping track of the numbering of key residues across organisms discussed in the remaining sections of this chapter, refer to Table 6.1 for an overview.

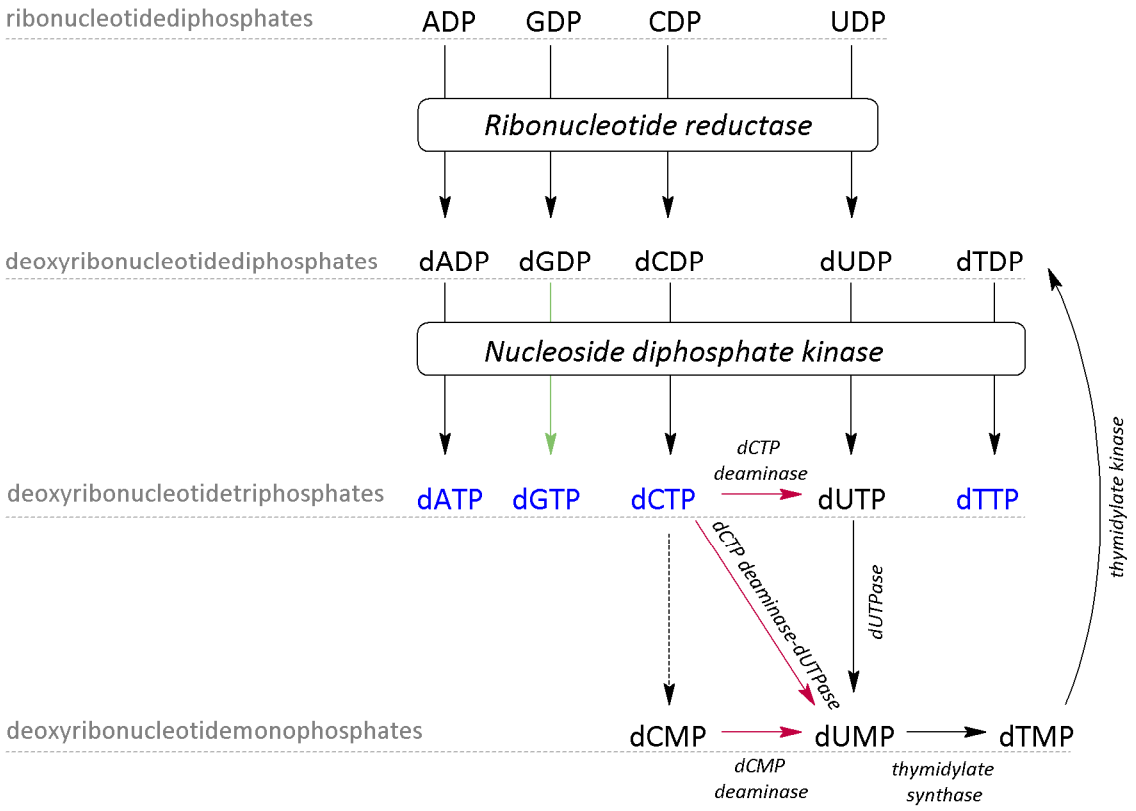


Figure 6.1: Schematic view of pathways for deoxyribonucleotide synthesis. Annotations: enzymes in italic, DNA building blocks in blue, dTTP inhibition as red arrows, dTTP activation as green arrow.

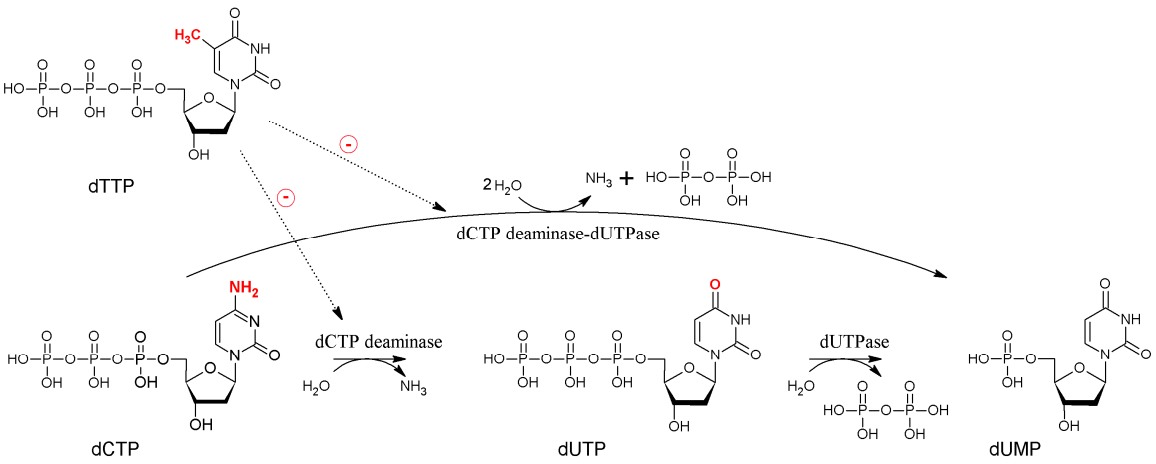


Figure 6.2: Schematic view of reactions catalyzed by dCTP deaminase, dUTPase and the bifunctional enzyme dCTP deaminase-dUTPase. dTTP is shown and dotted arrows illustrate the inhibition by dTTP on dCTP deaminase and dCTP deaminase-dUTPase. Differences between nucleotides are highlighted in red.

Table 6.1: Numbering of key residues in dCTP deaminase-dUTPase (DCD-DUT) from *M. tuberculosis* (*Mt*), *M. jannaschii* (*Mj*), monofunctional dCTP deaminase (DCD) from *E. coli* (*Ec*) and monofunctional dUTPase (DUT) from *Mt* and *Ec*. Residues involved in the DCD and DUT reactions are colored green and red, respectively, and the absolutely conserved serine is shown in blue.

Mt DCD-DUT	*Mj* DCD-DUT	*Ec* DCD	*Mt* DUT	*Ec* DUT
His112	His128	His121	Val76	Gly83
Gly116	Gly132	His125	Gly80	Gly87
Ser102	Ser118	Ser111	Ser65	Ser72
Arg106	Arg122	Arg115	Thr69	His76
Ala115	Ala131	Ala124	Pro79	Val86
Glu129	Glu145	Glu138	Ala93	Ser100
Asp119	Asp135	Asp128	Asp83	Asp90
Gln148	Gln163	Ala157	Gln113	Gln119

6.2 dCTP deaminase

dCTP deaminase catalyzes the deamination of dCTP thereby forming dUTP, see Figure 6.2. The structure of dCTP deaminase is a trimer with three active sites, each of which is composed of residues from two subunits of the trimer [88].

Kinetic studies show that the true substrate is dCTP·Mg^{2+} [89]. This is in agreement with the structure of dCTP deaminase from *E. coli* with bound reaction product, dUTP (pdb 1XS1), as well as with the structures of the E138A enzyme with bound substrate, dCTP (pdb 1XS4), or product, dUTP (pdb 1XS6). The structures all have a magnesium ion octahedrally coordinated to the α-, β- and γ-phosphate of the nucleotide and three water molecules. The magnesium ion does not play a catalytic role, but it is important for structurally arranging the nucleotide phosphates and for shielding their negative charge, thereby allowing the C-terminal to close over the active site nucleotide-Mg^{2+} complex [88].

Two suggestions regarding the catalytic mechanism of dCTP deaminase have been made, both of which involve a nucleophilic attack on the pyrimidine ring of dUTP by an activated water molecule [88]. The starting point of the reaction is based on the crystal structure of *Ec* DCD with dUTP bound (pdb 1XS1). Two hydrogen bonds (from Val136 and Gln182 to the pyrimidine ring of dUTP) anchor and orient the nucleotide pyrimidine ring upon binding to the active site. One water molecule in the active site is held in place by hydrogen bonds to Glu138, Ala124 and Ser111 and is activated by Glu138, see Figure 6.3A. Also, both mechanisms involve the release of ammonia and the completely conserved residues Glu138, Ala124 and Ser111, which are crucial for activity based on mutational studies [88,90].

dCTP deaminase can exist in two conformations: an active stabilized by substrate binding and an inactive stabilized by dTTP binding to the active site [91]. In the structure of E138A *Ec* DCD in complex with dTTP (pdb 2J4Q), amino acid residues 120-125 rearrange in the dTTP complex

compared with the dCTP and dUTP complexes. Furthermore, the entrance to the active site in the dTTP complex site is partly collapsed, as the region of α-helix 2 and β-strand 5 moves towards the active site in the absence of an ordered C-terminal fold. The nucleophilic water molecule, positioned between Glu138 and Ala124 in the dCTP and dUTP structure [88], appears to be expelled in the dTTP structure by the side chain of His121. Additionally, the side chains of Val122 and Thr123 in the dTTP structure move to new positions. The side chains of Val122 from one subunit and Thr123 from a neighboring molecule are likely to clash in the center of the trimer unless the subunits all are in the same conformation. Changing His121 to alanine results in a loss of activity, and the structure of H121A:dCTP almost superimposes with the E138A:dTTP complex in the region of residues 120-125. Both structures also have disordered C-terminals. These structural agreements support that the enzyme is inhibited through a change to an inactive conformation [91].

6.3 dUTPase

dUTPase catalyzes the hydrolysis of dUTP to dUMP. Depletion of toxic concentrations of dUTP in the cell is crucial to keep uracil out of DNA [92]. dUTPases have been well studied, since drugs designed for specific inhibition of dUTPase can be used for treatment of e.g. cancers, retroviral infections and tuberculosis [93,94]. Consequently, several structures of dUTPases from different organisms are available. Human dUTPase [95], bacterial (*Ec* [96] and *Mt* [93]) dUTPases and viral dUTPases (feline immunodeficiency [97] and equine infectious anaemia virus [98]) are trimeric enzymes with three active sites at the subunit interfaces. All subunits contribute to each binding site as the C-terminal extends to cover the opposite binding site. The C-terminal loop is essential for catalysis, as removal or point mutation of important residues in the motif results in a loss of activity [99,100]. dUTPase requires Mg^{2+} for the reaction to take place [99]. Mg^{2+} is octahedrally coordinated to the α-, β- and γ-phosphate of the nucleotide and three water molecules, which are held in place in the active site by hydrogen bonds to conserved active site residues Asp24, Asp28 and Arg140 in *Mt* dUTPase [93]. It has been proposed that Mg^{2+} plays a key role in ordering the C-terminal loop, both by direct coordination to the γ-phosphate and through a bifurcated hydrogen bond between the γ-phosphate and Arg140 [93].

In *Mt* dUTPase, a water molecule is positioned for a direct in-line nucleophilic attack and held in place by hydrogen bonds to the conserved residues Asp83 and Gln113. Asp83 can act as a general base for the water molecule by abstracting a proton, which results in a nucleophilic hydroxide ion. A hydrogen bond from Gln113 to the remaining hydrogen helps direct the electron lone pair of the water oxygen toward the α-phosphorous, see Figure 6.3B. The pyrophosphate leaving group is stabilized by hydrogen bonds to Gly66 and Arg64 as well as by a bifurcated hydrogen bond to Arg140. Negative charge on the α-β bridging oxygen can be stabilized by a hydrogen bond to a strictly conserved hydroxyl side group of Ser65 [93,101]. Point mutations of the conserved aspartate, acting as a general base (Asp84) in human endogenous retrovirus, and of the conserved serine in *E. coli* (Ser72), cause complete loss or great reduction of activity and thus emphasize the importance of these residues [94,102].

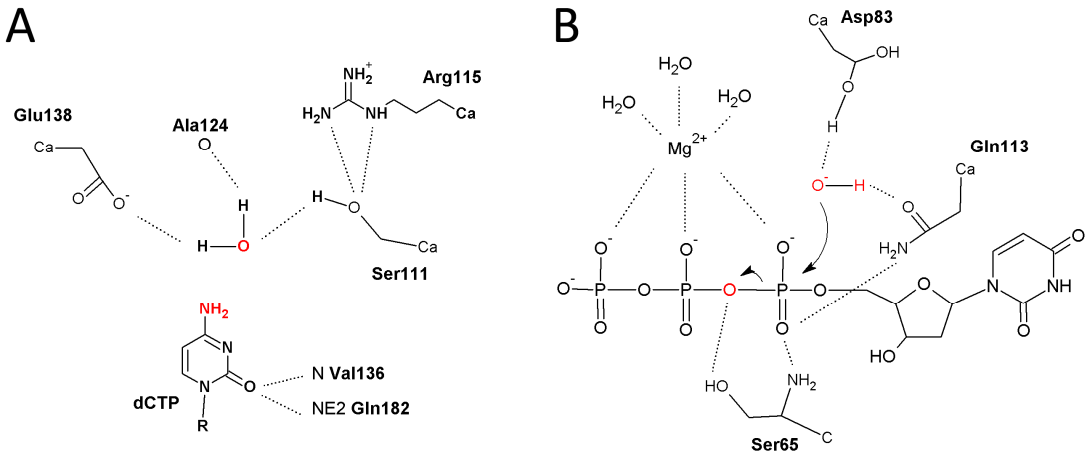


Figure 6.3: Schematic reaction starting points of: A: dCTP deaminase reaction based on the structure of *Ec* DCD [88], shown in red are the catalytic water oxygen and the amine group, which is replaced to expel ammonia. B: dUTPase reaction based on the structure of *Mt* dUTPase [93], shown in red are the nucleophilic hydroxide ion and the oxygen bridging the α-β-phosphorous, on which negative charge is formed during the course of the reaction.

6.4 dCTP deaminase-dUTPase

The bifunctional dCTP deaminase-dUTPase catalyzes deamination of dCTP and hydrolysis of dUTP to form dUMP and is inhibited by dTTP. The bifunctional enzyme is found in *M. jannaschii* [86,87,103] and in *M. tuberculosis* [85]. The two reactions are tightly coupled in the bifunctional enzymes and no dUTP is released during the course of the reaction. The enzymes are homotrimeric with three active sites consisting of residues from two subunits and require Mg^{2+} for activity. Residues important for dCTP deamination (section 6.2) are Ser102, Arg106, Ala115 and Glu129 in *Mt* DCD-DUT, which are conserved amongst enzymes with dCTP deamination activity. Residues important for dUTPase activity (section 6.3) in *Mt* DCD-DUT are Ser102, Asp119 and Gln148, which are conserved amongst enzymes with dUTPase activity. Glu129 corresponds to Glu145 in *Mj* DCD-DUT and changing this residue to Gln causes a complete lack of deaminase activity, while retaining 25 % dUTPase activity. Asp119 corresponds to Asp135 in *Mj* DCD-DUT and changing this residue to Asn causes a complete loss of activity [87]. DCD-DUT from *Mt* shows a higher affinity for dUTP compared with DCD-DUT from *Mj* [85], albeit the activity is still between 10- and 100-fold less than monofunctional dUTPases [99,104]. The reaction rate (k_{cat}) of the DUT reaction of *Mt* DCD-DUT is twice that of the bifunctional reaction, *i.e.* with dCTP as substrate, indicating deamination is rate limiting in the DCD-DUT reaction [85].

The crystal structure of *Mt* DCD-DUT:dTTP (pdb 2QXX) shows the pyrimidine moiety of dTTP forming hydrogen bonds with Arg106, Gln174, Thr127 and two water molecules. A magnesium ion is coordinated to the α-, β- and γ-phosphate of the nucleotide and three water molecules in an octahedral coordination sphere. The triphosphoryl is hydrogen bonded to Lys101, Ser102, Ser103, and Gln148 as well as to Tyr162 and Lys170 from the C-terminal lid covering the binding site. This lid is disordered in the apo structure of *Mt* DCD-DUT (pdb 2QLP). Another difference

between the apo and dTTP structures is a rearrangement of residues 109-118. The arrangement of these residues determines whether the enzyme is in its active or inactive conformation. Binding the inhibitor dTTP causes the enzyme to adopt the inactive conformation. The conformation of the key residues for catalytic activity is unchanged in the two conformations and the inability to cleave dTTP has been attributed to the lack of a key water molecule near Ala115 in the structure of *Mt* DCD-DUT:dTTP [85].

The crystal structure of DCD-DUT from *Mj* with the substrate analogue dUMPNPP bound (E145A variant, pdb 2HXD) shows the same conformation of the flexible region near the nucleotide binding side as *Ec* DCD with dUTP bound (pdb 1XS1). In addition, the ordered C-terminal interacts with the nucleotide phosphate moiety, which is held in place by coordination to Mg^{2+}, whereas the C-terminal is disordered in the structure of apo *Mj* DCD-DUT (pdb 2HXB, 1OGH). A conformational change is observed when substrate binds to the enzyme, which is in agreement with a mechanism where the enzyme exists in an active and inactive form and substrate binds to the active form of the enzyme [105].

6.5 Structural Comparison of the Enzymes

There is a high sequence and structural similarity between DCD, DCD-DUT and DUT enzymes. All structures are trimeric with active sites placed on the interfaces between molecules. *Mt* DCD-DUT shows a greater similarity with *Ec* DCD than with *Mt* DUT, both in sequence and structure. Figure 6.4 shows a structural overlay of monomers from the three enzymes illustrating the structural similarities and differences. A main structural difference between *Mt* DCD-DUT and *Mt* DUT is the C-terminal lid. In the DUT enzyme, the lid reaches across the monomer to form active sites composed of residues from all three subunits, whereas the active site in the bifunctional enzyme consists of residues from two neighboring subunits.

Figure 6.4: Ribbon view of structural overlay of *Mt* DUT:dUMPNPP (blue, pdb 1SIX), *Mt* DCD-DUT:dTTP (red, pdb 2QXX) and *Ec* DCD:dUTP (green, pdb 1XS1) with nucleotides shown in respectively colored sticks. The view is perpendicular to the 3-fold rotation axis. Arrows point to the close resemblance of the DCD (green) and DCD-DUT (red) structures, while the DUT (blue) structure shows some more variation.

Alignment based on structure and sequence is given in Figure 6.5 for the two bifunctional enzymes from *Mj* and *Mt* and the monofunctional *Ec* DCD and DUT from various sources. The two bifunctional enzymes include conserved residues important for both dCTP deaminase and dUTPase activity as described in section 6.4 and the monofunctional *Ec* DCD shows a key mutation at Ala157, which is a Gln in the structures with DUT function. Gly116 in *Mt* DCD-DUT has been suggested to be important in bifunctional enzymes [85] and is seen to be conserved also in enzymes with DUT activity.

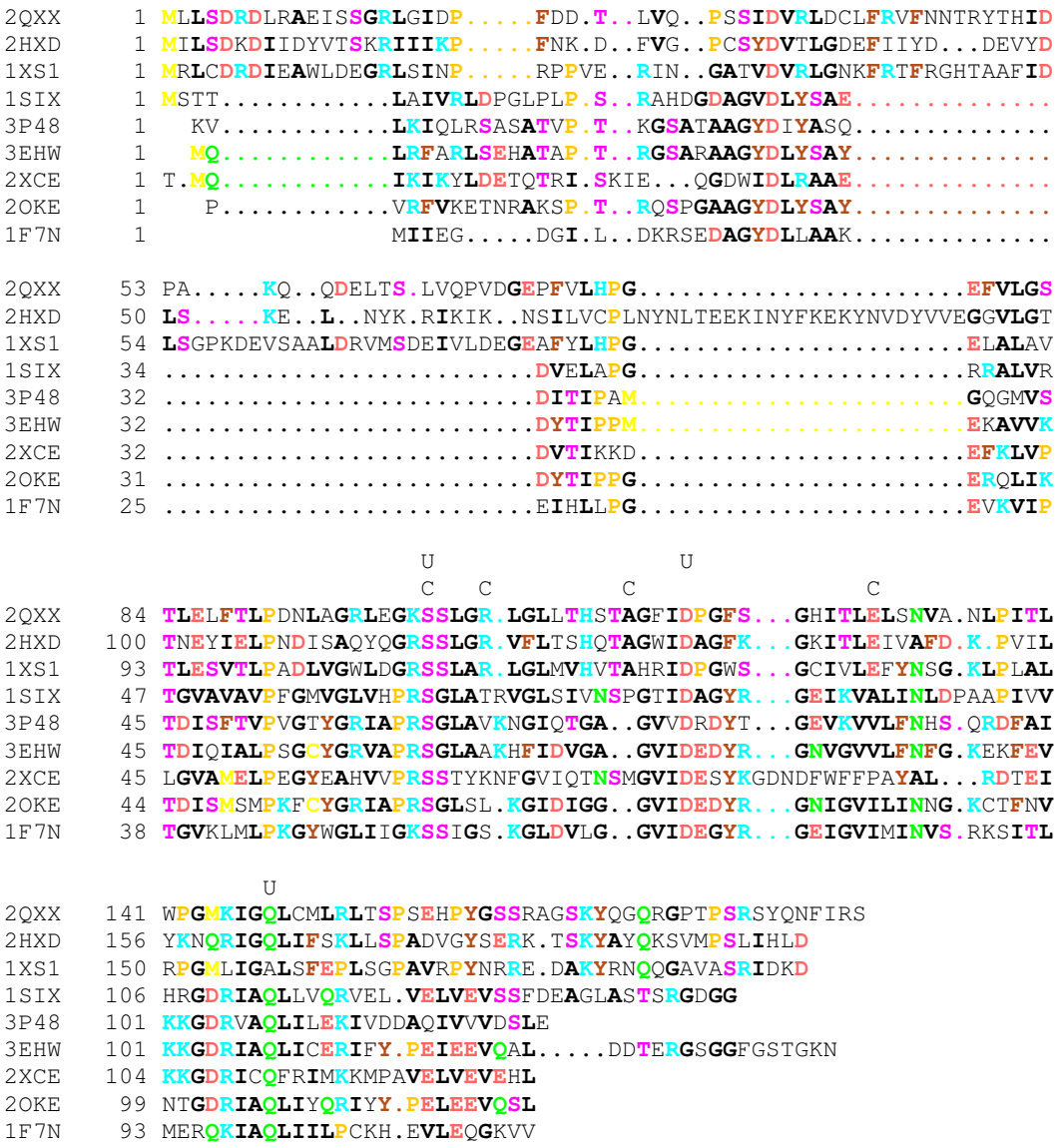


Figure 6.5: Alignment based on sequence and structure using *STRAP* [106,107] on DCD-DUT from *Mt* (pdb 2QXX) and *Mj* (pdb 2HXD excl. mutation), *Ec* DCD (pdb 1XS1) and DUT from *Mt* (pdb 1SIX), *Saccharomyces cerevisiae* (pdb 3P48), human (pdb 3EHW), *Bacillus subtilis* (pdb 2XCE), *Vaccina virus* (pdb 2OKE) and feline immonodeficiency virus (pdb 1F7N). C marks residues important for deamination and U marks residues important for dephosphorylation.

Chapter Seven

7 *Mycobacterium tuberculosis* dCTP deaminase-dUTPase: Mutational Background and Equations for Kinetic Studies

7.1 Structural reasoning behind the mutational study

The bifunctional *Mt* DCD-DUT enzymes are inhibited by dTTP, which is structurally very similar to the substrates and binds to the active site of the enzyme. The additional 5-methyl group of dTTP induces a structural reorganization near the active site of the enzyme. The inhibition is not fully relieved by an increase in substrate concentration suggesting a noncompetitive mechanism [85]. The residues involved directly in catalysis do not, however, rearrange in the active and inactive forms of the enzyme. Helt *et al.* ascribed the lack of dTTP hydrolysis to a water molecule, which was expelled from the WT structure with dTTP bound (pdb 2QXX) because of the Ala115 side chain [85]. This forms the basis of the mutational study presented here. The reasoning behind the A115V variant is that an increase in size of the side chain will prevent dTTP from binding because of a steric clash of V115 with the 5-methyl group of dTTP thereby leaving the enzyme unaffected by the presence of dTTP. The reasoning behind the A115G variant is an attempt to dephosphorylate dTTP. With dTTP bound, Gly115 would allow enough space for the water molecule suggested to be crucial for activity.

7.2 Steady state kinetics

Equations important for data treatment of steady state kinetic experiments are listed in this section. The Michaelis-Menten equation (7.1) is a central expression for steady state enzyme kinetics and is applied when the reaction rate, v, has a hyperbolic dependence of the substrate concentration, $[S]$.

$$v = \frac{V_{max} \times [S]}{K_m + [S]} \tag{7.1}$$

K_m is the substrate concentration that results in half maximal reaction rate, V_{max} (at $[S] = K_M$, $v = V_{max}/2$). The rate constant, k_{cat}, reflects summation of multiple chemical steps and provides the

maximal rate of an enzymatic reaction at given enzyme concentration, [*E*], and infinite availability of substrate *S*, equation (7.2).

$$V_{max} = k_{cat} \times [E] \tag{7.2}$$

Oligomeric enzymes, in which substrate binding at one active site influences the substrate binding affinity of the other active sites in the oligomer deviate from the hyperbolic dependence of the reaction rate on the substrate concentration. This cooperativity can be either positive or negative, reflecting an increase or a decrease, respectively, in the affinity of substrate binding to other active sites subsequent to substrate binding to one active site. Cooperativity influences the measured value of the reaction rate as described by the Hill equation (7.3). The Hill coefficient, *h*, describes cooperativity, while *K´* is related to the dissociation constant of the enzyme-substrate complex and factors describing the influence of substrate binding at one site on the affinity of the other sites. The Hill equation is also described by equation (7.4), but with k_{cat} and $S_{0.5}$ (the substrate concentration at half maximal reaction rate).

$$v = \frac{V_{max} \times [S]^h}{K' + [S]^h} \tag{7.3}$$

$$v = \frac{k_{cat} \times [S]^h}{S_{0.5}^h + [S]^h} \tag{7.4}$$

Equation (7.5) was used to describe ligand binding to an enzyme. *N* is the degree of ligand binding with dissociation constant K_d to an enzyme with a maximal number of binding sites N_{max} at a given concentration of the ligand, [*L*].

$$N = \frac{N_{max} \times [L]}{K_d + [L]} \tag{7.5}$$

Inhibition studies with constant substrate concentration and varying inhibitor concentration are described using equation (7.6) or (7.7) for a hyperbolic or sigmoidal dependence, respectively, of the reaction rate on the inhibitor concentration.

$$v = \frac{v_0}{1 + \left(\frac{[I]}{K_i}\right)} \tag{7.6}$$

$$v = \frac{v_0}{1 + \left(\frac{[I]}{K_i}\right)^n} \tag{7.7}$$

where v_0 is the reaction rate in the absence of the inhibitor, *I*, and K_i is the inhibitor concentration resulting in half reaction rate (at $[I] = K_i$, $v = v_0/2$).

Radioactive nucleotide labeling and liquid scintillation counting were used experimentally. Equation (7.8) was used to calculate reaction rates (s^{-1}) from the experimentally obtained

counts-per-minute (*CPM*) for dCTP and was used equivalently for experiments with dUTP as substrate.

$$v_{(s^{-1})} = \frac{[dCTP]_0 \times CPM_{MP}}{(CPM_{MP} + CPM_{TP})} \times \frac{F}{t_{(s)} \times c_{enzyme}} \tag{7.8}$$

$[dCTP]_0$ is the initial concentration of dCTP in mM, *MP* is monophosphate nucleotide, *TP* is triphosphate nucleotide, *F* is the factor corresponding to the total enzyme dilution, *t* is the reaction time in seconds and c_{enzyme} is the enzyme concentration in mM.

Equation (7.9) was used to calculate the degree of dTTP binding (N_{dTTP}) to the enzyme.

$$N_{dTTP} = [dTTP]_0 \times \frac{(CPM_0 - CPM_{free})}{CPM_{total}} \times \frac{F}{c_{enzyme}} \tag{7.9}$$

$[dTTP]_0$ is the initial concentration of dTTP in mM, CPM_0 is the initial counts-per-minute, while CPM_{free} is counts-per-minute for unbound ligand.

Chapter Eight

8 Purification of WT and Two Variants of *Mycobacterium tuberculosis* dCTP deaminase-dUTPase

8.1 Introduction

A prerequisite to protein characterization is acquiring purified protein and this chapter describes purification procedures of WT, A115G and A115V *Mt* DCD-DUT. All solutions were prepared with milli-q water. Protein solutions were kept on ice at all times and centrifugations were at 4 °C in a Sorvall R5C centrifuge with a SS-34 rotor unless otherwise stated. During the purifications, problems with protein precipitation were encountered several times, and a more suitable buffer was found through solubility experiments (results are given in section 8.3.5).

8.2 Experimental

8.2.1 Expression of the A115G variant

All instruments that came in contact with the cells were sterilized to avoid contamination. A glass pipette was used to transfer BL21 (DE3) *E. coli* cells over-expressing A115G *Mt* DCD-DUT into 5 mL LB media (containing 1 % tryptone, 0.5 % yeast extract, 0.5 % NaCl) with 100 µg/mL ampicillin, and the cells were incubated overnight at 37 °C. 2 mL of the overnight culture was added to approximately 250 mL media consisting of 225 mL LB media with 2.2 mM $MgCl_2 \cdot 6H_2O$, 0.11 mM $CaCl_2 \cdot 2H_2O$, 3.3 µM $FeCl_3 \cdot 6H_2O$) and 25 mL A media (containing 2 % $(NH_4)_2SO_4$, 7.5 % $Na_2HPO_4 \cdot 2H_2O$, 3 % KH_2PO_4, 3 % NaCl) and 2.5 mL 20 % glucose (or 0.25 mL 20 % glucose in a different expression) and 25 mg ampicillin (100 µg/mL final concentration) in a 1 L conically shaped flask. The culture was incubated with shaking at 37 °C and 250 µL 500 mM IPTG was added around OD_{436}=6. The culture was left overnight with shaking at 37 °C. The cells were transferred to a large centrifuge tube on ice and centrifuged for 15 minutes at 9,000 rpm using a Sorvall R5C centrifuge with a GSA rotor. The supernatant was discarded, and the cells were resuspended in 30 mL 0.9 % NaCl and transferred to two small centrifuge tubes. The cells were centrifuged for 15 minutes at 9,000 rpm. The supernatant was discarded and the cells were frozen at -20 °C.

8.2.2 Purifications of WT and the variants

A number of purifications were carried out on dCTP deaminase-dUTPase from *Mt* on WT, A115G and A115V variants. A combination of techniques was used with some variations due to precipitation and degradation issues. A standard purification is described below and detailed descriptions of purifications along with specifications of SDS-PAGE samples are given in Appendix D. An overview of the purifications is shown in Figure 8.1.

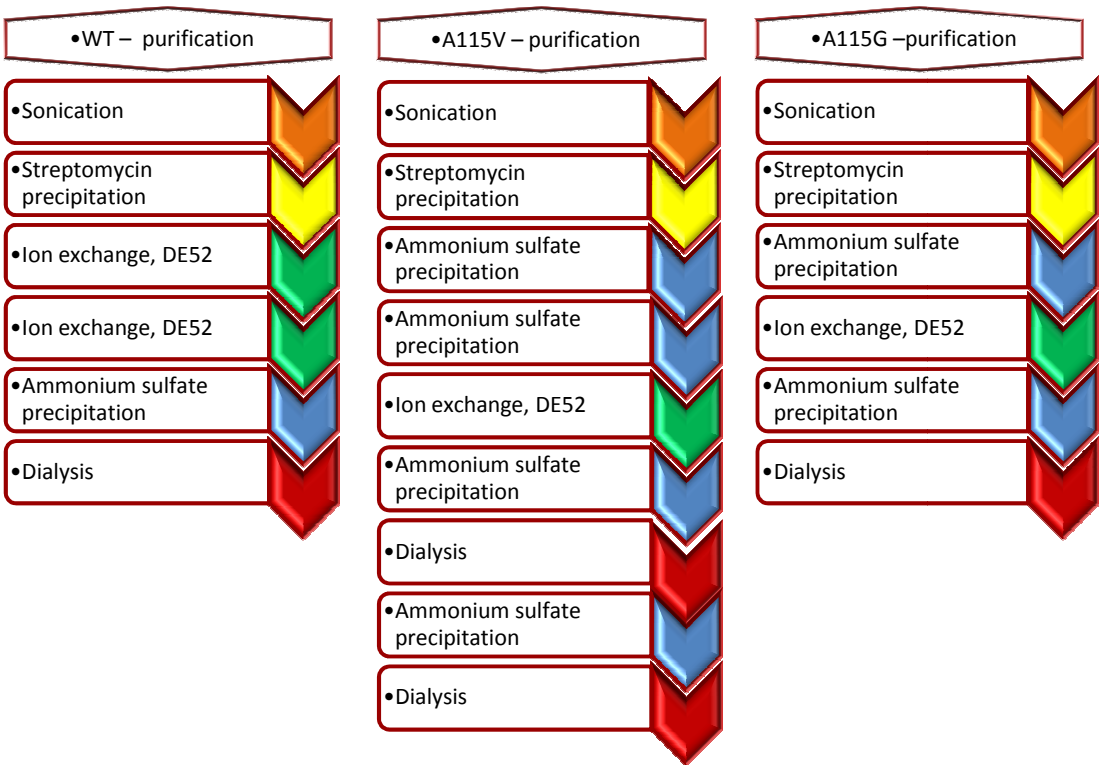


Figure 8.1: Diagram of the steps involved in purification of WT, A115G and A115V *Mt* DCD-DUT.

8.2.2.1 Standard Purification

Cell pellet from 250 mL culture was defrosted on ice. 20 mL 50 mM KH_2PO_4/K_2HPO_4 pH 6.8 was added and the cells resuspended. The cells were lysed by sonication twelve times for 30 seconds while kept on ice. The suspension was allowed to cool down for 30 seconds in between sonications. The suspension was centrifuged for 20 minutes at 10,000 rpm. The supernatant was decanted off and kept on ice. While stirred on ice, 10 (w/w)% streptomycin was slowly added to a final concentration of 1 %. The solution was left for 30 minutes with gentle stirring on ice and then centrifuged for 20 minutes at 14,000 rpm. The supernatant was decanted off and kept on ice. The solution was gently stirred on ice and solid $(NH_4)_2SO_4$ was slowly added to 40 % saturation. The solution was stirred for 5 minutes and then centrifuged for 20 minutes at 14,000

rpm. The supernatant was discarded and the pellet was dissolved in 10 mL 50 mM KH_2PO_4/K_2HPO_4 pH 6.8. The solution was dialyzed against 500 mL 50 mM KH_2PO_4/K_2HPO_4 pH 6.8.

The protein solution was loaded onto a 14 mm/30 cm DE52 anion exchange column equilibrated with 50 mM KH_2PO_4/K_2HPO_4 pH 6.8 and eluted using a linear salt gradient of 50 mM KH_2PO_4/K_2HPO_4 pH 6.8 to 0.4 M NaCl in 50 mM KH_2PO_4/K_2HPO_4 pH 6.8 over 4 CV. The absorbance at 280 nm and SDS-PAGE was used to determine the protein content and purity of the collected fractions. Selected fractions were pooled and solid $(NH_4)_2SO_4$ was added to 40 % saturation while stirring the solution on ice. The solution was stirred for 5 minutes and then centrifuged for 20 minutes at 14,000 rpm. The supernatant was discarded and the pellet was dissolved in 10 mL 20 mM HEPES pH 8.0 (the choice of buffer was based on solubility experiments on the A115V variant, results are given in section 8.3.5) and dialyzed against the same buffer.

8.2.3 SDS-PAGE

Samples for SDS-PAGE were mixed with equal volumes of sample buffer (80 mM Tris/HCl pH 6.8, 2 % Sodium Dodecyl Sulfate, 10 % Glycerol, 5 % Dithiothreitol, 0.02 % Bromophenol blue). Solid samples were dissolved in 95 µL sample buffer. Biorad Broad Range Marker was used on all gels and prepared by mixing 10 µL marker with 10 µL sample buffer. All samples were heated 5 minutes at 95 °C prior to loading onto the gel. Gel casting is described in Appendix D.6.

The mounted gel was placed in running buffer (25 mM Tris, 1.4 % Glycine, 0.1 % Sodium Dodecyl Sulfate) and the samples were loaded. An electric current of 25 mA was applied for approximately 1 hour until the samples had traveled an appropriate distance on the gel. The gel was stained with Coomassie Brilliant Blue R and then bleached with a bleaching solution (10 % Acetic Acid, 24 % Ethanol) for at least 2 hours prior to drying the gel under vacuum.

8.2.4 Solubility Experiments

Purified A115V *Mt* DCD-DUT (in 20 mM HEPES/NaOH pH 6.8 that initially was used as buffer) was defrosted on ice and centrifuged for 10 minutes at 20,000 rpm (Ole Dich centrifuge). The supernatant was decanted off and the solubility of the pellet was tested.

The pellet was dissolved in 300 µL 0.5 M KCl in 20 mM HEPES/NaOH pH 6.8. The absorbance at 280 nm was used to determine the protein content of the dissolved pellet and of the supernatant in 20 mM HEPES/NaOH pH 6.8.

80 µL 1 M KCl in 20 mM HEPES/NaOH pH 6.8 was gradually added to 200 µL protein in 20 mM HEPES/NaOH pH 6.8 to a final concentration of 0.3 M KCl in 20 mM HEPES/NaOH pH 6.8. To test if the dilution alone could dissolve the protein, 80 µL 20 mM HEPES/NaOH pH 6.8 was added to

200 µL protein in 20 mM HEPES/NaOH pH 6.8. Additionally 120 µL 20 mM HEPES/NaOH pH 6.8 was added to this tube.

The solubility of additional protein pellets were tested in 20 mM Tris/HCl pH 7.4, 200 mM HEPES/NaOH pH 8.0 and 20 mM HEPES/NaOH pH 8.0.

8.3 Results and Discussion

8.3.1 Expression of the A115G variant

E. coli (BL21) cells over-expressing A115G *Mt* DCD-DUT were grown twice with two different glucose concentrations, see Figure 8.2. IPTG was added at the final point on the curve and the cells were left overnight. OD_{436} was 11.9 at cell harvest when the cells were grown in media containing 0.02 % glucose, while OD_{436} was 17.8 at cell harvest when the cells were grown in 0.2 % glucose.

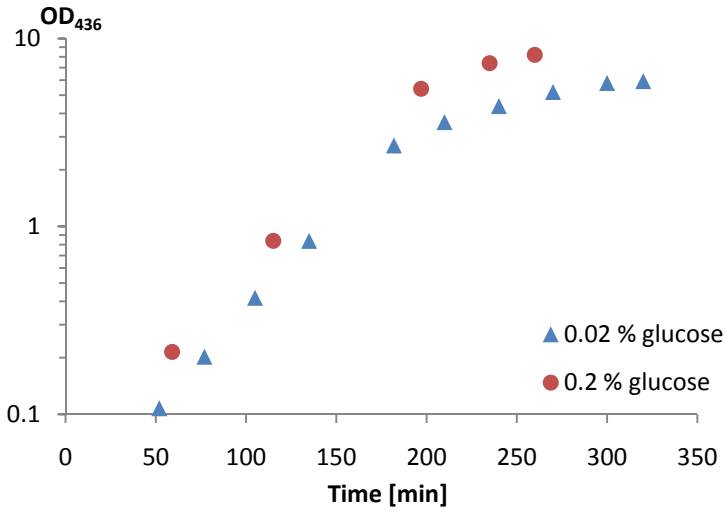


Figure 8.2: Growth of *E. coli* over-expressing A115G *Mt* DCD-DUT at two different glucose concentrations.

8.3.2 Purification of WT

The steps involved in the purification of WT *Mt* DCD-DUT are shown in Figure 8.1. No ammonium sulfate precipitation was carried out prior to the two anion exchange steps. SDS-PAGE and the absorbance at 280 nm of the fractions from both anion exchange steps are shown in Figure 8.3. Fractions 6-8 were collected during the first run, and fractions 10-12 were collected during the second run. The final ammonium sulfate precipitation was carried out to 45 % saturation and the precipitate dissolved in 2.5 mL 20 mM HEPES pH 8.0 and dialyzed against the same buffer.

Protein with a concentration of 2.5 mg/mL was apportioned into aliquots and stored at -20 °C. SDS-PAGE of purified WT *Mt* DCD-DUT is shown in Figure 8.6A.

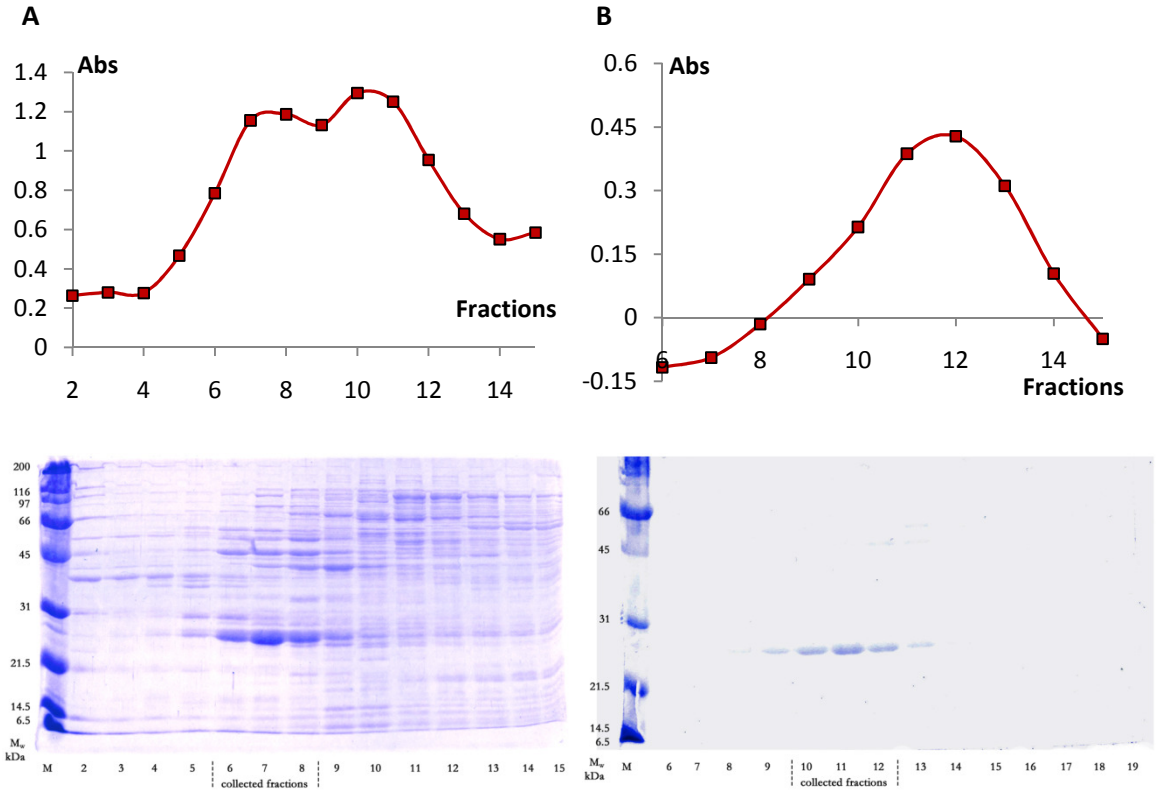


Figure 8.3: Purification of WT *Mt* DCD-DUT: SDS-PAGE with collected fractions marked and the absorbance at 280 nm (the connecting line serves as guide to the eye). A: fractions from the first anion exchange. B: fractions from the second anion exchange.

8.3.3 Purification of the A115V variant

The purification of A115V *Mt* DCD-DUT was followed using SDS-PAGE samples, and the results are shown in Figure 8.4A. The figure shows the presence of the protein in the supernatant after centrifuging the sample subsequent to sonication (V2). A slight loss of protein is seen in the pellet (V3). The protein is present in the supernatant after centrifuging the sample subsequent to streptomycin precipitation (V4) and in the dissolved pellets of the ammonium sulfate precipitations (V7 and V9). Solution V9 was quite pure, but because of a slight yellow discoloration of the solution, an anion exchange step was incorporated. The sample (V9) was loaded onto the DE52 anion exchange column, and SDS-PAGE and the absorbance at 280 nm of the fractions are shown in Figure 8.4B. Fractions 9-13 were collected and pooled. The solution was precipitated with ammonium sulfate and the pellet was readily dissolved in 5 mL 50 mM KH_2PO_4/K_2HPO_4 pH 6.8. The solution was dialyzed twice, during which it precipitated. The solution was centrifuged and the supernatant apportioned into aliquots. The aliquots had a tendency to precipitate during transfer on ice. Therefore, solubility experiments were made to

find a more suitable buffer (see section 8.3.5). These experiments showed that the solubility of the protein was enhanced at pH 8.0 compared with pH 6.8. Consequently, the A115V variant aliquots were pooled and precipitated with ammonium sulfate, and the precipitate was readily dissolved in 3 mL 20 mM HEPES/NaOH pH 8.0. The solution was dialyzed against the same buffer and the solution was apportioned into aliquots and stored at -20 °C. The final protein concentration was 4 mg/mL. SDS-PAGE of the purified protein is shown in Figure 8.6B.

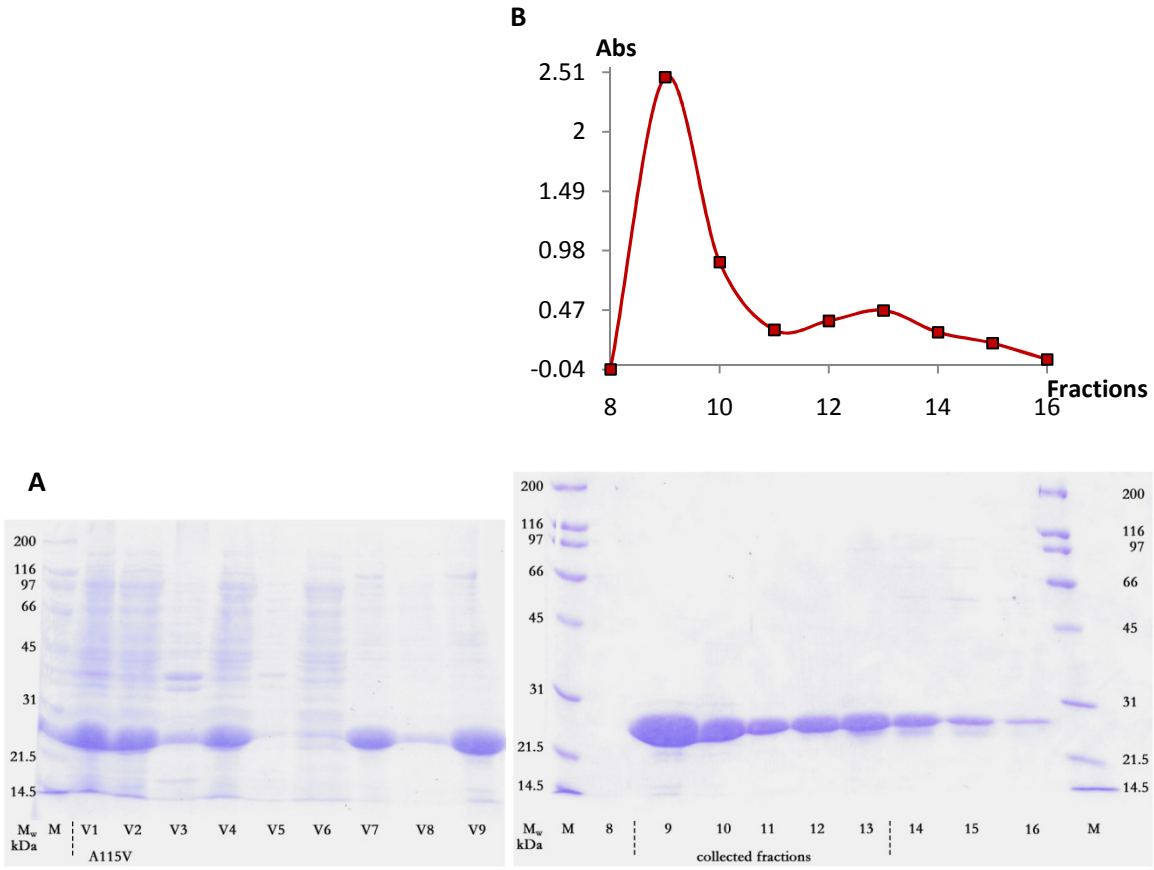


Figure 8.4: Purification of A115V *Mt* DCD-DUT. A: SDS-PAGE samples: V1: cell resuspension, V2: sonication supernatant, V3: sonication pellet, V4: streptomycin precipitation supernatant, V5: streptomycin pellet, V6: first ammonium sulfate precipitation supernatant, V7: first ammonium sulfate precipitation dissolved pellet, V8: second ammonium sulfate precipitation supernatant, V9: second ammonium sulfate precipitation dissolved pellet (see Appendix D.2 for additional details on the samples). B: SDS-PAGE of fractions from the anion exchange with collected fractions marked and the absorbance at 280 nm shown above (the connecting line serves as guide to the eye).

8.3.4 Purification of the A115G variant

Several problems with precipitation and degradation were encountered during work with A115G *Mt* DCD-DUT. This section describes the successful purification of A115G, while Appendix E.1

describes two additional purifications of the A115G variant, in which the protein degraded and was discarded.

The A115G variant protein was expressed (section 8.2.1) and purified with a modified procedure, see Figure 8.1. Pellet from the first ammonium sulfate precipitation proved difficult to dissolve (purification 2, Appendix E.1) and the volume of 50 mM KH_2PO_4/K_2HPO_4 was therefore doubled to 20 mL, in which the protein completely dissolved (G2). SDS-PAGE results of the supernatant after centrifuging the sample subsequent to sonication (G1) and of solution G2 are shown in Figure 8.5. Figure 8.5 also shows SDS-PAGE and the absorbance at 280 nm of fractions from the anion exchange. Fractions 5-9 were collected and pooled. The solution was precipitated with ammonium sulfate and the pellet was readily dissolved in 2 mL HEPES/NaOH pH 8.0. The solution was dialyzed against the same buffer, and the solution was apportioned into aliquots and stored at -20 °C. The final protein concentration was 3 mg/mL SDS-PAGE of the purified protein is shown in Figure 8.6B.

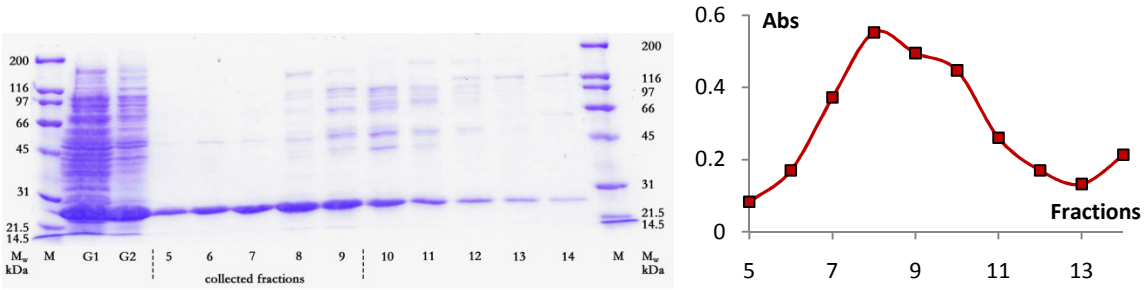


Figure 8.5: Purification of A115G *Mt* DCD-DUT. SDS-PAGE of: G1: sonication supernatant, G2: dissolved pellet from the first ammonium sulfate precipitation. Fractions 5-14 from anion exchange are shown. The collected fractions are marked, and the absorbance at 280 nm is shown to the right (the connecting line serves as guide to the eye). See Appendix D.3 for more details on the samples.

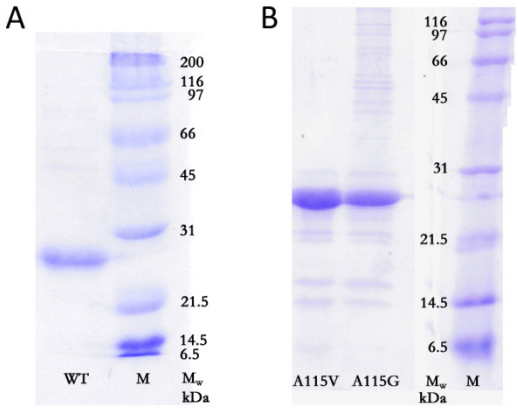


Figure 8.6: SDS-PAGE of the purified proteins. A: WT *Mt* DCD-DUT, B: A115V and A115G diluted four times.

8.3.5 Solubility Experiments

Solubility experiments were carried out to find more suitable conditions where the proteins were soluble. Increasing ion strength and changing pH were tested. However, the proteins were going to be used for crystallization and addition of salt should preferably be avoided.

Solubility experiments were performed on A115V *Mt* DCD-DUT that had precipitated during dialysis against 20 mM HEPES/NaOH pH 6.8 (Figure 8.1, A115V second to last dialysis). Supernatant and pellet were separated by centrifugation, and the total protein content was determined to be 3.8 mg/mL in the precipitated protein solution.

The protein was soluble in 0.3 M KCl in 20 mM HEPES/NaOH pH 6.8 at a protein concentration of 2.7 mg/mL. With a protein concentration of 2.7 mg/mL, the protein was not soluble in 20 mM HEPES/NaOH pH 6.8. Further dilution to a protein concentration of 1.9 mg/mL in 20 mM HEPES/NaOH pH 6.8 did not dissolve the protein. At the same protein concentrations, the protein was not soluble in 20 mM Tris/HCl pH 7.4.

The protein pellet proved easily soluble in both 200 mM HEPES/NaOH pH 8.0 and 20 mM HEPES/NaOH pH 8.0 up to concentrations of approximately 6 mg/mL. The preferred buffer was thus 20 mM HEPES/NaOH pH 8.0 and protein solutions were stored in this buffer at -20 °C. Defrosting the A115V variant enzyme and leaving it on ice for more than 20-30 minutes sometimes caused precipitation. This precipitation could be dissolved by diluting the protein to 2 mg/mL with 20 mM HEPES pH 8.0.

8.3.6 Comments on Stability

The successful purification of the intact A115G variant is distinguished from the two unsuccessful purifications of the A115G variant given in Appendix E.1 by being carried out in 3 days instead of 9 days, and precipitation issues were solved by changing pH during the final dialysis step. During purification 3 described in Appendix D.5, degradation was not observed until the protein was frozen at -20 °C prior to the final ammonium sulfate precipitation. Protein degradation was evident from the final SDS-PAGE shown in Figure E.3B.

Successful purification of the A115V variant was carried out in 9 days and no degradation was observed. Purified A115V protein used in the solubility experiments (section 8.2.4) was left refrigerated for an additional 9 days and the SDS-PAGE results are shown in Figure 8.7. SDS-PAGE lanes A, B, C, D refer to the protein in 0.3 M KCl in 20 mM HEPES pH 6.8, 20 mM HEPES pH 8.0, 20 mM HEPES pH 6.8 with precipitation and 20 mM HEPES pH 6.8 centrifuged supernatant without precipitation, respectively. It is seen that especially C shows degradation. The C solution was left with precipitation for the additional 9 days.

To avoid degradation, it is advantageous to minimize the time a solution is left with precipitation, and to carry out the purification quickly and without freezing the protein until the purification is complete – in particular when working with A115G *Mt* DCD-DUT.

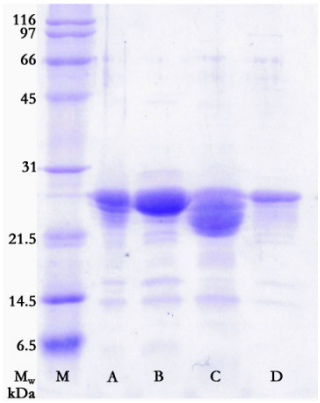


Figure 8.7: Purified A115V *Mt* DCD-DUT after 9 days storage at 4 °C (section 8.3.6). A: protein in 0.3 M KCl in 20 mM HEPES pH 6.8, B: protein in 20 mM HEPES pH 8.0, C: protein in 20 mM HEPES pH 6.8 with precipitation, D: protein in 20 mM HEPES pH 6.8 without precipitation.

8.4 Conclusion

WT *Mt* DCD-DUT and two variants (A115V and A115G) were successfully purified with yields of 35 mg/L, 45 mg/L and 77 mg/L cell culture, respectively. Problems with precipitation of A115G and A115V *Mt* DCD-DUT during purification were encountered and changing pH from 6.8 to 8.0 increased the solubility of the proteins. Purified WT, A115G and A115V *Mt* DCD-DUT were stored at -20 °C in 20 mM HEPES/NaOH pH 8.0 with concentrations of 2.5 mg/mL, 3 mg/mL and 4 mg/mL, respectively. SDS-PAGE results of the final product of purified WT, A115V and A115G *Mt* DCD-DUT are shown in Figure 8.6.

Chapter Nine

9 Characterization of WT and Two Variants of *Mycobacterium tuberculosis* dCTP deaminase-dUTPase

9.1 Introduction

As described in section 7.1, the two variants (A115V and A115G) of *Mt* DCD-DUT were produced to investigate the importance of steric hindrance by the side chain of Ala115 on a water molecule suggested to be crucial for activity [85]. The solubility of A115V *Mt* DCD-DUT was enhanced at pH 8.0 (section 8.3.5). In the kinetic studies, pH was therefore initially changed to 8.0 instead of pH 6.8, which was used in previous WT enzyme studies [85]. Kinetic studies were expanded to include pH 6.8.

This chapter includes characterization of WT, A115V and A115G *Mt* DCD-DUT and covers kinetic studies of all three enzymes as well as dTTP binding and inhibition experiments. Crystallization experiments and the crystal structure of A115V *Mt* DCD-DUT to 2.8 Å are also included in this chapter.

9.2 Experimental

9.2.1 Turnover Numbers

Experiments used to calculate the turnover numbers of WT, A115V and A115G *Mt* DCD-DUT were all carried out at 37 °C. 80 µL assay mix containing 2 mM $MgCl_2$ and 20 mM HEPES/NaOH pH 6.8 or 8.0 and a minor amount of ^3H labeled nucleotide (dCTP or dUTP) with a temperature of 37 °C was added 10 µL nucleotide and allowed to warm to 37 °C again. Initial assay concentrations of dCTP or dUTP were 1 mM. 10 µL appropriately diluted enzyme was added and samples were taken from the mixture at specific times and loaded directly onto a TLC PEI Cellulose F chromatogram. 1 µL 5 mM dUMP was added to each sample on the chromatogram as a marker. Chromatograms were treated and samples measured as described in section 9.2.5.

WT assays were carried out in duplicate with 91 µM enzyme with in-assay dilution factors of 100 and 200. A115V assays were carried out in duplicate with 191 µM enzyme with in-assay dilution factors of 30 and 60. A115G assays were carried out in duplicate with 144 µM enzyme with in-

assay dilution factors of 10 and 20. Samples were taken after 2, 5, 10 and 20 minutes in all assays with the exception of A115G and A115V dCTP pH 8.0 assays, which were carried out in duplicate with enzyme dilution factors of 20 and 40 and sampling after 2, 5 and 10 minutes.

9.2.2 Nucleotide Saturation

All nucleotide saturation experiments were carried out at 37 °C. Numerous nucleotide saturation assays with variations in nucleotide concentrations were carried out for WT, A115V and A115G *Mt* DCD-DUT with dCTP and dUTP at pH 8.0, as well as with dUTP at pH 6.8. The assay contents varied depending on pH and nucleotide type and concentration; specifications and corresponding in-assay nucleotide concentrations and enzyme dilutions are given in Appendix D, Table D.1. 10 µL nucleotide was added to 80 µL assay mix including a minor amount of ^{3}H labeled nucleotide (dCTP or dUTP) with a temperature of 37 °C and allowed to warm to 37 °C again. 10 µL appropriately diluted enzyme with a temperature of 37 °C was added to each assay and after a reaction time of 6 minutes, 10 µL samples were loaded directly onto a TLC PEI Cellulose F chromatogram. 1 µL 5 mM dUMP was added to each sample on the chromatogram as a marker. Chromatograms were treated and samples measured as described in section 9.2.5.

9.2.3 dTTP inhibition

All dTTP inhibition experiments were carried out at 37 °C. dTTP inhibition experiments were carried out with WT, A115V and A115G *Mt* DCD-DUT at pH 6.8 and pH 8.0 with fixed concentrations of dCTP or dUTP. Assays were made with variations in both initial substrate concentration and type, and also with variation in dTTP concentration. Details on assay mix contents are given in Appendix D, Table D.2 along with dTTP concentration range, total enzyme dilution in-assay and the applied nucleotide concentrations. 10 µL dTTP was added to 80 µL assay mix including a minor amount of ^{3}H labeled nucleotide (dCTP or dUTP) with a temperature of 37 °C and allowed to warm to 37 °C again. 10 µL appropriately diluted enzyme with a temperature of 37 °C was added to the assay and after a reaction time of 6 minutes, 10 µL samples were loaded directly onto a TLC PEI Cellulose F chromatogram. 1 µL 5 mM dUMP was added to each sample on the chromatogram as a marker. Chromatograms were treated and samples measured as described in section 9.2.5.

9.2.3.1 dTTP inhibition at three pH values

The activity of WT *Mt* DCD-DUT with dCTP or dUTP as substrate was tested at pH 6.8, 7.4 and 8.0 in 20 mM HEPES in the presence (500 µM) or absence of dTTP. 20 µL assay component (pH and dTTP steering) were added to 70 µL assay mix (2 mM $MgCl_2$, 300 µM dCTP or 300 µM dUTP and a minor amount of ^{3}H dCTP or ^{3}H dUTP) and allowed to reach 37 °C. Assays were carried out in duplicates. 10 µL appropriately diluted enzyme (100x or 200x total in-assay dilutions) with a

temperature of 37 °C was added to the mixture and after a reaction time of 6 minutes, a 10 µL sample was loaded directly onto a TLC PEI Cellulose F chromatogram. 1 µL 5 mM dUMP was added to each sample on the chromatogram as a marker. Chromatograms were treated and samples measured as described in section 9.2.5.

9.2.4 dTTP Binding

Ligand binding experiments were performed at ambient temperature. The experiments contained 2 mM $MgCl_2$, 20 mM HEPES/NaOH, dTTP at different concentrations including a small amount of ^{3}H dTTP (concentration in nM scale) and enzyme. Mixing WT *Mt* DCD-DUT enzyme in 20 mM HEPES/NaOH pH 8.0 with buffer at pH 6.8 caused a pH shift, and WT experiments were thus carried out at pH 7.1 with a total enzyme concentration of 33 µM. A115V and A115G *Mt* DCD-DUT experiments were carried out at pH 8.0 with total enzyme concentrations of 54 µM and 40-80 µM, respectively. The components were mixed in an eppendorf tube with a Microcon ultracel YM-30 membrane. A 5 µL sample was taken from each tube and added to a TLC PEI Cellulose F chromatogram. The tubes were centrifuged at 10,000 rpm (Ole Dich centrifuge) for 5 minutes and a 5 µL sample of the flow-through was added to the chromatogram. Chromatograms were treated and samples measured as described in section 9.2.5.

9.2.5 Liquid Scintillation Counting

Chromatograms were placed in a jar containing a shallow layer of 0.9 M CH_3COOH and 0.3 M LiCl, and the solvent was allowed to travel to the top of the chromatograms. The chromatograms were dried and placed under 254 nm UV light, and the samples were framed with a soft pencil. The framed pencil markings were cut out, placed in small plastic containers, added 2 mL 2 M ammonia and left to dissolve for one hour. 5 mL Optiphase HiSafe2 Liquid Scintillation Cocktail was added, the solutions were gently mixed and left to stand for an additional hour. The samples were measured using a 2200CA TRI-CARB® Liquid scintillation analyzer by United Technologies Packard®.

9.2.6 Differential Scanning Calorimetry

Differential scanning calorimetry experiments were performed using a Microcal VP-DSC with scan rates of 0.5 °C/min. Experiments were carried out with WT *Mt* DCD-DUT at pH 8.0 and pH 6.8 with apo enzyme, as well as at pH 6.8 in the presence of 500 µM dTTP. Experiments with A115V *Mt* DCD-DUT were carried out at pH 8.0 in the presence (500 µM) and absence of dTTP. Experiments with apo A115G *Mt* DCD-DUT were carried out at pH 8.0 and pH 6.8 as well as at pH 8.0 in the presence of 500 µM dTTP. Background measurements were carried out using the same buffer that was used in the reference cell during the subsequent measurements. Experimental details are given in Appendix D.3, Table D.3.

9.2.7 Crystallization

The initial crystallization conditions were based on crystallization of WT *Mt* DCD-DUT [85]. The crystallization conditions for apo WT *Mt* DCD-DUT were 20 % PEG 8000, 50 mM $MgCl_2$ and 100 mM HEPES pH 7.5, while the crystallization conditions for the enzyme in complex with dTTP were 45 % PEG 400, 200 mM $MgCl_2$ and 100 mM HEPES pH 7.5. The hanging-drop vapor-diffusion method was applied in 24 well plates during all crystallization experiments.

Substrate analogues used during crystallization were dUMPNPP or dUMPCPP, which are analogues to dUTP, but with the oxygen bridging the α-β phosphorous replaced by nitrogen or carbon, respectively. Crystallization trials were carried out with A115V *Mt* DCD-DUT in the apo form or in complex with dTTP, dUMPCPP or dUMPNPP. Crystallization trials of WT *Mt* DCD-DUT were carried out with apo enzyme or enzyme in complex with dUMPNPP or dUMPCPP. Crystallization trials of A115G *Mt* DCD-DUT were carried out with apo enzyme or enzyme in complex with dUMPCPP. A complete set of all tested conditions is not provided here. Instead, a variety of crystallization conditions are available for picture entries in the enclosed database (Appendix A, DCDDUT.xml, see Chapter 2 for a program description).

9.2.8 Data Collection and Processing

Data were collected to 2.8 Å resolution on A115V *Mt* DCD-DUT in complex with dTTP. Diffraction data were collected at ID14-4 on ESRF. Integration and scaling of the data were performed with XDS/XSCALE [108,109]. Data collection statistics are presented in Table 9.1.

Table 9.1: X-ray data-collection and refinement statistics.
Values in parentheses are for the outermost resolution shell.

Beamline	ESRF, ID14-4
Detector	ADSC Quantum Q315r
Wavelength (Å)	0.9765
Temperature (K)	100
Space group	$P2_1$
Sample-to-detector distance (mm)	377
No. of images	360
Oscillation angle (°)	0.5
Unit-cell parameters (Å, °)	a=56.07 (7), b=179.4 (2), c=96.22 (5), α=γ=90.0, β=96.4
Resolution range (Å)	99.724-2.900 (3.00-2.90)
No. of observed reflections	161,875 (15,832)
No. of unique reflections	43,288 (4,216)
Mosaicity (°)	0.3-0.7
Multiplicity	3.74 (3.76)
Completeness (%)	98.7 (99.1)
R_{merge}	0.177 (0.976)
I/σ(I)	7.01 (1.44)

9.2.9 Structure Determination and Refinement

Molecular replacement using MOLREP [70] was applied to solve the structure. Different search models were tested: monomer, dimer or trimer of WT *Mt* DCD-DUT (pdb 2QXX). The structure was solved with the peptide chains of two trimers of WT *Mt* DCD-DUT with dTTP bound as search model. 12 molecules were found in the asymmetric unit. Refinement was performed with REFMAC5 [71]: first a rigid body refinement with rigid domains in all chains (residues 1-189), then a restrained refinement with tight NCS restraints on residues 1-189 in all chains. Structural refinements with REFMAC5 and winCoot [61,62] were performed. 12 dTTP molecules and 12 Mg^{2+} were added to the structure. Two additional Mg^{2+} were added on interfaces between the molecules. Residue 189 was removed in five chains (B, D, E, J, K). NCS restraints were removed on resides 110-118 and kept tight on the remainder of the peptide chain and on dTTP during refinement. TLS refinement was used and the peptide chains were used as TLS groups [110]. Refinement using phenix.refine [111] with restraints on Ramachandran angles or with simulated annealing had no effect on R/R_{free}. Refinement statistics are given in Table 9.2. The Ramachandran plot [72] shows 8 residues in the outlier region (0.4 %). All these residues are placed in the flexible region at the nucleotide binding site and are probably outliers because of the strain in the structure from binding dTTP.

Table 9.2: Refinement statistics.

R factor	0.2530
R_{free}	0.2899
No. of subunits in asymmetric unit	12
No. of atoms in model	17912
No. of water molecules	12
Average *B* factors (Å^2)	
Main chain	31.23
Side chain	29.85
dTTP	25.01
Ramachandran plot [72]	
Favored (%)	96.8
Allowed (%)	2.8
Outliers (%)	0.4

9.3 Results

Modeling curves to data points was performed in GraphPad Prism [112], using the equations given in Chapter 7. Throughout this section, data points for rates (s^{-1}) were calculated using equation (7.8) and data points for the degree of dTTP binding were calculated using equation (7.9).

9.3.1 Turnover numbers

Turnover numbers for WT, A115V and A115G were determined for 1 mM dCTP and 1 mM dUTP at pH 8.0 and pH 6.8. Experiments were carried out as described in section 9.2.1. No detectable activity was measured for A115V and A115G at pH 6.8 with dCTP. Turnover numbers are given in Table 9.3 and the plots showing the turnover number vs. time are shown in Appendix E.1, Figure E.4. Both variant enzymes had reduced activities compared with the WT enzyme.

Table 9.3: Turnover numbers (s^{-1}) for WT, A115V and A115G *Mt* DCD-DUT with 1 mM dCTP or dUTP at pH 6.8 or 8.0. Plots of turnover number vs. time are shown in Appendix E.1, Figure E.4.

	dCTP pH 8.0	dUTP pH 8.0	dCTP pH 6.8	dUTP pH 6.8
WT	0.239 ± 0.007	0.29 ± 0.01	0.044 ± 0.002	0.195 ± 0.007
A115V	0.068 ± 0.007	0.016 ± 0.001	-	0.0215 ± 0.0009
A115G	0.014 ± 0.001	0.0120 ± 0.0006	-	0.0015 ± 0.0002

9.3.2 Substrate saturation

Substrate saturation experiments were carried out with dCTP and dUTP at pH 8.0 as well as with dUTP at pH 6.8, as described in section 9.2.2. Neither of the variant enzymes had any detectable activity with dCTP at pH 6.8. The determined kinetic constants are given in Table 9.4 and data points with the modeled curves are shown in Figure 9.1 for the WT enzyme, Figure 9.2 for the A115V variant and Figure 9.3 for the A115G variant.

Table 9.4: Kinetic constants calculated for substrate saturation with dCTP at pH 8.0 and dUTP at pH 8.0 as well as pH 6.8 for WT, A115V and A115G *Mt* DCD-DUT.

	Michaelis-Menten equation (7.1)		Hill equation (7.4)		
	k_{cat}	K_m	k_{cat}	$S_{0.5}$	h
WT					
dCTP pH 8.0	0.35 ± 0.01 s^{-1}	40 ± 6 µM			
dUTP pH 8.0	0.29 ± 0.01 s^{-1}	57 ± 10 µM			
dUTP pH 6.8			0.188 ± 0.007 s^{-1}	118 ± 12 µM	1.7 ± 0.2
A115V					
dCTP pH 8.0	0.22 ± 0.02 s^{-1}	1.3 ± 0.2 mM			
dUTP pH 8.0	0.015 ± 0.001 s^{-1}	0.6 ± 0.2 µM			
dUTP pH 6.8	0.0204 ± 0.0008 s^{-1}	15 ± 3 µM			
A115G					
dCTP pH 8.0			0.020 ± 0.001 s^{-1}	171 ± 16 µM	2.5 ± 0.5
dUTP pH 8.0	0.0157 ± 0.0007 s^{-1}	24 ± 5 µM			
dUTP pH 6.8	0.010 ± 0.003 s^{-1}	1.4 ± 0.7 mM			

Cooperativity was observed for the WT enzyme at pH 6.8 (Figure 9.1B) with kinetic constants similar to what has been observed previously [85]. In contrast, no significant cooperativity was observed at pH 8.0 with dCTP or dUTP as substrates (Figure 9.1C).

In comparison with the WT enzyme, A115V showed some major differences in dCTP and dUTP saturation kinetics. Contrary to WT, A115V displayed hyperbolic saturation kinetics at pH 6.8 with dUTP, as shown in Figure 9.2B. Hyperbolic saturation was also observed at pH 8.0 with both dCTP and dUTP (Figure 9.2C+D), albeit with large differences in the kinetic constants compared with the WT enzyme. k_{cat} was about 1.5 and 20 fold lower for dCTP and dUTP, respectively, for A115V compared with the WT enzyme. K_M was about 25 fold higher and 100 fold lower for dCTP and dUTP, respectively, for A115V compared with the WT enzyme.

The A115G enzyme also showed major differences compared with the WT enzyme. Hyperbolic dUTP saturation was observed at both pH 6.8 and 8.0 (Figure 9.3), whereas cooperativity was observed with dCTP at pH 8.0 (Figure 9.3C). k_{cat} and K_m were about 20 and 2.5 fold lower, respectively, for dUTP at pH 8.0 for the A115G enzyme compared with the WT enzyme.

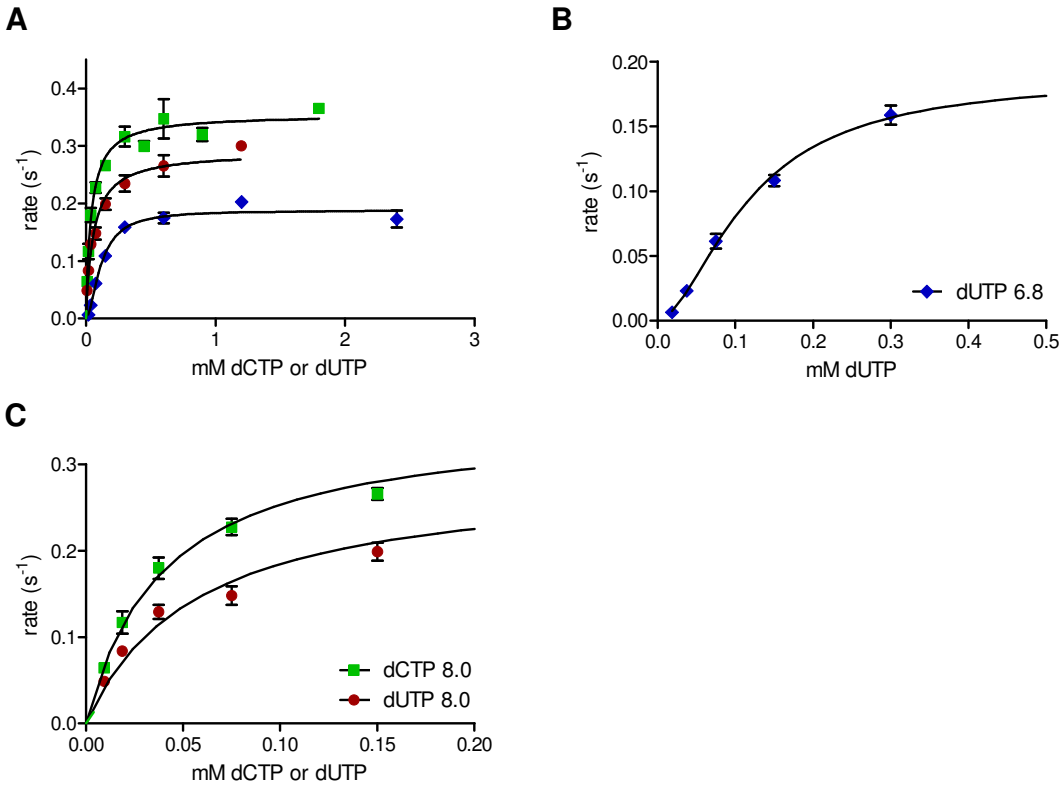


Figure 9.1: Substrate saturation curves fitted to data points of WT *Mt* DCD-DUT. A: curves and data points. B: zoomed plot of curve and data points at pH 6.8. C: zoomed plot of curves and data points at pH 8.0. Calculated kinetic constants and references to the applied equations can be found in Table 9.4. Each point was determined at least twice at different enzyme concentrations.

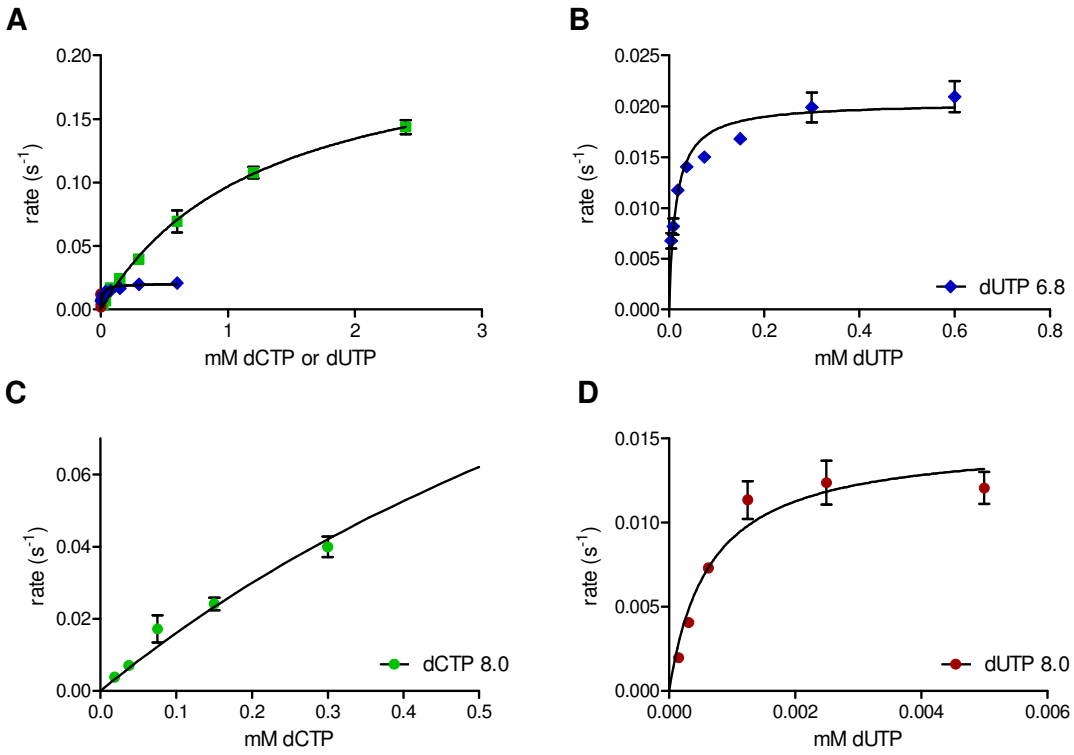


Figure 9.2: Substrate saturation curves fitted to data points of A115V *Mt* DCD-DUT. A: curves and data points. B: zoomed plot of curve and data points at pH 6.8. C+D: zoomed plot of curves and data points at pH 8.0. Calculated kinetic constants and references to the applied equations can be found in Table 9.4. Each point was determined at least twice at different enzyme concentrations.

The full range of measured rates at different nucleotide concentrations is available in Appendix E.3, Figure E.5. It should be noted that some of the data points at high nucleotide concentrations (typically above 2 mM) were omitted during modeling of the curves to the data points, as seen in Figure 9.1, Figure 9.3 and in particular for the A115V variant shown in Figure 9.2. The observed behavior of a slow increase (seen for A115V dUTP pH 6.8 and pH 8.0) or a decrease (seen for A115V dCTP pH 8.0) could not be modeled satisfactorily using the equations given in section 7.2 and data were therefore cut off to yield reasonable fits.

9.3.3 dTTP Inhibition

Inhibition by dTTP of WT, A115V and A115G *Mt* DCD-DUT was examined at pH 6.8 and pH 8.0. Experiments were carried out as described in section 9.2.3 and the calculated kinetic constants, as well as references to the applied equations, are presented in Table 9.5. Modeling of curves to the data points are shown in Figure 9.4.

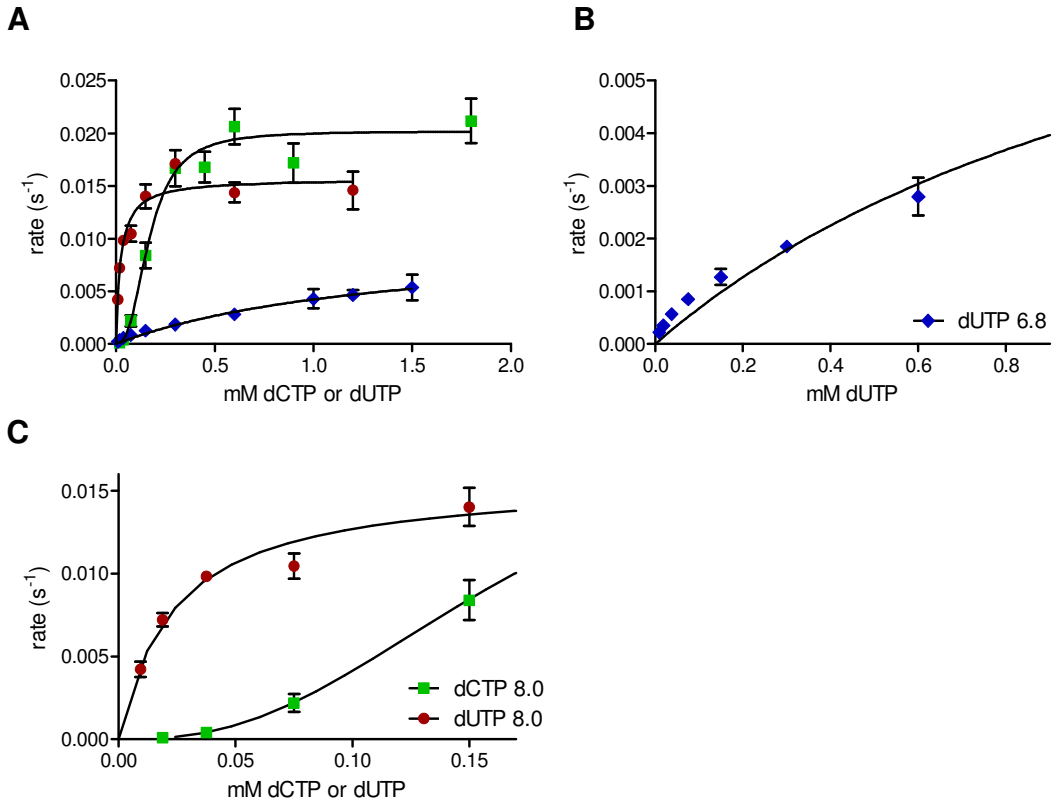


Figure 9.3: Substrate saturation curves fitted to data points of A115G *Mt* DCD-DUT. A: curves and data points. B: zoomed plot of curves and data points at pH 6.8. C: zoomed plot of curves and data points at pH 8.0. Calculated kinetic constants and references to the applied equations can be found in Table 9.4. Each point was determined at least twice at different enzyme concentrations.

Table 9.5: Kinetic constants calculated for dTTP inhibition of WT *Mt* DCD-DUT at pH 6.8, A115V *Mt* DCD-DUT at pH 6.8 and pH 8.0, and A115G *Mt* DCD-DUT at pH 8.0.

	Inhibition (hyperbolic) equation (7.6)	Inhibition (sigmoidal, cooperative) equation (7.7)	
	K_i	K_i	h
WT			
300 µM dCTP pH 6.8		125 ± 9 µM	1.6 ± 0.1
300 µM dUTP pH 6.8		183 ± 14 µM	1.4 ± 0.1
A115V			
2.4 mM dCTP pH 8.0	323 ± 44 µM		
300 µM dCTP pH 8.0	72 ± 9 µM		
300 µM dUTP pH 6.8	1.4 ± 0.3 mM		
A115G			
300 µM dCTP pH 8.0		4.9 ± 0.9 µM	2.2 ± 0.8
800 µM dCTP pH 8.0		11.2 ± 0.6 µM	2.2 ± 0.2

Unexpectedly, the WT enzyme was insensitive to the presence of dTTP at pH 8.0 even at dCTP concentrations around K_M (Figure 9.4A). The inhibition by dTTP at pH 6.8 [85] was verified with dCTP and dUTP as substrates (Figure 9.4D). The apparent pH effect of dTTP inhibition was further examined by measuring activity with dCTP or dUTP at pH 6.8, 7.4 and 8.0 in the presence or absence of dTTP for the WT enzyme. A clear tendency of attenuation of dTTP inhibition at pH above 6.8 is observed, results are shown in Figure 9.5.

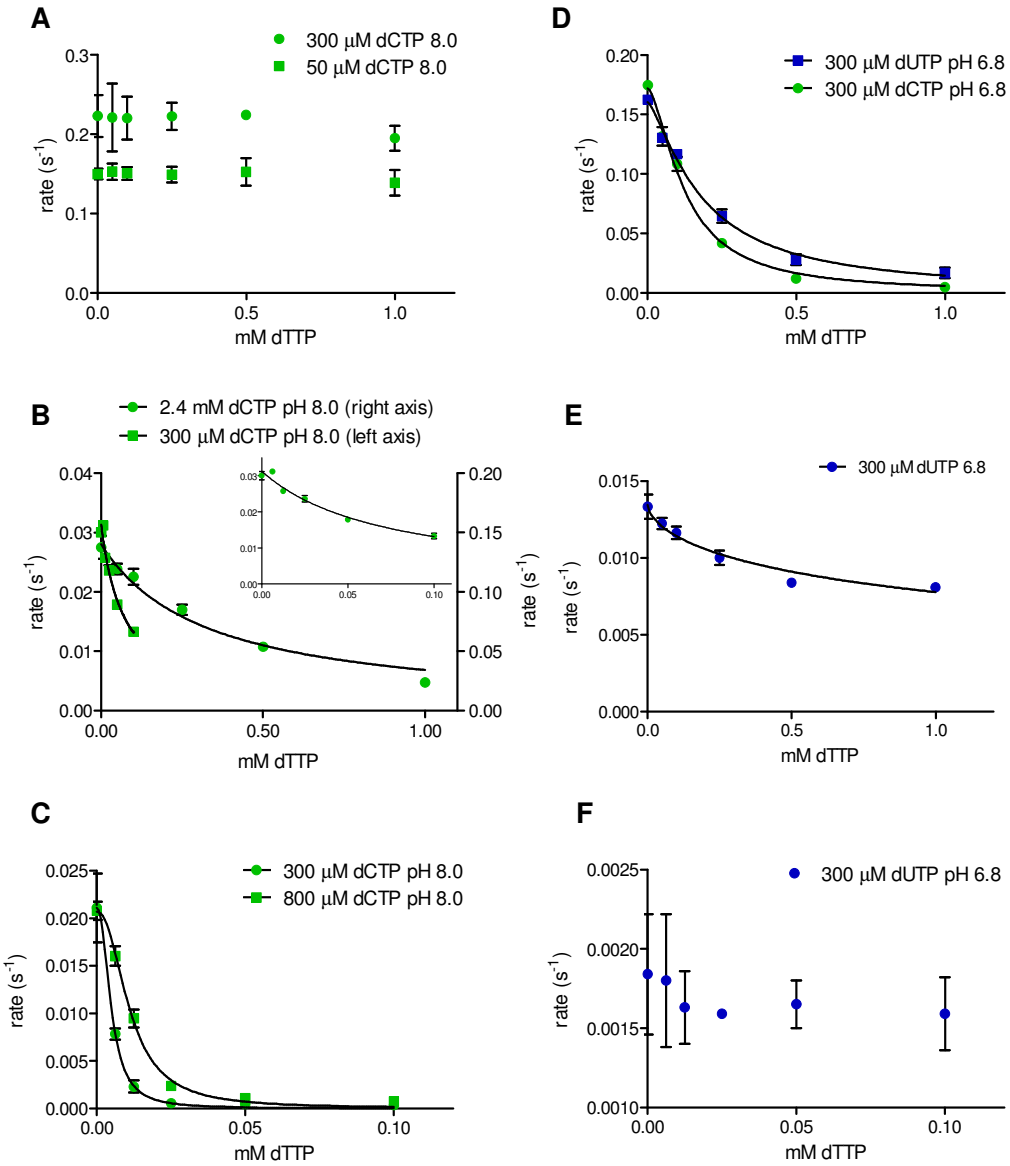


Figure 9.4: Inhibition by dTTP at constant substrate concentration at pH 8.0 of A: WT, B: A115V (zoomed insert: 300 µM dCTP pH 8.0) and C: A115G. dTTP inhibition at pH 6.8 of D: WT, E: A115V and F: A115G. Each point was determined twice at different enzyme concentrations. Calculated kinetic constants and references to the applied equations can be found in Table 9.5.

The A115V and A115G enzymes were constructed to evaluate the importance of steric hindrance by the side chain of Ala115 on a water molecule suggested to be crucial for activity [85] as described in section 7.1. Neither of the variants displayed the expected properties (discussed in section 9.4). In contrast to WT, the A115V enzyme was inhibited by dTTP at both pH 8.0 and 6.8 as shown in Figure 9.4B+E. A115V showed increased sensitivity to dTTP at pH 8.0 compared with pH 6.8 (K_i was about 20 times lower at pH 8.0 compared with pH 6.8). The A115G enzyme showed dTTP inhibition with pH effects contrary to the WT enzyme. A115G showed a high sensitivity to dTTP at pH 8.0 with a complete loss of activity at 0.1 mM dTTP (Figure 9.4C). K_i was 25 fold lower for A115G at pH 8.0 compared with WT at pH 6.8. In contrast, no significant inhibition was observed at pH 6.8 (Figure 9.4F).

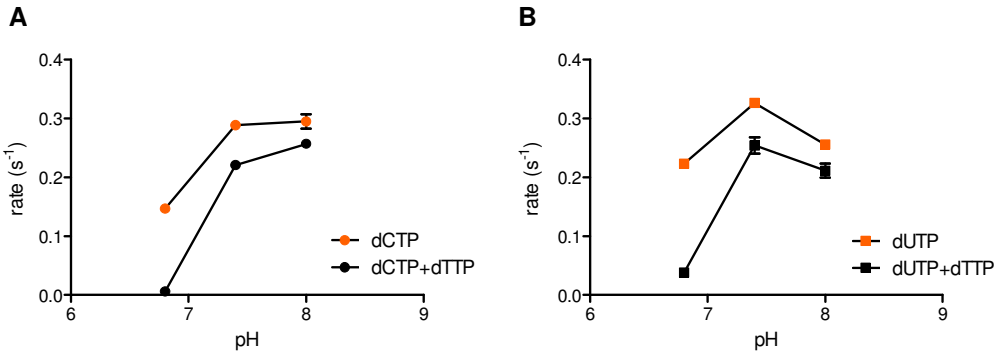


Figure 9.5: Reaction rates for WT *Mt* DCD-DUT at pH 6.8, 7.4 and 8.0 with A: 300 µM dCTP and B: 300 µM dUTP as substrates. Black and orange data points are for assays in the presence (500 µM) and absence of dTTP, respectively. Straight lines were drawn between the data points as guides to the eye.

9.3.4 dTTP Binding

dTTP equilibrium binding experiments were carried out with WT, A115V and A115G *Mt* DCD-DUT as described in section 9.2.4. The WT dTTP binding experiments were designed to be carried out at pH 6.8. However, the shift in pH from mixing a substantial amount of protein at pH 8.0 with the buffer at pH 6.8 caused the final pH to be 7.1. At pH 7.1, WT does not bind dTTP, as shown in Figure 9.6A. pH of the WT binding experiments was changed compared with the variant dTTP binding experiments (carried out at pH 8.0) because the WT enzyme had proven insensitive to the presence of dTTP at pH 8.0 (Figure 9.4A). dTTP binding to the enzyme was therefore not expected at pH 8.0.

A115V binds dTTP at pH 8.0 as shown in Figure 9.6B. dTTP binding was modeled using equation (7.5) resulting in a dissociation constant K_d = 7 ± 2 µM. dTTP binding to A115V at pH 8.0 is stronger than dTTP binding to the monofunctional dCTP deaminase from *E. coli* at pH 6.8. K_d was around 7 for A115V:dTTP compared with K_d around 35 for *Ec* DCD:dTTP [91]. A115G shows a very unusual dependence of the degree of dTTP binding on the concentration of dTTP. Initially, the degree of binding increases with the concentration of dTTP, but the enzyme never reaches

full occupation (*N*=1) and it appears that the dTTP binding site occupation decreases at dTTP concentrations above approximately 100 mM.

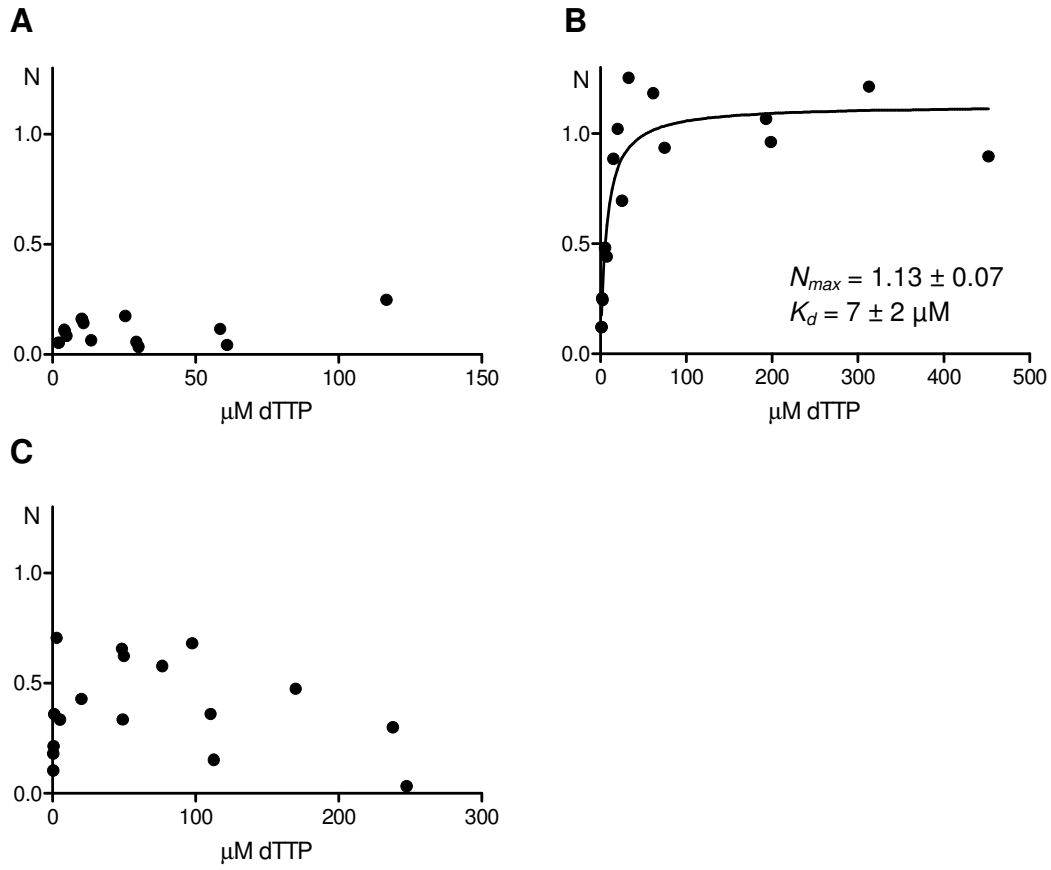


Figure 9.6: Degree of dTTP binding (*N*) to A: WT at pH 7.1, B: A115V at pH 8.0, curve was modeled using equation (7.5) and C: A115G *Mt* DCD-DUT at pH 8.0.

9.3.5 Differential Scanning Calorimetry

Weak signals for irreversible transitions were observed for DSC experiments, and transition temperatures should therefore be considered as approximates. Transition temperatures are listed in Table 9.6 and DSC measurements are given in Appendix E.4.

A115V *Mt* DCD-DUT is destabilized compared with the WT enzyme. The transition temperature for the A115V variant enzyme is 20 °C lower than for the WT enzyme (67 °C vs. 87 °C) in the presence of dTTP at pH 8.0.

Table 9.6: Transition temperatures for DSC experiments with WT, A115V or A115G *Mt* DCD-DUT.

	pH	c_{enzyme}	Apo	500 µM dTTP
WT	8.0	20 µM	77 °C	
	6.8	20 µM	84 °C	87 °C
A115V	8.0	24 µM	inconclusive	67 °C
A115G	8.0	12 µM	81 °C	83 °C
	8.0	24 µM	83 °C	85 °C
	6.8	24 µM	inconclusive	

9.3.6 Crystallization and Diffraction Tests

9.3.6.1 WT enzyme

Crystallization of WT *Mt* DCD-DUT was carried out mainly in the presence of substrate analogues in an attempt to get a structure of WT *Mt* DCD-DUT in its active conformation. Selected crystallization results are given in Figure 9.7 and a range of crystallization pictures and conditions are available in the enclosed database (Appendix A, DCDDUT.xml, see Chapter 2 for a program description). dUMPNPP spontaneously hydrolyzes at the same rate whether or not WT *Mt* DUT is present [102]. A crystal of WT *Mt* DCD-DUT that had grown in the presence of dUMPNPP over a period of 4 months proved to contain enzyme in complex with monophosphate or diphosphate [113]. The crystals in Figure 9.7C+D had grown after 7-8 months at 15 °C and were therefore probably formed without an intact nucleotide.

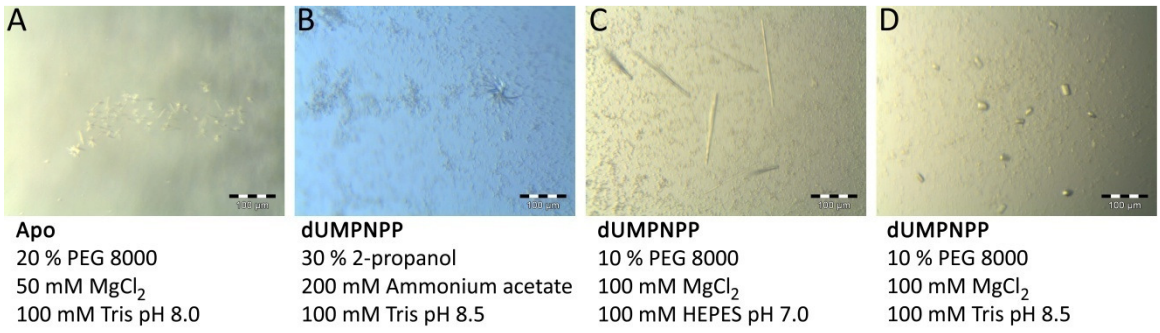


Figure 9.7: Crystallization of 2 mg/mL WT *Mt* DCD-DUT with specified reservoir conditions. For dUMPNPP: the enzyme used for crystallization setup was in solution with 2 mM dUMPNPP and 14 mM $MgCl_2$.

9.3.6.2 A115V variant enzyme

Crystallization of A115V *Mt* DCD-DUT was carried out with enzyme in its apo form or in complex with dTTP, dUMPNPP or dUMPCPP. Selected crystallization results are given in Figure 9.8 and additional pictures and conditions can be found in the enclosed database (Appendix A, DCDDUT.xml). Data were collected to 2.8 Å resolution (section 9.2.8) on a crystal from Figure 9.8C: A115V in complex with dTTP.

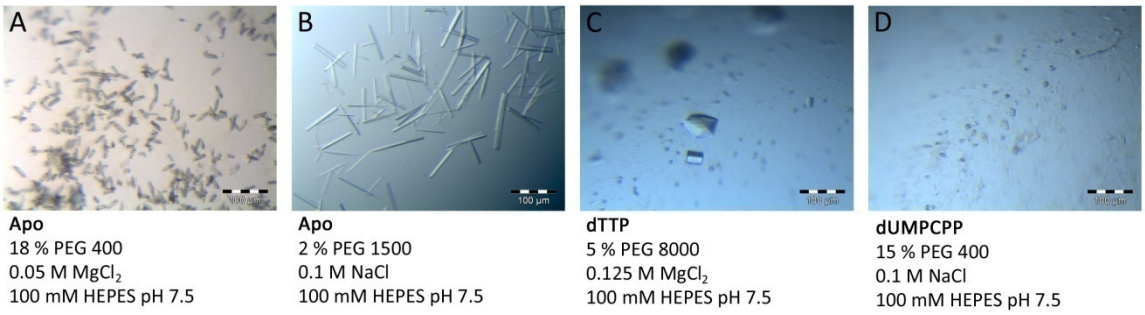


Figure 9.8: Crystallization of 2-4 mg/mL A115V *Mt* DCD-DUT with specified reservoir conditions. For dTTP: the enzyme used for crystallization setup was in solution with 5 mM dTTP and 20 mM $MgCl_2$. For dUMPCPP: the enzyme was in solution with 2 mM dUMPCPP and 10 mM $MgCl_2$.

9.3.6.3 A115G variant enzyme

Crystallization of A115G *Mt* DCD-DUT was carried out mainly with enzyme in its apo form because the enzyme showed unusual dTTP binding (section 9.3.4). Selected crystallization results are given in Figure 9.9 and additional pictures and conditions can be found in the enclosed database (Appendix A, DCDDUT.xml). Diffraction properties were tested on the largest of the apo crystals (Figure 9.9D), but no diffraction was observed at ESRF. The morphology of these crystals was similar to apo WT enzyme [113].

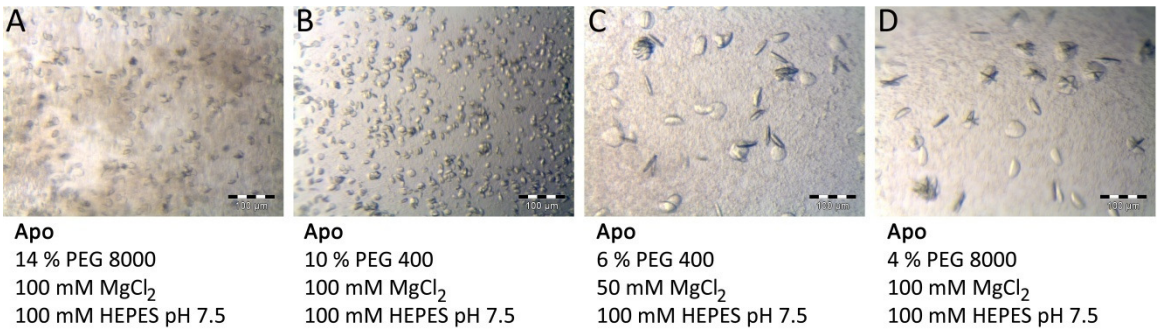


Figure 9.9: Crystallization of 2 mg/mL A115G *Mt* DCD-DUT with specified reservoir conditions.

9.3.7 Crystal structure of A115V:dTTP

The structure of A115V with 12 molecules in the asymmetric unit was refined using tight NCS restraints except for residues 110-118 that constitute a flexible region near the nucleotide binding pocket. dTTP is bound in all 12 active sites with clear electron density, see Figure 9.10. Hydrogen bonds from the peptide chain to dTTP in A115V *Mt* DCD-DUT (Figure 9.11) are generally the same as observed in the WT structure [85]. However, a total of 7 fewer protein-protein hydrogen bonds are observed in the 110-118 region of the A115V structure compared with the WT structure.

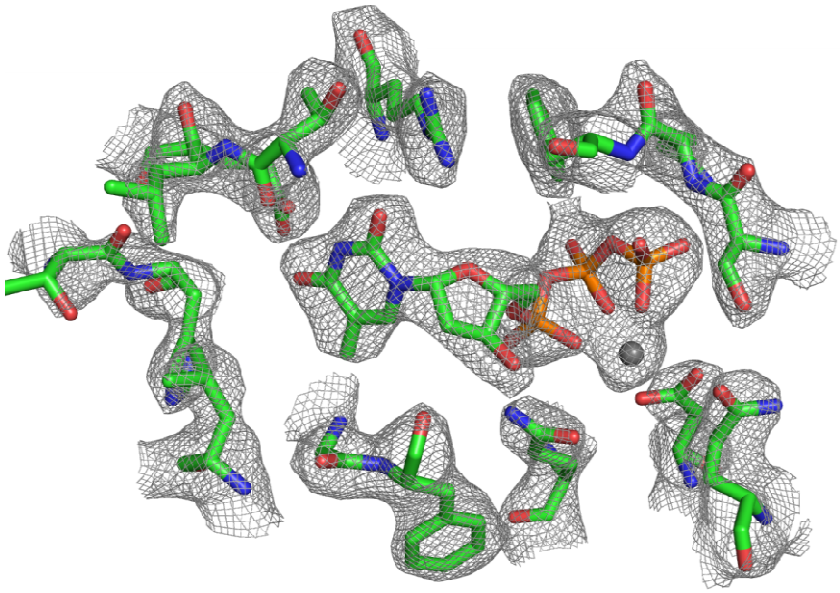


Figure 9.10: A115V *Mt* DCD-DUT with dTTP bound. dTTP from chain J and selected residues from the HJL trimer with $2F_{obs} - F_{calc}$ σ = 1.0 electron density map are shown.

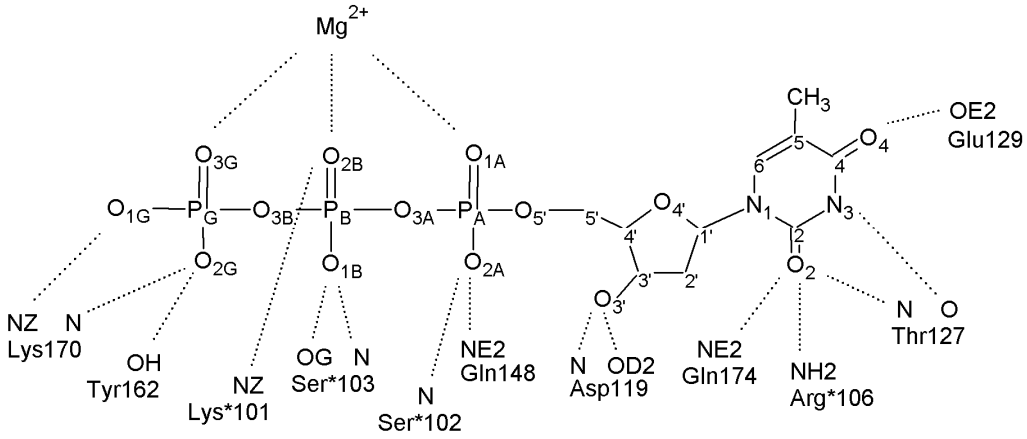


Figure 9.11: Schematic view of hydrogen bonds to dTTP in the structure of A115V *Mt* DCD-DUT. * denote residues from a neighboring monomer.

The structure of the flexible loop of residues 110-118 in A115V:dTTP more closely resembles apo WT and the bifunctional enzyme from *Mj* with a substrate analogue bound (Figure 9.12B) than WT:dTTP (Figure 9.12A). The flexible loop in the A115V:dTTP structure is thus very similar to the active conformation, whereas the WT:dTTP structure is in the inactive conformation. However, some differences are observed between A115V:dTTP and *Mj* DCD-DUT:dUMPNPP (Figure 9.13). In order to accommodate the 5-methyl group of dTTP, Gly116 in A115V:dTTP has moved 2.4 Å, which furthermore results in a 1 Å wider loop measured from His112 CA to Val115 CA.

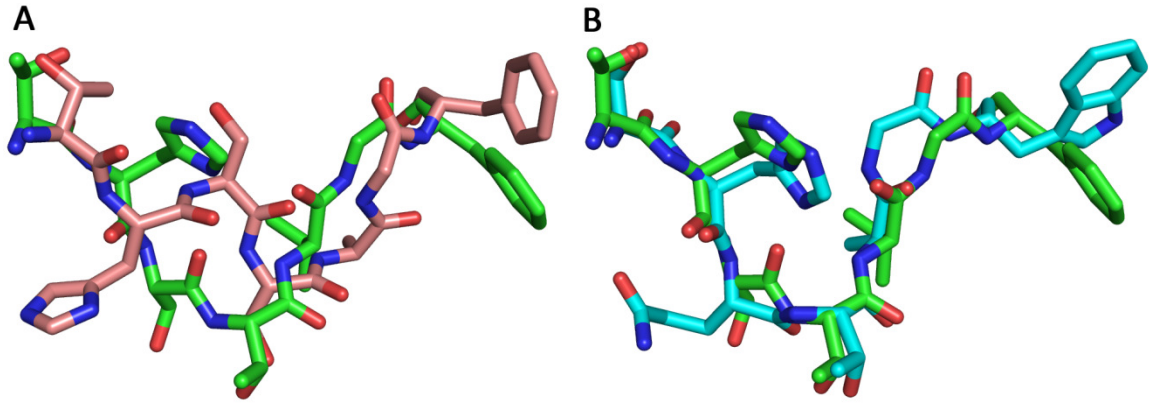


Figure 9.12: Superposed structures of A115V *Mt* DCD-DUT (green) and A: WT *Mt* DCD-DUT:dTTP (red, pdb 2QXX), B: *Mj* DCD-DUT:dUMPNPP (blue, pdb 2HXD[5]).

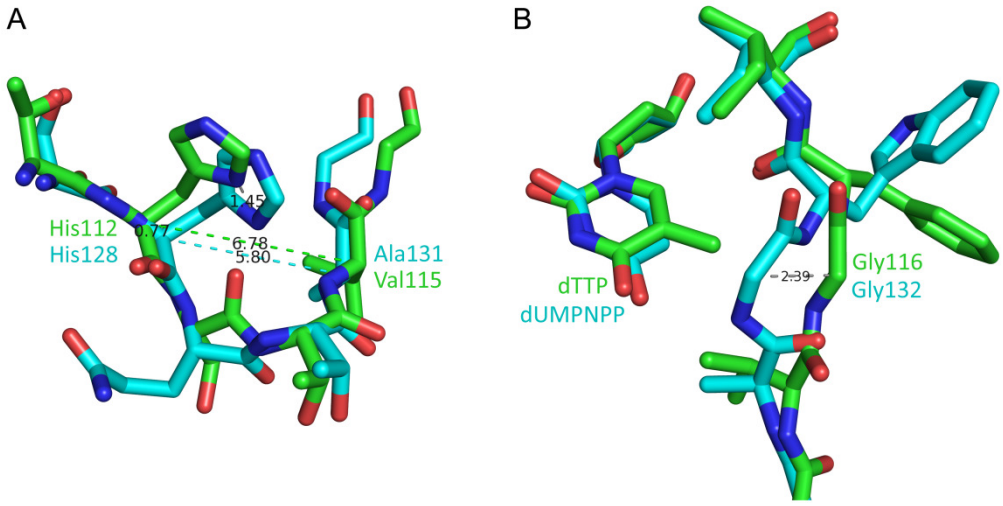


Figure 9.13: Differences between the structures of A115V:dTTP (green) and *Mj* DCD-DUT:dUMPNPP (blue, pdb 2HXD[5]). A: the 1 Å wider loop from His CA to Ala/Val CA is observed along with 1.45 Å movement of Nπ. B: to accommodate the 5-methyl group of dTTP, the backbone of Gly116 in A115V:dTTP has moved 2 Å.

[5] His128 was flipped because it faced the wrong way based on hydrogen bonds

The A115V enzyme is unable to exist in the inactive conformation with dTTP bound because of the larger size of the valine side chain compared with alanine. In the inactive conformation, the valine side chain would clash with the 5-methyl group of dTTP. An approximately 2.5 Å movement of V115 is observed in order for A115V to bind dTTP, see Figure 9.14 that shows a structural overlay of WT:dTTP and A115V:dTTP. This 2.5 Å backbone movement is also seen at Thr114, and a rearrangement of Thr111 and His112 in a neighboring molecule is necessary to avoid clashing of the residues (Figure 9.14, for stereo view see Appendix E, Figure E.9). His112 in the A115V structure is placed in the active conformation and hydrogen bonds of this residue are shown in Figure 9.15 for the structures of A115V:dTTP and WT:dTTP and for the corresponding His128 in *Mj* DCD-DUT:dUMPNPP.

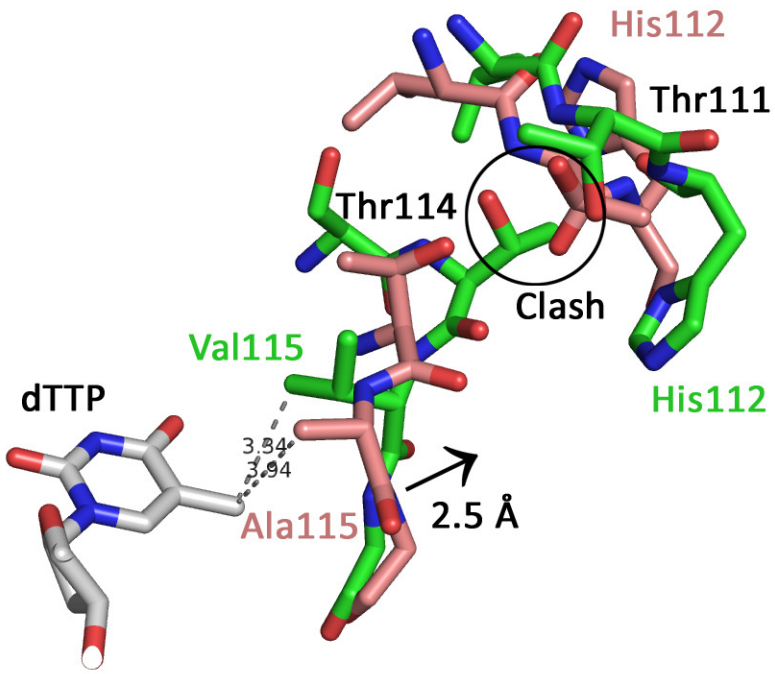


Figure 9.14: Overlay of A115V:dTTP (green) and WT:dTTP (red, pdb 2QXX). A shift in the backbone causes the valine to take up less space in the nucleotide binding pocket and allows room for dTTP binding. A rearrangement of the backbone at Thr111 and His112 prevents the side chain of Thr114 from clashing with Thr111 of a neighboring monomer.

The water molecule suggested to be important for dephosphorylation [85] is present in chain H in the structure of A115V:dTTP even with the larger side chain of valine compared with alanine, see Figure 9.16.

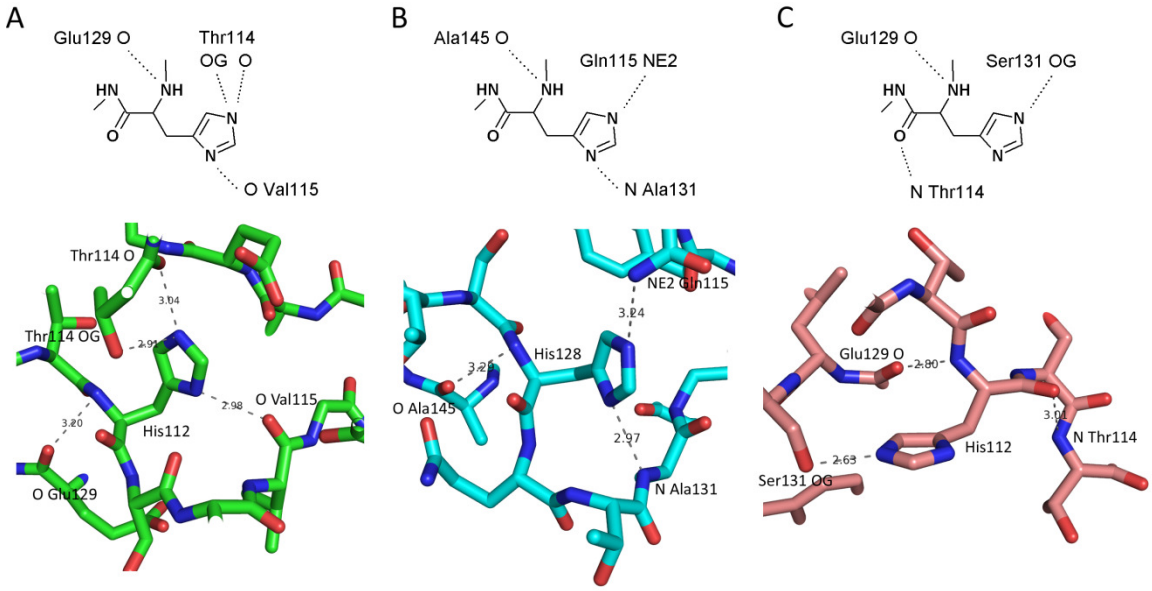


Figure 9.15: Schematic and structural views of the hydrogen bonding network of a conserved histidine in A: His112 in A115V *Mt* DCD-DUT:dTTP, B: His128 in *Mj* DCD-DUT:dUMPNPP (pdb 2HXD[5]) and C: His112 in WT *Mt* DCD-DUT:dTTP (pdb 2QXX).

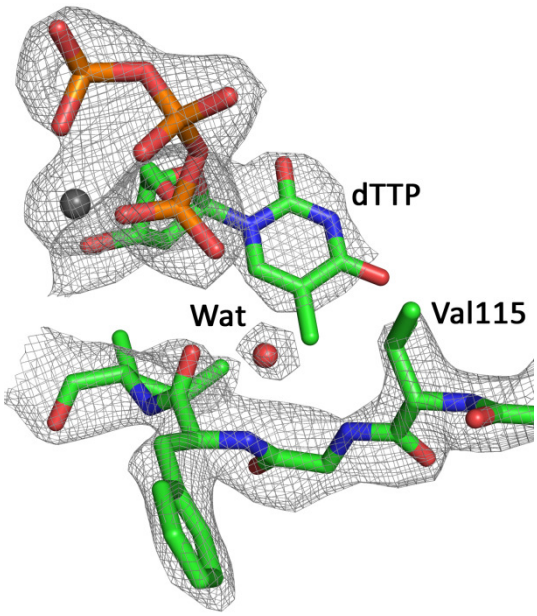


Figure 9.16: A115V:dTTP bound (chain H) shown with $2F_{obs} - F_{calc}$ σ = 1.0 electron density map. Electron density clearly shows the presence of the water molecule suggested to be important for dephosphorylation [85].

A rotation of the nucleotide pyrimidine moiety is observed between the substrate analogue in *Mj* DCD-DUT (pdb 2HXD) and dTTP in A115V and WT (this work and pdb 2QXX). A 23° angle is seen between the pyrimidine planes of superposed nucleotide structures of dTTP in A115V and

dUMPNPP in *Mj* DCD-DUT, see Figure 9.17. The dTTP methyl group in the structure of A115V would apparently clash with Val115 and Ser102 if the pyrimidine moiety attained the same conformation as dUMPNPP in *Mj* DCD-DUT. The reduced distances from these residues to dTTP are shown in Figure 9.18. This rotation causes changes in the hydrogen bond pattern; a hydrogen bond present between O4 of the nucleotide pyrimidine moiety and Gly116 in the structure of *Mj* DCD-DUT:dUMPNPP, is not observed in the structures of A115V and WT with dTTP bound. Instead, the structures with dTTP bound have an additional hydrogen bond from O2 of the dTTP pyrimidine moiety to Nη1 of the conserved Arg106, see Figure 9.19.

Figure 9.17: Stereo view of nucleotide structures from A115V *Mt* DCD-DUT:dTTP (green) and *Mj* DCD-DUT:dUMPNPP (blue, pdb 2HXD) showing a 23° turn of the nucleotide pyrimidine ring in the two structures.

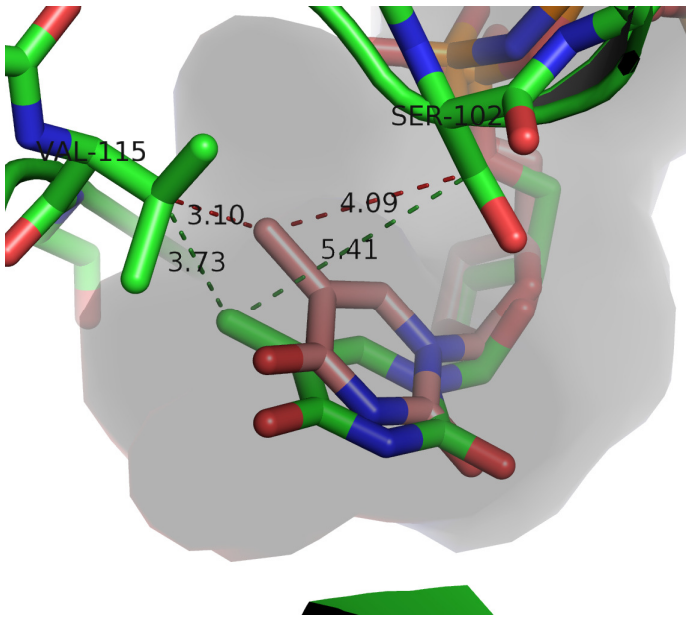


Figure 9.18: Structure of the A115V:dTTP complex shown in green with the surface of the nucleotide binding pocket shown in grey. dTTP shown in red was forced to obtain the same conformation as seen for dUMPNPP in *Mj* DCD-DUT (pdb 2HXD). The figure was made using PDB2PQR [114] and the APBS plugin [115] to PyMol [5].

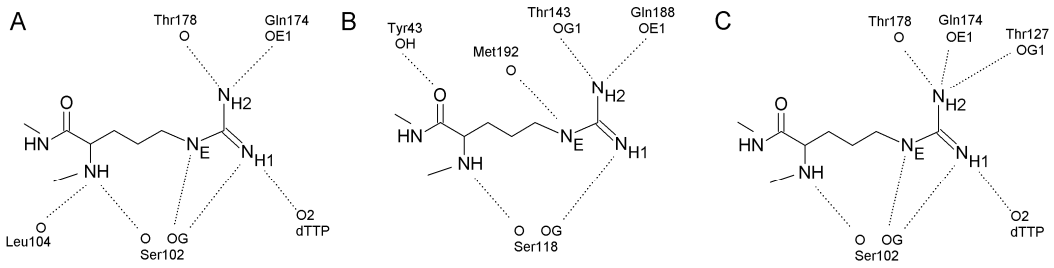


Figure 9.19: Schematic view of the hydrogen bonding network of A: Arg106 in A115V:dTTP, B: Arg122 in *Mj* DCD-DUT:dUMPNPP (pdb 2HXD) and C: Arg106 in WT:dTTP (pdb 2QXX).

A different conformation of the Arg97 side chain is seen in A115V:dTTP (Figure 9.20A) compared with WT:dTTP (Figure 9.20B). In two of the four trimers in the A115V structure, a Mg^{2+} ion has been modeled at the center of the trimer. The distance from Mg^{2+} to Glu99 Oε is 3-3.2 Å and water molecules are likely to occupy some of the space in the center of the trimer. However, the resolution is insufficient for modeling of these. Also, it is likely that Mg^{2+} is present in the other trimers, but that the resolution is insufficient for modeling these. The positive charge of Mg^{2+} stabilizes the side chains of Glu99. The conformation of the nearby His112 also does not affect the different conformation of Arg97, since apo WT shows the same conformation of His112 as A115V, but with Arg97 facing the center of the trimer similarly to WT:dTTP, see Figure 9.20C.

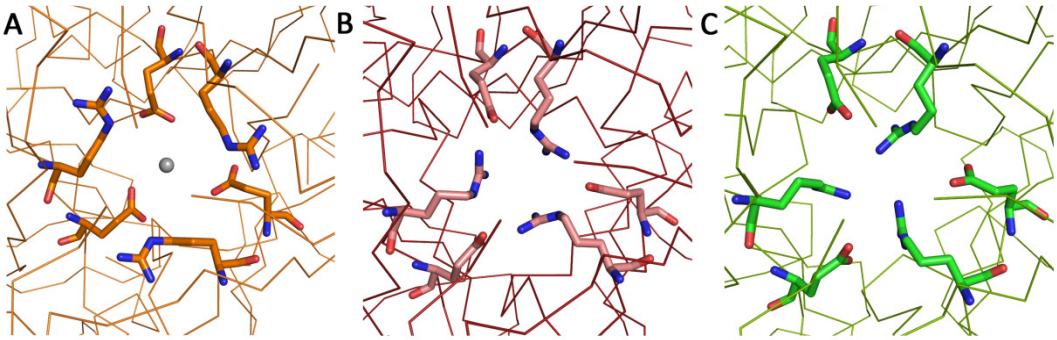


Figure 9.20. Center of trimers from A: A115V:dTTP (orange, trimer HJL, Mg^{2+} grey), B: WT:dTTP (red, pdb 2QXX), C: apo WT (green, pdb 2QLP) with Glu99 and Arg97 shown in ball and stick.

9.4 Discussion

As described in section 7.1, A115V and A115G variants of *Mt* DCD-DUT were created to investigate the importance of steric hindrance by the side chain of Ala115 on a water molecule suggested to be crucial for activity [85]. Both variants proved to have some unexpected properties.

9.4.1 Properties of WT and A115V enzymes

The flexible region consisting of residues 110-118 in the A115V variant with dTTP bound does not have the same conformation as the WT enzyme with dTTP bound. Instead, the conformation of the A115V variant is very similar to the active conformation found in *Mj* DCD-DUT:dUMPNPP. The water molecule suggested to be important for dephosphorylation [85] is present in chain H in the structure of A115V:dTTP, see Figure 9.16. The A115V variant does not dephosphorylate dTTP even though the water molecule is present in the structure and it can therefore not be the missing piece in explaining why dTTP is not dephosphorylated.

Steric hindrance forces a rotation of the nucleotide pyrimidine moiety between dUMPNPP in *Mj* DCD-DUT (pdb 2HXD) and dTTP in A115V and WT (pdb 2QXX), see Figure 9.17. This rotation causes changes in the hydrogen bond pattern of a conserved arginine as shown in Figure 9.19. Changes in the hydrogen bonds of this conserved arginine (*Mt* DCD-DUT Arg106, *Mj* DCD-DUT Arg122) also affect a conserved serine (*Mt* DCD-DUT Ser102, *Mj* DCD-DUT Ser118) because of their shared hydrogen bonds in all structures (Figure 9.19). The serine is crucial for activity and has been suggested to play an important role in dephosphorylation by balancing negative charge on the oxygen bridging the α,β-phosphorous [93,102]. Arginine is a hydrogen bond donor to serine and thus aid in balancing negative charge formed on the serine. The ability of arginine to function as a hydrogen bond donor to serine is impaired when dTTP is bound because Arg106 Nη1 is involved in an additional hydrogen bond to O2 of the dTTP pyrimidine moiety. Also, the additional hydrogen bond between arginine and dTTP restrains the flexibility of the arginine. The more restrained arginine may be less capable of stabilizing negative charge formed on the flexible serine side chain.

A115V:dTTP is unable to attain in the same conformation as WT:dTTP because of the introduced valine side chain. The larger size of the valine side chain forces a rearrangement of the flexible 110-118 residue range (Figure 9.14) to adopt a conformation very similar to the active conformation observed in *Mj* DCD-DUT:dUMPNPP (Figure 9.12B) at the expense of 7 hydrogen bonds in the flexible region. The overall structures are generally very similar, and the loss of 7 hydrogen bonds in the flexible region may explain the much reduced thermal stability of A115V:dTTP compared with WT:dTTP (transition temperatures of 67 vs. 87 °C, section 9.3.5). The inability of A115V *Mt* DCD-DUT to exist in the inactive conformation is supported by the hyperbolic saturation curve with dUTP at pH 6.8 (Figure 9.2B), which indicates that no structural rearrangement is necessary and the enzyme is active.

WT binds dTTP in an inactive conformation (pdb 2QXX), see Figure 9.12A. WT:dTTP is unable to exist in the active conformation observed in *Mj* DCD-DUT:dUMPNPP unless the backbone of Ala115 and Gly116 moves to accommodate the 5-methyl group of dTTP, see Figure 9.21. This movement is accompanied by a rearrangement of His112. This conformational change along with hydrogen bond patterns of His112 is shown in Figure 9.15. There is no structural hindrance for the WT enzyme to be able to bind dTTP in a similar conformation as seen for A115V:dTTP. However, the WT:dTTP structure in the inactive conformation has 7 additional hydrogen bonds

in the flexible region compared with A115V:dTTP in the active conformation. Therefore, WT favors binding dTTP in the inactive conformation.

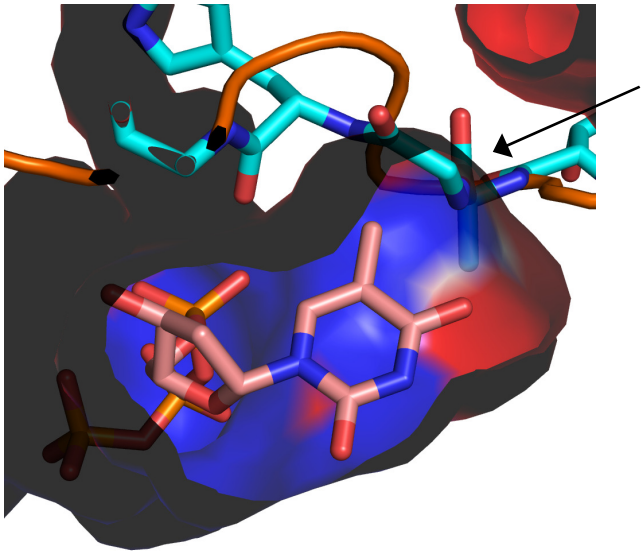


Figure 9.21: Electrostatic surface of protein backbone (orange) of WT *Mt* DCD-DUT with dTTP shown (pdb 2QXX). The backbone structure of *Mj* DCD-DUT:dUMPNPP (pdb 2HXD) is shown in blue. The arrow points to the clash between Ala131 and Gly132 in *Mj* (corresponding to Ala115 and Gly116 in *Mt*) and dTTP in *Mt* DCD-DUT. The figure was made using PDB2PQR [114] and the APBS plugin [115] to PyMol [5].

Unexpectedly, the WT enzyme was uninhibited by dTTP at pH 8.0 (Figure 9.4). The WT enzyme was even uninhibited by dTTP at dCTP concentrations around K_M (50 µM, Figure 9.4A) and the enzyme showed no binding of dTTP at pH 7.1 (Figure 9.6A). The pH dependence of dTTP binding and inhibition may be explained by looking into the protonation state of a conserved histidine near the nucleotide binding site and its effects on available structures with active and inactive conformations (*Mt* His112, Figure 9.15). Histidine has a pK_a value of approximately 6, albeit it is likely to be affected by the protein environment[6]. With the observed differences, the histidine may be partly protonated at pH 6.8, while it is primarily deprotonated at pH 8.0. Therefore, the difference in dTTP inhibition is probably related to the protonation state of His112. WT *Mt* DCD-DUT is likely to adopt the same active conformation as seen in *Mj* DCD-DUT, which is very similar to the conformation seen in A115V:dTTP. Protonation of His112 stabilizes the inactive conformation because of the strain caused by the additional proton in the 110-118 loop. In addition, the hydrogen bond between His112 Nπ and Ala115 N is broken because the protonated histidine can no longer act as a hydrogen bond acceptor (Figure 9.15B shows the

[6] Calculating the pK_a value of His112 in WT:dTTP (pdb 2QXX) using PROPKA [59,60,118] gives values of approximately 3. The decrease is attributed to desolvation, however, un-modeled blobs in the $F_{obs} - F_{calc}$ electron density map at the center of the trimer strongly suggest the presence of disordered solvent near the histidine. Therefore, the desolvation correction by PROPKA is likely to be exaggerated.

histidine hydrogen bond pattern and loop conformation for the active conformation of *Mj* DCD-DUT). A complete deprotonation of His112 in WT *Mt* DCD-DUT causes the enzyme to be present solely in the active conformation where it is unable to bind dTTP because of the placement of Ala115 and Gly116 (Figure 9.21) [85]. This may explain why no dTTP inhibition was observed at pH 8.0. The hyperbolic dCTP and dUTP saturation curves at pH 8.0 support that no structural rearrangement is needed for the enzyme to be active. Conversely, this structural rearrangement is needed at pH 6.8, where His112 is partly protonated and the WT enzyme exhibits cooperative substrate saturation [85].

Unlike the WT enzyme, A115V *Mt* DCD-DUT was able to bind dTTP at pH above 6.8. Also, dTTP inhibition was observed both at pH 6.8 and 8.0 for A115V. Although the structure of A115V:dTTP is very similar to *Mj* DCD-DUT:dUMPNPP, the strained loop in A115V:dTTP is more loose because the V115 side chain and the 5-methyl group of dTTP causes movement of the backbone (Figure 9.13). The additional space at His112 (approximately 1.5 Å movement of His112 Nπ) in A115V:dTTP may allow protonated His112 in the active conformation. A115V dTTP inhibition was stronger at pH 8.0 compared with pH 6.8 (Figure 9.4B), which can be explained by deprotonated His112 leaving more space for dTTP binding than protonated His112.

The determined K_M for the A115V variant was very low for dUTP at pH 8.0 (0.6 ± 0.2 µM). K_M values in the 0.1-0.5 µM range have been reported for *E. coli* dUTPase [101,116], but it is unlikely that the A115V variant contains a residual amount of the dUTPase subsequent to purification. *Ec* dUTPase is brought to 60 % saturation with ammonium sulfate during purification [117], whereas the bifunctional enzyme is only brought to 40-45 % saturation. Also, the pH effect on K_M and inhibition by dTTP seen for A115V is not seen for *E. coli* dUTPase [116].

The WT enzyme shows a slightly higher k_{cat} (1.2 fold) and a slightly lower K_M (1.5 fold) for dCTP compared with dUTP at pH 8.0, see Table 9.4. This is contrary to the rates determined by Helt *et al.* [85] at pH 6.8, where a lower reaction rate for dCTP was observed compared with dUTP. The protonation state of His112 thus seems to have an influence on whether dCTP deamination or dUTP dephophorylation appears to be the rate determining reaction.

Figure 9.22 presents a schematic overview of the discussed conformations of A115V and WT *Mt* DCD-DUT as well as the protonation state of His112 and how it influences the nucleotide binding.

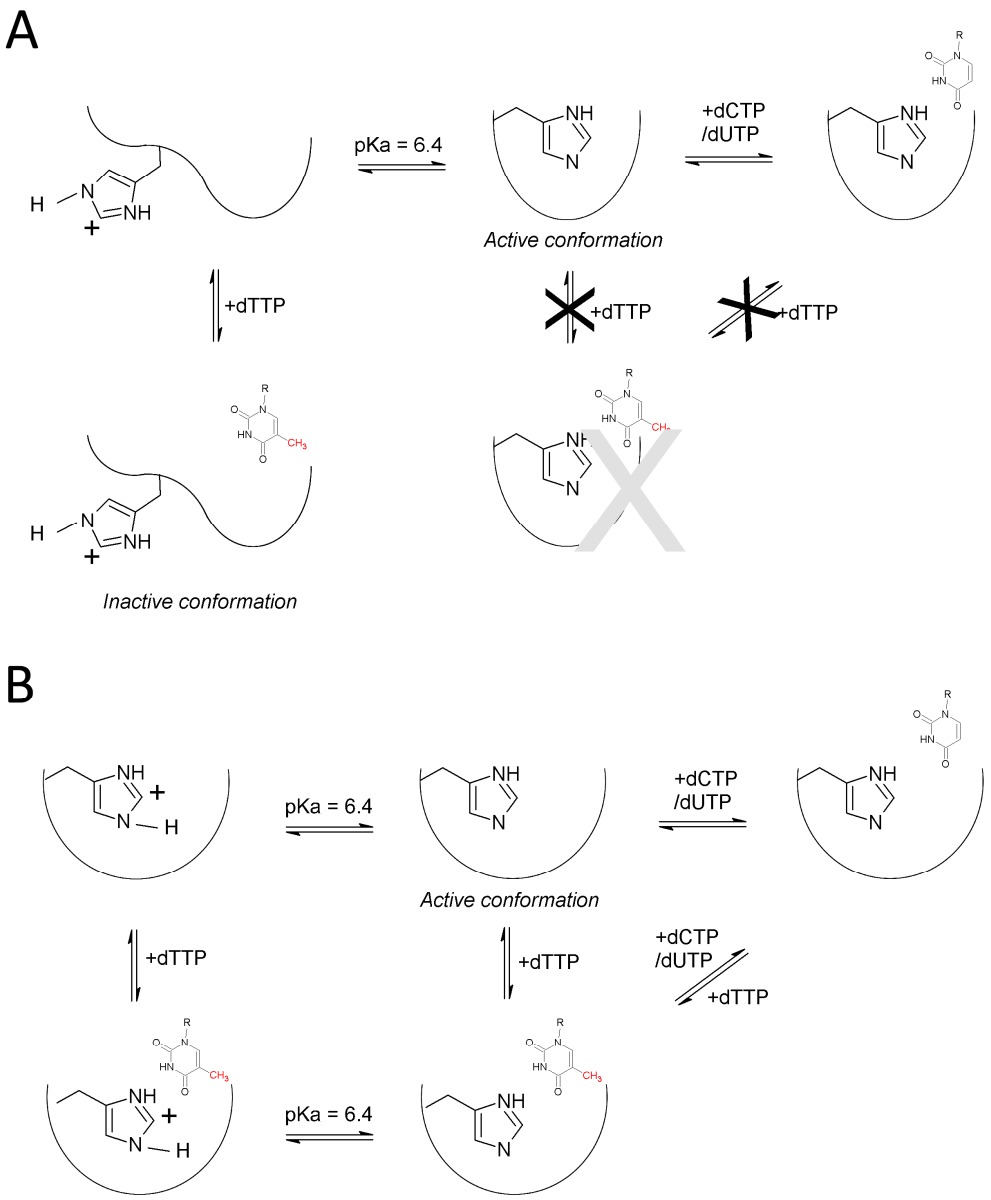


Figure 9.22: Schematic overview of the protonation/deprotonation of His112 and relation to the active/inactive conformations of the protein. A: WT *Mt* DCD-DUT does not bind dTTP in the active conformation and assumes the inactive conformation when His112 is protonated. B: A115V *Mt* DCD-DUT only exists in a slightly wider conformation, very similar to the active conformation, where it binds dTTP regardless of the protonation state of His112.

9.4.2 Properties of the A115G variant

The A115G variant was investigated as a potential enzyme for cleavage of dTTP as discussed in section 7.1. However, the enzyme was unable to cleave dTTP and had about 20 fold reduction in k_{cat} for dCTP at pH 8.0 as well as for dUTP at pH 8.0 and 6.8.

Change of Ala115 to a glycine is very likely to introduce additional flexibility to the already flexible 110-118 region as judged from available structures of bifunctional enzymes (Figure 9.12 and Figure 9.15). There is no structural hindrance for the A115G variant to be able to exist in both the active and inactive conformations. The similar morphology of the apo crystals of WT and A115G suggests that the A115G enzyme adopts the same conformation as the WT enzyme. The introduced flexibility is situated very close to the interfaces between the molecules in the trimer. The observed A115G dTTP binding curve (Figure 9.6C), where dTTP occupation decreases at higher dTTP concentrations, could possibly be explained by the trimer falling apart when dTTP attempts to bind to the second or third site of the trimer, thereby releasing dTTP to the solution. The increased flexibility may thus result in a destabilization of the trimer when dTTP binds, causing a breakdown of the A115G quaternary structure.

While the dUTP saturation curves for WT are hyperbolic at pH 8.0 and sigmoidal at pH 6.8, dUTP saturation curves for A115G are hyperbolic at both pH 8.0 and pH 6.8. Furthermore, the A115G variant shows cooperativity for dCTP saturation at pH 8.0, while the WT enzyme shows hyperbolic dCTP saturation. The A115G variant thus appears to distinguish between dCTP and dUTP as substrates at pH 8.0 (Figure 9.3C). dUTP readily binds and is dephosphorylated, whereas a rearrangement is necessary for dCTP deamination to take place. Ala115 is not conserved among dUTPases (Chapter 6) and, hence, it may be speculated that changing Ala115 to a glycine may have little effect on the conformation of residues important for dUTPase activity. However, the glycine may still affect the ability of the enzyme to bind and convert dUTP due to the introduced flexibility near the active site. In the WT enzyme, the backbone oxygen of Ala115 binds to the catalytic water molecule in the deamination reaction and plays an important role for dCTP deaminase activity (described in section 6.2). Glycine lacks Cβ and thus has additional allowed backbone conformations compared with other residues [72]. This may be related to the observed cooperativity for A115G with dCTP at pH 8.0, where an adjustment of the Gly115 backbone in the flexible loop may be needed for the reaction to take place.

An opposite pH effect of dTTP inhibition is observed for the A115G variant compared with the WT enzyme (Figure 9.4). A115G shows strong dTTP inhibition at pH 8.0 and no significant inhibition at pH 6.8. However, the maximal dTTP concentration was 0.1 mM and additional experiments are needed to determine whether the enzyme is completely uninhibited at pH 6.8. The pH dependence of the A115G dTTP inhibition was contrary to that of the WT enzyme. The A115G enzyme was strongly inhibited by dTTP at pH 8.0, whereas the WT enzyme was insensitive to dTTP at pH 8.0. The additional space obtained by replacing alanine with glycine could be a simple explanation for dTTP binding at pH 8.0. The effect associated with the unusual dTTP binding curve at pH 8.0 discussed above is also likely to contribute to the loss of activity observed at pH 8.0. Protonation of His112 at pH 6.8 might have a stabilizing effect on the flexible

loop through favorable interactions with backbone oxygens thereby reducing the detrimental effect of dTTP. For an illustration of the loop, see Figure 9.15A.

9.5 Conclusion

Substrate saturation and dTTP inhibition studies have been carried out for WT, A115V and A115G *Mt* DCD-DUT, and the structure of A115V in complex with dTTP has been solved.

The two variant enzymes were constructed to investigate the importance of steric hindrance on a water molecule suggested to play a key role in dephosphorylation. However, this water molecule was present in the structure of A115V:dTTP. The A115V variant did not dephosphorylate dTTP and the water molecule can therefore not explain the lack of activity with dTTP bound. The pyrimidine moiety of dTTP in the structures of WT and A115V is positioned at a different angle than the pyrimidine moiety of dUMPNPP in *Mj* DCD-DUT. This rotation allows a conserved arginine to form a hydrogen bond with O2 of the dTTP pymidine moiety. The conserved arginine also forms hydrogen bonds with a conserved serine that plays an important role in dephosphorylation, possibly by stabilizing negative charge formed on the oxygen bridging the α-β-phosphorous of the substrate. The additional hydrogen bond between arginine and dTTP impairs the ability of the arginine to be a hydrogen bond donor to the serine and thereby also impairs the stabilization of negative charge formed on the oxygen bridging the α-β-phosphorous in the course of the reaction. The hydrogen bond between the arginine and dTTP is also likely to restrain the position of the arginine and affect the ability of the arginine to bond to the flexible serine side chain.

WT enzyme proved insensitive to the presence of dTTP at pH 8.0, while dTTP inhibition at pH 6.8 was confirmed. This can be explained by the protonation state of His112 situated close to the nucleotide binding site. The flexible loop consisting of residues 110-118 can exist in the active or the inactive conformation. The active conformation of the WT enzyme is likely to be the same as is observed in *Mj* DCD-DUT:dUMPNPP, in which the WT enzyme is unable to bind dTTP due to steric hindrance by Ala115 and Gly116. The WT enzyme is stabilized in the active conformation at pH 8.0 because of the completely deprotonated His112. This can explain the insensitivity to dTTP at pH 8.0. The hyperbolic substrate saturation of the WT enzyme at pH 8.0 for both dCTP and dUTP supports that the enzyme is present solely in the active conformation. In contrast, protonation of His112 at pH 6.8 strains the loop in the active conformation and allows the enzyme to bind dTTP in the inactive conformation.

The conformation of the flexible 110-118 region of the A115V:dTTP structure is very similar to the active conformation observed in *Mj* DCD-DUT:dUMPNPP, whereas the WT:dTTP structure is clearly in a different inactive conformation. The A115V variant enzyme is unable to exist in the inactive conformation due to steric hindrance caused by the introduced valine side chain. The flexible loop in the inactive conformation of WT:dTTP has 7 additional hydrogen bonds compared with the A115V:dTTP structure, and the WT enzyme thus favors binding dTTP in the inactive conformation. The difference in hydrogen bonds can explain the reduced thermal

stability of the A115V enzyme compared with the WT enzyme in the presence of dTTP (transition temperatures of 67 °C vs. 87 °C, respectively).

The A115V variant is inhibited by dTTP at both pH 6.8 and 8.0, albeit with a stronger inhibition at pH 8.0. The flexible loop in the A115V:dTTP structure is more loose compared with the loop in the active structure of *Mj* DCD-DUT:dUMPNPP. This additional space may account for dTTP binding even with protonated His112. The stronger inhibition at pH 8.0 can be explained by deprotonated His112 leaving additional space for dTTP binding compared with protonated His112.

The A115G variant was very sensitive to dTTP at pH 8.0, while no substantial inhibition was observed at pH 6.8. The pH dependence of the A115G dTTP inhibition was contrary to that of the WT enzyme. This can be explained by the additional space for dTTP binding, which is a result of replacing alanine with glycine. However, the unusual dTTP binding curve at pH 8.0 suggests that dTTP has a detrimental effect on the protein because of the introduced flexibility near the nucleotide binding site. This effect appears to be much reduced at pH 6.8, where the A115G shows significantly lower sensitivity to dTTP. This could be caused by protonated His112 having a stabilizing effect through a hydrogen bond to Gly115 O, thereby introducing a structural restraint on the very flexible 110-118 region.

The A115G variant showed cooperativity with dCTP as substrate and not with dUTP as substrate at pH 8.0. This behavior is not seen for the WT enzyme, which shows hyperbolic saturation with both substrates at pH 8.0. The difference in substrate saturation for dCTP and dUTP for the A115G variant at pH 8.0 may be caused by the additional flexibility introduced by the glycine in the already flexible loop. The additional allowed backbone conformations of glycine may have an effect on the observed cooperativity with dCTP at pH 8.0. The backbone of Gly115 may have to rearrange for the deamination to take place. This rearrangement may be indifferent for the dUTPase reaction, as this residue is involved in deamination and not in dephosphorylation.

Chapter Ten

10 Concluding Remarks

This thesis presents work on two different projects: the all cysteinyl coordinated D14C variant of the hyperthermostable *Pyrococcus furiosus* ferredoxin as well as WT and two variants of the bifunctional dCTP deaminase-dUTPase from *Mycobacterium tuberculosis*. The ferredoxin work includes expression and purification of D14C [Fe_4S_4] *Pf* Fd, as well as cluster conversion and subsequent purification, crystallization and crystal structure of D14C [Fe_3S_4] *Pf* Fd. The work on the bifunctional enzyme and variants includes purifications, substrate saturation at pH 8.0 and pH 6.8, dTTP inhibition studies, dTTP binding, DSC and crystallization setups with apo enzyme or enzyme in complex with dTTP, dUMPNPP or dUMPCPP. The crystal structure of A115V *Mt* DCD-DUT with dTTP bound has been solved. To keep track of crystallization trials and results, a program named *MyCrystals* has been developed and used extensively throughout this work. *MyCrystals* combines pictures with crystallization conditions and is able to sort the pictures based on selected conditions to give an overview of the crystallization results.

Changing the cluster coordinating aspartate to cysteine in *Pf* Fd proved to impair the ease with which D14C [Fe_4S_4] Fd converted to D14C [Fe_3S_4] Fd. Purification of the [Fe_3S_4] containing D14C Fd was furthermore challenged by the formation of a disulfide bonded dimer when the free cysteine was not protonated. Formation of the disulfide bonded dimer was observed at pH 8.0, whereas only the monomer was present at pH 5.8. The crystal structure of D14C [Fe_3S_4] was solved to 2.8 Å despite difficulties in crystallization due to inhomogeneity of the protein solution. The crystal structure is the first structure with a [Fe_3S_4] cluster, in which a cysteine from a full cysteine binding motif is unprotected and facing away from the cluster. The structure is in close resemblance with the WT [Fe_3S_4] structure. The crystal packing in both D14C and WT [Fe_3S_4] Fd showed extended β-sheet dimers. These dimers were not observed in solution and were probably formed as a result of the high protein concentration in the crystals.

WT, A115V and A115G *Mt* DCD-DUT have been successfully purified. The A115G variant was prone to degradation and in order to obtain intact protein, the purification procedure was carried out quickly and with minimal protein precipitation. Precipitation issues were solved by changing pH of the buffer from 6.8 to 8.0. Distinct differences in the enzyme kinetics were observed at these two pH values, and structural analyses were used as a basis for explaining this behavior. The variants were created to investigate the importance of steric hindrance on a water molecule suggested to play a key role in dephosphorylation. However, this water molecule was present in the structure of A115V:dTTP and the variant did not dephosphorylate dTTP. A rotation of the dTTP pyrimidine moiety is observed in the structures of A115V and WT compared with the pyrimidine moiety of dUMPNPP in *Mj* DCD-DUT. This causes changes in the hydrogen bonding pattern of conserved residues in the active site and may give rise to less stabilization of the

negative charge formed on the oxygen bridging the α-β-phosphorous in the course of the dephosphorylation reaction. The structure of A115V *Mt* DCD-DUT with dTTP bound is in close resemblance with DCD-DUT from *M. jannaschii* with dUMPNPP bound and thus contributes with important information on the conformation of conserved residues. The flexible region made up of residues 110-118 near the nucleotide binding site rearranges between an active and an inactive conformation. dTTP binds to the WT enzyme in the inactive conformation (pdb 2QXX), while the active conformation is seen in the structure of *Mj* DCD-DUT with a substrate analogue bound (pdb 2HXD). The flexible region in the solved structure of A115V:dTTP is very similar to the active conformation. Unlike the WT enzyme, the A115V variant is unable to bind dTTP in the inactive conformation due to steric hindrance caused by the introduced valine side chain. The differences seen between A115V:dTTP and *Mj* DCD-DUT:dUMPNPP are caused by the additional methyl group on dTTP and the larger size of the valine side chain. The WT enzyme could exist in the same conformation as A115V:dTTP with dTTP bound, but the WT enzyme has 7 additional hydrogen bonds in the 110-118 region in the inactive conformation and thus favors binding dTTP in the inactive conformation. This difference in hydrogen bonds may also explain the much reduced thermal stability of the A115V enzyme compared with the WT enzyme in the presence of dTTP (transition temperatures of 67 vs. 87 °C, respectively).

WT enzyme was inhibited by dTTP at pH 6.8 and unexpectedly proved insensitive to dTTP at pH 8.0. The protonation state of the conserved His112 in the flexible loop is likely to play an important role in the differences in dTTP binding and inhibition observed with varying pH. At pH 8.0, His112 is completely deprotonated and stabilized in the active conformation, which does not allow dTTP binding because of the Ala115 and Gly116 backbones. Protonated His112 strains the 110-118 active conformation and allows the WT enzyme to bind dTTP in the inactive conformation at pH 6.8.

The A115V variant was inhibited by dTTP at both pH 6.8 and 8.0 and thus allowed dTTP binding even with partly protonated His112. The structure of A115V:dTTP reveals a more loose loop of the 110-118 region compared with *Mj* DCD-DUT:dUMPNPP because of steric hindrance caused by the introduced valine side chain and the methyl group of dTTP. This may allow sufficient room for protonated histidine even with dTTP bound in a conformation very similar to the active conformation.

The A115G variant was very sensitive to dTTP at pH 8.0, while no substantial inhibition was observed at pH 6.8. dTTP binding at pH 8.0 may be explained by the additional flexibility and space caused by the introduced glycine. In contrast to both WT and A115V, the protonated histidine in A115G may induce stability by binding to the backbone oxygen of Gly115, thereby causing a structural restraint on the very flexible loop and reducing dTTP inhibition. The A115G variant appeared to distinguish between dCTP and dUTP as substrates at pH 8.0, where substrate saturation curves were sigmoidal and hyperbolic, respectively. This may be related to Ala115 (for WT) being involved in dCTP deamination. Glycine has additional allowed backbone conformations, and A115G may thus require adjustment of the protein chain at Gly115 in order for the deamination to take place. In contrast, no dependence on the Gly115 conformation was seen for dephosphorylation, emphasizing that this residue is not involved in dephosphorylation.

Chapter Eleven

11 Outlook

The successful formation of D14C [Fe_3S_4] *Pf* Fd is a preceding step to the formation of heterometallic clusters. Subsequent to the developed oxidation procedure presented in this thesis, the D14C Fd has been synthesized with two heterometallic clusters. [$ZnFe_3S_4$] and [$AgFe_3S_4$] were produced, purified and characterized by Maja Martic [63]. Additional studies of other types of heterometallic clusters as well as studies of the possible catalytic abilities of these ferredoxins are among the perspectives.

Evolution is constantly on-going and studying evolution is of interest to provide substantial, fundamental knowledge on how organisms function and evolve. In unraveling the evolution of base pairing in DNA, the study of an intermediate step in dTTP synthesis in this work is but a small contribution. However, all contributions are important when adding up information across molecular families.

It is difficult to ascertain what kind of mutational study could entail dephosphorylation of dTTP by the bifunctional *Mt* DCD-DUT. Removing side chains introduces additional flexibility and has a detrimental effect on the protein nucleotide binding, as was seen for the A115G variant. However, the steric hindrance of particularly Ala115 and Gly116 backbones cause dTTP to bind with the pyrimidine moiety in a different angle, thereby causing the hydrogen bond pattern and the flexibility of conserved residues to change and possibly disrupt activity.

pH proved to play an important role for activity and inhibition, and further studies of the pH effect on the properties of other members of the enzyme family can contribute to the understanding of the factors influencing substrate and/or inhibitor binding.

BIBLIOGRAPHY

[1] Aono S, Bryant FO and Adams MWW. A Novel and Remarkably Thermostable Ferredoxin from the Hyperthermophilic Archaebacterium *Pyrococcus furiosus*. *J. Bacteriol.* **(1989)** 171: 3433-9.

[2] Conover RC, Kowal AT, Fu W, Park JB, Aono S, Adams MWW and Johnson MK. Spectroscopic Characterization of the Novel Iron-Sulfur Cluster in *Pyrococcus furiosus* Ferredoxin. *J. Biol. Chem.* **(1990)** 265: 8533-41.

[3] Lu Y, Yeung N, Sieracki N and Marshall NM. Design of Functional Metalloproteins. *Nature* **(2009)** 460: 855-62.

[4] World Health Organization. www.who.int/tb. *accessed 26 April 2011*.

[5] DeLano WL. The PyMOL Molecular Graphics System. *www.pymol.org, Schrödinger* **(2002)**.

[6] Løvgreen MN, Løvgreen M, Christensen HEM and Harris P. *MyCrystals* - A Simple Visual Data Management Program for Laboratory-Scale Crystallization Experiments. *J. Appl. Cryst.* **(2009)** 42: 741-2.

[7] Meining W. *XtalBase* - A Comprehensive Data Management System for Macromolecular Crystallography. *J. Appl. Cryst.* **(1999)** 39: 759-66.

[8] Brodersen DE, Jenner LB, Andersen GR and Nyborg J. *XAct:* A Program for Construction, Automated Setup and Bookkeeping of Crystallization Experiments. *J. Appl. Cryst.* **(1999)** 32: 1012-6.

[9] Harris M and Jones TA. *Xtrack* - A Web-Based Crystallographic Notebook. *Acta Cryst.* **(2002)** D58: 1889-91.

[10] Fulton KF, Ervine S, Faux N, Forster R, Jodun RA, Ly W, Robiliard L, Sonsini J, Whelan D, Whisstock JC and Buckle AM. *CLIMS*: Crystallography Laboratory Information Management System. *Acta Cryst.* **(2004)** D60: 1691-3.

[11] Haebel PW, Arcus VL, Baker EN and Metcalf P. *LISA*: An Intranet-Based Flexible Database for Protein Crystallography Project Management. *Acta Cryst.* **(2001)** D57: 1341-3.

[12] Morris C, Pajon A, Griffiths SL, Daniel E, Savitsky M, Lin B, Diprose JM, Wilter da Silva A, Pilicheva K, Troshin P, van Niekerk J, Isaacs N, Naismith J, Nave C, Blake R, Wilson KS, Stuart DI, Henrick K and Esnouf RM. The Protein Information Management System (*PiMS*): A Generic Tool for Any Structural Biology Research Laboratory. *Acta Cryst.* **(2011)** D67: 249-60.

[13] Meyer J. Iron-Sulfur Proteins Folds, Iron-Sulfur Chemistry, and Evolution. *J. Biol. Inorg. Chem.* **(2008)** 13: 157-70.

[14] Beinert H. Iron-sulfur Proteins: Ancient Structures, Still Full of Surprises. *J. Biol. Inorg. Chem.* **(2000)** 5: 2-15.

[15] Sheftel A, Stehling O and Lill R. Iron-Sulfur Proteins in Health and Disease. *Trends Endocrinol. Metab.* **(2010)** 21: 302-14.

[16] Eck RV and Dayhoff MO. Evolution of the Structure of Ferredoxin Based on Living Relics of Primitive Amino Acid Sequences. *Science* **(1966)** 152: 363-6.

[17] Hall DO, Cammack R and Rao KK. Role for Ferredoxins in the Origin of Life and Biological Evolution. *Nature* **(1971)** 233: 136-8.

[18] Beinert H, Holm RH and Münck E. Iron-Sulfur Clusters: Nature's Modular, Multipurpose Structures. *Science* **(1997)** 277: 653-9.

[19] Rees DC and Howard JB. The Interface Between the Biological and Inorganic Worlds: Iron-Sulfur Metalloclusters. *Science* **(2003)** 300: 929-31.

[20] Kerfeld CA, Salmeen AE and Yeates TO. Crystal Structure and Possible Dimerization of the High-Potential Iron-Sulfur Protein from *Chromatium purpuratum*. *Biochemistry* **(1998)** 37: 13911-7.

[21] Backes G, Mino Y, Loehr TM, Meyer TE, Cusanovich MA, Sweeney WV, Adman ET and Sanders-Loehr J. The Environment of Fe_4S_4 Clusters in Ferredoxins and High-Potential Iron Proteins. New Information from X-ray Crystallography and Resonance Raman Spectroscopy. *J. Am. Chem. Soc.* **(1991)** 113: 2055-64.

[22] Giastas P, Pinotsis N, Efthymiou G, Wilmanns M, Kyritsis P, Moulis JM and Mavridis IM. The Structure of the 2[4Fe-4S] Ferredoxin from *Pseudomonas aeruginosa* at 1.32-Å Resolution: Comparison with other High-Resolution Structures of Ferredoxins and Contributing Structural Features to Reduction Potential Values. *J. Biol. Inorg. Chem.* **(2006)** 11: 445-58.

[23] Saridakis E, Giastas P, Efthymiou G, Thoma V, Moulis JM, Kyritsis P and Mavridis I. Insights Into the Protein and Solvent Contributions to the Reduction Potentials of $[4Fe\text{-}4S]^{2+/+}$ Clusters: Crystal Structures of the *Allochromatium vinosum* Ferredoxin Variants C57A and V13G and the Homologous *Escherichia coli* Ferredoxin. *J. Biol. Inorg. Chem.* **(2009)** 14: 783-99.

[24] Adman E, Watenpaugh KD and Jensen LH. NH--S Hydrogen Bonds in *Peptococcus aerogenes* Ferredoxin, *Clostridium pasteurianum* Rubredoxin, and *Chromatium* High Potential Iron Protein. *Proc. Nat. Acad. Sci. USA* **(1975)** 72: 4854-8.

[25] Dey A, Jenney Jr FE, Adams MWW, Babini E, Takahashi Y, Fukuyama K, Hodgson KO, Hedman B and Solomon EI. Solvent Tuning of Electrochemical Potentials in the Active Sites of HiPIP Versus Ferredoxin. *Science* **(2007)** 318: 1464-8.

[26] Niu S and Ichiye T. Insight into Environmental Effects on Bonding and Redox Properties of [4Fe-4S] Clusters in Proteins. *JACS* **(2009)** 131: 5724-5.

[27] Robbins AH and Stout CD. The Structure of Aconitase. *Proteins Struct. Funct. Genet.* **(1989)** 5: 289-312.

[28] Robbins AH and Stout CD. Structure of Activated Aconitase: Formation of the [4Fe-4S] Cluster in the Crystal. *Proc. Natl. Acad. Sci. USA* **(1989)** 86: 3639-43.

[29] Beinert H, Kennedy MC and Stout CD. Aconitase as Iron-Sulfur Protein, Enzyme, and Iron-Regulatory Protein. *Chem. Rev.* **(1996)** 96: 2335-73.

[30] Fukuyama K. Ferredoxins Containing One [4Fe-4S] Center. Ferredoxins Containing One [4Fe-4S] Center. *Wiley,* **(2001)** New York, 543-52.

[31] Busch JLH, Breton JL, Bartlett BM, Armstrong FA, James R and Thomson AJ. [3Fe-4S] « [4Fe-4S] Cluster Interconversion in *Desulfovibrio africanus* Ferredoxin III: Properties of an Asp^{14} ® Cys mutant. *Biochem. J.* **(1997)** 323: 95-102.

[32] Butt JN, Fawcett SEJ, Breton J, Thomson AJ and Armstrong FA. Electrochemical Potential and pH Dependences of [3Fe-4S] « [M3Fe-4S] Cluster Transformations (M = Fe, Zn, Co, and Cd) in Ferredoxin III from *Desulfovibrio africanus* and Detection of a Cluster with M = Pb. *J. Am. Chem. Soc.* **(1997)** 119: 9729-37.

[33] Moura I, Moura JJG, Munck E, Papaefthymiou V and Legall J. Evidence for the Formation of a $CoFe_3S_4$ Cluster in *Desulfovibrio gigas* Ferredoxin II. *J. Am. Chem. Soc.* **(1986)** 108: 349-51.

[34] Conover RC, Park JB, Adams MWW and Johnson MK. Formation and Properties of a $NiFe_3S_4$ Cluster in *Pyrococcus furiosus* Ferredoxin. *J. Am. Chem. Soc.* **(1990)** 112: 4562-4.

[35] Srivastava KKP, Surerus KK, Conover RC, Johnson MK, Park JB, Adams MWW and Munck E. Mossbauer Study of $ZnFe_3S_4$ and $NiFe_3S_4$ Clusters in *Pyrococcus furiosus* Ferredoxin. *Inorg. Chem.* **(1993)** 32: 927-36.

[36] Fu WG, Telser J, Hoffman BM, Smith ET, Adams MWW, Finnegan MG, Conover RC and Johnson MK. Interaction of Tl^+ and Cs^+ with the [Fe_3S_4] Cluster of *Pyrococcus furiosus* Ferredoxin - Investigation by Resonance Raman, MCD, EPR, and ENDOR Spectroscopy. *J. Am. Chem. Soc.* **(1994)** 116: 5722-9.

[37] Finnegan MG, Conover RC, Park JB, Zhou ZH, Adams MWW and Johnson MK. Electronic, Magnetic, Redox, and Ligand-Binding Properties of [MFe_3S_4] Clusters (M=Zn, Co, Mn) in *Pyrococcus furiosus* Ferredoxin. *Inorg. Chem.* **(1995)** 34: 5358-69.

[38] Staples CR, Dhawan IK, Finnegan MG, Dwinell DA, Zhou ZH, Huang HS, Verhagen MFJM, Adams MWW and Johnson MK. Electronic, Magnetic, and Redox Properties of [MFe_3S_4] Clusters (M = Cd, Cu, Cr) in *Pyrococcus furiosus* Ferredoxin. *Inorg. Chem.* **(1997)** 36: 5740-9.

[39] Calzolai L, Gorst CM, Zhou ZH, Teng Q, Adams MWW and La Mar GN. ^1H NMR Investigation of the Electronic and Molecular Structure of the Four-Iron Cluster Ferredoxin from the Hyperthermophile *Pyrococcus furiosus*. Identification of Asp 14 as a Cluster Ligand in Each of the Four Redox States. *Biochemistry* **(1995)** 34: 11373-84.

[40] Gorst CM, Zhou ZH, Ma K, Teng Q, Howard JB, Adams MWW and La Mar GN. Participation of the Disulfide Bridge in the Redox Cycle of the Ferredoxin from the Hyperthermophile *Pyrococcus furiosus*: ^{1}H Nuclear Magnetic Resonance Time Resolution of the Four Redox States at Ambient Temperature. *Biochemistry* **(1995)** 34: 8788-95.

[41] Nielsen MS, Harris P, Ooi BL and Christensen HEM. The 1.5 Å Resolution Crystal Structure of $[Fe_3S_4]$-Ferredoxin from the Hyperthermophilic Archaeon *Pyrococcus furiosus*. *Biochemistry* **(2004)** 43: 5188-94.

[42] Sham S, Calzolai L, Bren KL, Teng Q, Jenney Jr FE, Brereton PS, Howard JB, Adams MWW and La Mar GN. A Solution NMR Molecular Model for the Aspartate-Ligated, Cubane Cluster Containing Ferredoxin from the Hyperthermophilic Archeaon *Pyrococcus furiosus*. *Biochemistry* **(2002)** 41: 12498-508.

[43] Nielsen MS, Harris P and Christensen HEM. Crystallization of $[Fe_3S_4]$-Ferredoxin from the Hyperthermophile Archaeon *Pyrococcus furiosus*. *Acta Cryst.* **(2003)** D59: 2325-7.

[44] Hasan MN, Hagedoorn PL and Hagen WR. *Pyrococcus furiosus* Ferredoxin is a Functional Dimer. *FEBS Lett.* **(2002)** 531: 335-8.

[45] Hasan MN, Kwakernaak C, Sloof WG, Hagen WR and Heering HA. *Pyrococcus furiosus* 4Fe-Ferredoxin, Chemisorbed on Gold, Exhibits Gated Reduction and Ionic Strength Dependent Dimerization. *J. Biol. Inorg. Chem.* **(2006)** 11: 651-62.

[46] Blamey JM, Chiong M, López C and Smith ET. Purification and Characterization of Ferredoxin from the Hyperthermophilic *Pyrococcus woesei*. *Anaerobe* **(2000)** 6: 285-90.

[47] Duderstadt RE, Staples CR, Brereton PS, Adams MWW and Johnson MK. Effects of Mutations in Aspartate 14 on the Spectroscopic Properties of the $[Fe_3S_4]^{+,0}$ Clusters in *Pyrococcus furiosus* Ferredoxin. *Biochemistry* **(1999)** 38: 10585-93.

[48] Zhou ZH and Adams MWW. Site-Directed Mutations of the 4Fe-Ferredoxin from the Hyperthermophilic Archaeon *Pyrococcus furiosus*: Role of the Cluster-Coordinating Aspartate in Physiological Electron Transfer Reactions. *Biochemistry* **(1997)** 36: 10892-900.

[49] Brereton PS, Duderstadt RE, Staples CR, Johnson MK and Adams MWW. Effect of Serinate Ligation at Each of the Iron Sites of the $[Fe_4S_4]$ Cluster of *Pyrococcus furiosus* Ferredoxin on the Redox, Spectroscopic, and Biological Properties. *Biochemistry* **(1999)** 38: 10594-605.

[50] Brereton PS, Verhagen MFJM, Zhou ZH and Adams MWW. Effect of Iron-Sulfur Cluster Environment in Modulating the Thermodynamic Properties and Biological Function of Ferredoxin from *Pyrococcus furiosus*. *Biochemistry* **(1998)** 37: 7351-62.

[51] Calzolai L, Gorst CM, Bren KL, Zhou ZH, Adams MWW and La Mar GN. Solution NMR Study of the Electronic Structure and Magnetic Properties of Cluster Ligation Mutants of the Four-Iron Ferredoxin from the Hyperthermophilic Archaeon *Pyrococcus furiosus*. *J. Am. Chem. Soc.* **(1997)** 119: 9341-50.

[52] Calzolai L, Zhao ZH, Adams MWW and La Mar GN. Role of Cluster-Ligated Aspartate in Gating Electron Transfer in the Four-Iron Ferredoxin from the Hyperthermophilic Archaeon *Pyrococcus furiosus*. *J. Am. Chem. Soc.* **(1996)** 118: 2513-4.

[53] Kim C, Brereton PS, Verhagen MFJM and Adams MWW. Ferredoxin from *Pyrococcus furiosus*. *Meth. Enzymol.* **(2001)** 334: 30-40.

[54] Metalloprotein Chemistry and Engineering Research Group. Department of Chemistry, Technical University of Denmark. *Unpublished Results*.

[55] Moura JJ, Moura I, Kent TA, Lipscomb JD, Huynh BH, LeGall J, Xavier AV and Münck E. Interconversions of [3Fe-3S] and [4Fe-4S] Clusters. Mossbauer and Electron Paramagnetic Resonance Studies of *Desulfovibrio gigas* Ferredoxin II. *J. Biol. Chem.* **(1982)** 257: 6259-67.

[56] Kolthoff IM and Tomsicek WJ. The Oxidation Potential of the System Potassium Ferrocyanide-Potassium Ferricyanide at Various Ionic Strengths. *J. Phys. Chem.* **(1935)** 39: 945-54.

[57] Hume DN and Kolthoff IM. A Revision of the Oxidation Potentials of the Orthophenanthroline- and Dipyridyl-Ferrous Complexes. *J. Am. Chem. Soc.* **(1943)** 65: 1895-7.

[58] Alphalyse. Molecular Weight Determination Report. *Order 11334, Odense, February 12* **(2008)**.

[59] Bas DC, Rogers DM and Jensen JH. Very Fast Prediction and Rationalization of pK_a Values for Protein-Ligand Complexes. *Proteins: Struct. , Funct. , Bioinf.* **(2008)** 73: 765-83.

[60] Li H, Robertson AD and Jensen JH. Very Fast Empirical Prediction and Rationalization of Protein pK_a Values. *Proteins: Struct. , Funct. , Bioinf.* **(2005)** 61: 704-21.

[61] Emsley P, Lohkamp B, Scott WG and Cowtan K. Features and Development of *Coot*. *Acta Cryst.* **(2010)** D66: 486-501.

[62] Lohkamp B, Emsley P and Cowtan K. Coot News. *CCP4 Newsletter* **(2005)** 42: 7.

[63] Martic M. Ph.D. Thesis: Ferredoxins containing Heterometallic Clusters - Synthesis and Characterization. *Technical University of Denmark, Department of Chemistry* **(2011)**.

[64] Løvgreen MN, Martic M, Windahl MS, Christensen HEM and Harris P. Crystal Structures of the All-Cysteinyl-Coordinated D14C Variant of *Pyrococcus furiosus* Ferredoxin: [4Fe–4S] ↔ [3Fe–4S] Cluster Conversion. *J. Biol. Inorg. Chem.* **(2011)** 16: 763-75.

[65] Kissinger CR, Sieker LC, Adman ET and Jensen LH. Refined Crystal Structure of Ferredoxin II from *Desulfovibrio gigas* at 1.7 Å. *J. Mol. Biol.* **(1991)** 219: 693-715.

[66] Shirakawa T, Takahashi Y, Wada K, Hirota J, Takao T, Ohmori D and Fukuyama K. Identification of Variant Molecules of *Bacillus thermoproteolyticus* Ferredoxin: Crystal Structure Reveals Bound Coenzyme A and an Unexpected [3Fe-4S] Cluster Associated with a Canonical [4Fe-4S] Ligans Motif. *Biochemistry* **(2005)** 44: 12402-10.

[67] Leslie AGW. Recent Changes to the MOSFLM Package for Processing Film and Image Plate Data. *Jnt. CCP4/ESF-EAMCB Newsl. Protein Crystallogr.* **(1992)** 26.

[68] Evans P. Scaling and Assessment of Data Quality. *Acta Cryst.* **(2006)** D62: 72-82.

[69] Collaborative Computational Project, Number 4. The CCP4 suite: Programs for Protein Crystallography. *Acta Cryst.* **(1994)** D50: 760-3.

[70] Vagin A and Teplyakov A. Automated Program for Molecular Replacement. *J. Appl. Cryst.* **(1997)** 30: 1022-5.

[71] Murshudov GN, Shubák P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, Winn MD, Long F and Vagin AA. REFMAC5 for the Refinement of Macromolecular Crystal Structures. *Acta Cryst.* **(2011)** D67: 355-69.

[72] Lovell SC, Davis IW, Adrendall WB, de Bakker PIW, Word JM, Prisant MG, Richardson JS and Richardson DC. Structure Validation by C alpha Geometry: phi,psi and C beta Deviation. *Proteins: Struct. Funct. Genet.* **(2003)** 50: 437-50.

[73] Krissinel E and Henrick K. Secondary-Structure Matching (SSM), A New Tool for Fast Protein Structure Alignment in Three Dimensions. *Acta Cryst.* **(2004)** D60: 2256-68.

[74] Thompson JD, Higgins DG and Gibson TJ. CLUSTAL W: Improving the Sensitivity of Progressive Multiple Sequence Alignment through Sequence Weighting, Position Specific Gap Penalties and Weight Matrix Choice. *Nucleic Acids Res.* **(1994)** 22: 4673-80.

[75] The UniProt Consortium. Ongoing and Future Developments at the Universal Protein Resource. *Nucleic Acids Res.* **(2011)** 39: D214-9.

[76] Jain E, Bairoch A, Duvaud S, Phan I, Redaschi N, Suzek BE, Martin MJ, McGarvey P and Gasteiger E. Infrastructure for the Life Sciences: Design and Implementation of the UniProt Website. *BMC Bioinformatics* **(2009)** 10: 136.

[77] Macedo-Ribeiro S, Darimont B, Sterner R and Huber R. Small Structural Changes Account for the High Thermostability of 1[4Fe-4S] Ferredoxin from the Hyperthermophilic Bacterium *Thermotoga maritima*. *Structure* **(1996)** 4: 1291-301.

[78] Smith ET, Blamey JM, Zhou ZH and Adams MWW. A Variable-Temperature Direct Electrochemical Study of Metalloproteins from Hyperthermophilic Microorganisms Involved in Hydrogen-Production from Pyruvate. *Biochemistry* **(1995)** 34: 7161-9.

[79] Sery A, Housset D, Serre L, Bonicel J, Hatchikian C, Frey M and Roth M. Crystal Structure of the Ferredoxin-I from *Desulfovibrio africanus* at 2.3 Angstrom Resolution. *Biochemistry* **(1994)** 33: 15408-17.

[80] Hatchikian EC, Cammack R, Patil DS, Robinson AE, Richards AJM, George S and Thomson AJ. Spectroscopic Characterization of Ferredoxin-I and Ferredoxin-II from *Desulfovibrio africanus*. *Biochim. Biophys. Acta* **(1984)** 784: 40-7.

[81] Fukuyama K, Okada T, Kakuta Y and Takahashi Y. Atomic Resolution Structures of Oxidized [4Fe-4S] Ferredoxin from *Bacillus thermoproteolyticus* in Two Crystal Forms: Systematic Distortion of [4Fe-4S] Cluster in the Protein. *J. Mol. Biol.* **(2002)** 315: 1155-66.

[82] Mullinger RN, Cammack R, Rao KK, Hall DO, Dickson DPE, Johnson CE, Rush JD and Simopoulos A. Physicochemical Characterization of 4-Iron-4-Sulfide Ferredoxin from *Bacillus stearothermophilus*. *Biochem. J.* **(1975)** 151: 75-83.

[83] Berg JM, Tymoczko JL and Stryer L. Biochemistry, Fifth Edition. *W.H. Freeman and Company* **(2002)** New York.

[84] O'Donovan GA and Neuhard J. Pyrimidine Metabolism in Microorganisms. *Bacteriol. Rev.* **(1970)** 34: 278-343.

[85] Helt SS, Thymark M, Harris P, Aagaard C, Dietrich J, Larsen S and Willemoës M. Mechanism of dTTP Inhibition of the Bifunctional dCTP deaminase:dUTPase Encoded by *Mycobacterium tuberculosis*. *J. Mol. Biol.* **(2008)** 376: 554-69.

[86] Björnberg O, Neuhard J and Nyman PO. A Bifunctional dCTP deaminase-dUTP nucleotidohydrolase from the Hyperthermophilic Archaeon *Methanoocaldococcus jannaschii*. *J. Biol. Chem.* **(2003)** 278: 20667-72.

[87] Li H, Graham DE and White RH. The *Methanococcus jannaschii* dCTP deaminase is a Bifunctional deaminase and diphosphatase. *J. Biol. Chem.* **(2003)** 278: 11100-6.

[88] Johansson E, Fanø M, Bynck JH, Neuhard J, Larsen S, Sigurskjold BW, Christensen U and Willemoës M. Structures of dCTP deaminase from *Escherichia coli* with Bound Substrate and Product. *J. Biol. Chem.* **(2005)** 280: 3051-9.

[89] Neuhard J. dCTP deaminase from *Salmonella typhimurium*. *Meth. Enzymol.* **(1978)** 51: 418-23.

[90] Thymark M, Johansson E, Larsen S and Willemoës M. Mutational Analysis of the Nucleotide Binding Site of *Escherichia coli* dCTP deaminase. *Arch. Biochem. Biophys.* **(2008)** 470: 20-6.

[91] Johansson E, Thymark M, Bynck JH, Fanø M, Larsen S and Willemoës M. Regulation of dCTP deaminase from *Escherichia coli* by Nonallosteric dTTP binding to an Inactive Form of the Enzyme. *FEBS J.* **(2007)** 274: 4188-98.

[92] Shlomai J and Kornberg A. Deoxyuridine triphosphatase of *Escherichia coli* -Purification, Properties and Use as a Reagent to Reduce Uracil Incorporation into DNA. *J. Biol. Chem.* **(1978)** 253: 3305-12.

[93] Chan S, Segelke B, Lekin T, Krupka H, Cho US, Kim M, So M, Kim C-, Naranjo CM, Rogers YC, Park MS, Waldo GS, Pashkov I, Cascio D, Perry JL and Sawaya MR. Crystal Structure of the *Mycobacterium tuberculosis* dUTPase: Insights into the Catalytic Mechanism. *J. Mol. Biol.* **(2004)** 341: 503-17.

[94] Harris JM, McIntosh EM and Muscat GEO. Structure/Function Analysis of a dUTPase: Catalytic Mechanism of a Potential Chemotherapeutic Target. *J. Mol. Biol.* **(1999)** 288: 275-87.

[95] Mol CD, Harris JM, McIntosh EM and Tainer JA. Human dUTP pyrophosphatase: Uracil Recognition by a β hairpin and Active Sites Formed by Three Separate Subunits. *Structure* **(1996)** 4: 1077-92.

[96] Cedergren-Zeppezauer ES, Larsson G, Nyman PO, Dauter Z and Wilson KS. Crystal Structure of a dUTPase. *Nature* **(1992)** 355: 740-3.

[97] Prasad GS, Stura EA, McRee DE, Laco GS, Hasselkus-Light C, Eldar JH and Stout CD. Crystal Structure of dUTP pyrophosphatase from Feline Immunodeficiency Virus. *Protein Sci.* **(1996)** 5: 2429-37.

[98] Dauter Z, Persson R, Rosengren AM, Nyman PO, Wilson KS and Cedergren-Zeppezauer ES. Crystal Structure of dUTPase from Equine Infectious Anaemia Virus; Active Site Metal Binding in a Substrate Analogue Complex. *J. Mol. Biol.* **(1999)** 285: 655-73.

[99] Shao H, Robek MD, Threadgill DS, Mankowski LS, Cameron CE, Fuller FJ and Payne SL. Characterization and Mutational Studies of Equine Infectious Anemia Virus dUTPase. *Biochim. Biophys. Acta* **(1997)** 1339: 181-91.

[100] Vertessy BG. Flexible Glycine Rich Motif in *Escherichia coli* deoxyuridine triphosphate nucleotidohydrolase is Important for Functional but not for Structural Integrity of the Enzyme. *Proteins: Struct. Funct. Genet.* **(1997)** 28: 568-79.

[101] Barabás O, Pongrácz V, Kovári J, Wilmanns M and Vértessy B. Structural Insights into the Catalytic Mechanism of Phosphate Ester Hydrolysis by dUTPase. *J. Biol. Chem.* **(2004)** 279: 42907-15.

[102] Palmén LG, Becker K, Bülow L and Kvassman J. A Double Role for a Strictly Conserved Serine: Further Insights into the dUTPase Catalytic Mechanism. *Biochemistry* **(2008)** 47: 7863-74.

[103] Johansson E, Björnberg O, Nyman PO and Larsen S. Structure of the Bifunctional dCTP deaminase-dUTPase from *Methanocaldococcus jannaschii* and its Relation to Other Homotrimeric dUTPases. *J. Biol. Chem.* **(2003)** 278: 27916-22.

[104] Nord J, Larsson G, Kvassman JO, Rosengren AM and Nyman PO. dUTPase from the Retrovirus Equine Infectious Anemia Virus: Specificity, Turnover and Inhibition. *FEBS Lett.* **(1997)** 414: 271-4.

[105] Siggaard JHB, Johansson E, Vognsen T, Helt SS, Harris P and Willemoës M. Concerted Bifunctionality of the dCTP deaminase-dUTPase from *Methanocaldococcus jannaschii*: A Structural and Pre-Steady State Kinetic Analysis. *Arch. Biochem. Biophys.* **(2009)** 490: 42-9.

[106] Gille C. Structural Interpretation of Mutations and SNPs using *STRAP-NT*. *Protein Sci.* **(2006)** 15: 208-10.

[107] Gille C and Frömmel C. *STRAP*: editor for STRuctual Alignments of Proteins. *Bioinformatics* **(2001)** 17: 377-8.

[108] Kabsch W. XDS. *Acta Cryst.* **(2010)** D66: 125-32.

[109] Kabsch W. Integration, Scaling, Space-Group Assignment and Post-Refinement. *Acta Cryst.* **(2010)** D66: 133-44.

[110] Winn MD, Isupov MN and Murshudov GN. Use of TLS Parameters to Model Anisotropic Displacements in Macromolecular Refinement. *Acta Cryst.* **(2001)** D57: 122-33.

[111] Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung L, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC and Zwart PH. PHENIX: A Comprehensive Python-Based System for Macromolecular Structure Solution. *Acta Cryst.* **(2010)** D66: 213-21.

[112] GraphPad Software. GraphPad Prism version 5.04 (trial) for Windows, San Diego California USA. *www.graphpad.com*.

[113] Helt SS. Ph.D. Thesis: Studies of Ferredoxins and dCTP deaminase:dUTPase. *Technical University of Denmark, Department of Chemistry* **(2008)**.

[114] Dolinsky TJ, Nielsen JE, McCammon JA and Baker NA. PDB2PQR: An Automated Pipeline for the Setup, Execution, and Analysis of Poisson-Boltzmann Electrostatics Calculations. *Nucleic Acids Res.* **(2004)** 32: W665-7.

[115] Lerner M. APBS PyMol plugin. *http://www.poissonboltzmann.org/apbs/*.

[116] Larsson G, Nyman PO and Kvassman JO. Kinetic Characterization of dUTPase from *Escherichia coli*. *J. Biol. Chem.* **(1996)** 271: 24010-6.

[117] Hoffmann I, Widström J, Zeppezauer M and Nyman PO. Overproduction and Large-Scale Preparation of Deoxyuridine Triphosphate Nucleotidohydrolase from *Escherichia coli*. *Eur. J. Biochem.* **(1987)** 164: 45-51.

[118] Olsson MHM, Sondergaard CR, Rostkowski M and Jensen JH. PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical pK(a) Predictions. *J. Chem. Theory Comput.* **(2011)** 7: 525-37.

APPENDICES

Appendix A

A *MyCrystals* and Databases

Enclosed is a DVD with the program *MyCrystals*, which is described in Chapter 2. The DVD also contains three databases with crystallization results and corresponding folders including crystallization pictures. The DVD is placed on the last page of the thesis for practical reasons.

Available databases:

Ferredoxin.xml	Database with 461 entries of crystallization results for D14C [Fe_3S_4] *Pyrococcus furiosus* ferredoxin (section 5.3.1).
DCDDUT.xml	Database with 379 entries of crystallization results for WT, A115V and A115G *Mycobacterium tuberculosis* dCTP deaminase-dUTPase. There is an option to sort database results based on enzyme variant (section 9.3.6).
NDKext.xml	Database with 383 entries of crystallization results for an extended version of nucleoside diphosphate kinase from *Sulfolobus solfataricus* (Appendix F.3.1).

Appendix B

B FERREDOXIN EXPERIMENTAL PROCEDURES

Subsequent to the oxidation procedure described in detail below, the protein solution was exchanged into 20 mM Tris/HCl pH 8.0 by ultrafiltration and then purified by anion exchange at pH 8.0 unless otherwise specified. Each solution was loaded onto a 16 mm/10 cm Source 30Q column equilibrated with 20 mM Tris/HCl pH 8.0. The column was washed with 2 CV 20 mM Tris/HCl pH 8.0 and elution was carried out using a linear salt gradient of 0.15 M to 0.4 M NaCl in 20 mM Tris/HCl pH 8.0 over 17 CV.

Ultrafiltration was carried out at 4 °C in a stirred Amicon cell with a PLBC NMWL 3 kDa membrane. During the anaerobic experiments, all buffer solutions were flushed with argon to remove oxygen (1 hour per 100 mL using the manifold and 1 hour per L using the degassing equipment connected to the HPLC). All flasks and tubes were sealed off with Suba-Seal silicon rubber septa (Sigma-Aldrich) and filled with argon through teflon tubings fitted with stainless steel needles. Syringes used for transfer of protein and solutions were also repeatedly purged with argon prior to use.

B. 15 min, 0.3 M NaCl in 20 mM Tris/HCl pH 8.0

15x molar excess of $K_3[Fe(CN)_6]Cl_3$ was used. 5 mL ferricyanide in 20 mM Tris/HCl pH 8.0 was added to 10 mg protein in 5 mL 0.3 M NaCl in 20 mM Tris/HCl pH 8.0. The solution was incubated at room temperature for 15 minutes with gentle stirring.

C. 15 min, 100 mM Tris/HCl pH 7.8

15x molar excess of $K_3[Fe(CN)_6]Cl_3$ was used. 5 mL ferricyanide solution in 100 mM Tris/HCl pH 7.8 was added to 10 mg protein in 5 mL 100 mM Tris/HCl pH 7.8. The solution was incubated at room temperature for 15 minutes with gentle stirring.

D/E. 45 min and 3 hrs, 100 mM Tris/HCl pH 7.8

15x molar excess of $K_3[Fe(CN)_6]Cl_3$ was used. The ferricyanide solution in 100 mM Tris/HCl pH 7.8 was added to 17 mg protein solution (also in 100 mM Tris/HCl pH 7.8) and the volume adjusted with buffer to 17 mL. 8.5 mL were extracted after an incubation time of 45 minutes and the remaining was incubated for a total of 3 hours with gentle stirring.

F. 15 min, 100 mM BisTrisPropane/HCl pH 6.5, anaerobic purification

15x molar excess of ferricyanide was used. 5 mL ferricyanide in 100 mM BisTrisPropane/HCl pH 6.5 was added to 10.2 mg protein in 5 mL 100 mM BisTrisPropane/HCl pH 6.5. The solution was

incubated at room temperature for 15 minutes with gentle stirring. Ferricyanide in 5 mL buffer. The solution was exchanged into 20 mM Tris/HCl pH 8.0 buffer by ultrafiltration. Anion exchange purification was carried out anaerobically using the same purification procedure as described above.

G. 45 min, EDTA, 20 mM Tris/HCl pH 8.0

15x molar excess of ferricyanide was used. The ferricyanide solution in 20 mM Tris/HCl pH 8.0 was added to 8 mg protein solution (also in 20 mM Tris/HCl pH 8.0) and 1.5x molar excess EDTA was added. The volume adjusted with buffer to 8 mL. The solution was incubated at room temperature for 45 minutes with gentle stirring.

H. 50 min, EDTA, 50 mM Sodium Acetate/Acetic Acid pH 4.5

15x molar excess of ferricyanide was used. 5 mg protein in 50 mM Sodium Acetate/Acetic Acid pH 4.5 was added the ferricyanide solution and 1.5x molar excess of EDTA. The volume was adjusted to approximately 5 mL with 50 mM Sodium Acetate/Acetic Acid pH 4.5. Incubated at room temperature for 50 minutes with stirring.

I. 3 hrs and overnight at 4 °C, EDTA, 20 mM BisTris/HCl pH 5.8

15x molar excess of ferricyanide was used. 4.5 mg protein (in 20 mM BisTris/HCl pH 5.8) was added the ferricyanide solution and 1.5x molar excess of EDTA. The volume was adjusted to 4.5 mL with 20 mM BisTris/HCl pH 5.8. The solution was incubated for 3 hours with stirring and left overnight at 4 °C with mild stirring.

J. Overnight, EDTA, 20 mM BisTris/HCl pH 5.8

15x molar excess of ferricyanide was used. The ferricyanide solution and 1.5x molar excess of EDTA were added to 11.7 mg protein in 20 mM BisTris/HCl pH 5.8. The volume was adjusted to 13 mL with 20 mM BisTris/HCl pH 5.8. The solution was incubated for overnight with stirring.

The experiment was repeated; 15x molar excess ferricyanide solution and 1.5x molar excess EDTA were added to 16.9 mg protein in 20 mM BisTris/HCl pH 5.8. The volume was adjusted to 16.9 mL with 20 mM BisTris/HCl pH 5.8 and the solution was incubated for overnight with stirring. Purified in two identical anion exchange runs at pH 8.0.

K. 3 hrs, EDTA, 20 mM BisTris/HCl pH 5.8, anion exchange pH 5.8

15x molar excess of ferricyanide was used. 4.3 mg protein (in 20 mM BisTris/HCl pH 5.8) was added the ferricyanide solution and 1.5x molar excess of EDTA. The volume was adjusted to 4.5 mL with 20 mM BisTris/HCl pH 5.8. The solution was incubated for 3 hours with stirring. The solution was exchanged into 20 mM BisTris/HCl pH 5.8 by ultrafiltration and subsequently loaded onto a 16 mm/10 cm Source 30Q column equilibrated with 20 mM BisTris/HCl pH 5.8. The column was washed with 2 CV 20 mM Tris/HCl pH 5.8 and elution was carried out using a linear salt gradient of 0 M to 0.3 M NaCl in 20 mM Tris/HCl pH 5.8 over 20.4 CV.

Additional anion exchange runs were performed on fractions from the initial purification and on D14C [Fe_4S_4] *Pf* Fd. Column and procedure were identical with the above mentioned purifications at either pH 5.8 or 8.0 in 20 mM Tris/HCl. Prior to additional runs, the fractions were washed into the respective anion exchange buffer by ultrafiltration.

L. 3 hrs, EDTA, BisTris/HCl pH 5.8, anion exchange pH 5.8

15x molar excess of ferricyanide was used. 8.2 mg protein (in 20 mM BisTris/HCl pH 5.8) was added the ferricyanide solution and 1.5x molar excess of EDTA. The volume was adjusted to 8.2 mL with 20 mM BisTris/HCl pH 5.8. The solution was incubated for 3 hours. The solution was exchanged into 20 mM BisTris/HCl pH 5.8 by ultrafiltration and subsequently loaded onto a 16 mm/10 cm Source 30Q column equilibrated with 20 mM BisTris/HCl pH 5.8. The column was washed with 2 CV 20 mM Tris/HCl pH 5.8 and elution was carried out using a linear salt gradient of 0.05 M to 0.28 M NaCl in 20 mM Tris/HCl pH 5.8 over 15.6 CV.

Additional anion exchange runs were performed on fractions from the initial purification and on a mix of fractions from the initial purification and purified D14C [Fe_4S_4] *Pf* Fd. The fractions were washed into 20 mM Tris/HCl pH 8.0 by ultrafiltration. Column and procedure were identical with the above mentioned purification at pH 8.0 in 20 mM Tris/HCl.

Also, an anaerobic purification was performed at pH 8.0 on fractions from the initial purification. The protein fractions were washed into 20 mM Tris/HCl pH 8.0 by ultrafiltration and deoxygenized under argon atmosphere prior to the anion exchange run.

M. Overnight, EDTA, BisTris/HCl pH 5.8, anion exchange pH 5.8

15x molar excess of ferricyanide was used. The ferricyanide solution and 1.5x molar excess of EDTA were added to 14.5 mg protein in 20 mM BisTris/HCl pH 5.8. The volume was adjusted to 14.5 mL with 20 mM BisTris/HCl pH 5.8. The solution was incubated overnight with stirring at room temperature. The solution was exchanged into 20 mM BisTris/HCl pH 5.8 by ultrafiltration and subsequently loaded onto a 16 mm/10 cm Source 30Q column equilibrated with 20 mM BisTris/HCl pH 5.8. The column was washed with 2 CV 20 mM Tris/HCl pH 5.8 and elution was carried out using a linear salt gradient of 0.05 M to 0.28 M NaCl in 20 mM Tris/HCl pH 5.8 over 15.6 CV.

N. 15 min, 100 mM Tris/HCl pH 7.8, anion exchange pH 7.8

15x molar excess of ferricyanide was used. The ferricyanide solution was added to 14.5 mg protein in 100 mM Tris/HCl pH 7.8. The volume was adjusted to 14.5 mL with 100 mM Tris/HCl pH 7.8. The solution was incubated at room temperature for 15 minutes with gentle stirring and then exchanged into 20 mM Tris/HCl pH 7.8 by ultrafiltration. The protein solution was loaded onto a 16 mm/10 cm Source 30Q column equilibrated with 20 mM Tris/HCl pH 7.8. The column was washed with 2 CV 20 mM Tris/HCl pH 7.8 and elution was carried out using a linear salt gradient of 0.15 M to 0.4 M NaCl in 20 mM Tris/HCl pH 7.8 over 17 CV.

O. Attempt to insert Co

15x molar excess of ferricyaide was used. The ferricyanide solution in 20 mM Tris/HCl pH 8.0 was added to 10 mg protein solution (also in 20 mM Tris/HCl pH 8.0) and 10x molar excess $CoCl_2$ was added. The volume adjusted with buffer to 10 mL. The solution was incubated at room temperature for 45 minutes with gentle stirring.

P. Attempts with dithionite, conversion suggested by Moura *et al.* [A1]

7.5 mg protein (in 1125 µL 20 mM Tris/HCl pH 8.0) was added 750 µL 0.8 M Tris/HCl pH 7.6 and then added 100x molar excess of dithionite. The volume was adjusted to 7.5 mL with 0.8 M Tris/HCl pH 7.6.

7.5 mg protein (in 1125 µL 20 mM Tris/HCl pH 8.0) was added 750 µL 0.8 M NaCl in 0.1 M Tris/HCl pH 7.6 and then added 100x molar excess of dithionite. The volume was adjusted to 7.5 mL with 0.1 M Tris/HCl + 0.8 M NaCl pH 7.6.

Both solutions were incubated 2.5 hours and EDTA in 1.5x molar excess was added prior to ultrafiltration.

Q. Attempt with $[Fe(phen)_3]^{3+}$

15x molar excess of $[Fe(phen)_3]^{3+}$ was used. 3 mL $[Fe(phen)_3]^{3+}$ solution in water was added to 5 mg protein (in 2 mL 20 mM Tris/HCl pH 8.0). The solution was incubated for 20 minutes.

$[Fe(phen)_3]^{2+}$ was prepared by mixing $Fe(NH_4)_2(SO_4)_2$) with 1,10-phenanthroline monohydrate in aqueous solution. The product was oxidized to $[Fe(phen)_3]^{3+}$ by addition of $(NH_4)_2Ce(NO_3)_6$.$NaClO_4$ was added and the solution was cooled on ice, filtered and dried to form solid $Fe(phen)_3(ClO_4)_3$ [A2,3].

Appendix C

C FERREDOXIN RESULTS

General denotations for all figures in this Appendix: chromatogram show elution followed at 280, 390 and 408 nm (blue, red and pink, respectively) and collected fractions are marked with a black line, whereas vertical lines on UV-vis spectra mark 390 nm and/or 408 nm.

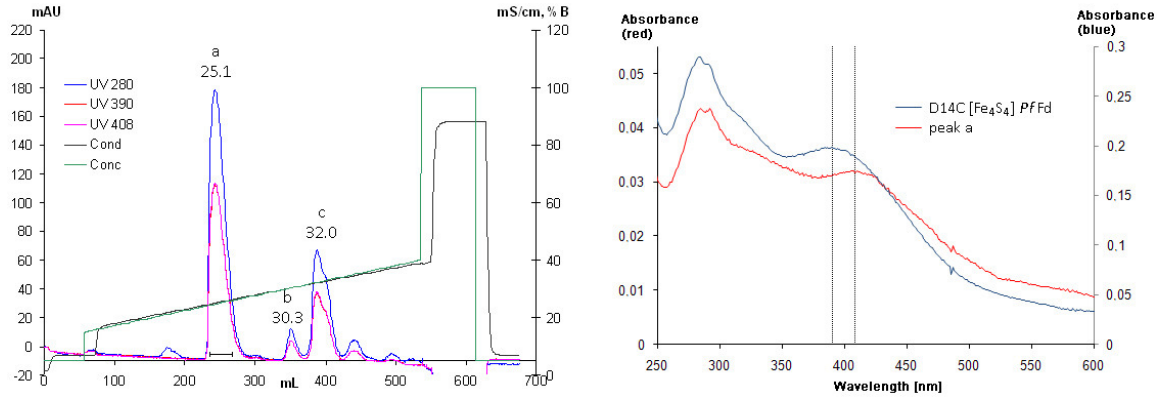


Figure C.1: Oxidation of D14C *Pf* Fd; chromatogram and UV-vis spectra from experiment B.

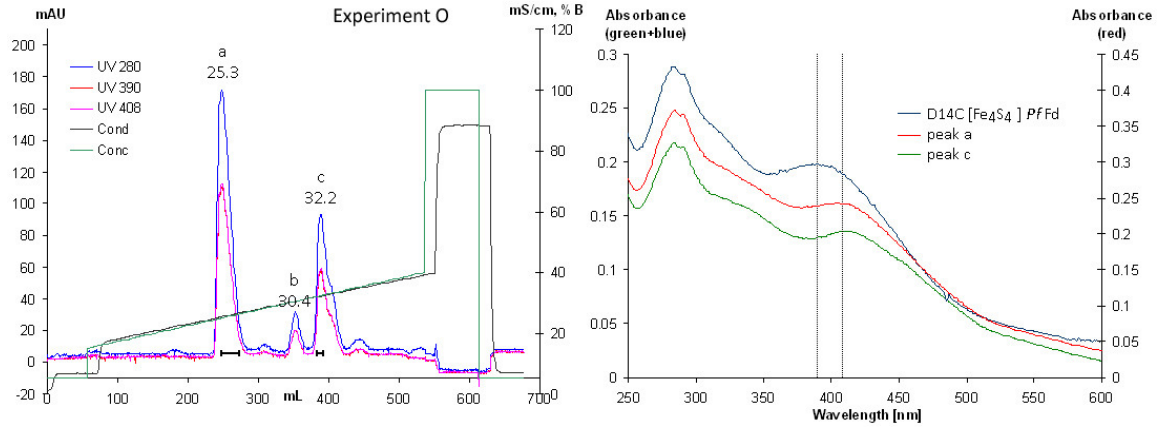


Figure C.2: Oxidation of D14C *Pf* Fd; chromatogram and UV-vis spectrum from experiment O.

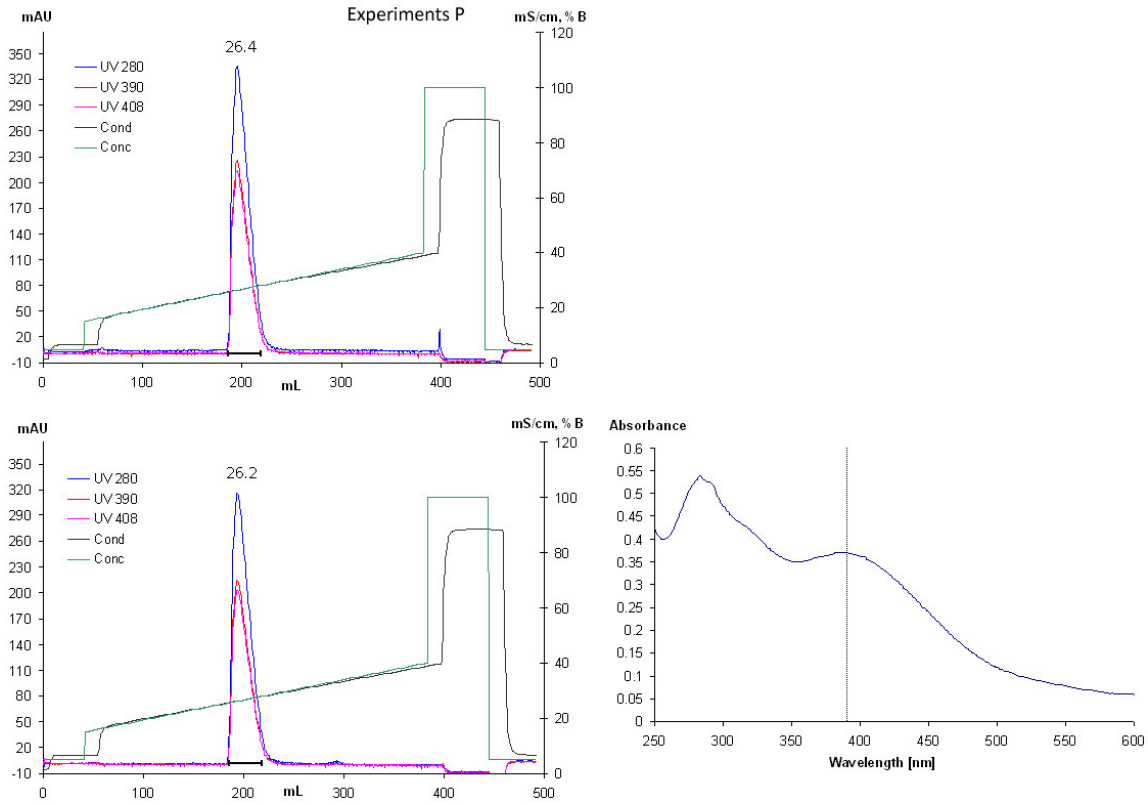


Figure C.3: Treatment of D14C *Pf* Fd; chromatograms and UV-vis spectrum from experiments P. Top: treated in 0.8 M Tris/HCl pH 7.6. Bottom: treated in 0.8 M NaCl in 0.1 M Tris/HCl pH 7.6.

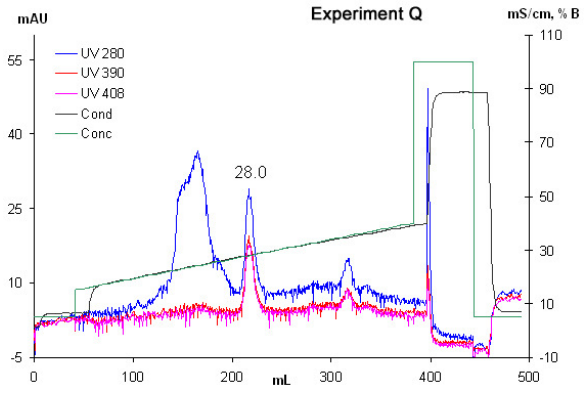


Figure C.4: Oxidation of D14C *Pf* Fd; chromatogram from experiment Q.

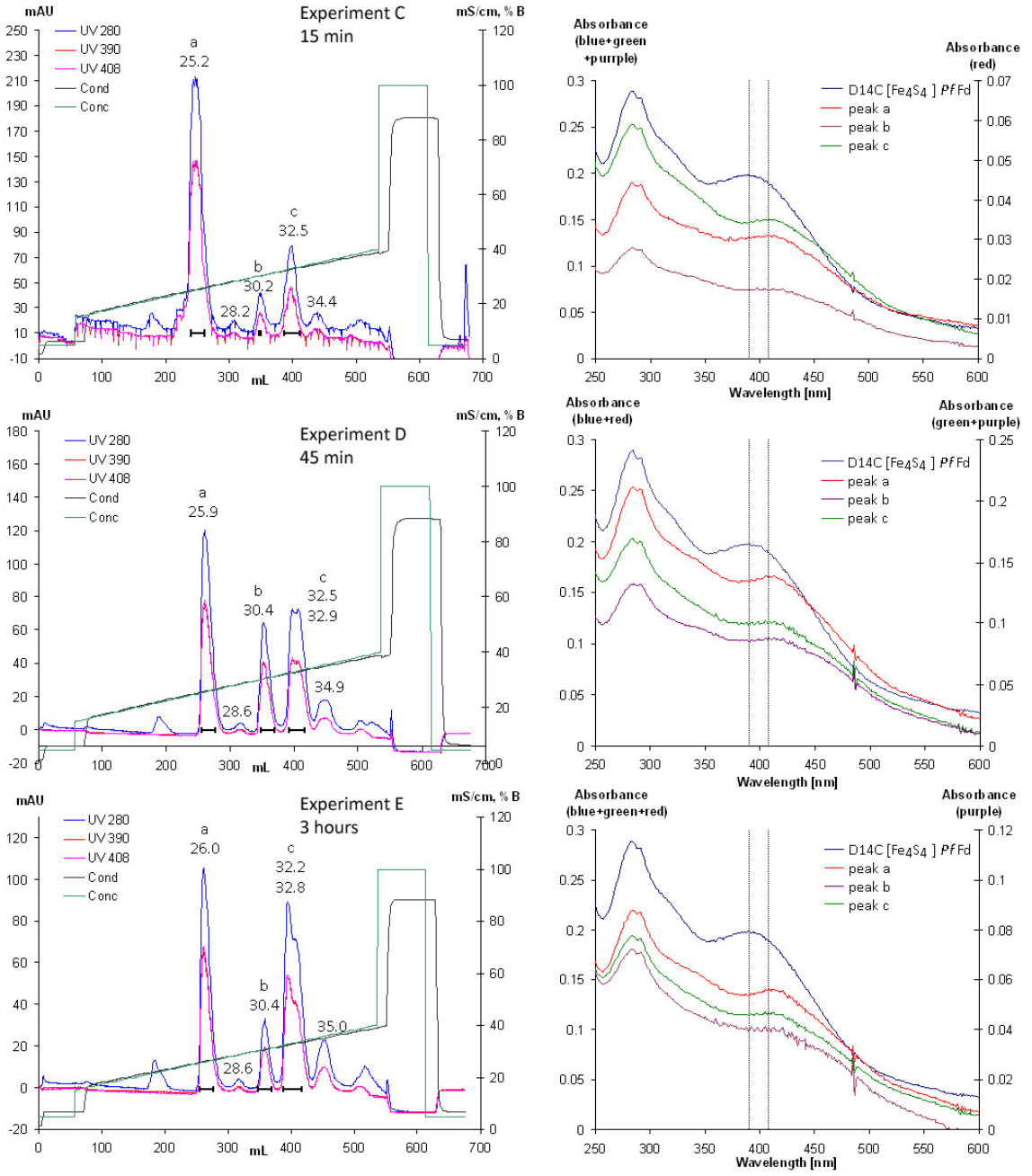


Figure C.5: Oxidation of D14C *Pf* Fd; chromatograms from experiments C, D and E. Incubation times are given and the corresponding UV-vis spectra are given to the right of the chromatograms.

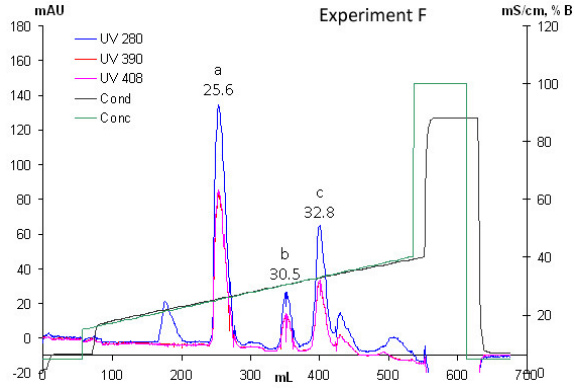


Figure C.6: Oxidation of D14C *Pf* Fd; chromatogram from experiment F.

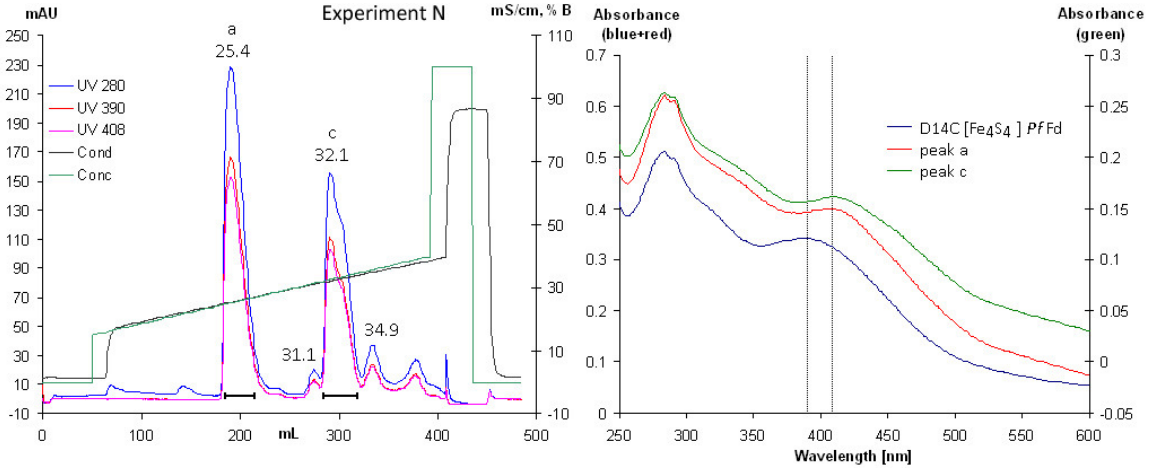


Figure C.7: Oxidation of D14C *Pf* Fd; chromatogram and UV-vis spectrum from experiment N.

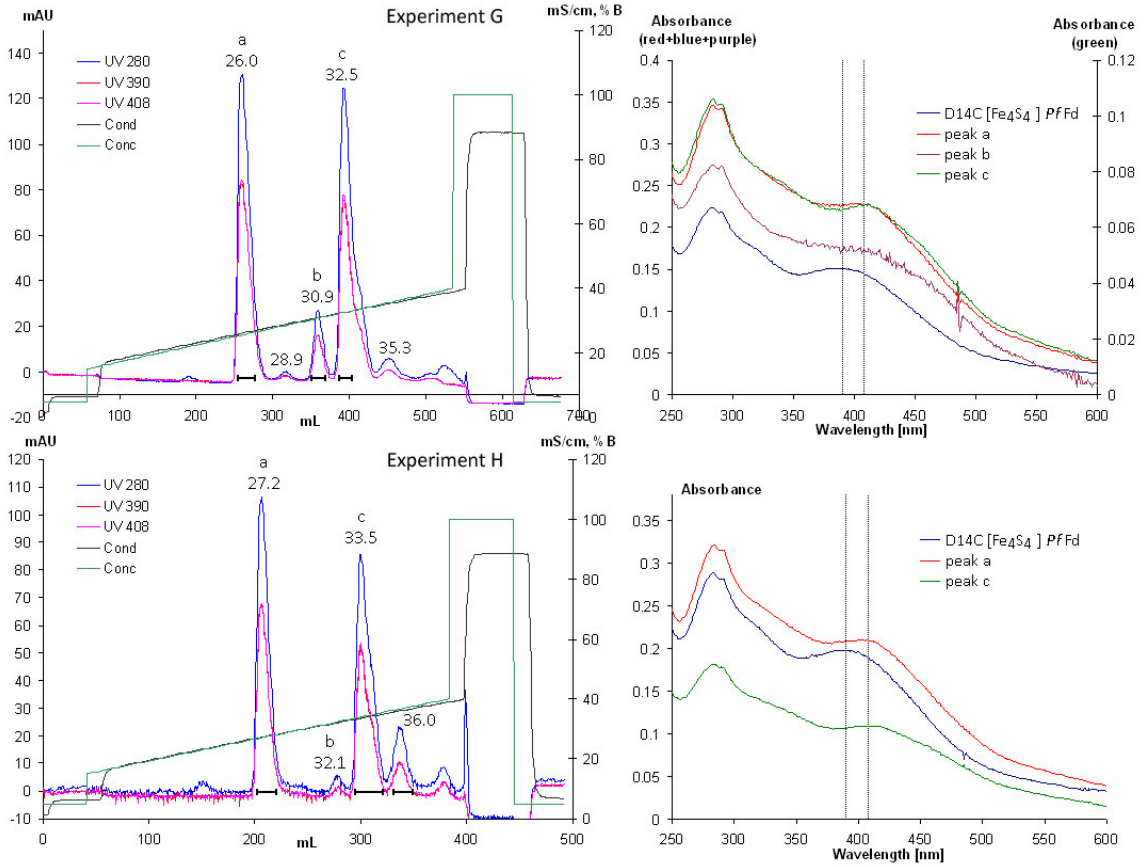


Figure C.8: Oxidation of D14C *Pf* Fd; chromatogram from experiment G and H.

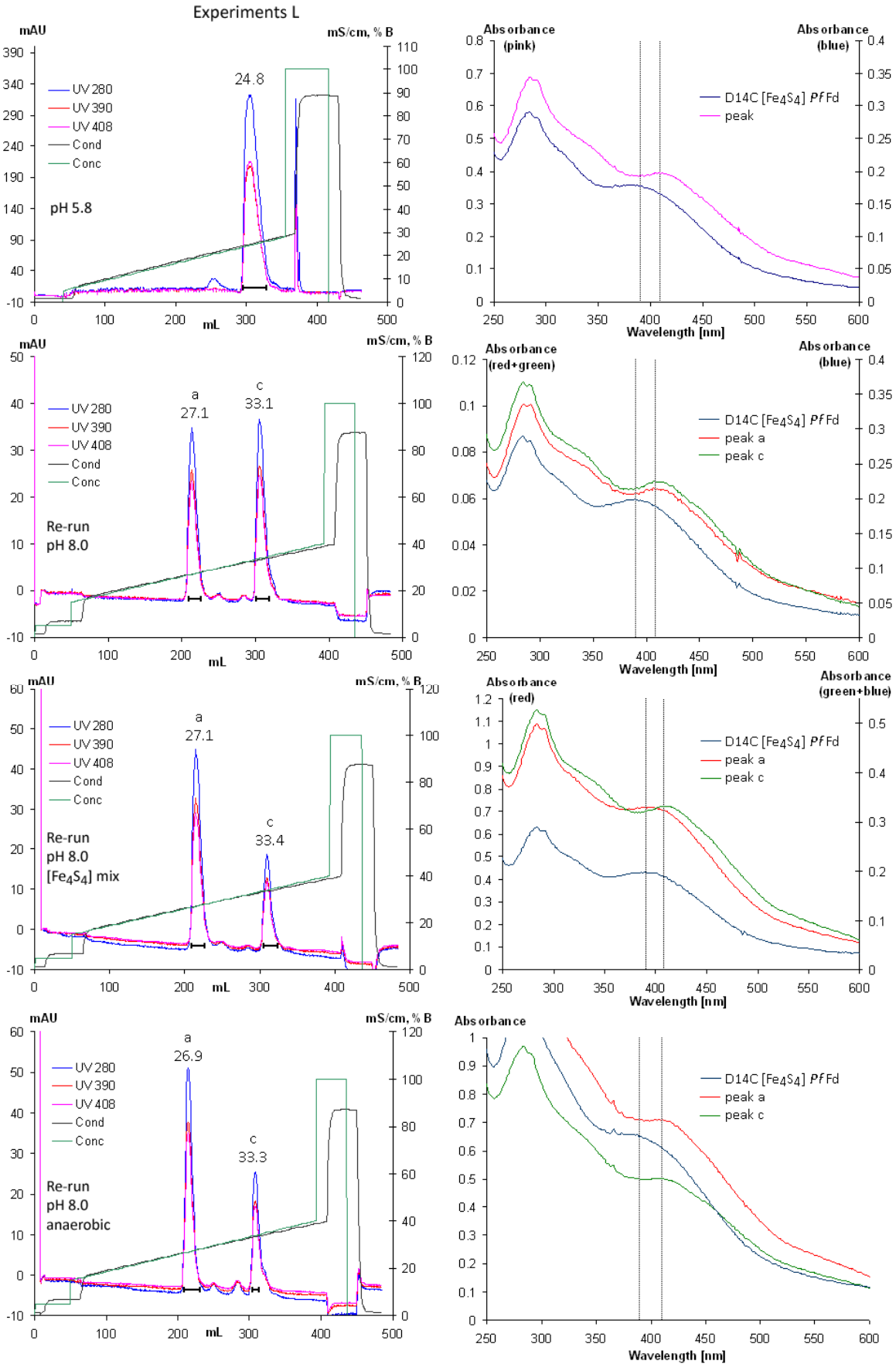


Figure legend on the next page

Figure C.9: Oxidation of D14C *Pf* Fd; chromatograms and UV-vis spectra from experiments L. Top: chromatogram, purification at pH 5.8. Second to the top: aliquot of collected fractions from the top chromatogram re-run at pH 8.0. Second to the bottom: aliquot of collected fraction from top chromatogram mixed with D14C [Fe_4S_4] *Pf* Fd and re-run at pH 8.0. Bottom: aliquot of collected fractions from top chromatogram re-run at pH 8.0 anaerobically.

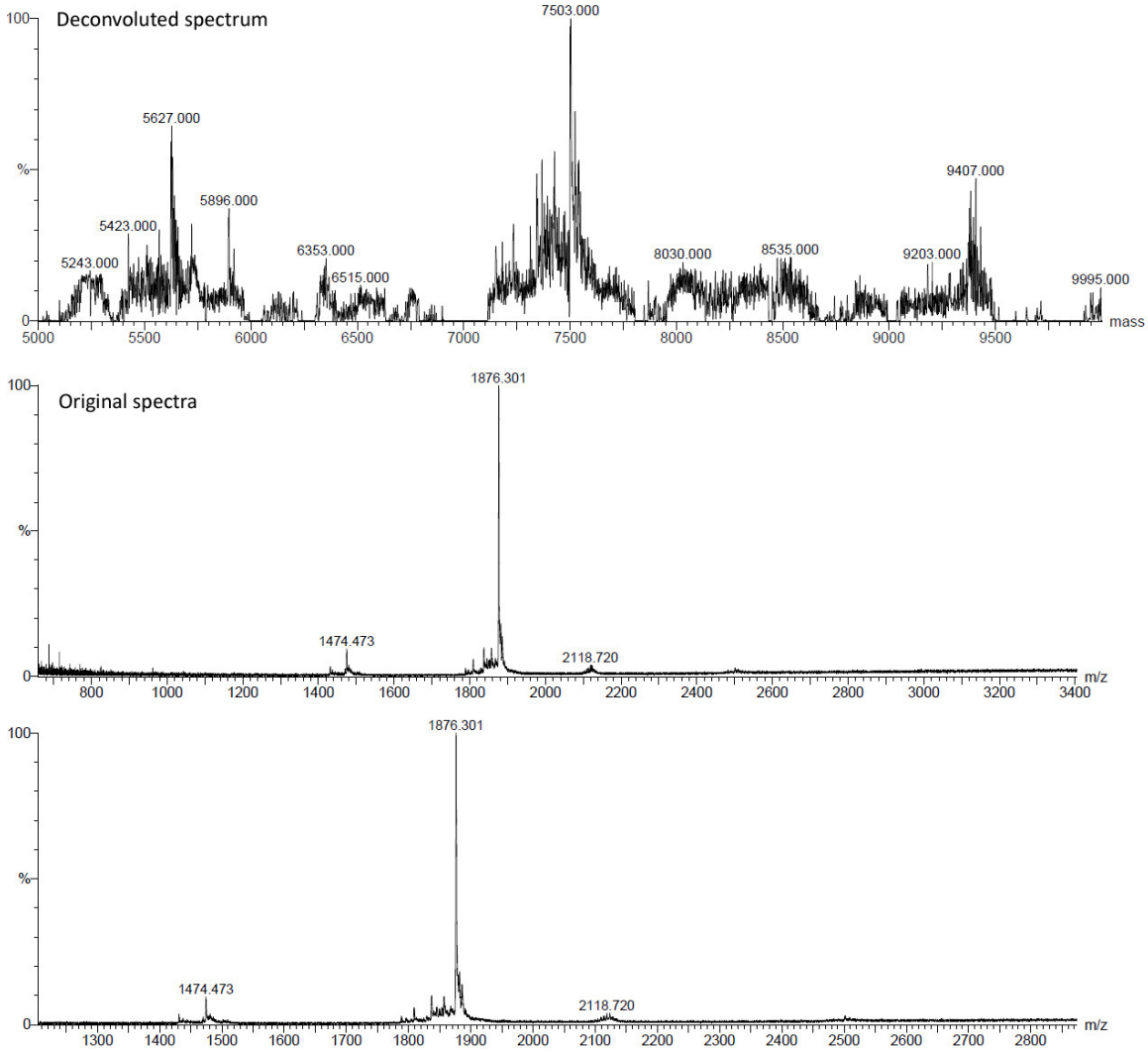


Figure C.10: Mass spectrometric analysis by Alphalyse A/S. Deconvoluted spectrum and original spectra of protein in *peak a* from experiment G.

Appendix D

D dCTP deaminase-dUTPase Experimental Procedures

Centrifugations were at 4 °C in a Sorvall R5C centrifuge with a SS-34 rotor.

D.1 Purification of WT enzyme

Cell pellet from 250 mL culture was defrosted on ice. 20 mL 50 mM KH_2PO_4/K_2HPO_4 pH 6.8 was added and the cells resuspended. The cells were lysed by sonication 8 times for 1 minute while kept on ice. The suspension was kept on ice for 30 seconds in between sonications. The suspension was centrifuged for 20 minutes at 10,000 rpm. The supernatant was decanted off and kept on ice. While stirred on ice, 10 (w/w)% streptomycin was slowly added to the solution to a final concentration of 1 %. The solution was left for 30 minutes with gentle stirring on ice. The solution was centrifuged for 20 minutes at 10,000 rpm. The supernatant was decanted off and dialyzed against 500 mL 50 mM KH_2PO_4/K_2HPO_4 pH 6.8 at 4 °C overnight.

The protein solution was loaded onto a 14 mm/30 cm DE52 anion exchange column equilibrated with 50 mM KH_2PO_4/K_2HPO_4 pH 6.8. The column was washed with 1 CV 50 mM KH_2PO_4/K_2HPO_4 pH 6.8 and eluted using a linear salt gradient of 50 mM KH_2PO_4/K_2HPO_4 pH 6.8 to 0.4 M NaCl in 50 mM KH_2PO_4/K_2HPO_4 pH 6.8 over 5 CV. The absorbance at 280 nm and SDS-PAGE were used to determine the protein content and purity of the collected fractions. The selected fractions were pooled, diluted two times and loaded onto the re-equilibrated column. The anion exchange step was then repeated. The selected fractions were pooled and solid $(NH_4)_2SO_4$ was added to 45 % saturation while stirring the solution on ice. The solution was stirred for 5 minutes and then centrifuged for 20 minutes at 14,000 rpm. The supernatant was discarded and the pellet was dissolved in 2.5 mL 20 mM HEPES/NaOH pH 8.0. The solution was dialyzed twice against 500 mL 20 mM HEPES/NaOH pH 8.0. The protein content was measured using the absorbance at 280 nm and the protein stored at -20 °C in 150 µL aliquots.

D.2 Purification of the A115V variant

The A115V variant of *Mt* DCD-DUT was purified using the same techniques as used for the WT enzyme described in section D.1. The most significant change was an additional ammonium sulfate precipitation prior to the anion exchange.

Cell pellet from 250 mL culture was defrosted on ice. 20 mL 50 mM KH_2PO_4/K_2HPO_4 pH 6.8 was added and the cells resuspended. A 15 µL sample (V1) of the resuspension was drawn for SDS-PAGE. The cells were lysed by sonication 12 times for 30 seconds while kept on ice. The suspension was kept on ice for 30 seconds in between sonications. The suspension was centrifuged for 20 minutes at 10,000 rpm. The supernatant was decanted off and kept on ice. A 15 µL sample of the supernatant (V2) and a sample of the pellet (V3) was drawn for SDS-PAGE. While stirred on ice, 10 (w/w)% streptomycin was slowly added to the solution to a final concentration of 1 %. The solution was left for 30 minutes with gentle stirring on ice. The solution was centrifuged for 20 minutes at 14,000 rpm. The supernatant was decanted off and kept on ice. A 15 µL sample of the supernatant (V4) and a sample of the pellet (V5) were drawn for SDS-PAGE. The solution was gently stirred on ice and solid $(NH_4)_2SO_4$ was slowly added to 40 % saturation. The solution was stirred for 5 minutes on ice and then centrifuged for 20 minutes at 14,000 rpm. The supernatant was discarded and the pellet was dissolved in 10 mL 50 mM KH_2PO_4/K_2HPO_4 pH 6.8. 15 µL samples of the supernatant (V6) and of the dissolved pellet (V7) were drawn for SDS-PAGE. Solid $(NH_4)_2SO_4$ was slowly added to 40 % saturation and the solution was gently stirred on ice for 5 minutes. The solution was stirred for 5 minutes and then centrifuged for 20 minutes at 14,000 rpm. The supernatant was discarded and the pellet was dissolved in 10 mL 50 mM KH_2PO_4/K_2HPO_4 pH 6.8. 15 µL samples of the supernatant (V8) and of the dissolved pellet (V9) were drawn for SDS-PAGE. The protein solution was loaded onto a 14 mm/30 cm DE52 anion exchange column equilibrated with 50 mM KH_2PO_4/K_2HPO_4 pH 6.8. The column was washed with 2 CV 50 mM KH_2PO_4/K_2HPO_4 pH 6.8 and eluted using a linear salt gradient of 50 mM KH_2PO_4/K_2HPO_4 pH 6.8 to 0.4 M NaCl in 50 mM KH_2PO_4/K_2HPO_4 pH 6.8 over 4 CV. The absorbance at 280 nm and SDS-PAGE were used to determine the protein content and purity of the collected fractions. The selected fractions were pooled and solid $(NH_4)_2SO_4$ was added to 40 % saturation while stirring the solution on ice. The solution was stirred for 5 minutes and then centrifuged for 20 minutes at 14,000 rpm. The supernatant was discarded and the pellet was dissolved in 5 mL 50 mM KH_2PO_4/K_2HPO_4 pH 6.8. The solution was dialyzed twice against 500 mL 50 mM KH_2PO_4/K_2HPO_4 pH 6.8 and centrifuged for 20 minutes at 14,000 rpm. The supernatant was decanted off and stored at -20 °C in 200 µL aliquots.

At this point, the solubility experiments described in section 8.2.4 were carried out. Based on the results (section 8.3.5), the A115V protein aliquots were defrosted on ice and pooled. Solid $(NH_4)_2SO_4$ was slowly added to 40 % saturation. The solution was stirred for 5 minutes on ice and then centrifuged for 20 minutes at 14,000 rpm. The supernatant was discarded and the pellet dissolved in 20 mM HEPES/NaOH pH 8.0. The protein solution was dialyzed against 500 mL 20 mM HEPES/NaOH pH 8.0 and frozen in aliquots of 150 µL at -20 °C.

D.3 Purification of the A115G variant

Purification of A115G *Mt* DCD-DUT was carried out similarly to the purification described in section D.4, albeit leaving out dialysis prior to ion exchange in order to save time.

Two tubes of cell pellet from a total of 250 mL culture were defrosted on ice. 10 mL 50 mM KH_2PO_4/K_2HPO_4 pH 6.8 was added to each tube and the cells resuspended. The cells were lysed

by sonication twelve times for 30 seconds while kept on ice. The suspension was allowed to cool down for 30 seconds in between sonications. The suspension was centrifuged for 20 minutes at 10,000 rpm. The supernatant was decanted off and kept on ice. A 50 µL sample of the supernatant (G1) was drawn for SDS-PAGE. While stirred on ice, 10 (w/w)% streptomycin was slowly added to the solution to a final concentration of 1 %. The solution was left for 30 minutes with gentle stirring on ice. The solution was centrifuged for 20 minutes at 14,000 rpm. The supernatant was decanted off and kept on ice.The solution was gently stirred on ice and solid $(NH_4)_2SO_4$ was slowly added to 40 % saturation. The solution was stirred for 5 minutes on ice and then centrifuged for 20 minutes at 14,000 rpm. The supernatant was discarded and the pellet was dissolved in 20 mL 50 mM KH_2PO_4/K_2HPO_4 pH 6.8. A sample of the dissolved pellet (G2) was drawn for SDS-PAGE. The protein solution was loaded onto a 14 mm/30 cm DE52 anion exchange column equilibrated with 50 mM KH_2PO_4/K_2HPO_4 pH 6.8 and eluted using a linear salt gradient of 50 mM KH_2PO_4/K_2HPO_4 pH 6.8 to 0.4 M NaCl in 50 mM KH_2PO_4/K_2HPO_4 pH 6.8 over 4 CV. The absorbance at 280 nm and SDS-PAGE were used to determine the protein content and purity of the collected fractions. The selected fractions were pooled and solid $(NH_4)_2SO_4$ was added to 40 % saturation while stirring the solution on ice. The solution was stirred for 5 minutes and then centrifuged for 20 minutes at 14,000 rpm. The supernatant was discarded and the pellet was dissolved in 2 mL 20 mM HEPES/NaOH pH 8.0. The solution was dialyzed against 500 mL 20 mM HEPES/NaOH pH 8.0 and a sample for SDS-PAGE was drawn. The protein solution was stored in 150 µL aliquots at -20 °C.

D.4 Purification 2 of the A115G variant

Cell pellet from 250 mL culture was defrosted on ice. 20 mL 50 mM KH_2PO_4/K_2HPO_4 pH 6.8 was added and the cells resuspended. The cells were lysed by sonication twelve times for 30 seconds while kept on ice. The suspension was allowed to cool down for 30 seconds in between sonications. The suspension was centrifuged for 20 minutes at 10,000 rpm. The supernatant was decanted off and kept on ice. While stirred on ice, 10 (w/w)% streptomycin was slowly added to the solution to a final concentration of 1 %. The solution was left for 30 minutes with gentle stirring on ice. The solution was centrifuged for 20 minutes at 14,000 rpm. The supernatant was decanted off and kept on ice. The solution was gently stirred on ice and solid $(NH_4)_2SO_4$ was slowly added to 40 % saturation. The solution was stirred for 5 minutes and then centrifuged for 20 minutes at 14,000 rpm. The supernatant was discarded and the pellet was dissolved in 10 mL 50 mM KH_2PO_4/K_2HPO_4 pH 6.8. The solution was dialyzed against 500 mL 50 mM KH_2PO_4/K_2HPO_4 pH 6.8. The solution was centrifuged 20 minutes at 14,000 rpm. The supernatant (S) was decanted off and the pellet (B) resuspended in 10 mL 50 mM KH_2PO_4/K_2HPO_4 pH 6.8. SDS-PAGE samples were drawn from both solutions.

A glass pipette was used to collect the clear protein solution when the pellet of solution B had settled. A sample for SDS-PAGE was drawn (B2). The clear solution was added 10 mL 50 mM KH_2PO_4/K_2HPO_4 pH 6.8. Solid $(NH_4)_2SO_4$ was slowly added to 40 % saturation and the solution was left for 5 minutes with gentle stirring on ice. The solution was centrifuged for 20 minutes at

14,000 rpm. The supernatant was discarded and the pellet was dissolved in 5 mL 50 mM KH_2PO_4/K_2HPO_4 pH 6.8 (B3).

The supernatant protein solution (S) was loaded onto a 14 mm/30 cm DE52 anion exchange column equilibrated with 50 mM KH_2PO_4/K_2HPO_4 pH 6.8 and eluted using a linear salt gradient of 50 mM KH_2PO_4/K_2HPO_4 pH 6.8 to 0.4 M NaCl in 50 mM KH_2PO_4/K_2HPO_4 pH 6.8 over 4 CV. The absorbance at 280 nm and SDS-PAGE were used to determine the protein content and purity of the collected fractions. The selected fractions were pooled and solid $(NH_4)_2SO_4$ was added to 40 % saturation while gently stirring the solution on ice. The solution was gently stirred for 5 minutes and then centrifuged for 20 minutes at 14,000 rpm. The supernatant was discarded and the pellet was dissolved in 10 mL 50 mM KH_2PO_4/K_2HPO_4 pH 6.8 (S2).

The two protein solutions (B3 and S2) were both dialyzed twice against 500 mL 50 mM KH_2PO_4/K_2HPO_4 pH 6.8. The solutions were both centrifuged for 20 min at 14,000 rpm. The supernatants were decanted off and added 10 % azid to a final concentration of 0.02 % (B4 and S3).

D.5 Purification 3 of the A115G variant

Purification 3 of A115G *Mt* DCD-DUT was carried out identically with the WT purification described in Appendix D.1, except the protein was frozen at -20 °C for approximately 3 weeks prior to the final ammonium sulfate precipitation.

D.6 SDS-PAGE gel casting

SDS-PAGE gels were cast with 12.5 % Acrylamide. A rack designed for gel casting was used and the running gel was cast first and allowed to polymerize for a few minutes before the stacking gel was cast on top of it.

SDS-PAGE running gel	
Water	3000 µL
1.5 M Tris/HCl pH 8.8	2400 µL
Acrylamid/Bis-acrylamid (30/0.8 w/w %)	4000 µL
10 % Sodium dodecyl sulfate	2400 µL
20 % Ammonium persulfate	5 µL
N,N,N',N'-tetramethylethylene diamine (TEMED)	96 µL

SDS-PAGE stacking gel	
Water	3440 µL
1 M Tris/HCl pH 6.8	625 µL
Acrylamid/Bisacrylamid (30/0.8 w/w %)	830 µL
10 % Sodium dodecyl sulfate	50 µL
20 % Ammonium persulfate	50 µL
N,N,N',N'-tetramethylethylene diamine (TEMED)	4 µL

D.1 Nucleotide Saturation

Experimental details for nucleotide saturation assays with WT, A115V and A115G *Mt* DCD-DUT carried out as described in section 9.2.2 are given in Table D.1.

Table D.1: Assay mix contents and corresponding in-assay nucleotide concentrations and enzyme dilution factors (in-assay) during nucleotide saturation experiments with WT, A115V and A115G Mt DCD-DUT. All given concentrations are in-assay initial concentrations. Labeled nucleotide was added in minor amounts; 3H dCTP to dCTP nucleotide mixtures and 3H dUTP to dUTP nucleotide mixtures.

Assay mix		Assay components		
$[MgCl_2]$	HEPES/NaOH	[dCTP]	[dUTP]	Total enzyme dilution
WT				$c_{initial}$ = 91 µM
≤ 2 mM	20 mM pH 8.0	9.4-3600 µM	-	100-400x
≤ 2 mM	20 mM pH 8.0	-	9.4-2400 µM	100-400x
≤ 2 mM	20 mM pH 6.8	-	19-2400 µM	100-200x
A115V				$c_{initial}$ = 191 µM
2-5.5 mM	20 mM pH 8.0	0.019-4.8 mM	-	30-60x
2-3 mM	20 mM pH 8.0	-	0.16-2400 µM	30-2000x
2-3 mM	20 mM pH 6.8	-	5-2400 µM	30-600x
A115G				$c_{initial}$ = 144 µM
≤ 2 mM	20 mM pH 8.0	19-3600 µM	-	10-40x
≤ 2 mM	20 mM pH 8.0		9.4-2400 µM	10-80x
≤ 2 mM	20 mM pH 6.8	-	9.4-2400 µM	10-20x

D.2 dTTP Inhibition

Experimental details for dTTP inhibition assays carried out with WT, A115V and A115G *Mt* DCD-DUT as described in section 9.2.3 are given in Table D.2.

D.3 Differential Scanning Calorimetry

Experimental details for DSC experiments with WT, A115V and A115G *Mt* DCD-DUT carried out as described in section 9.2.6 are given in Table D.3.

Table D.2: Assay mix contents and corresponding in-assay dTTP concentration range and enzyme dilution factors (in-assay) during dTTP inhibition experiments with WT, A115V and A115G *Mt* DCD-DUT. All given concentrations are in-assay initial concentrations. Labeled nucleotide was added in minor amounts; ^{3}H dCTP to dCTP nucleotide mixtures and ^{3}H dUTP to dUTP nucleotide mixtures.

Assay mix				Assay components	
[$MgCl_2$]	HEPES/NaOH	[dCTP]	[dUTP]	[dTTP]	Total enzyme dilution
WT					$c_{initial}$ = 91 µM
2 mM	20 mM pH 8.0	300 µM	-	0-1 mM	100x, 200x
2 mM	20 mM pH 6.8	300 µM	-	0-1 mM	100x, 200x
2 mM	20 mM pH 6.8	-	300 µM	0-1 mM	100x, 200x
A115V					$c_{initial}$ = 191 µM
2 mM	20 mM pH 8.0	300 µM	-	0-0.1 mM	30x, 60x
4 mM	20 mM pH 8.0	2.4 mM	-	0-1 mM	30x, 60x
2 mM	20 mM pH 6.8	-	300 µM	0-1 mM	30x, 60x
A115G					$c_{initial}$ = 144 µM
2 mM	20 mM pH 8.0	300 µM	-	0-0.1 mM	10x, 20x
2 mM	20 mM pH 8.0	800 µM	-	0-0.1 mM	10x, 20x
2 mM	20 mM pH 6.8	-	300 µM	0-0.1 mM	10x, 20x

Table D.3: Enzyme- and dTTP concentrations as well as pH for DSC experiments.

	20 mM HEPES/NaOH	[dTTP]	[$MgCl_2$]	Enzyme conc.
WT	pH 8.0	-	-	20 µM
	pH 6.8	0.5 mM	1 mM	20 µM
	pH 6.8	-	-	20 µM
A115V	pH 8.0	-	-	24 µM
	pH 8.0	0.5 mM	1 mM	24 µM
A115G	pH 8.0	-	-	12 µM
	pH 8.0	-	-	24 µM
	pH 8.0	0.5 mM	1 mM	12 µM
	pH 8.0	0.5 mM	1 mM	24 µM
	pH 6.8	-	-	24 µM

Appendix E

E dCTP deaminase-dUTPase Results

E.1 Purification of the A115G variant

An overview of the steps involved in the two purifications of A115G *Mt* DCD-DUT that resulted in protein degradation is shown in Figure E.1.

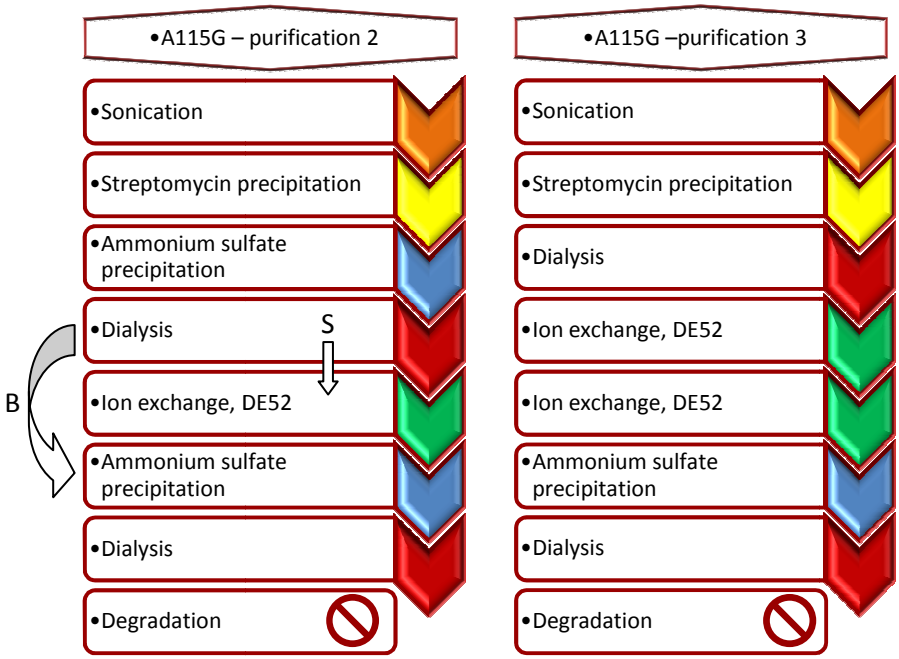


Figure E.1: Diagram of the steps involved in purifications of A115G *Mt* DCD-DUT that resulted in protein degradation. S and B refer to two fractions after dialysis.

Purification 2 of the A115G variant

During this purification of A115G *Mt* DCD-DUT, several problems with precipitation were encountered. After the first ammonium sulfate precipitation and subsequent centrifugation, the pellet did not completely dissolve in 10 mL 50 mM KH_2PO_4/K_2HPO_4 pH 6.8. The blurred solution was dialyzed and then centrifuged and split into the supernatant (S) and the pellet was resuspended in 10 mL buffer, forming a blurred solution with white precipitation (B). SDS-PAGE results of the two samples are shown in Figure E.2A. Both solutions proved to contain protein and sample B contained very pure protein.

The supernatant of sample B (B2) contained pure protein and was collected (SDS-PAGE lane B2, Figure E.2B) and the remaining precipitate was discarded. Supernatant S was purified by anion exchange and Figure E.2B shows SDS-PAGE and the absorbance at 280 nm (fractions 10-12 were collected and pooled). Proteins in both solutions (S and B2) were precipitated with ammonium sulfate. The pellet from B2 readily dissolved in 5 mL 50 mM KH_2PO_4/K_2HPO_4 pH 6.8 (B3), while the pellet from S did not completely dissolve in 10 mL 50 mM KH_2PO_4/K_2HPO_4 pH 6.8 (S2). Samples S2 and B3 were dialyzed twice. Both solutions precipitated and were centrifuged. The supernatants were collected and azid was added to a final concentration of 0.02 % (B4 and S3). SDS-PAGE results of the samples are shown in Figure E.2C. It is seen that B4 contained little protein and protein in S3 had degraded. Both solutions (B4 and S3) were discarded.

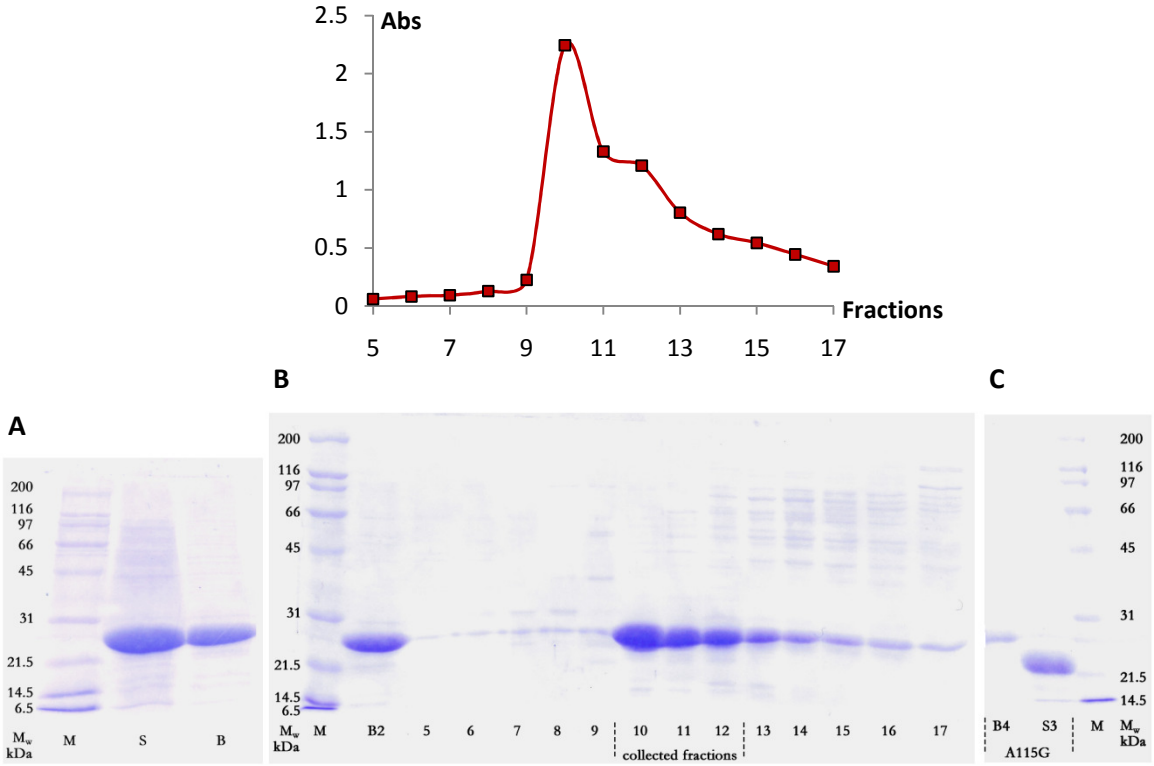


Figure E.2: Purification 2 of A115G *Mt* DCD-DUT. SDS-PAGE of A: S and B (two solutions after first dialysis step). B: solution B2 and fractions from anion exchange. The collected fractions are marked and the absorbance at 280 nm is shown above (the connecting line serves as guide to the eye). C: purified samples of the A115G variant protein. See Appendix D.4 for further details on samples.

Purification 3 of the A115G variant

Purification 3 of A115G *Mt* DCD-DUT was carried out identically with the WT purification (for details, see Appendix D.5). SDS-PAGE results of fractions from the second anion exchange are shown in Figure E.3A. The collected fractions are marked and were pooled for further purification. SDS-PAGE of the final product from purification is shown in Figure E.3B. Degradation is observed and the protein was discarded.

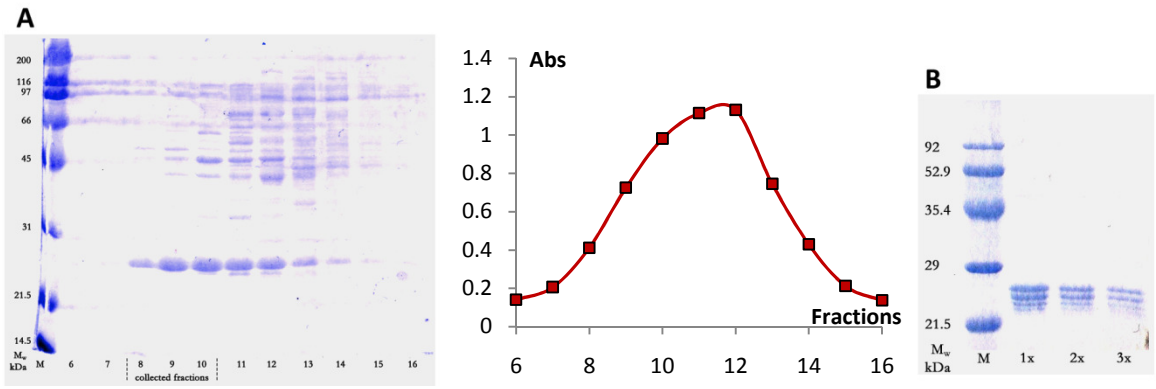


Figure E.3: Purification 3 of A115G *Mt* DCD-DUT. SDS-PAGE of A: fractions from anion exchange. The collected fractions are marked and the absorbance at 280 nm is shown to the right (the connecting line serves as guide to the eye). B: purified protein (concentrated protein (1x), diluted two times (2x), diluted three times (3x)). See Appendix D.5 for purification details.

E.2 Turnover numbers

Experiments used to calculate turnover numbers (with 1 mM dCTP or dUTP at pH 6.8 or pH 8.0) were carried out as described in section 9.2.1 and results for WT, A115V and A115G *Mt* DCD-DUT are given in Figure E.4.

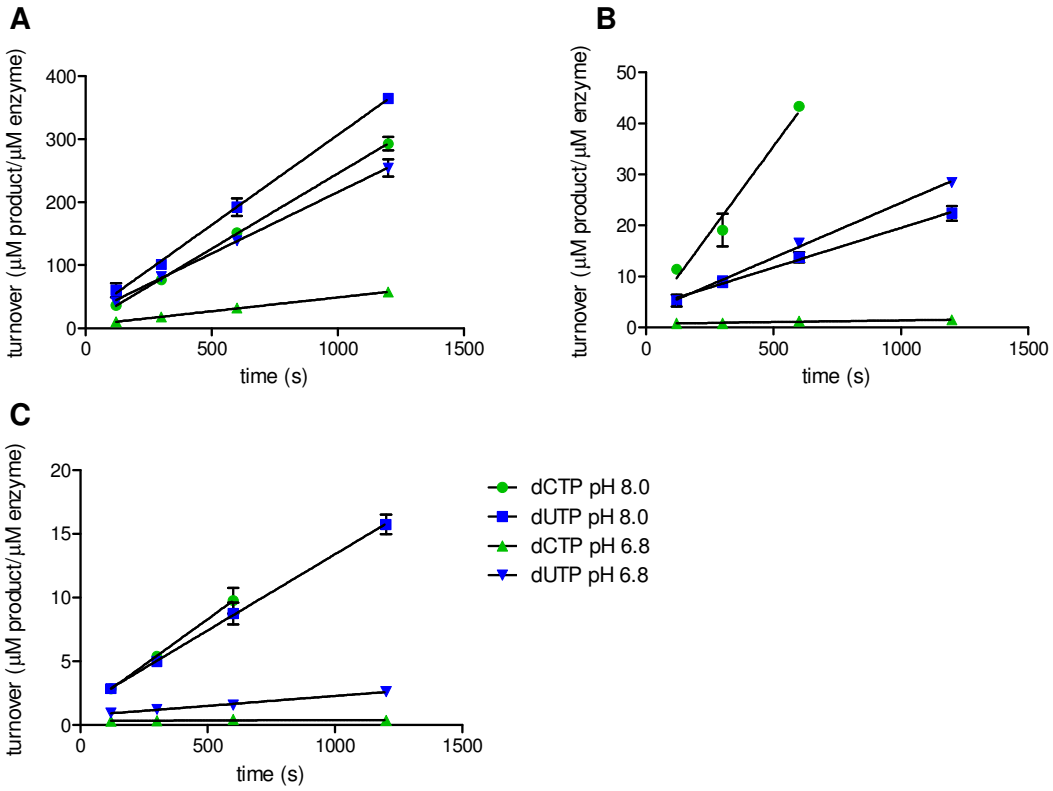


Figure E.4: Measured reaction rates with 1 mM dCTP or dUTP at pH 6.8 or pH 8.0 plotted against the time for A: WT, B: A115V and C: A115G *Mt* DCD-DUT. Each point was determined twice at different enzyme concentrations.

E.3 Nucleotide saturation

Results for dUTP saturation kinetic studies carried out as described in section 9.2.2 are given in Figure E.5.

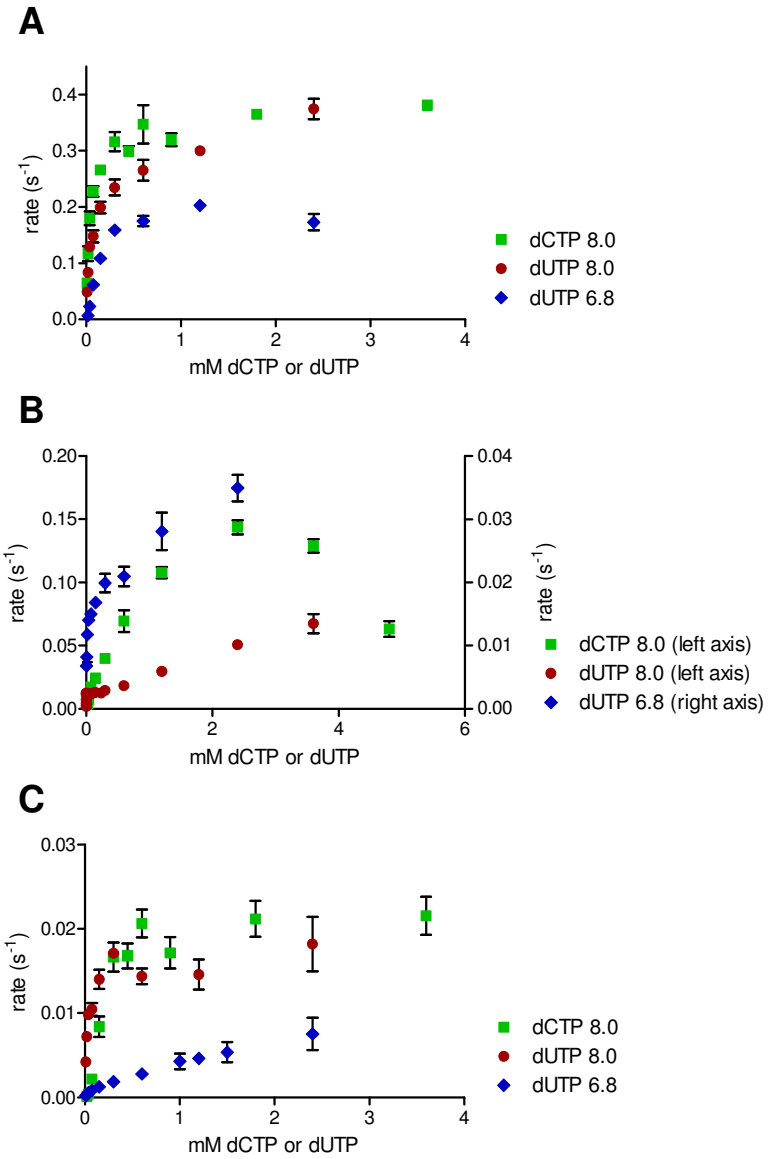


Figure E.5: Data points for nucleotide saturation experiments. A: WT, B: A115V, C: A115G *Mt* DCD-DUT. Each point has been determined for at least twice at different enzyme concentrations.

E.4 Differential Scanning Calorimetry

Results for DSC carried out as described in section 3.2.8 are given in Figure E.6, Figure E.7 and Figure E.8 for WT, A115V and A115G Mt DCD-DUT, respectively. Weak signals were observed for DSC measurements.

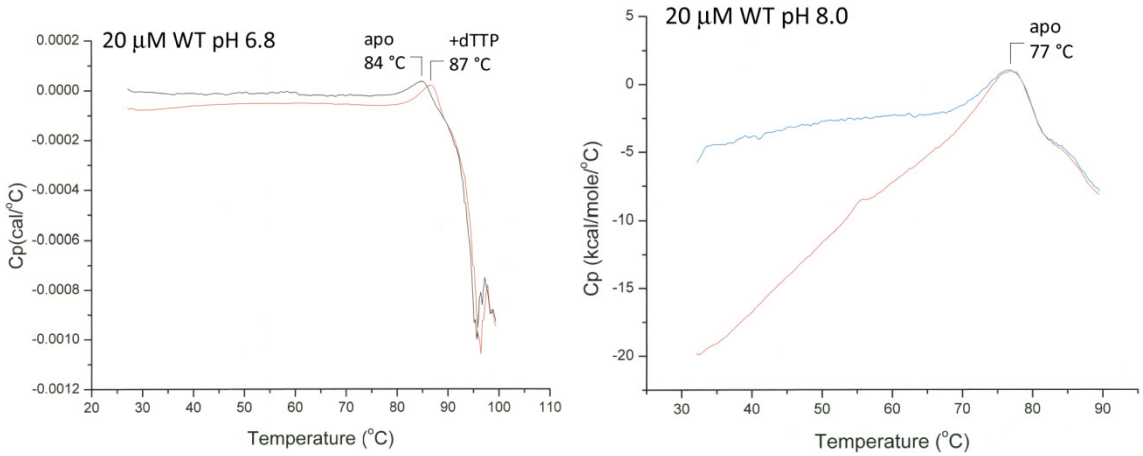


Figure E.6: DSC of WT *Mt* DCD-DUT at pH 6.8 and pH 8.0 in the presence or absence of dTTP. Two experiments were made with the apo enzyme. Irreversible transition temperatures are given.

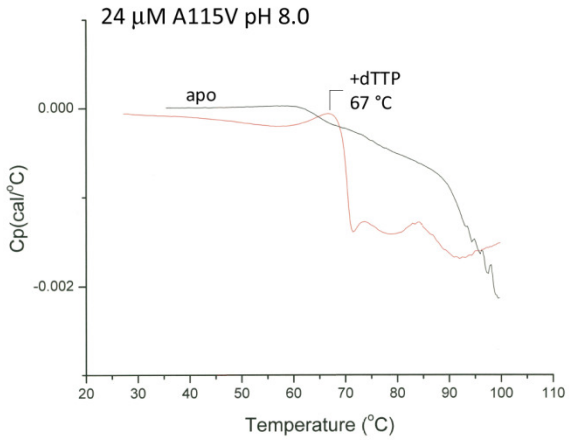


Figure E.7: DSC of A115V *Mt* DCD-DUT at pH 8.0. Irreversible transition temperature for the enzyme in the presence of dTTP is given. The experiment failed for the apo enzyme.

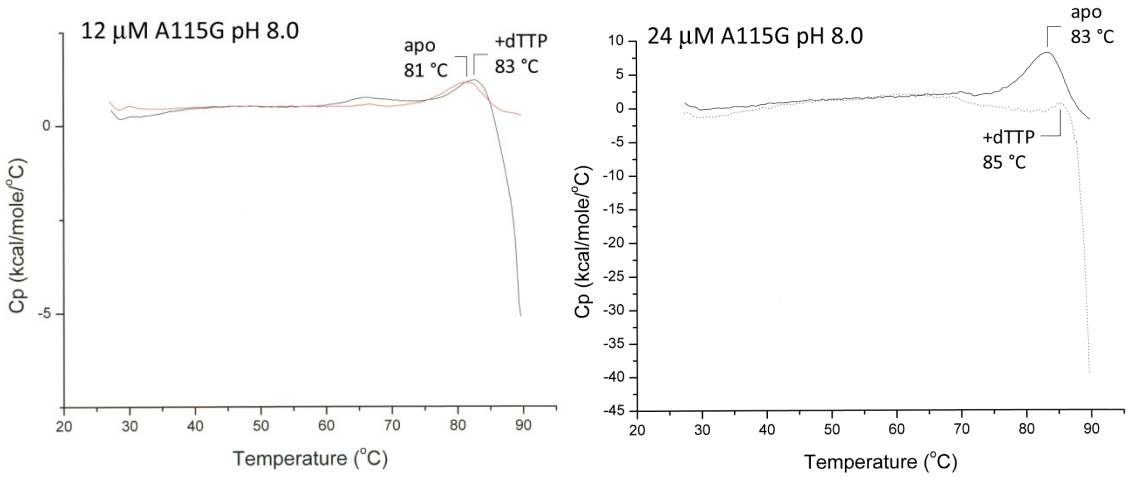


Figure E.8: DSC of A115G *Mt* DCD-DUT at pH 8.0 at two different enzyme concentrations in the presence or absence of dTTP. Irreversible transition temperatures are given.

E.5 Crystal Structure of A115V *Mt* DCD-DUT:dTTP

Pictures related to the crystal structure of A115V *Mt* DCD-DUT in complex with dTTP are given here.

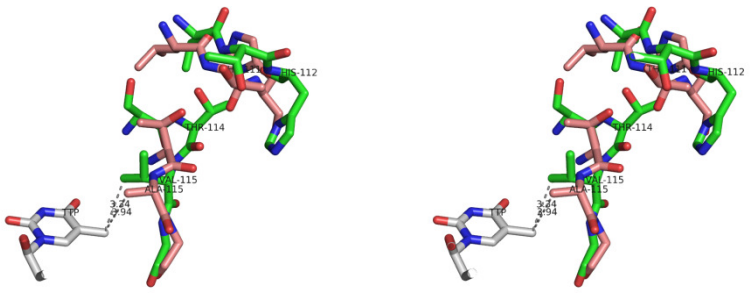


Figure E.9: Stereo view of Figure 9.14**:** Overlay of A115V:dTTP (green) and WT:dTTP (red, pdb 2QXX). A shift in the backbone causes the valine to take up less space in the nucleotide binding pocket and allows room for dTTP binding. A rearrangement of the backbone at Thr111 and His112 prevents the side chain of Thr114 from clashing with Thr111 of a neighboring monomer.

Appendix F

F Crystallization of Extended *Sulfolobus solfataricus* Nucleoside Diphosphate Kinase

F.1 Introduction

Nucleoside diphosphate kinase (NDK) catalyzes dephosphorylation of nucleoside triphosphates and subsequent phosphorylation of nucleoside diphosphates (see Figure 6.1, page 38). The γ-phosphate of a nucleoside triphosphate (often ATP) is transferred to an active site histidine and a nucleoside diphosphate is released (often ADP). A different nucleoside diphosphate binds to the phosphorylated enzyme and the reverse reaction produces nucleoside triphosphate and free enzyme.

Many crystal structures of NDK from different organisms, ranging from bacteria to human, are available. The subunit fold of NDK is the same and sequence identity of more than 40 % is observed between NDK from different species [A4]. Most quaternary structures of NDK are hexameric, but some bacterial NDK are tetrameric, and a halophilic NDK has been identified as a dimer [A4,5]. The same dimer is used in the assembly of hexameric and tetrameric structures. The subunit fold of NDK is a βαβ βαβ fold, also known as the ferredoxin fold because it was first seen in the ferredoxin from *Pseudomonas aerogenes*. The fold comprises a center of a four-stranded antiparallel β-sheet and two connecting α-helices. The two β-strands in the center are very hydrophobic and the least variable part of the sequence [A4]. Assembly into hexamers is likely to play an important role in the stability of NDK because of the increased stability of hexamers compared with isolated subunits [A6].

In general, an extended C-terminal is observed for NDK from mesophilic organisms compared with NDK from thermophilic organisms, see Figure F.1. The C-terminal could thus have a destabilizing effect on the enzyme. Contrary to this is the finding that increasing the length of the C-terminal of NDK from the hexameric thermoacidophilic archaea *Sulfolobus solfataricus* (*Ss*) has a stabilizing effect by increasing its transition temperature by 2-3 °C[7]. The C-terminal extension consists of 12 residues from NDK from the hexameric mesophilic eukaryote *Dictyostelium discoideum*, see Figure F.1. The crystal structure of WT *SS* NDK was recently

[7] Unpublished results. Martin Willemoës, Department of Biology, University of Copenhagen.

solved[8], see Figure F.2. The crystal structure of the extended NDK (NDKext) from *Ss* could explain the increased thermostability. This appendix describes NDKext crystallization and attempts to solve the crystal structure.

```
                                10        20        30        40        50        60
                      ....|....|....|....|....|....|....|....|....|....|....|....|
Chloroflexus aurantiacus      -------MERALLILKPDAVQRGLIGAIISRFEQRGLKFQGLKLMQVDEALARRHYAEHE
Thermus aquaticus             -------MERTFVMVKPDGFRRGLVGEILARFERKGFRIVGLKALRISQELAEKHYAEHR
Thermus thermophilus          -------MERTFVMIKPDGVRRGLVGEILARFERKGFRIAALKLMQISQELAERHYAEHR
Oceanithermus profundus       -------MERTFAMIKPDGVRRGLTGKIIQRLEDKGFKIVALKKMRISFDLAEEHYGEHK
Meiothermus ruber             -------MERTYIMVKPDGVRRGLTGEIINRIERKGFKIVAMKKMVIPRETAETHYGEHR
Meiothermus silvanus          -------MERTYIMVKPDGVRRGLTGEIISRIERKGFKIVAMKKMLISQQTAETHYGEHK
Sulfolobus solfataricus       -----MVAQRTFVMIKPDGVKRGLIGEIISRFEKRGLKIVSLKMVKMSRDTAEKLYEEHK
Dictyostelium discoideum      MSTNKVNKERTFLAVKPDGVARGLVGEIIARYEKKGFVLVGLKQLVPTKDLAESHYAEHK
Saccharomyces cerevisiae      MSS---QTERTFIAVKPDGVQRGLVSQILSRFEKKGYKLVAIKLVKADDKLLEQHYAEHV
Drosophila melanogaster       MAA---NKERTFIMVKPDGVQRGLVGKIIERFEQKGFKLVALKFTWASKELLEKHYADLS
Homo sapiens                  -MA---NLERTFIAIKPDGVQRGLVGEIIKRFEQKGFRLVAMKFLRASEEHLKQHYIDLK
Escherichia coli              -----MAIERTFSIIKPNAVAKNVIGNIFARFEAAGFKIVGTKMLHLTVEQARGFYAEHD
Myxococcus xanthus            -----MAIERTLSIIKPDGLEKGVIGKIISRFEEKGLKPVAIRLQHLSQAQAEGFYAVHK
Arabidopsis thaliana          -------MEQTFIMIKPDGVQRGLIGEVICRFEKKGFTLKGLKLISVERSFAEKHYEDLS

                                70        80        90       100       110       120
                      ....|....|....|....|....|....|....|....|....|....|....|....|
Chloroflexus aurantiacus      GKSFFNGLVSYITSAPVVVAVVAGKPGTVELVRAMVGATNPAKAAPGTIRGDFGVDIGR-
Thermus aquaticus             EKPFFPSLVGFITSGPVVAMVLEGPN-AVAEVRKMMGATHPKDALPGTIRGDYATTIDE-
Thermus thermophilus          EKPFFPGLVRFITSGPVVAMVLEGPG-VVAEVRKMMGATHPKDALPGTIRGDFATTIDE-
Oceanithermus profundus       EKPFFKPLVEFITSGPVVAMVLEGPG-VIAELRKMMGATNPADALPGTIRGDFATTIDE-
Meiothermus ruber             GKPFFEGLVNFITSGPVVAMVVEGPG-VIAEMRRLMGATRPWEAAPGTIRADFATTVDE-
Meiothermus silvanus          GKPFFEGLVRFITSGPVVAMVVEGPQ-AVSEMRRLMGATRPWEAAPGTIRADYATTVDE-
Sulfolobus solfataricus       GKSFFEELVNYVTSGPVVCMVIEGDD-VVQVIRRMIGNTDPKEAPPGTIRGDYALSKSE-
Dictyostelium discoideum      ERPFFGGLVSFITSGPVVAMVFEGKG-VVASARLMIGVTNPLASAPGSIRGDFGVDVGR-
Saccharomyces cerevisiae      GKPFFPKMVSFMKSGPILATVWEGKD-VVRQGRTILGATNPLGSAPGTIRGDFGIDLGR-
Drosophila melanogaster       ARPFFPGLVNYMNSGPVVPMVWEGLN-VVKTGRQMLGATNPADSLPGTIRGDFCIQVGR-
Homo sapiens                  DRPFFPGLVKYMNSGPVVAMVWEGLN-VVKTGRVMLGETNPADSKPGTIRGDFCIQVGR-
Escherichia coli              GKPFFDGLVEFMTSGPIVVSVLEGEN-AVQRHRDLLGATNPANALAGTLRADYADSLTE-
Myxococcus xanthus            ARPFFKDLVQFMISGPVVLMVLEGEN-AVLANRDIMGATNPAQAAEGTIRKDFATSIDK-
Arabidopsis thaliana          SKSFFSGLVDYIVSGPVVAMIWEGKN-VVLTGRKIIGATNPAASEPGTIRGDFAIDIGR-

                                130       140       150       160
                      ....|....|....|....|....|....|....|....|..
Chloroflexus aurantiacus      --NLIHASDSPESGERETAIFFQPHELIGEWNRALDNWIYE------
Thermus aquaticus             --NVIHGSATLEDAQREIALFFRPEELL--------------------
Thermus thermophilus          --NVIHGSATLEDAQREIALFFRPEELL--------------------
Oceanithermus profundus       --NVIHGSANEADAEREIALFFRPEEFVS--------------------
Meiothermus ruber             --NVIHGSDSPESAQREIGIFFKPEEIIG--------------------
Meiothermus silvanus          --NVIHGSDSPESAAREIGIFFKPEEIIG--------------------
Sulfolobus solfataricus       --NVIHASDSIEKAQREMSLFFDKSDL----------------------
Dictyostelium discoideum      --NIIHGSDSVESANREIALWFKPEELLT-EVKPNPNLYE-------
Saccharomyces cerevisiae      --NVCHGSDSVDSAEREINLWFKKEELVDWESNQAKWIYE-------
Drosophila melanogaster       --NIIHGSDAVESAEKEIALWFNEKELVTWTPAAKDWIYE-------
Homo sapiens                  --NIIHGSDSVKSAEKEISLWFKPEELVDYKSCAHDWVYE-------
Escherichia coli              --NGTHGSDSVESAAREIAYFFGEGEVC--PRTR--------------
Myxococcus xanthus            --NTVHGSDSLENAKIEIAYFFRETEIHSYPYQK--------------
Arabidopsis thaliana          --NVIHGSDSVESARKEIALWF-PDGPVNWQSSVHPWVYET------
```

Figure F.1: Sequence alignment of NDK from thermophilic organisms (red names) and mesophilic organisms (black names), made using ClustalW [A7] and BioEdit [A8] with sequence information from UniProt [A9,10]. Highlighted in yellow are the 12 residues of the C-terminal of NDK *Dd* that were attached to NDK *Ss* to form the extended NDK (NDKext).

[8] Unpublished structure solved by Maria Blanner, special course at DTU Chemistry

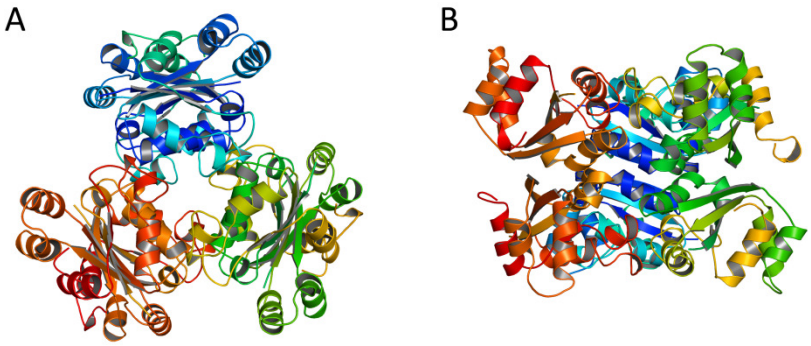


Figure F.2: Two views of the unpublished structure of *Ss* NDK. Three dimers make up the hexameric structure. Each dimer is shown in red, blue or green.

F.2 Experimental

F.2.1 Crystallization

Crystallization conditions for NDKext were based on initial crystallization screens and optimization by Tina Vognsen[9]. Protein with a concentration of 2.9 mg/mL was kindly supplied by Martin Willemoës. Hanging-drop vapor-diffusion in 24 well plates was applied during all crystallization experiments. Crystallization conditions were mainly 35-43 % MPD and 0-0.1 M $CaCl_2$. A selection of buffers and additives (Hampton Research Additive Screens 1, 2 and 3) were used. Crystallization with $(NH_4)_2SO_4$ with 2-5 % different PEG or MPEG was also tested with various buffers. A complete set of all tested conditions is not given, but a variety of crystallization conditions are available for picture entries in the enclosed database (Appendix A, NDKext.xml, see Chapter 2 for a program description).

Crystallization conditions on crystals used for data collection were 38-39 % MPD, 0.05-0.07 M $CaCl_2$ in 0.1 M sodium acetate pH 4.6.

F.2.2 Data Collection

Prior to diffraction tests and data collection, crystals were mounted in a loop and cryo-cooled directly in liquid nitrogen. Data were collected to 3 Å on a large rod-shaped crystal (approximately 60x200 µm): 152 frames with an oscillation angle of 0.5° at MAX-lab (dataset 1). At frame 152, diffraction was poor and annealing of the crystal was attempted. Two data sets were collected at ESRF on similar large rod-shaped crystals. One dataset was collected to 2.5 Å (dataset 2, 240 frames) and the other dataset to 2.7 Å (dataset 3, 250 frames) with oscillation angles of 0.5°.

[9] Tina Vognsen, Master Thesis, Department of Chemistry, Technical University of Denmark, September 2008

F.3 Results and Discussion

F.3.1 Crystallization

Several crystals were produced, some of which are shown in Figure F.3. An elaborate summary of crystallization results can be found in the enclosed database (Appendix A, NDKext.xml) that includes 383 picture entries with matching crystallization conditions. When crystals grew large, they had a tendency to form irregular ends.

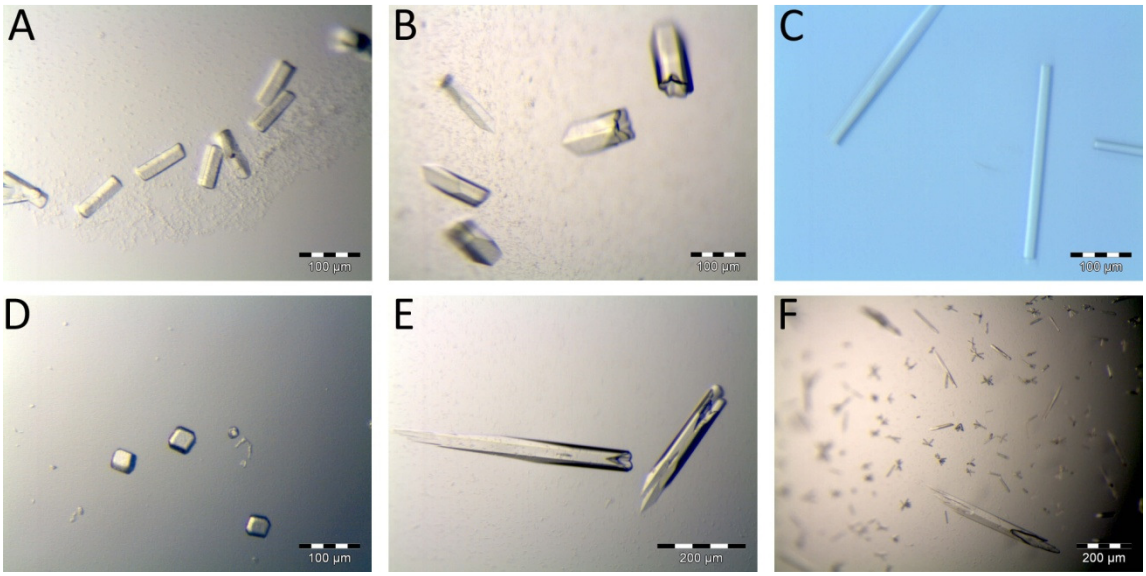


Figure F.3: A: 42 % MPD, 0.04 M $CaCl_2$, 100 mM BisTris pH 5.5 (protein:reservoir ratio 1:1), B: 38 % MPD, 0.04 M $CaCl_2$, 100 mM BisTris pH 5.5 (protein:reservoir ratio 1:1), C: 38 % MPD, 0.07 M $CaCl_2$, 100 mM sodium acetate pH 4.6, (protein:reservoir ratio 3:2 + 0.6 µL 40 % 2,5 hexanediol), D: 40 % MPD, 0.04 M $CaCl_2$, 100 mM Tris pH 8.5 (protein:reservoir ratio 3:2), E: 41 % MPD, 0.06 M $CaCl_2$, 100 mM sodium acetate pH 4.6 (protein:reservoir ratio 3:2), F: 39 % MPD, 0.05 M $CaCl_2$, 100 mM sodium acetate pH 4.6 (protein:reservoir ratio 3:2).

F.3.2 Data Collection and Processing

Diffraction of the crystals shown in Figure F.3A-D was very poor or non-existing. Crystals with morphology as shown in Figure F.3E-F were used for data collection at MAX-lab and ESRF. A tendency is observed where the largest crystals have the best diffraction properties despite their cracked appearance.

Data were collected to 3 Å at MAX-lab (dataset1). The poor diffraction at frame 152 did not improve with annealing and data collection was stopped. Two data sets were collected at ESRF, one to 2.7 Å and one to 2.5 Å. Merging of the data proved to be a challenging task and several attempts were made by Pernille Harris to merge the data in order to solve and refine the

structure (difficulties in solving and refining the structure are described in section F.4). Data from MAX-lab and from one of the crystals at ESRF merged in hexagonal space group ($P6_322$), whereas the other crystal at ESRF only merged in monoclinic (P2) space group.

Data processing proved very problematic and the difficulties met are best explained by a pseudomerohedral twinning of the crystals. In merohedral twinning, a 2-fold axis along a unit cell axis relates two parts within the crystal. If the two parts are equally represented within the crystal, a higher degree of symmetry will appear as reflections from each 'subcrystal' seem symmetry related. This is in agreement with a higher degree of symmetry (hexagonal) observed for two crystals (dataset 1+2), whereas the third crystal (dataset 3) has lower symmetry (monoclinic). The diffraction of the crystals was furthermore observed as a combination of well-defined and streaky diffraction spots, see Figure F.4.

Even though the monoclinic crystal could be a single crystal, no reasonable structure solution was found likely because of the low quality of the data. Data collection statistics and processing used for structure solution attempts are given in Table F.1.

Table F.1: Data collection statistics.[a]

	Dataset1	Dataset2	Dataset3
Beamline	MAX-LAB, 911-2	ESRF, ID14-2	ESRF, ID14-2
Detector	mar165	ADSC Q4 CCD	ADSC Q4 CCD
Wavelength (Å)	1.04	0.933	0.933
Temperature (K)	100	100	100
Crystal space group	$P6_322$	$P6_322$	$P2_1$
Unit cell parameters			
a (Å)	71.59	71.662	71.49
b (Å)	71.59	71.662	198.43
c (Å)	198.12	199.072	71.47
β (°)			120.087
Resolution (Å)	29.6-3.0 (3.07-2.99)	45.3-2.6 (2.81-2.60)	45-2.9 (3.08-2.9)
No. of reflections	54,592 (3,608)	138,944 (27,095)	94,755 (15,853)
No. of unique reflections	6,590 (440)	9,977 (1,990)	32,400 (5,519)
Redundancy	8.3 (8.2)	13.9 (13.6)	2.9 (2.9)
Mosaicity (°)	0.5-0.8	0.2-0.3	0.4-0.7
Completeness (%)	99.3 (95.2)	99.8 (99.4)	84.6 (87.5)
$I/\sigma(I)$	8.05 (1.46)	14.45 (2.04)	6.32 (0.95)
R_{merge}	29.3 (181.0)	14.6 (120.6)	14.8 (128.9)

[a] Values for the outermost resolution shell are given in parenthesis.

$R_{merge} = \Sigma_i \left| I_i - \langle I_i \rangle \right| / \Sigma_i I_i$

F.4 Attempts to Solve the Structure and Discussion

The structure of NDKext was attempted solved with dataset 1 (see Table F.1 for statistics) using molecular replacement with WT NDK *Ss* as a search model and two molecules in the asymmetric unit. R_{free} was 0.396 after a rigid body refinement and a restrained refinement and improved slightly to 0.353 when refined to 3.1 Å instead of 3 Å. Any other change; decreasing the resolution, refinement of positions in winCoot, addition of any residues to the extended C-terminal or the use of NCS, resulted in a notable increase in R_{free} (R_{free} in the range of 0.4-0.45). The side chain of Glu97 was removed because it clashed with a symmetry generated molecule, however, this also increased R_{free}.

With improved data from ESRF (dataset 2, see Table F.1 for statistics), the structure was attempted solved with molecular replacement using MOLREP [A11]. A total of 14 NDK crystal structures from different organisms were tested as search models for molecular replacement. Solutions were only possible with 2 molecules in the asymmetric unit. The best result with hexagonal space groups was achieved with space group $P6_322$ and molecular replacement using WT NDK *Ss* as a search model. Phenix.refine [A12] or REFMAC5 [A13] was used for refinement. However, R_{free} did not drop below 0.4 at any time during refinement. A range of phenix.refine options were tested: Ramachandran restrained refinement, restrained refinement with NCS, reference model steering refinement, ordering solvent and simulated annealing. A segment of NDKext (residues 108-114) from refinement is shown in Figure F.4, and even though electron density appeared sufficient to model the approximate positions of the residues, R_{free} had a very high value of 0.522.

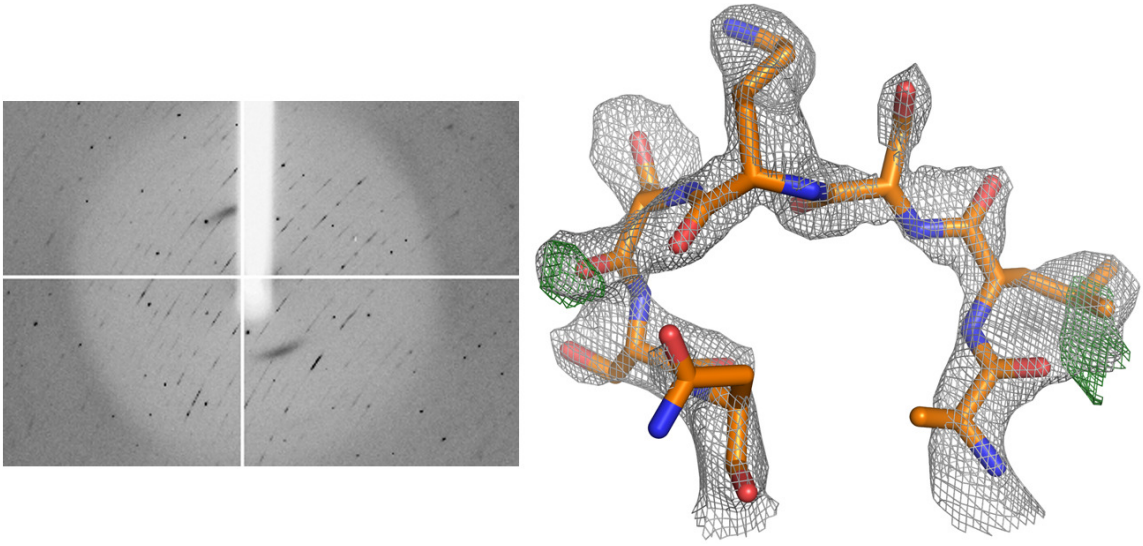


Figure F.4: Left: segment of diffraction image from data collection (Table F.1 dataset 2). Right: residues 108-114 shown with $2F_{obs} - F_{calc}$ = 1.4 (grey) and $F_{obs} - F_{calc}$ = 3.0 (green) during an attempt to solve the structure of NDKext with dataset 2 (R_{free} was 0.522).

F.5 Conclusion

The crystallization conditions of NDKext have been optimized and several crystals have been made. Larger crystals with cracked edges had better diffraction properties than smaller crystals with sharp edges. Data were collected on three crystals: one at MAX-lab and two at ESRF. However, data processing and refinement did not produce any reasonable structure solution likely because of pseudomerohedral twinning of the crystals. New crystallization trials should be made in search of a single crystal with improved diffraction properties.

Appendix G

G PUBLICATIONS

Løvgreen MN, Martic M, Windahl MS, Christensen HEM and Harris P. Crystal structures of the all cysteinyl coordinated D14C variant of *Pyrococcus furiosus* ferredoxin: [4Fe-4S]↔[3Fe-4S] cluster conversion, *J. Biol. Inorg. Chem.* **(2011)**

Løvgreen MN, Løvgreen M, Christensen HEM and Harris P. *MyCrystals* – a simple visual data management program for laboratory-scale crystallization experiments. *J. Appl. Cryst.* **(2009)** 42, 741-742.

J Biol Inorg Chem (2011) 16:763–775
DOI 10.1007/s00775-011-0778-7

ORIGINAL PAPER

Crystal structures of the all-cysteinyl-coordinated D14C variant of *Pyrococcus furiosus* ferredoxin: [4Fe–4S] ↔ [3Fe–4S] cluster conversion

Monika Nøhr Løvgreen · Maja Martic · Michael S. Windahl · Hans E. M. Christensen · Pernille Harris

Received: 7 December 2010 / Accepted: 25 March 2011 / Published online: 12 April 2011

Abstract The structure of the all-cysteinyl-coordinated D14C variant of [4Fe–4S] ferredoxin from the hyperthermophilic archaeon *Pyrococcus furiosus* has been determined to 1.7 Å resolution from a crystal belonging to space group $C222_1$ with two types of molecules, A and B, in the asymmetric unit. A and B molecules have different crystal packing and intramolecular disulfide bond conformation. The crystal packing reveals a β-sheet interaction between A molecules in adjacent asymmetric units, whereas B molecules are packed as monomers in a less rigid position next to the A–A extended β-sheet dimers. The A molecules contain an intramolecular disulfide bond in a double conformation with 60% occupancy left-handed and 40% occupancy right-handed spiral conformation, whereas B molecules have an intramolecular disulfide bond in a right-handed spiral conformation. The cluster in D14C [4Fe–4S] *P. furiosus* ferredoxin was chemically oxidized at pH 5.8 to [3Fe–4S]. For purification at pH 8.0, two forms of the protein are obtained. Mass spectrometric analysis shows that the two forms are the D14C [3Fe–4S] *P. furiosus* ferredoxin monomer and a disulfide-bonded dimer of D14C [3Fe–4S] *P. furiosus* ferredoxin. When oxidization and purification are carried out at pH 5.8, only the monomer is obtained. The crystal structure of D14C [3Fe–4S] *P. furiosus* ferredoxin monomer was determined to 2.8 Å resolution from a crystal belonging to space group $P2_12_12_1$ with two molecules in the asymmetric unit. The molecules resemble molecule A of D14C [4Fe–4S] *P. furiosus* ferredoxin and electron density clearly shows the presence of a [3Fe–4S] cluster.

Keywords Ferredoxin · Crystal structure · *Pyrococcus furiosus* · Disulfide bond · Dimer

Abbreviations

BisTris	[Bis(2-hydroxyethyl)amino] tris(hydroxymethyl)methane
CV	Column volume
PDB	Protein Data Bank
RMS	Root mean square
Tris	Tris(hydroxymethyl)aminomethane
WT	Wild type

Introduction

Iron–sulfur proteins are present in all kinds of organisms and cellular compartments and assume a range of functions, predominantly electron transfer, but also catalysis and regulatory and sensing functions. Iron–sulfur proteins can contain several active site iron–sulfur clusters in a variety of frameworks: from simple clusters with one to four irons up to large clusters—some of which are

An interactive 3D complement page in Proteopedia is available at http://proteopedia.org/wiki/index.php/Journal:JBIC:10.

Electronic supplementary material The online version of this article (doi:10.1007/s00775-011-0778-7) contains supplementary material, which is available to authorized users.

M. N. Løvgreen · M. Martic · H. E. M. Christensen (✉) · P. Harris (✉)
Department of Chemistry, Technical University of Denmark, Building 207, 2800 Kgs. Lyngby, Denmark
e-mail: hemc@kemi.dtu.dk

P. Harris
e-mail: ph@kemi.dtu.dk

M. S. Windahl
Carlsberg Laboratory, Gamle Carlsberg Vej 10, 1799 Copenhagen V, Denmark

heterometallic with nickel or molybdenum as part of or bound to the cluster [1].

Ferredoxins contain iron–sulfur clusters coordinated most often by cysteine residues to the protein chain. A common cluster in bacterial ferredoxins is the cuboidal [4Fe–4S] cluster bound to the polypeptide chain by cysteine ligands with a typical binding motif: –Cys–X_2–Cys–X_2–Cys–X_n–Cys–. These [4Fe–4S] clusters can reversibly interconvert to [3Fe–4S] clusters. Common small ferredoxins with only a single cluster often have two additional cysteines forming an intramolecular disulfide bond, which replace a second cluster in more ancient ferredoxins [2, 3]. In the binding motif of *Pyrococcus furiosus* ferredoxin, the second cysteine is replaced by aspartate and [4Fe–4S] to [3Fe–4S] cluster conversion is readily observed [4]. Incomplete cysteine coordination is also seen in one of two [4Fe–4S] clusters in *Desulfovibrio africanus* ferredoxin III that readily converts to [3Fe–4S] [5]. Site-directed change of cluster-coordinating aspartate to cysteine has been reported to impair the ease with which [4Fe–4S] converts to [3Fe–4S] [5].

The other reported crystal structures of [3Fe–4S] ferredoxins with complete cysteine binding motifs are *Desulfovibrio gigas* (Protein Data Bank, PDB, entry 1FXD [6]) and *Bacillus thermoproteolyticus* (PDB entry 1WTF [7]). These both have additional chemical groups bound that protect the free cysteine; hence, the structure of the D14C [3Fe–4S] *P. furiosus* ferredoxin is the first crystal structure in which a cysteine in the binding motif is unmodified and facing away from the cluster.

The small electron-transfer ferredoxin from the hyperthermophilic archaeon *P. furiosus* [8] consists of 66 amino acids (7.5 kDa) and contains a single $[4Fe–4S]^{2+,+}$ cluster and five cysteine residues, two of these forming an intramolecular, redox-active disulfide bond [4, 9]. This ferredoxin is very thermostable, with no denaturation after 12 h at 95 °C [8]. The protein fold of thermophilic ferredoxins often shows polypeptide chain extensions [1] and this is also seen in *P. furiosus* ferredoxin, which compared with other ferredoxins has structural extensions in the β-sheet of the two termini and in α-helix 2 [10]. *P. furiosus* ferredoxin is further stabilized by an extensive hydrogen-bonding network as well as a hydrophobic interaction between Trp2 and Tyr46. The crystal structure of *P. furiosus* ferredoxin shows crystal packing as a dimer formed by extended β-sheet interactions [10]. It has been proposed that dimerization of wild type (WT) *P. furiosus* ferredoxin occurs depending on ionic strength [11, 12]. Although we have observed crystal packing as extended β-sheet dimers, these dimers were not observed in solution and could be an artifact of the high protein concentration in the crystal.

A single exchange of the cluster coordinating Asp14 to cysteine (D14C) in *P. furiosus* ferredoxin results in a complete cysteine binding motif and causes some changes compared with WT *P. furiosus* ferredoxin; however, the thermostability is maintained [13]. An additional Fe–S charge transfer band in D14C compared with WT *P. furiosus* ferredoxin causes an increase in the absorbance ratio and the extinction coefficient: $A_{390}/A_{280} = 0.73$ and $\varepsilon_{390} = 20.2\ \mathrm{mM^{-1}\ cm^{-1}}$ for D14C compared with $A_{390}/A_{280} = 0.56$ and $\varepsilon_{390} = 17\ \mathrm{mM^{-1}\ cm^{-1}}$ for the WT [14, 15]. The reduction potential of the D14C variant is 59 mV more negative than that of the WT ferredoxin: −427 versus −368 mV [16]. The electrochemical properties of ferredoxins have been reported to be sensitive to changes in the cluster environment: reduced surface/solvent accessibility, reduced hydrogen-bonding network, or reduced volume of the cluster cavity lower the reduction potential and the conformation of the ligands also has an influence [17–22].

We have expressed and purified D14C $[4Fe–4S]^{2+}$ *P. furiosus* ferredoxins in a form suitable for crystallization and structure determination. D14C [3Fe–4S] *P. furiosus* ferredoxin was produced by oxidation. Purification by anion exchange at pH 5.8 results in a single peak, whereas for purification at pH 8.0 two peaks are observed. Mass spectrometric analysis reveals a monomer at pH 5.8, whereas both a monomer and a disulfide-bonded dimer are formed at pH 8.0. In this article we report the three-dimensional crystal structure of D14C [4Fe–4S] *P. furiosus* ferredoxin to 1.7 Å resolution. D14C [3Fe–4S] *P. furiosus* ferredoxin proved difficult to crystallize because of protein solution inhomogeneity, and the three-dimensional structure is reported to 2.8 Å resolution.

Materials and methods

Expression strain and cultivation

The *P. furiosus* ferredoxin gene was amplified by PCR and cloned into the pET3a vector. The D14C mutation was created by site-directed mutagenesis using a QuikChange kit (Stratagene). The plasmid encoding the D14C *P. furiosus* ferredoxin was transformed into *E. coli* BL21(DE3) cells (Novagen) for expression. The expression strain was grown in 650 mL cultures.

Purification of D14C [4Fe–4S] *P. furiosus* ferredoxin

Approximately 30 mL 20 mM tris(hydroxymethyl)aminomethane (Tris)/HCl pH 8.0 buffer was added to six tubes each containing cells from 650 mL culture and the cells were resuspended. Eighty microliters of 1 M sodium dithionite was added to each tube. Subsequently, the cells were lysed by sonication three times for 40 s. After sonication, the suspension was incubated for 10 min in a 70 °C

water bath, then centrifuged at 4 °C for 20 min at 18,500*g*, and the clear supernatant decanted off. The supernatant was diluted four times with 20 mM Tris/HCl pH 8.0, 2 mM sodium dithionite. The protein solution was loaded onto a 50 mm/7 cm Q Sepharose Fast Flow column equilibrated with 20 mM Tris/HCl pH 8.0. The column was washed with 2 column volumes (CV) 20 mM Tris/HCl pH 8.0 and 2 CV 20 mM Tris/HCl pH 8.0, 0.15 M NaCl. The protein was eluted isocratically with 20 mM Tris/HCl pH 8.0, 0.4 M NaCl. The volume was reduced to approximately 5 mL by ultrafiltration using a stirred Amicon cell with a PLBC NMWL 3 kDa membrane. A maximum of 5 mL protein solution was then loaded onto a HiLoad 16 mm/60 cm Superdex 75 column equilibrated with 20 mM Tris/HCl pH 8.0, 0.15 M NaCl and eluted using the same buffer with a flow rate of 1 mL/min. Fractions containing D14C *P. furiosus* [4Fe–4S] ferredoxin were pooled and exchanged into 20 mM Tris/HCl pH 8.0 by ultrafiltration. A maximum of 15 mg was loaded onto a 16 mm/10 cm Source 30Q column equilibrated with 20 mM Tris/HCl pH 8.0. The column was washed with 2 CV 20 mM Tris/HCl pH 8.0 and elution was carried out using a linear salt gradient of 0.15–0.4 M NaCl in 20 mM Tris/HCl pH 8.0 over 17 CV. The protein was exchanged into 20 mM Tris/HCl pH 8.0 by ultrafiltration.

Oxidation to and purification of D14C [3Fe–4S] *P. furiosus* ferredoxin

Experiments were carried out with buffer A corresponding to 20 mM Tris/HCl pH 8.0 or 20 mM [bis(2-hydroxyethyl) amino]tris(hydroxymethyl)methane (BisTris)/HCl pH 5.8. Fifteen times molar excess ferricyanide and 1.5 times molar excess EDTA were added to purified D14C *P. furiosus* [4Fe–4S] ferredoxin in 20 mM BisTris/HCl pH 5.8 and the mixture was left overnight with stirring at room temperature. The excess of ferricyanide and EDTA was removed by ultrafiltration using buffer A. The protein solution was filtered before being loaded onto a 16 mm/10 cm Source 30Q column equilibrated with buffer A. The column was washed with 2 CV buffer A and elution was carried out using a linear salt gradient of 0.15–0.4 M NaCl in buffer A over 17 CV. The protein was exchanged into buffer A by ultrafiltration.

Mass spectrometry

Prior to mass spectrometric analysis all protein samples were desalted, and the buffer was exchanged into 100 mM ammonium acetate/acetic acid pH 5.8 using Micro Bio-Spin P-6 columns (Bio-Rad). Samples were diluted to a protein concentration of approximately 20 μM. Mass spectra were obtained with a nano-electrospray ionization time-of-flight instrument (LCT Premier, Waters). The data were acquired in negative ion mode. The instrument was calibrated using 2 mg/mL NaI in 50% 2-propanol solution. Au/Pd-coated capillaries (Proxeon) were used. The source temperature was maintained at 50 °C during the experiments. The capillary voltage applied was typically between 800 and 900 V, while the cone voltage applied was 50 V. The MaxEnt1 algorithm provided with MassLynx^{TM} version 4.1 (Waters) was used to produce mass spectra from the observed *m*/*z* data.

Crystallization and data collection

For crystallization, the conditions for the WT *P. furiosus* ferredoxin [23] were optimized. Platelike crystals of D14C [4Fe–4S] *P. furiosus* ferredoxin were obtained at 4 °C by hanging-drop vapor diffusion with a reservoir containing 35% PEG600 and a 0.1 M malonic acid/imidazole/boric acid buffer system pH 7.5 [24]. The drop contained 2 μL reservoir solution, 2 μL 6.7 mg/mL D14C [4Fe–4S] *P. furiosus* ferredoxin in 20 mM Tris/HCl pH 8.0, and 0.3 μL 0.1 M hexaamminecobalt(III) chloride. The crystals had grown after 3 days and were cryocooled directly in liquid nitrogen on the eighth day for subsequent data collection. Diffraction data were collected: 180 frames with an oscillation angle of 1°. Integration and scaling of the data were performed with XDS/XSCALE [25, 26].

Crystallization of D14C *P. furiosus* [3Fe–4S] ferredoxin was much more difficult. Numerous crystallization conditions were tested and at pH below 7 precipitation generally occurred. This precipitation was most often gray, indicating that the protein had lost its iron–sulfur cluster. Drops were generally inhomogeneous, with a mixture of several combinations of phase separation, precipitation, protein films, and crystalline clusters with very high mosaicity. Conditions from crystallization screens were tested; however; crystals were only obtained with conditions similar to those for crystallization of the WT [23]. The best crystals were obtained at room temperature with hanging-drop vapour diffusion with a reservoir containing 35% PEG1500 and 0.1 M Tris/HCl pH 8.5. The drop contained 2 μL reservoir solution, 2 μL 6.4 mg/mL D14C [3Fe–4S] *P. furiosus* ferredoxin in 20 mM Tris/HCl pH 8.0, and 0.3 μL 0.1 M hexaamminecobalt(III) chloride. Diffraction data were collected: 360 frames with an oscillation angle of 0.5°. Processing and space group determination were difficult because of anisotropic mosaicity. Processing was performed with MOSFLM [27] and space group determination and averaging were achieved with Pointless and SCALA [28]. The space group determined was in agreement with the findings of previous diffraction experiments.

Data collection statistics for both structures are given in Table 1.

Structure determination and refinement of D14C *P. furiosus* [4Fe–4S] ferredoxin

The CCP4 [30] suite was used for structure determination and refinement. MOLREP [31] was used to solve the structure with molecular replacement. The search model was the peptide chain of WT [3Fe–4S] *P. furiosus* ferredoxin (PDB entry 1SJ1) [10]. Two molecules, type A and B, were found in the asymmetric unit. Refinement was performed with REFMAC5 [32]: first a rigid-body refinement and then restrained refinements with isotropic *B* factors and geometric restraints as defined by Engh and Huber [33] and a limited *B* value range of 2–200. Inspection of the $2F_{obs} - F_{calc}$ and the $F_{obs} - F_{calc}$ sigma-A weighted difference maps was done using winCoot [34, 35]. Electron density appeared clearly from two [4Fe–4S] clusters and four $[Co(NH_3)_6]^{3+}$. Two $[Co(NH_3)_6]^{3+}$ were added to each of the two molecules in the asymmetric unit. One cobalt atom in molecule A was disordered and split into two cobalt atoms with 50% occupancy. One $[Co(NH_3)_6]^{3+}$ in molecule B was placed on a twofold rotation axis and thus assigned 50% occupancy. In molecule A, the following residues were split in alternate conformations (occupancies were refined manually by inspection of $2F_{obs} - F_{calc}$ and $F_{obs} - F_{calc}$ sigma-A weighted difference maps and are given in parentheses): Thr10 (50/50), Cys17 (95/5), Ser19 (50/50), Cys21 (40/60), Leu45 (50/50), Cys48 (40/60), Glu54 (50/50), Glu65 (50/50). Molecule B displayed alternate conformation only at Ser19 (50/50). A total of 155 water molecules were included in the structure in a combined use of REFMAC5 [32] and manually using winCoot [34, 35]; all water molecules were inspected manually. The structure of D14C *P. furiosus* [4Fe–4S] ferredoxin is available in the PDB with entry 2Z8Q (see Table 2 for refinement statistics).

Structure determination and refinement of D14C *P. furiosus* [3Fe–4S] ferredoxin

The CCP4 [30] suite was used for structure determination and refinement. MOLREP [31] was used to solve the structure with molecular replacement. The search model was the peptide chain of type A D14C [4Fe–4S] *P. furiosus* ferredoxin. Two molecules were found in the asymmetric unit. Molecular replacement was also carried out with the search model of type B D14C [4Fe–4S] *P. furiosus* ferredoxin. The *R* factor and score after molecular replacement were 0.622 and 0.455, compared with 0.592 and 0.534 for type A. An increase in R_{free} was observed during refinement when using a type B molecule compared with a type A molecule as the starting model. Molecule A of D14C [4Fe–4S] *P. furiosus* ferredoxin is thus a considerably better model for molecular replacement than molecule B. Refinement was performed with REFMAC5 [32]: first a rigid-body refinement and then restrained refinements with isotropic *B* factors, tight non-crystallographic symmetry restrains on residues 1–25 and 28–66, and geometric restraints as defined by Engh and Huber [33] and a limited *B* value range of 2–200. Inspection of the $2F_{obs} - F_{calc}$ and the $F_{obs} - F_{calc}$ sigma-A weighted difference maps was done using winCoot [34, 35]. Electron density appeared clearly from two [3Fe–4S] clusters and four cobalt atoms; electron density was not visible for cobalt

Table 1 Data collection statistics

	D14C *Pyrococcus furiosus* [4Fe–4S] ferredoxin	D14C *P. furiosus* [3Fe–4S] ferredoxin
Beamline [29]	ESRF, ID14-3	ESRF, ID14-4
Detector	ADSC Quantum 4	ADSC Quantum Q315r
Wavelength (Å)	0.931	0.9765
Temperature (K)	100	100
Crystal space group	$C222_1$	$P2_12_12_1$
Unit cell parameters		
a (Å)	51.4	47.4
b (Å)	116.8	49.8
c (Å)	47.7	51.23
Resolution (Å)	26–1.7 (1.8–1.7)	35.7–2.8 (2.95–2.8)
No. of reflections	113,257 (17,828)	17,078 (2,504)
No. of unique reflections	16,226 (2,509)	3,254 (456)
Redundancy	6.98 (7.11)	5.2 (5.5)
Mosaicity (°)	0.2–0.4	Fixed 0.8
Completeness (%)	99.9 (100.0)	99.9 (99.9)
$I/\sigma(I)$	11.41 (4.93)	3.6 (1.9)
R_{merge}	0.112 (0.258)	0.252 (0.755)

Values for the outermost resolution shell are given in *parentheses*.
$R_{merge} = \Sigma_i |I_i - \langle I_i \rangle| / \Sigma_i I_i$

Table 2 Refinement statistics for D14C *P. furiosus* ferredoxin

	[4Fe–4S]	[3Fe–4S]
R	0.164	0.279
R_{free}	0.194	0.318
No. of atoms in model		
Protein	988	988
FeS clusters	16	14
Cobalt complexes	24.5[a]	4
Water oxygens	155	0
Mean temperature (B) factors (Å^2)		
Protein chain A/B	7.5/18	39/39
FeS cluster chain A/B	3.9/9.8	34/28
Cobalt complexes chain A/B	13/16	34/34
Water	23	–
Ramachandran plot [36]		
No. of residues in favored region	127	116
No. of residues in allowed region	1	12
No. of residues in outlier region	0	0

[a] One cobalt complex is placed on a twofold rotation axis and assigned 50% occupancy

ligands. Two cobalt atoms were added to each molecule in the asymmetric unit. Molecule A displayed alternate conformation (occupancies were refined manually by inspection of $2F_{obs} - F_{calc}$ and $F_{obs} - F_{calc}$ sigma-A weighted difference maps and are given in parentheses) at Met27 (50/50) and molecule B displayed alternate conformation at Cys14 (50/50). The structure of D14C *P. furiosus* [3Fe–4S] ferredoxin is available in the PDB with entry 3PNI (see Table 2 for refinement statistics).

Programs used for the analysis in "Results and discussion"

Structural figures were prepared using PyMOL [37]. Solvent accessibilities were calculated using AREAIMOL [38, 39]. Root mean square (RMS) deviations were calculated using the program superpose [40]. The omit map was calculated using Sfcheck [41]. pK_a was estimated using PROPKA [42, 43] Sequence alignment was done using ClustalW [44].

Results and discussion

Purification and cluster conversion

Aerobic purification of D14C *P. furiosus* ferredoxin produces the ferredoxin with an intact [4Fe–4S] cluster that is eluted at 25–26 mS/cm during anion exchange at pH 8.0. Aerobic purification of WT *P. furiosus* ferredoxin produces both [4Fe–4S] and [3Fe–4S] ferredoxins, which are eluted in the form of two peaks during anion exchange at pH 8.0.

Oxidation of [4Fe–4S] to [3Fe–4S] D14C *P. furiosus* ferredoxin produces a free cysteine, which corresponds to Asp14 in WT [3Fe–4S] *P. furiosus* ferredoxin [10]. Oxidation was initially carried out at pH 7.8 as reported in [45], but multiple peaks were observed during subsequent anion exchange at pH 8.0 and the pH was therefore changed. At pH 5.8, a single peak is observed during anion exchange (see Fig. 1). At pH 8.0, the free Cys14 in D14C [3Fe–4S] *P. furiosus* ferredoxin is partly deprotonated and elution in two peaks is observed (see Fig. 1). This can be explained by elution of a monomer and a disulfide-bonded dimer ferredoxin. The dimer is formed when two deprotonated monomers form an intermolecular disulfide bond (see Scheme 1). The pK_a of Cys14 was estimated with PDB entry 1SJ1 (WT [3Fe–4S] *P. furiosus* ferredoxin [10]), where Asp14 was mutated to cysteine. The pK_a was estimated to be 8.8, which is an increase compared with the pK_a of 8.3 of normal cysteine as is expected for a cysteine residue in a negatively charged environment. The calculated pK_a value is consistent with the presence of both protonated (present as monomers) and deprotonated (present as dimers) D14C [3Fe–4S] *P. furiosus* ferredoxin at pH 8.0 (see Scheme 1).

Mass spectrometry

Mass spectrometric analysis identified the protein purified at pH 5.8 (see Fig. 1, chromatogram labeled pH 5.8) to be D14C $[3Fe–4S]^+$ *P. furiosus* ferredoxin. Its m/z spectrum is shown in Fig. 2 (spectrum A1). Seven protein charge states are observed—from charge state 9− to charge state 3−. The m/z range from 800 to 1,300 was used to calculate the mass spectrum. The apparent mass of the protein, derived assuming that all negative charge is from proton deficit [46], is 7,447.6 Da (see Fig. 2, spectrum A2). The calculated mass is calculated from the sum of average molecular masses of all amino acids present in the protein sequence plus the average molecular mass of the iron–sulfur cluster components [46, 47]. The calculated mass has to be lowered by the mass of two protons for each disulfide bond, and by the mass of protons equal in number to the oxidation state of the metal cluster present [46]. The calculated mass for the protein with one disulfide bond (C21–C48) and 1+ oxidation state of the metal cluster is 7,447.8 Da, which perfectly matches the observed apparent mass (Fig. 2, spectrum A2).

Subsequently, the two observed peaks in protein repurification at pH 8.0 (see Fig. 1) were analyzed by mass spectrometry. Protein from the first peak was identified to be D14C $[3Fe–4S]^+$ *P. furiosus* ferredoxin, whereas protein from the second peak was identified to be a disulfide-bonded

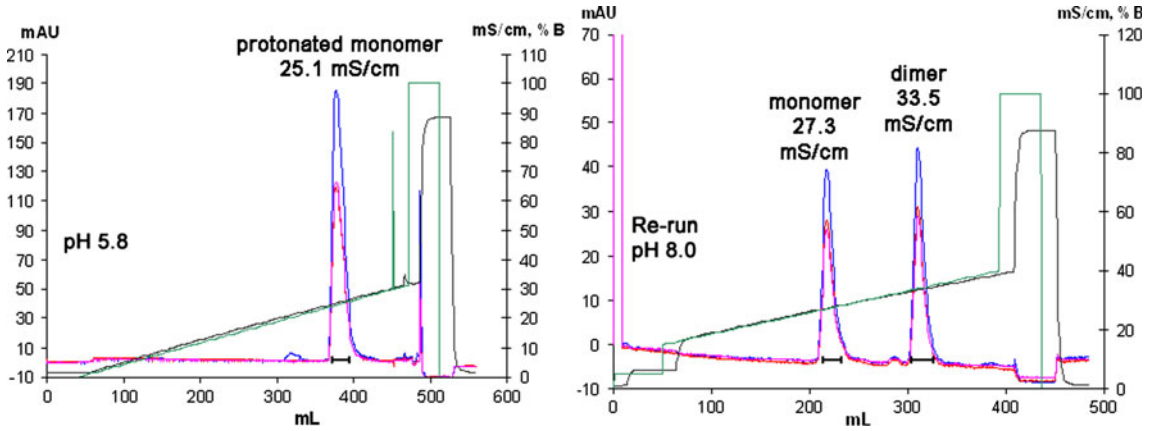


Fig. 1 Purification of D14C [3Fe–4S] *Pyrococcus furiosus* ferredoxin at pH 5.8 (20 mM BisTris/HCl pH 5.8) and a re-run at pH 8.0 (20 mM Tris/HCl pH 8.0)—on a Source 30Q anion exchange column. Elution was followed at 280, 390, and 408 nm (*blue*, *red*, and *pink*, respectively). Conductivity (mS/cm) is shown in *brown* and salt concentration (*% B*; buffer with 1 M NaCl) is shown in *green*. *Dimer* refers to a disulfide-bonded dimer of D14C [3Fe–4S] *P. furiosus* ferredoxin

$$\text{Fd-Cys14-SH} \underset{\text{pK}_a = 8.8}{\rightleftharpoons} \text{Fd-Cys14-S}^-$$

$$2\ \text{Fd-Cys14-S}^- \xrightarrow[-2e^-]{} \text{Fd-Cys14-S-S'-Cys'14-Fd'}$$

Scheme 1 pH-dependent equilibrium of D14C [3Fe–4S] *Pyrococcus furiosus* ferredoxin between protonated and deprotonated monomers and formation of a disulfide-bonded dimer from deprotonated monomers. *Fd* ferredoxin

dimer: D14C $[3Fe–4S(Cys–S)–(S'–Cys')3Fe'–4S']^{2+}$ *P. furiosus* ferredoxin. From its *m/z* spectrum (Fig. 2, spectrum B1), seven protein charge states are observed—from charge state 12− to charge state 6−. The *m/z* range from 1,200 to 1,900 was used to calculate the mass spectrum. The apparent mass of the disulfide-bonded dimer is 14,893.1 Da (see Fig. 2, spectrum B2). The calculated mass for the protein with two C21–C48 intramolecular disulfide bonds, a C14–C14′ intermolecular disulfide bond, and two metal clusters with oxidation state of 1+ is 14,893.7 Da, which matches the observed apparent mass (Fig. 2, spectrum B2).

Description of the structures

The overall structure of D14C *P. furiosus* [4Fe–4S] ferredoxin is illustrated in Fig. 3. The D14C variant [4Fe–4S] ferredoxin incorporates the same secondary structure elements as seen in the [3Fe–4S] *P. furiosus* ferredoxin structure [10]. The extended β-sheet dimer formation of *P. furiosus* ferredoxin seen in the 1.5 Å-resolution structure of *P. furiosus* [3Fe–4S] ferredoxin [10] is also seen in D14C *P. furiosus* [4Fe–4S] ferredoxin between adjacent A molecules (see Fig. 4). The A molecules form an extended β-sheet by four main-chain intermolecular hydrogen bonds: Ile63 N–Glu65 O and Glu65 N–Ile63 O in both molecules. The B molecules do not form this extended β-sheet dimer and no main-chain intermolecular hydrogen bonds are observed between B molecules. The hexaammineco-balt(III) complexes are situated between negatively charged residues on the ferredoxin surfaces (Table S1) and thus play an important electrostatic role in the crystal packing. The overall charge of *P. furiosus* ferredoxin at neutral pH is −12 and the cobalt complex is also necessary during crystallization of WT *P. furiosus* ferredoxin [23] and is present in the structure of WT [3Fe–4S] *P. furiosus* ferredoxin [13] (see Fig. 5).

The structure of D14C [3Fe–4S] *P. furiosus* ferredoxin is similar to that of the A molecules of D14C [4Fe–4S] *P. furiosus* ferredoxin (see Fig. 3) and to that of the WT [3Fe–4S] *P. furiosus* ferredoxin. The resolution is not sufficient to make elaborate comparisons; however, electron density clearly shows the presence of a [3Fe–4S] cluster in the structure (see Fig. 6). D14C [3Fe–4S] *P. furiosus* ferredoxin with an omit map also clearly shows the presence of a [3Fe–4S] cluster, and when we tried to model the same structure with a [4Fe–4S] cluster, it was clear from the $F_o - F_c$ map that the fourth iron does not belong in the model (Fig. S1). Crystal packing in D14C [3Fe–4S] ferredoxin is as extended β-sheet dimers of adjacent molecules, which is the same as in WT [3Fe–4S] *P. furiosus* ferredoxin [10] even though the space groups are different (see Fig. 5). The crystal packing of D14C [3Fe–4S] emphasizes a strong resemblance to that of WT [3Fe–4S] *P. furiosus* ferredoxin and to that of molecule A of D14C [4Fe–4S] *P. furiosus* ferredoxin, which is a better search model for molecular replacement than molecule B, which does not form this type of extended β-sheet dimer. Crystals were grown at pH 8.5 and no disulfide-bonded dimers are observed in the structure. Monomers and

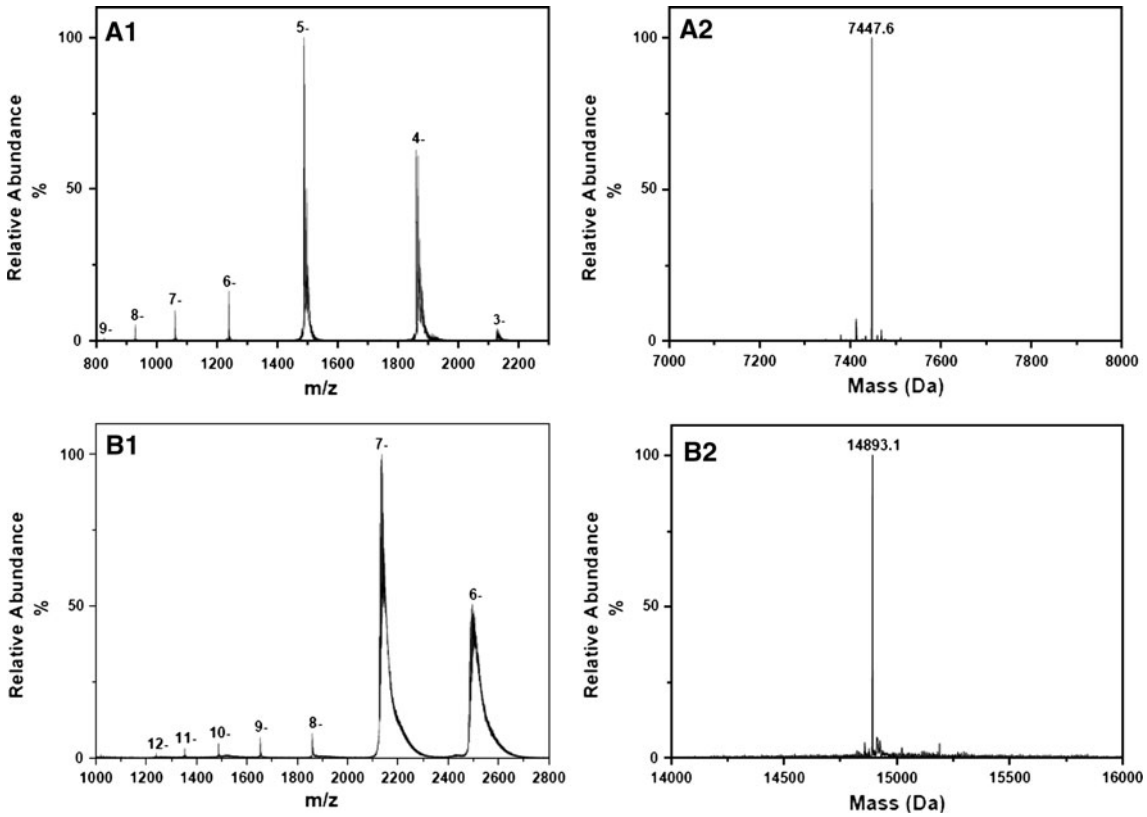


Fig. 2 Mass spectrometric analysis spectra. *A1 m/z* spectrum and *A2* mass spectrum of protein purified at pH 5.8 (Fig. 1, peak at 25.1 mS/cm); *B1 m/z* spectrum and *B2* mass spectrum of protein purified at pH 8.0 (Fig. 1, peak at 33.5 mS/cm)

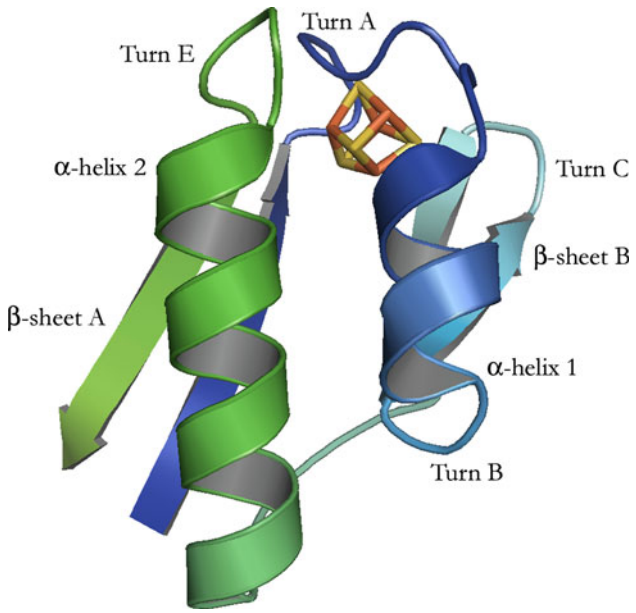


Fig. 3 Overall structure of D14C *P. furiosus* [4Fe–4S] ferredoxin molecule A represented as ribbons. The two α-helices (1 and 2) and two *β*-sheets (A and B) as well as turns A, B, C, and E are marked. The iron–sulfur cluster is shown as *sticks*

disulfide-bonded dimers should be present at pH 8.5 (see Scheme 1); however, the intrinsic crystal properties appear to favor monomers, likely because of the large flexibility around the intermolecular disulfide bond. The inhomogeneous solution of monomers and disulfide-bonded dimers used for crystallization can explain the difficulty to grow crystals and the very high mosaicity of the crystals.

Superposing the A and B molecules of D14C [4Fe–4S] ferredoxin gives an RMS deviation of 0.8 Å, whereas superposing WT [3Fe–4S] ferredoxin onto molecule A and molecule B, respectively, gives an RMS deviation of 0.3 and 0.8 Å, respectively. D14C [4Fe–4S] molecule A thus shows a greater resemblance to WT [3Fe–4S] ferredoxin than to D14C [4Fe–4S] molecule B. However, the differences are subtle compared with those obtained from superposing the structures onto ferredoxins from *Thermotoga maritima* (PBD entry 1VJW) or *D. gigas* (PDB entry 1FXD), which gives RMS deviations of 6–8 Å.

Superposing the structure of D14C [3Fe–4S] *P. furiosus* ferredoxin onto molecule A and molecule B of D14C [4Fe–4S] *P. furiosus* ferredoxin gives an RMS deviation of 0.5 and 0.8 Å, respectively. Superposing D14C [3Fe–4S] *P. furiosus* ferredoxin onto WT [3Fe–4S] *P. furiosus*

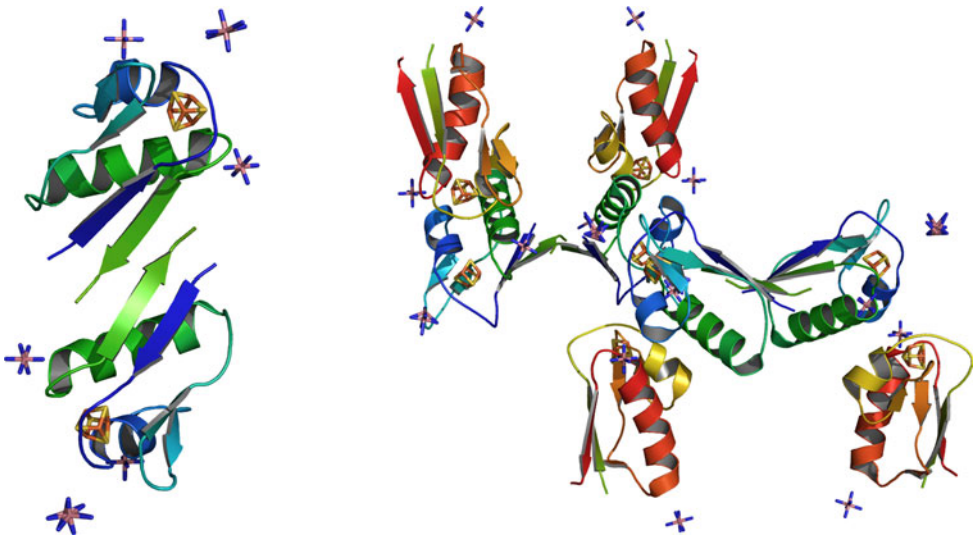


Fig. 4 D14C [4Fe–4S] *P. furiosus* ferredoxin A molecules are shown in *blue* and *green*, whereas B molecules are shown in *red* and *yellow*. Hexaamminecobalt(III) complexes are shown in *blue* and *orange*. The extended *β*-sheet dimer of A–A molecules from two different asymmetric units is shown to the *left* and A–A extended *β*-sheet dimers forming a layer with B molecules on each side is shown to the *right*

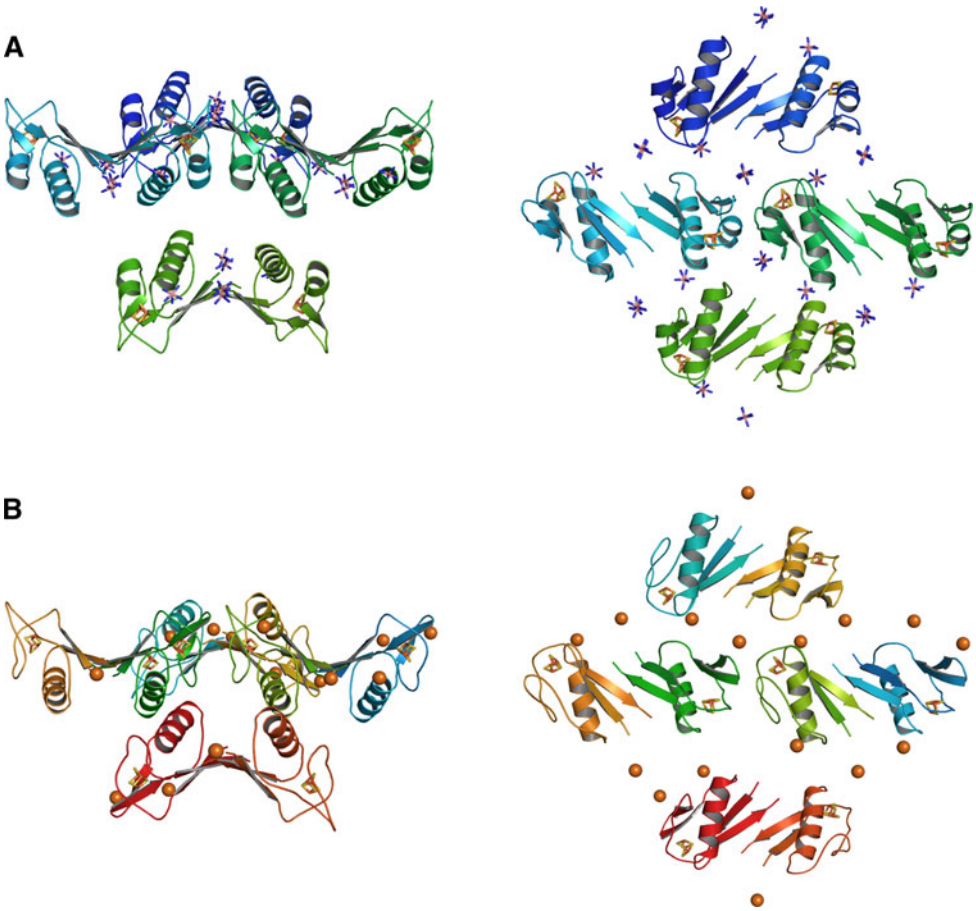


Fig. 5 Crystal packing: side and top views of **a** wild-type [3Fe–4S] *P. furiosus* ferredoxin with hexaamminecobalt(III) complexes shown in *blue* and *orange* (PDB entry 1SJ1 [10]) and **b** D14C [3Fe–4S] *P. furiosus* ferredoxin with cobalt atoms shown in *orange*

ferredoxin (PDB entry 1SJ1) gives an RMS deviation of 0.4 Å. Fewer differences are thus observed between D14C [3Fe–4S] *P. furiosus* ferredoxin and WT [3Fe–4S] ferredoxin than with D14C [4Fe–4S] ferredoxin.

Two crystal structures of [3Fe–4S] ferredoxins with complete cysteine binding motifs –Cys–X_2–Cys–X_2–Cys–X_n–Cys– are available: *D. gigas* [3Fe–4S] ferredoxin (PDB entry 1FXD) and *B. thermoproteolyticus* [3Fe–4S] ferredoxin variant V–II (PDB entry 1WTF). Both of these structures have an additional chemical group bound to the non-cluster-coordinating cysteine and thus no free cysteine is seen [6, 7]. The structure of D14C *P. furiosus* ferredoxin is the first available structure where a cysteine from a complete cysteine binding motif is free and facing away from the cluster.

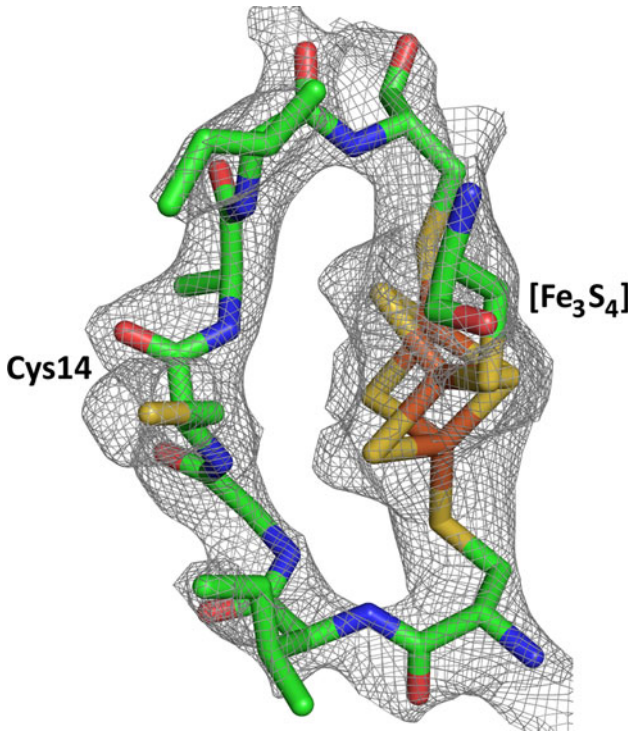


Fig. 6 D14C [3Fe–4S] *P. furiosus* ferredoxin with $2F_{obs} - F_{calc}$ $\sigma = 1.3$ electron density map clearly showing the presence of a [3Fe–4S] cluster and Cys14 facing away from the cluster

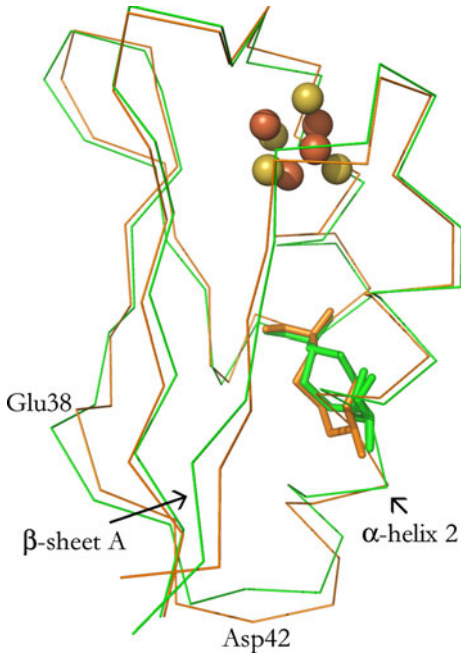


Fig. 7 Superposed structure of D14C *P. furiosus* [4Fe–4S] ferredoxin molecule A (*green*) and molecule B (*orange*). The intramolecular disulfide bonds and the cluster are also shown

The structure of D14C [4Fe–4S] *P. furiosus* ferredoxin

A pronounced difference between the A and B molecules in the asymmetric unit of D14C [4Fe–4S] *P. furiosus* ferredoxin is at the backbone of Asp42, Glu43, and Glu44, where molecule B is displaced 1–2.6 Å from molecule A outwards, leading to α-helix 2 being displaced approximately 1 Å (see Fig. 7). Superposing molecule A and B and leaving out residues with pronounced differences (38–45 and 63–66) gives an RMS deviation of 0.5 Å as compared with 0.8 Å for full-length comparison. Some differences are observed, but generally the molecules are very similar.

Another difference between the two molecules of D14C [4Fe–4S] *P. furiosus* ferredoxin is that molecule B shows a higher mean temperature *B* factor of 18 $Å^2$ compared with molecule A, which has a mean *B* factor of 7.5 $Å^2$ (see Table 2). The difference in mean *B* factors may be an intrinsic property of the molecules or it may be explained by the less restrained position of molecule B compared with molecule A. This is illustrated in Fig. 4, which shows the crystal packing of the molecules in layers: A–A extended β-sheet dimers form a layer, whereas the B molecules are placed in a less rigid position on each side of the extended β-sheet dimer layer.

Also, an analysis of hydrogen bonds in the two molecules in the asymmetric unit of D14C [4Fe–4S] *P. furiosus* ferredoxin shows that molecule A of D14C [4Fe–4S] *P. furiosus* ferredoxin is more similar to WT *P. furiosus* [3Fe–4S] ferredoxin [10] than to molecule B. Table 3 lists the hydrogen bonds that are different between the ferredoxin molecules. The hydrogen bond between Val4 N and Glu38 O is only seen in molecule B and has been suggested to be present in the native *P. furiosus* ferredoxin from NMR analysis [48]. This bond, along with the interactions between Trp2 and Ile40, links β-sheet A in molecule B to a third and shorter strand (see Fig. 8). Hydrogen bond Val4 N–Glu38 O is not seen in molecule A and the backbones of molecules A and B at Glu38 are displaced by 1.5 Å (see Fig. 7). A number of hydrogen-bond differences between molecules A and B in α-helix 2 are listed in Table 3, and these are consistent with an approximately 1 Å displacement of the α-helix.

The intramolecular disulfide bond between Cys21 and Cys48 is in a double conformation in D14C [4Fe–4S] molecule A (40% occupancy of the right-handed and 60% occupancy of the left-handed spiral conformation), whereas in molecule B, it is in a single right-handed spiral conformation. This disulfide bond is also found in a double conformation in the WT [3Fe–4S] ferredoxin. Here, however, the right-handed conformation has 60% occupancy, whereas the left-handed conformation has 40% occupancy [10]. Table 4 lists intramolecular distances related to the disulfide bond. Differences are observed for two of the shorter steric interactions (emphasized in bold in Table 4): in the left-handed conformation the Cys48 Sγ–Leu20 O

Table 3 Possible main-chain hydrogen bonds in D14C [4Fe–4S] *P. furiosus* ferredoxin and comparison with those for wild-type (*WT*) *P. furiosus* [3Fe–4S] ferredoxin

Position	Residues		N···O (Å), this study		N···O (Å), WT [10]
	N	O	Chain A	Chain B	Averaged chains
β-sheet A					
	Val4	Glu38	–	2.8	–
Helix 2					
	Ala52	Cys48a	3.0	–	3.1
		b	2.9		2.9
	Leu45 a	Asp42	3.2	–	3.2
	b		3.2		
	Asn47	Glu44	–	3.2	–
	Lys50	Cys48 a	–	–	–
		b	3.2		3.1
	Glu51	Cys48 a	–	–	–
		b	3.0		2.8
Turn E					
	Ala60	Val58	(3.5)	3.2	(3.5)

a and b refer to alternate conformations of the residue

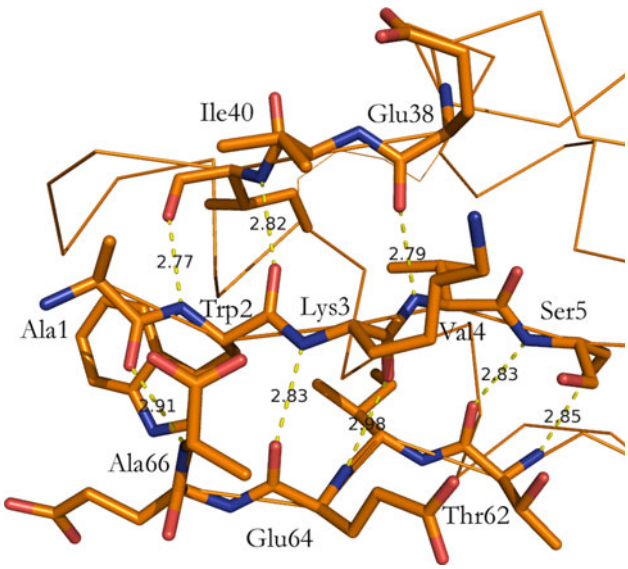


Fig. 8 Hydrogen bonds in β-sheet B of D14C *P. furiosus* [4Fe–4S] ferredoxin molecule B and the three bonds linking the sheet to the backbone at residues 38 and 40

distance is 3.2 Å in D14C molecule A compared with 2.8 Å in WT ferredoxin and in the right-handed conformation the Cys21 Sγ–Val24 Cγ2 distance is 3.1 Å in molecule A compared with 3.3 Å in WT ferredoxin. It appears that the conformation is favored when the shorter distance is more relaxed.

It is also seen in Table 4 that the intramolecular distances related to the disulfide bond generally are more relaxed in molecule B, which is possible since α-helix 2 is shifted approximately 1 Å. The more favorable single right-handed disulfide bond thus involves a loss of hydrogen bonds in α-helix 2 as mentioned earlier. This is accompanied by a displacement of the backbone of Asp42, Glu43, and Glu44 1–2.6 Å outwards compared with molecule A. β-sheet A, which links the A molecules to form an extended β-sheet dimer, also shows a minor displacement (see Fig. 7). Whether the disordered intramolecular disulfide bond and the more tightly packed α-helix 2 and β-sheet A is a consequence of the extended β-sheet dimerization is difficult to judge.

Structure–reduction potential relationship

Sequence alignment of ferredoxins from D14C *P. furiosus*, *T. maritima*, *D. africanus* I, and *B. thermoproteolyticus* is shown in Fig. 9. These ferredoxins all contain a [4Fe–4S] cluster coordinated to the protein by four cysteine residues and were selected because of similarity and the availability of crystal structures with an intact [4Fe–4S] cluster.

The structures of conserved cluster-coordinating cysteine residues are shown in Fig. 10 for [4Fe–4S] ferredoxins from D14C *P. furiosus* molecule A (this work), *D. africanus* I [49], *T. maritima* [50], and *B. thermoproteolyticus* [51]. Figure 10 shows only minor changes in the conformation of cluster-coordinating cysteines and even though residues in between the cysteine residues differ, the backbones show little displacement.

Sγ of the second cysteine in the coordination motif (Cys14 in *P. furiosus* ferredoxin) is the least buried atom of the [4Fe–4S] cluster including ligands [18], and solvent accessibilities of these atoms are given in Table 5 for [4Fe–4S] ferredoxins from D14C *P. furiosus*, *T. maritima*, *D. africanus* I, and *B. thermoproteolyticus* along with reduction potentials of the ferredoxins. In agreement with the literature [17–19], a correlation is seen where less buried [4Fe–4S] clusters have higher reduction potentials. Table 5 also lists the N–H···S bonds of Sγ of cluster-coordinating cysteines, which has been discussed in the literature to have an influence on the reduction potential of [4Fe–4S] clusters [18, 20]. The ferredoxins have equal numbers of bonds with comparable distances. A shift is seen in the reduction potential of D14C versus WT *P. furiosus* ferredoxin as the D14C ferredoxin has a more negative reduction potential in its [4Fe–4S] form and a more positive reduction potential in its [3Fe–4S] form when it is compared with WT *P. furiosus* ferredoxin (see Table 5). Also, in ferredoxin III from *D. africanus*, changing a cluster-coordinating aspartate to cysteine results in a more negative reduction potential in the

Table 4 Intramolecular distances related to the double conformation of the disulfide bond in D14C [4Fe–4S] ferredoxin molecule A and WT *P. furiosus* [3Fe–4S] ferredoxin molecule A and distances in D14C [4Fe–4S] ferredoxin molecule B

	60% left D14C A	40% right D14C A	40% left WT A	60% right WT A	100% right D14C B
Cys48 O–Cys21 Sγ	3.0	3.8	3.1	3.6	3.6
Cys21 Sγ–Val24 Cγ2	5.2	**3.1**	5.6	**3.3**	3.6
Cys48 Sγ–Leu20 O	**3.2**	4.3	**2.8**	4.1	4.4

Shorter distances relevant for shift in occupancies are shown in *bold*

```
D14C P.furiosus   MAWKVSVDQDTCIGCAICASLCPDVFEMNDEGKAQP----KVEVIEDE-ELYNCAKEAME
T.maritima        --MKVRVDADACIGCGVCENLCPDVFQLGDDG--------KAKVLQPE-TDLPCAKDAAD
D.africanus I     MARKFYVDQDECIACESCVEIAPGAFAMDPEIE-------KAYVKDVEGASQEEVEEAMD
B.thermoprot.     -PKYTIVDKETCIACGACGAAAPDIYDYDEDGIAYVTLDDNQGIVEVPDILIDDMMDAFE
                        ** : **.   *   .*. :  . :         :  : :          :* :

D14C P.furiosus   ACPVSAITIEEA----------
T. maritima       SCPTGAISVEE----------
D. africanus I    TCPVQCIHWEDE---------
B.thermoprot.     GCPTDSIKVADEPFDGDPNKFE
                   **. .*   :
```

Fig. 9 Sequence alignment of ferredoxins from D14C *P. furiosus*, *Thermotoga maritima*, *Desulfovibrio africanus* I, and *Bacillus thermoproteolyticus*. Cluster-coordinating cysteines are marked in *yellow* and conserved residues are marked in *blue*

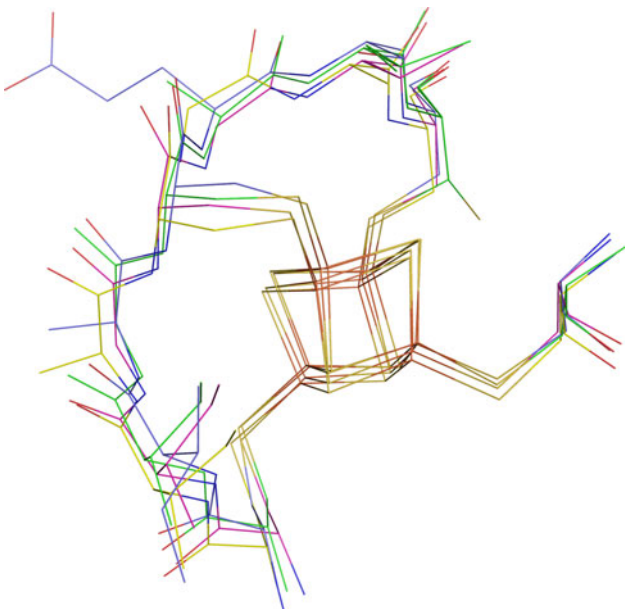


Fig. 10 Superposed structures of the –Cys–X_2–Cys–X_2–Cys– segment in [4Fe–4S] ferredoxins from D14C *P. furiosus* molecule A (*green*, this work), *D. africanus* I (*blue*, PDB entry 1FXR [49]), *T. maritima* (*red*, PDB entry 1VJW [50]), and *B. thermoproteolyticus* (*yellow*, PDB entry 1IQZ [51])

[4Fe–4S] form and a more positive reduction potential in the [3Fe–4S] form [5]. For the [4Fe–4S] cluster-containing ferredoxins, this can be explained by aspartate being a better electron-withdrawing group than cysteine. For the [3Fe–4S] cluster-containing ferredoxins, it can be explained by the pK_a difference between aspartate and cysteine: aspartate is completely deprotonated, whereas cysteine is partly protonated and the increased charge density of the WT ferredoxin with a free aspartate results in a lower reduction potential in the [3Fe–4S] form compared with the D14C ferredoxin with a free (partly protonated) cysteine near the cluster.

Conclusion

When oxidized to contain a [3Fe–4S] cluster, the D14C variant of *P. furiosus* ferredoxin has a free cysteine residue: Cys14. The [3Fe–4S]-containing ferredoxin shows a pH-dependent equilibrium between protonated and deprotonated monomers and formation of a disulfide-bonded dimer at pH 8.0. At pH 5.8, only the protonated Cys14–SH monomer is present as determined by mass spectrometric analysis. Mass spectrometric analysis also confirmed the presence of both the monomer and a dimer consisting of two [3Fe–4S]-containing ferredoxins with an additional intermolecular disulfide bond at pH 8.0. This inhomogeneity is likely to be the reason behind the difficulty in crystallizing D14C [3Fe–4S] *P. furiosus* ferredoxin; however, the structure clearly shows the presence of a [3Fe–4S] cluster.

Molecules A and B in the D14C variant of [4Fe–4S] *P. furiosus* ferredoxin show some distinct differences that can be related to differences seen in the crystal packing and conformation of the intramolecular disulfide bond.

Table 5 Solvent accessibility of Sγ of the second cluster-coordinating cysteine, number of N–H···S bonds to Sγ of cluster-coordinating cysteines and their average distances, and reduction potentials of ferredoxins

	Protein Data Bank entry	E^0 (mV)	Solvent accessibility of second Cys Sγ	No. of N–H···S (average distance Å)
[4Fe–4S] ferredoxin				
P. furiosus		−368 [16]		
D14C *P. furiosus*	2Z8Q	−427 [16]	A/B: 0/0	A/B: 10/10 (3.40/3.44)
Thermotoga maritima	1VJW [50]	−388 [52]	0.6	10 (3.47)
Desulfovibrio africanus I	1FXR [49]	−385 [53]	0.1	10 (3.58)
Bacillus thermoproteolyticus	1IQZ [51]	−280[a] [54]	3.9	10 (3.49)
[3Fe–4S] ferredoxin				
P. furiosus	1SJ1 [10]	−203 [16]		
D14C *P. furiosus*	3PNI	−155[b]		

No structure is available for *P. furiosus* [4Fe–4S] ferredoxin

[a] The reduction potential is for the very closely related *Bacillus stearothermophilus* ferredoxin

[b] Measured at pH 8.0, unpublished results

A β-sheet interaction is present between A molecules in adjacent asymmetric units and these A–A extended β-sheet dimers form a layer in the crystal packing. The B molecules do not show the same interaction and are packed in a less rigid position. Molecule B has the intramolecular disulfide bond in a single right-handed conformation, whereas molecule A has it in a double conformation. More relaxed intramolecular distances related to the disulfide bond in molecule B compared with molecule A are possible owing to a shift through α-helix 2. The shift of α-helix 2 also causes a shift of β-sheet A to a lesser extent and this shift could be related to molecule B not forming an extended β-sheet dimer. Generally, D14C [4Fe–4S] molecule A and WT [3Fe–4S] ferredoxin show a larger resemblance than the A and B molecules in the D14C [4Fe–4S] variant. The conformations of the highly conserved cluster-coordinating cysteine residues in the –Cys–X_2–Cys–X_2–Cys– binding motif of ferredoxins from D14C *P. furiosus*, *T. maritima*, *D. africanus* I, and *B. thermoproteolyticus* are in close resemblance.

Acknowledgments We acknowledge the European Synchrotron Radiation Facility for provision of synchrotron radiation beam time at ID14-3 and ID14-4. This work was supported by grants from the Danish Natural Science Research Council (21-04-0392), the Carlsberg Foundation (ANS-1622/40), and the Danish Council for Independent Research in Technology and Production Sciences (FTP 274-07-0444). Financial support of the work by M.N.L. from the Siemens Foundation is acknowledged. DanScatt is acknowledged for financial support.

References

1. Meyer J (2008) J Biol Inorg Chem 13:157–170
2. Fukuyama K, Matsubara H, Tsukihara T, Katsube Y (1989) J Mol Biol 210:383–398
3. Fukuyama K (2001) In: Messerschmidt A, Huber R, Wieghardt K, Poulos T (eds) Handbook of metalloproteins. Wiley, New York, pp 543–552
4. Conover RC, Kowal AT, Fu W, Park J-B, Aono S, Adams MWW, Johnson MK (1990) J Biol Chem 265:8533–8541
5. Busch JLH, Breton JL, Bartlett BM, Armstrong FA, James R, Thomson AJ (1997) Biochem J 323:95–102
6. Kissinger CR, Sieker LC, Adman ET, Jensen LH (1991) J Mol Biol 219:693–715
7. Shirakawa T, Takahashi Y, Wada K, Hirota J, Takao T, Ohmori D, Fukuyama K (2005) Biochemistry 44:12402–12410
8. Aono S, Bryant FO, Adams MWW (1989) J Bacteriol 171:3433–3439
9. Gorst CM, Zhou ZH, Ma K, Teng Q, Howard JB, Adams MWW, La Mar GN (1995) Biochemistry 34:8788–8795
10. Nielsen MS, Harris P, Ooi BL, Christensen HEM (2004) Biochemistry 43:5188–5194
11. Hasan MN, Hagedoorn PL, Hagen WR (2002) FEBS Lett 531:335–338
12. Hasan MN, Kwakernaak C, Sloof WG, Hagen WR, Heering HA (2006) J Biol Inorg Chem 11:651–662
13. Calzolai L, Zhou ZH, Adams MWW, La Mar GN (1996) J Am Chem Soc 118:2513–2514
14. Zhou ZH, Adams MWW (1997) Biochemistry 36:10892–10900
15. Brereton PS, Duderstadt RE, Staples CR, Johnson MK, Adams MWW (1999) Biochemistry 38:10594–10605
16. Brereton PS, Verhagen MFJM, Zhou ZH, Adams MWW (1998) Biochemistry 37:7351–7362
17. Saridakis E, Giastas P, Efthymiou G, Thoma V, Moulis JM, Kyritsis P, Mavridis I (2009) J Biol Inorg Chem 14:783–799
18. Giastas P, Pinotsis N, Efthymiou G, Wilmanns M, Kyritsis P, Moulis JM, Mavridis IM (2006) J Biol Inorg Chem 11:445–458
19. Backes G, Mino Y, Loehr TM, Meyer TE, Cusanovich MA, Sweeney WV, Adman ET, Sanders-Loehr J (1991) J Am Chem Soc 113:2055–2064
20. Adman E, Watenpaugh KD, Jensen LH (1975) Proc Natl Acad Sci USA 72:4854–4858
21. Dey A, Jenney FE Jr, Adams MWW, Babini E, Takahashi Y, Fukuyama K, Hodgson KO, Hedman B, Solomon EI (2007) Science 318:1464–1468
22. Niu S, Ichiye T (2009) J Am Chem Soc 131:5724–5725
23. Nielsen MS, Harris P, Christensen HEM (2003) Acta Crystallogr D 59:2325–2327

24. Newman J (2003) Acta Crystallogr D 60:610–612
25. Kabsch W (2010) Acta Crystallogr D 66:125–132
26. Kabsch W (2010) Acta Crystallogr D 66:133–144
27. Leslie AGW (1992) Jt CCP4/ESF-EAMCB Newsl Protein Crystallogr 26
28. Evans P (2006) Acta Crystallogr D 62:72–82
29. McCarthy AA, Brockhauser S, Nurizzo D, Theveneau P, Mairs T, Spruce D, Guijarro M, Lesourd M, Ravelli RBG, McSweeney S (2009) J Synchrotron Radiat 16:803–812
30. Collaborative Computational Project, Number 4 (1994) Acta Crystallogr D 50:760–763
31. Vagin A, Teplyakov A (1997) J Appl Crystallogr 30:1022–1025
32. Murshudov GN, Vagin AA, Dodson EJ (1997) Acta Crystallogr D 53:240–255
33. Engh RA, Huber R (1991) Acta Crystallogr A 47:392–400
34. Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Acta Crystallogr D 66:486–501
35. Lohkamp B, Emsley P, Cowtan K (2005) Coot News CCP4 Newsl 42:7
36. Lovell SC, Davis IW, Arendall WB III, de Bakker PIW, Word JM, Prisant MG, Richardson JS, Richardson DC (2003) Proteins Struct Funct Genet 50:437–450
37. DeLano WL (2002) PyMOL. DeLano Scientific. http://www.pymol.org
38. Lee B, Richards FM (1971) J Mol Biol 55:379–400
39. Saff EB, Kuijlaars ABJ (1997) Math Intell 19:5–11
40. Krissinel E, Henrick K (2004) Acta Crystallogr D 60:2256–2268
41. Vagin AA, Richelle J, Wodak SJ (1999) Acta Crystallogr D 55:191–205
42. Li H, Robertson AD, Jensen JH (2005) Proteins Struct Funct Bioinform 61:704–721
43. Bas DC, Rogers DM, Jensen JH (2008) Proteins Struct Funct Bioinform 73:765–783
44. Thompson JD, Higgins DG, Gibson TJ (1994) Nucleic Acids Res 22:4673–4680
45. Duderstadt RE, Staples CR, Brereton PS, Adams MWW, Johnson MK (1999) Biochemistry 38:10585–10593
46. Johnson KA, Verhagen M, Brereton PS, Adams MWW, Amster IJ (2000) Anal Chem 72:1410–1418
47. de Hoffmann E, Stroobant V (2007) Mass spectrometry principles and applications: analytical information. Wiley, Chichester
48. Sham S, Calzolai L, Bren KL, Teng Q, Jenney FE Jr, Brereton PS, Howard JB, Adams MWW, La Mar GN (2002) Biochemistry 41:12498–12508
49. Séry A, Housset D, Serre L, Bonicel J, Hatchikian C, Frey M, Roth M (1994) Biochemistry 33:15408–15417
50. Macedo-Ribeiro S, Darimont B, Sterner R, Huber R (1996) Structure 4:1291–1301
51. Fukuyama K, Okada T, Kakuta Y, Takahashi Y (2002) J Mol Biol 315:1155–1166
52. Smith ET, Blamey JM, Zhou ZH, Adams MWW (1995) Biochemistry 34:7161–7169
53. Hatchikian EC, Cammack R, Patil DS, Robinson AE, Richards AJM, George S, Thomson AJ (1984) Biochim Biophys Acta 784:40–47
54. Mullinger RN, Cammack R, Rao KK, Hall DO (1975) Biochem J 151:75–83

Journal of
Applied Crystallography
ISSN 0021-8898

Received 17 February 2009
Accepted 28 April 2009

MyCrystals – a simple visual data management program for laboratory-scale crystallization experiments

Monika Nøhr Løvgreen,* Mikkel Løvgreen, Hans E. M. Christensen and Pernille Harris

Department of Chemistry, Technical University of Denmark, Kemitorvet 206, DK-2800 Kgs. Lyngby, Denmark. Correspondence e-mail: monika@kemi.dtu.dk

MyCrystals is designed as a user-friendly program to display crystal images and list crystallization conditions. The crystallization conditions entry fields can be customized to suit the experiments. *MyCrystals* is also able to sort the images by the entered crystallization conditions, which presents a unique opportunity to easily assess the effect of, for example, changing pH or concentration and thus establish the best conditions to be used for optimization.

1. Introduction

Macromolecular structure determination by X-ray diffraction requires a successful crystallization. This can be a difficult task where numerous crystallization experiments lead to a large amount of data. Crystallization projects can be worked on by several people over long periods of time, and keeping track of, as well as sharing, results can be made much easier by the use of a database.

There are many available systems and programs that can be used to store crystallization data. In connection with crystallization robots there is usually commercial software available to handle the large amount of data. Some of the free programs are *XtalBase* (Meining, 2006), *XAct* (Brodersen *et al.*, 1999), *Xtrack* (Harris & Jones, 2002), *CLIMS* (Fulton *et al.*, 2004) and *LISA* (Haebel *et al.*, 2001). *XtalBase* (Meining, 2006) is a comprehensive system that amongst many other features can manage and display images of the results, and *CLIMS* (Fulton *et al.*, 2004) is able to store visual observations of the drops. Images combined with scores afford a simple way to assess crystallization results and simplify the optimization procedure because of the transparency of the results. Laboratory information management systems like the *Protein Information Management System* (*PiMS*; https://www.compbio.dundee.ac.uk/pims2_0/) are designed to manage various types of data, such as target, construct and experimental data, through a web interface. We have developed a program named *MyCrystals* with the main purpose of managing and storing images and crystallization conditions and sorting them to give an overview of the results.

MyCrystals is simple to use, easy to navigate and requires minimal time investment by the researcher to utilize the software. The unique feature of *MyCrystals* is its ability to display and sort crystallization images. The images are stored locally or on external media, and the program creates a database with file paths to identify the images. *MyCrystals* is also able to sort the images by the entered data to view the effect of, for example, changing the pH or precipitant. This makes it easier to establish the best conditions and to assess how they can be optimized. *MyCrystals* can also be used as a notebook program to keep track of attempted crystallization conditions without necessarily using the image-viewing part. Furthermore, *MyCrystals* may be useful in other experimental optimization procedures, such as purification, where a lot of slightly different procedures are usually tested. These may be stored in the database and linked to an image file containing the resulting chromatogram.

2. Program description and use

MyCrystals is primarily designed for viewing and sorting images in a simple and intuitive user interface.

The `File` menu is used for creating new databases and loading or saving existing databases. Creating a new database in the `File` menu will auto-generate a standard set of conditions.

The user interface of *MyCrystals* is divided into three simple main tabs: `Filter Data`, `Edit Data` and `Data Table`.

2.1. Data structure

The data consist of two parts. The first part is an XML database that contains all entered conditions including the appropriate image path for each set of conditions (data entries). The second part is the images themselves, which are placed in a directory on the hard drive. *MyCrystals* recognizes graphics file formats JPEG, GIF, PNG and TIFF.

The full path to each image consists of two parts, a project directory and a relative image path. The image path is unique for each data entry, while the project directory is shared. This arrangement facilitates the moving of all images to another location, as only the project directory needs to be changed.

2.2. Filter Data

The `Filter Data` tab is used for viewing and sorting images entered into the database. The sorting options are available in a drop-down menu and consist of all entered data. First, a condition is selected, and then a filter option is applied. Text strings can be sorted as absolutes, while numerical strings also can be sorted using $>$, $<$, $\geq$, $\leq$ and $\neq$. It is possible to apply 15 filters simultaneously. A list of entries that match the criteria is generated and displayed in the tab (see Fig. 1).

The information shown for each data entry in the list can be customized using the field `Filter text`. Writing `[1]` will list the data entered in the first field in the `Data Entry` area of the `Edit Data` tab, `[2]` will list the data in the second field, and so forth.

Selecting one of the listed entries will display the matching image to the right of the tab, as shown in Fig. 1.

2.3. Edit Data

The database can be edited in the Edit Data tab. The tab consists of three parts: Project Directory, Column Properties and Data Entry.

The project directory is the path to a directory where all the images used in the database are placed. The rest of the path is stored as a property (image path) for each data entry.

Column Properties makes it possible to change the data columns of the database. It is possible to delete or add columns, change the names and types, and switch the positions of two columns. An intuitive user interface renders this possible (see Fig. 2).

The Data Entry area is used to add, change and remove data entries in the database. The Selected Data Entry area operates on an existing data entry. To add a new data entry into the database, use the Add Data Entry button, which creates a new data entry with a consecutive index number. The data entry fields are empty when the first data are typed in, and after Add Data Entry has been used, the entry fields will contain the information of the latest data entry. Drop-down menus of previously entered data are available when entering or correcting database entries. The Selected Data Entry area can advantageously be used when entering data, as an entire set of previously entered conditions can be loaded by giving the index number. Minor changes can be applied to the set of conditions and Add Data Entry will generate a new entry in the database.

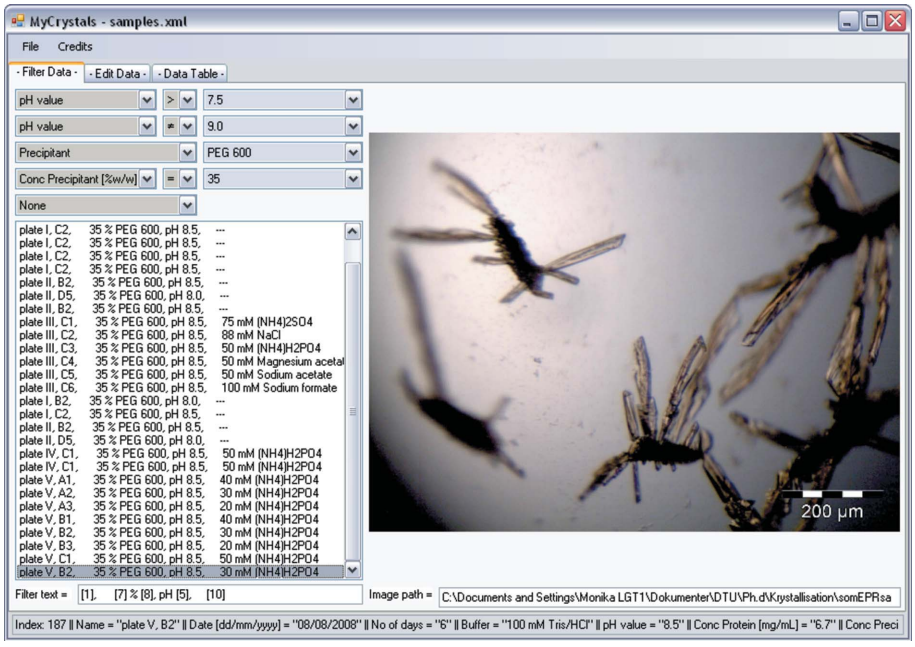


Figure 1
A screenshot showing how the Filter Data tab of *MyCrystals* can appear. Data with pH > 7.5 and pH $\neq$ 9.0 in 35% PEG 600 have been selected. For the displayed image, the conditions are pH = 8.5 and 30 m*M* $NH_4H_2CO_4$. The scalebar is part of the depicted image.

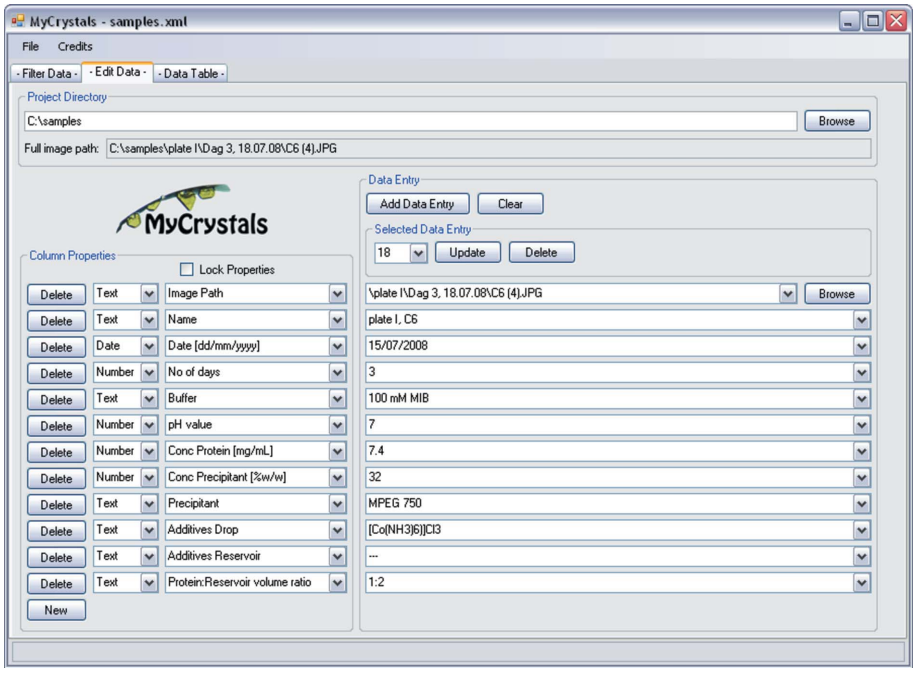


Figure 2
A screenshot showing how the Edit Data tab of *MyCrystals* can appear. Checking the Lock Properties option will simplify the Column Properties by only showing the name.

2.4. Data Table

The Data Table tab lists all entered data to give an overview of the contents of the database.

3. Platforms and availability

The program is written in C# as a Windows Form in Microsoft Visual Studio. The program consists of a single executable file, and no installation is needed. *MyCrystals* runs under Windows and has no special hardware requirements.

MyCrystals is available free of charge for download from http://xray.kemi.dtu.dk/mycrystals/.

4. Conclusions

The unique main feature of the program *MyCrystals*, presented here, is the displaying and sorting of crystallization images based on the entered crystallization conditions. This enables the user to view the effect of changing conditions, which helps to identify parameters for optimization.

References

Brodersen, D. E., Jenner, L. B., Andersen, G. R. & Nyborg, J. (1999). *J. Appl. Cryst.* **32**, 1012–1016.
Fulton, K. F., Ervine, S., Faux, N., Forster, R., Jodun, R. A., Ly, W., Robilliard, L., Sonsini, J., Whelan, D., Whisstock, J. C. & Buckle, A. M. (2004). *Acta Cryst.* D**60**, 1691–1693.
Haebel, P. W., Arcus, V. L., Baker, E. N. & Metcalf, P. (2001). *Acta Cryst.* D**57**, 1341–1343.
Harris, M. & Jones, T. A. (2002). *Acta Cryst.* D**58**, 1889–1891.
Meining, W. (2006). *J. Appl. Cryst.* **39**, 759–766.

APPENDIX BIBLIOGRAPHY

[A1] Moura JJ, Moura I, Kent TA, Lipscomb JD, Huynh BH, LeGall J, Xavier AV and Münck E. Interconversions of [3Fe-3S] and [4Fe-4S] Clusters. Mossbauer and Electron Paramagnetic Resonance Studies of *Desulfovibrio gigas* Ferredoxin II. *J. Biol. Chem.* **(1982)** 257: 6259-67.

[A2] Shakhashiri BZ and Gordon G. The Oxidation of Tris(1,10-phenanthroline)iron(II) Ion by Aqueous Chlorine. *Inorg. Chem.* **(1968)** 7: 2454-6.

[A3] Schilt AA and Taylor RC. Infra-red Spectra of 1:10-phenanthroline Metal Complexes in the Rock Salt Region below 2000 cm^{-1}. *J. Inorg. Nucl. Chem.* **(1959)** 9: 211-21.

[A4] Janin J, Dumas C, Moréra S, Xu Y, Meyer P, Chiadmi M and Cherfils J. Three-Dimensional Structure of Nucleoside Diphosphate Kinase. *J. Bioenerg. Biomembr.* **(2000)** 32: 215-25.

[A5] Yonezawa Y, Izutsu Ki, Tokunaga H, Maeda H, Arakawa T and Tokunaga M. Dimeric Structure of Nucleoside Diphosphate Kinase from Moderately Halophilic Bacterium: Contrast to the Tetrameric *Pseudomonas* Counterpart. *FEMS Microbiol. Lett.* **(2007)** 268: 52-8.

[A6] Lascu I, Giartosio A, Ransac S and Erent M. Quaternary Structure of Nucleoside Diphosphate Kinases. *J. Bioenerg. Biomembr.* **(2000)** 32: 227-36.

[A7] Thompson JD, Higgins DG and Gibson TJ. CLUSTAL W: Improving the Sensitivity of Progressive Multiple Sequence Alignment through Sequence Weighting, Position Specific Gap Penalties and Weight Matrix Choice. *Nucleic Acids Res.* **(1994)** 22: 4673-80.

[A8] Hall T. BioEdit: A User-Friendly Biological Sequence Alignment Editor and Analysis Program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.* **(1999)** 41: 95-8.

[A9] The UniProt Consortium. Ongoing and Future Developments at the Universal Protein Resource. *Nucleic Acids Res.* **(2011)** 39: D214-9.

[A10] Jain E, Bairoch A, Duvaud S, Phan I, Redaschi N, Suzek BE, Martin MJ, McGarvey P and Gasteiger E. Infrastructure for the Life Sciences: Design and Implementation of the UniProt Website. *BMC Bioinformatics* **(2009)** 10: 136.

[A11] Vagin A and Teplyakov A. Automated Program for Molecular Replacement. *J. Appl. Cryst.* **(1997)** 30: 1022-5.

[A12] Afonine PV, Grosse-Kunstleve RW and Adams PD. *CCP4 Newsletter* **(2005)** 42, contribution 8.

[A13] Murshudov GN, Shubák P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, Winn MD, Long F and Vagin AA. REFMAC5 for the Refinement of Macromolecular Crystal Structures. *Acta Cryst.* **(2011)** D67: 355-69.

DTU Chemistry
Department of Chemistry