Studies of ferredoxins and dCTP deaminase:dUTPase

Helt, Signe Smedegaard; Harris, Pernille; Christensen, Hans Erik Mølager; Ooi, Bee Lean

Publication date:
2008

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

Studies of ferredoxins and dCTP deaminase:dUTPase

Part I:
Characterization and crystallization trials of ancient ferredoxins and *Pyrococcus furiosus* ferredoxin

Part II:
Mycobacterium tuberculosis dCTPdeaminase:dUTPase:
Structural implications for inhibition and mechanism of hydrolysis

Ph.D. Thesis by
Signe Smedegaard Helt

Department of Chemistry
Technical University of Denmark

March 2008

Publications

Helt S. S., 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.* 376 (2008) 554-569.

Siggaard J. H. B., Johansson E., Helt S. S., Harris P., Larsen S. and Willemoës M., Structural evidence for a concerted bifunctionality in dCTP deaminase-dUTPase from *Methanocaldococcus jannaschii,* submitted to *J. Mol. Biol.* 2008 (Part II, appendix A.4).

Conference contributions

Christophersen S. (Helt S. S.), Structure of the bifunctional dCTP deaminase:dUTPase from *Mycobacterium tuberculosis*, talk at the 9th Dansync Meeting and 37th Danish Crystallographic Meeting, Denmark, Copenhagen, May 2007.

Christophersen S. (Helt S. S.), Nørgaard H., Jankovics H., Ooi B. L., Harris P. and Christensen H. E. M., Reconstruction and characterization of the evolutionary origin of iron-sulfur proteins: The oldest known protein and its relation to the origin of life, poster at the 8th Dansync Meeting and 36th Danish Crystallographic Meeting, Denmark, Odense, May 2006.

Preface

This dissertation constitutes the result of my three years as a Ph.D.-student at Department of Chemistry, Technical University of Denmark (DTU) from February 2005 to March 2008. Results obtained during a 10 weeks stay at Institute of Structural Biology (IBS, Grenoble, France) in the autumn of 2006 have also been included. The Ph.D.-scholarship was funded by DANSYNC (now part of DANSCATT). Associate Professor Pernille Harris has functioned as main supervisor while Associate Professors Hans E. M. Christensen and Bee Lean Ooi have functioned as co-supervisors; all three from Department of Chemistry, DTU.

The dissertation has been divided into two parts. Part I: Characterization and crystallization trials of ancient ferredoxins and *Pyrococcus furiosus* ferredoxin. Part II: *Mycobacterium tuberculosis* dCTPdeaminase:dUTPase: Structural implications for inhibition and mechanism of hydrolysis.

First of all, I would like to thank my supervisors for their help and guidance throughout this project, special thanks to Pernille Harris for her keen interest in the project and support whenever needed. Also thanks to DANSYNC for providing all of the financial means to perform this project.

I would also like to thank Associate Professor Martin Willemoës (Department of Biology, University of Copenhagen) for including me in his ongoing research on *Mycobacterium tuberculosis* dCTP deaminase:dUTPase.

During a period of 10 weeks I worked in the laboratory of Juan-Carlos Fontecilla-Camps (Director of Laboratory, IBS), who is acknowledged for providing me with the opportunity to work with anaerobic chambers. I would further like to thank the members of his research group for making my stay scientifically interesting and enjoyable. Professor Sine Larsen (Director of Research, European Synchrotron Radiation Facility, ESRF, Grenoble, France) is acknowledged for arranging contact to IBS and arranging accommodations for the first few weeks of my stay.

I would like to thank Professor Jens Ulstrup (Department of Chemistry, DTU) for placing an anaerobic chamber at my disposal.

Managing Laboratory Technician Flemming Hansen (Department of Chemistry, University of Copenhagen) is acknowledged for collecting diffraction data on protein crystals at the ESRF.

Several students and laboratory technicians have assisted me in the laboratory or performed some of the presented experiments under my supervision and I owe them many thanks; Hanne Nørgaard, Lise-Lotte Jespersen, Stefanie Boy, Miriam Grønlund Pedersen, Mette Jensen, Tina Vognsen, Ditte Skibstrup, Zühal Kahveci, Sabine Fantoni and Ivalo Linnebjerg.

Finally, I would like to thank all the members of the Analytical and Bioinorganic Chemistry Group and the Structural Chemistry Group (Department of Chemistry, DTU) for providing a nice and friendly working environment.

Signe Smedegaard Helt
Kgs. Lyngby, 31st of March 2008

Abstract

Part I: Characterization and crystallization trials of ancient ferredoxins and *Pyrococcus furiosus* ferredoxin

Ferredoxins are iron-sulfur proteins that are believed to have evolved during the very early stages of evolution. The first part of this dissertation primarily concerns reconstruction and characterization of a ferredoxin which is believed to be the Last Common Ancestor (LCA) of todays ferredoxins and a ferredoxin traced even further back to the origin of iron-sulfur proteins (Ori-ISP). The sequences of these ferredoxins have previously been determined based on phylogenetic analysis and a theory concerning co-evolution of amino acids and the genetic code. Ori-ISP only contains 23 amino acid residues and has been argued to be at least 3.5 billion years old. Ori-ISP and LCA ferredoxin constitute model systems for gaining further knowledge about cluster-protein interactions in extant ferredoxins as well as in relation to the origin of life and early evolution.

In vitro iron-sulfur cluster reconstitution into the fusion protein Ub-Ori-ISP (Ub for ubiquitin) was performed using different iron sources under strict anaerobic conditions. The reaction products were purified and analysed. When Fe(III) was used, two species were obtained; a monomeric $[4Fe\text{-}4S]^{2+}$ protein and a dimeric protein possibly containing one or two $[3Fe\text{-}4S]^{1+}$ clusters or two different cluster types. However, when Fe(II) was used, two monomeric species were obtained. The dominating species was a $[4Fe\text{-}4S]^{1+}$ protein and a protein containing the oxidized cluster, $[4Fe\text{-}4S]^{2+}$, was also obtained. Attempts to crystallize holo-Ub-Ori-ISP were not successful.

A Ub protease was expressed in *E. coli* cells and a purification procedure was developed. A method for proteolytic cleavage of the fusion protein, Ub-Ori-ISP, and simultaneous cluster reconstitution into Ori-ISP was developed. Using Fe(III) as the iron source yielded results comparable to cluster reconstitution into Ub-Ori-ISP.

Two approaches for obtaining LCA were investigated. In the first approach holo-Ub-LCA ferredoxin was expressed in *E. coli* cells, purified, Ub was cleaved off *in vitro* and holo-LCA ferredoxin was then further purified. In the second approach holo-LCA ferredoxin was purified directly from *E. coli* cells co-expressing holo-Ub-LCA and a Ub protease. The yields of holo-LCA ferredoxin from the two methods were comparable, but very low. The second method is preferable since it was less time-consuming. It was found that holo-LCA ferredoxin contained at least one $[4Fe\text{-}4S]^{2+}$ cluster.

It was finally attempted to crystallize the $[4Fe\text{-}4S]^{2+}$ ferredoxin from the hyperthermophilic archaeon *Pyrococcus furiosus* under strict anaerobic conditions. Crystals in the form of needles were obtained. However, they were too small for diffraction experiments. The

crystallization conditions were difficult to optimize and the temperature appeared to be an important parameter for crystal formation.

Part II: *Mycobacterium tuberculosis* dCTPdeaminase:dUTPase: Structural implications for inhibition and mechanism of hydrolysis

The second part of this dissertation concerns crystallization and structure determination of *Mycobacterium tuberculosis* dCTP deaminase:dUTPase. The bifunctional enzyme, dCTP deaminase:dUTPase, is involved in the biosynthesis of the nucleotide, dTTP, and it catalyses the conversion of dCTP to dUMP. It belongs to a family of enzymes also including the monofunctional enzymes; dCTP deaminase and dUTPase. Studying these enzymes will provide knowledge on the evolutionary relations of the members of this family. Since the bifunctional enzyme is not found in eukaryotes it could be considered a possible drug target in the treatment of e.g. tuberculosis.
dTTP acts as an inhibitor of the enzyme, which was crystallized as the enzyme-dTTP complex and as the apo-enzyme and data were collected to 2.0 Å and 2.4 Å, respectively. Each crystal structure was solved using molecular replacement. The enzyme was also crystallized in the presence of the substrate analogue, α,β-imido dUTP, and data was collected to 3.0 Å. However, the quality of the obtained model was poor and α,β-imido dUTP did not appear in the model.
Comparing the structures of the apo-enzyme and enzyme-dTTP complex, major differences were apparent. Upon binding of dTTP in the active site, the otherwise disordered C-terminal arranges as a lid covering the active site. As a result of structural changes in the active site induced by the methyl group bound to C5 of the thymine moiety of dTTP, the enzyme adapts to an inactive conformation. dTTP is not hydrolysed due to the absence of a water molecule otherwise hydrogen-bonded to O2 of the α-phosphoryl. The inhibition mode of dTTP has been inferred; this gives new insight into the reaction mechanism of bifunctional dCTP deaminase:dUTPases as well as monofunctional dUTPases.

Dansk resumé

Del I: Karakterisering og krystallisationsforsøg af oldgamle ferredoxiner og *Pyrococcus furiosus* ferredoxin

Ferredoxiner tilhører klassen jern-svovl proteiner og menes at være opstået under den tidlige evolution. Den første del af denne afhandling omhandler primært rekonstruktion og karakterisering af to oldgamle ferredoxiner; LCA ferredoxin (”Last Common Ancestor ferredoxin”, sidste fælles forfader ferredoxin) og Ori-ISP (”Origin of Iron-Sulfur Proteins”, ophav til jern-svovl proteiner). Sekvenserne for disse ferredoxiner er tidligere blevet udledt; LCA ferredoxin baseret på fylogenetisk analyse og Ori-ISP baseret på LCA ferredoxin-sekvensen og en teori vedrørende samtidig udvikling af aminosyrer og den genetiske kode. Ori-ISP indeholder kun 23 aminosyrer og hævdes at være mindst 3,5 milliarder år gammel. Ori-ISP og LCA ferredoxin udgør modelsystemer til at opnå mere viden om vekselvirkninger mellem jern-svovl-klynge og protein i nutidige ferredoxiner såvel som i relation til livets oprindelse og tidlig evolution.

Jern-svovl-klyngerekonstruktion i fusionsproteinet Ub-Ori-ISP (Ub for ubiquitin) blev udført *in vitro* med forskellige jernkilder under strengt anaerobe betingelser. Reaktionsprodukterne blev oprenset og analyseret. Med Fe(III) som jernkilde blev der opnået to specier; et monomert $[4Fe\text{-}4S]^{2+}$ protein og et dimert protein indeholdende en eller to $[3Fe\text{-}4S]^{1+}$ klynger eller to forskellige klyngetyper. Når Fe(II) blev brugt, blev der opnået to monomere specier. Den dominerende specie var et $[4Fe\text{-}4S]^{1+}$ protein, og et protein indeholdende den oxiderede klynge, $[4Fe\text{-}4S]^{2+}$, blev også opnået. Forsøg på at krystallisere holo-Ub-Ori-ISP gav ikke nogen brugbare resultater.

En Ub protease blev udtrykt i *E. coli* celler og en oprensningsprocedure udarbejdet. Endvidere blev en metode til at udføre proteolytisk kløvning af fusionsproteinet, Ub-Ori-ISP, og samtidig klyngerekonstruktion i Ori-ISP udarbejdet. Med Fe(III) som jernkilde var de opnåede resultater sammenlignelige med klyngerekonstruktion i Ub-Ori-ISP.

To fremgangsmåder til at opnå LCA ferredoxin blev undersøgt. I den første fremgangsmåde blev holo-Ub-LCA ferredoxin udtrykt i *E. coli* celler, oprenset, fusionsproteinet blev kløvet af *in vitro* og holo-LCA ferredoxin blev videre oprenset. I den anden fremgangsmåde blev holo-LCA ferredoxin oprenset direkte fra *E. coli* celler, der på samme tid udtrykte holo-Ub-LCA ferredoxin og en Ub protease. Udbyttet af LCA ferredoxin fra de to metoder var sammenligneligt, men meget lavt. Sidstnævnte metode er at fortrække, idet den er hurtigere. Holo-LCA ferredoxin indeholdt mindst en $[4Fe\text{-}4S]^{2+}$ klynge.

Det blev endelig forsøgt at krystallisere $[4Fe\text{-}4S]^{2+}$ ferredoxin fra den hypertermofile archaea *Pyrococcus furiosus* under strengt anaerobe betingelser. Krystaller i form af nåle blev opnået,

men de var for små til diffraktionseksperimenter. Krystallisationsbetingelserne var svære at optimere, og temperaturen synes at være en vigtig parameter for krystaldannelse.

Del II: *Mycobacterium tuberculosis* dCTPdeaminase:dUTPase: Strukturelle implikationer for hæmning og hydrolysemekanisme

Den anden del af denne afhandling omhandler krystallisation og strukturbestemmelse af *Mycobacterium tuberculosis* dCTPdeaminase:dUTPase. Det bifunktionelle enzym, dCTP deaminase:dUTPase, er involveret i reaktionsvejen til dannelse af nukleotidet, dTTP, og det katalyserer omdannelsen af dCTP til dUMP. Det tilhører en familie af enzymer, som også inkluderer de monofunktionelle enzymer; dCTP deaminase og dUTPase. Studier af disse enzymer vil tilvejebringe viden om de evolutionære sammenhænge blandt medlemmerne af denne familie. Idet det bifunktionelle enzym ikke findes i eukaryoter, kan det betragtes som et muligt medicinsk mål i behandlingen af f.eks. tuberkulose.
dTTP fungerer som hæmmer for enzymet, der blev krystalliseret som enzym-dTTP kompleks og som apo-enzym. Data blev opsamlet til henholdsvis 2,0 Å og 2,4 Å. Hver krystalstruktur blev løst vha. molecular replacement. Enzymet blev også ko-krystalliseret med det substratlignende molekyle, α,β-imido dUTP, og data blev opsamlet til 3,0 Å. Kvaliteten af modellen var dårlig, og α,β-imido dUTP kunne ikke bygges i elektrontætheden.
Ved sammenligning af strukturerne for apo-enzymet og enzym-dTTP komplekset var store forskelle synlige. Når dTTP bindes i det aktive center, danner den ellers uordnede C-terminal et låg, der dækker det aktive center. Som et resultat af strukturelle ændringer i det aktive center forsaget af methyl gruppen bundet til C5 af thymin delen i dTTP indtager enzymet en inaktiv konformation. dTTP hydrolyseres ikke pga. fraværet af et vandmolekyle, der ellers er hydrogenbundet til O2 fra α-fosforyl. dTTPs strukturelle hæmning af enzymet har givet ny indsigt i reaktionsmekanismen for de bifunktionelle dCTP deaminase:dUTPaser såvel som for de monofunktionelle dUTPaser.

List of abbreviations

Amp	Ampicillin
asu	Asymmetric unit
AU	Absorbance unit
Bis-Tris propane	1,3-bis(tris(hydroxymethyl)methylamino)propane
C. acidiurici	*Clostridium acidiurici*
Cam	Chloroamphenicol
CT	Central trunk
CV	Column volume
Da	Dalton
D. gigas	*Desulfovibrio gigas*
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
E. coli	*Escherichia coli*
EIAV	Equine infectious anaemia virus
EPR	Electron paramagnetic resonance
ESRF	European Synchrotron Radiation Facility
FeS	Iron-sulfur
FIV	Feline immunodeficiency virus
f-Met	N-formylmethionine
HEPES	4-(2-Hydroxyethyl)piperazine-1-ethanesulfonic acid
HPLC	High performance liquid chromatography
IBS	Institute of Structural Biology (Grenoble, France)
IPTG	Isopropyl-*β*-D-thiogalactopyranoside
Kan	Kanamycin sulfate
LB	Luria Broth
LCA	Last Common Ancestor
LMCT	Ligand-to-metal-charge-transfer
milli-q	18.2 MΩcm QPAK Milli-Q water (Millipore)
M. jannaschii	*Methanocaldococcus jannaschii*
MPD	2-methyl-2,4-pentanediol
M. tuberculosis	*Mycobacterium tuberculosis*
MWCO	Molecular weight cut off
NTSB	2-nitro-5-(thiosulfo)-benzoate
OD_{600}	Optical density at 600 nm
Ori-ISP	Origin of Iron-Sulfur Proteins

PEG	Polyethylene glycol
P. furiosus	*Pyrococcus furiosus*
pI	Isoelectric point
RCC	Reductive citrate cycle
rmsd	Root mean square deviation
RNA	Ribonucleic acid
RPC	Reductive pentose phosphate cycle
rpm	Revolutions per minute
S	Spin
SDS-PAGE	Sodium dodecyl sulfate polyacrylaminde gel electrophoresis
Sec	Selenocysteine
SHE	Standard hydrogen electrode
S. typhimurium	*Salmonella typhimurium*
TB	Terrific Broth
Tris	Tris(hydroxymethyl)-aminoethane
tRNA	Transfer RNA
Ub	Ubiquitin
Ub-LCA	Protein fusion of Ub and LCA
Ub-Ori-ISP	Protein fusion of Ub and Ori-ISP
UBP	Ubiquitin specific proteases
UCH	Ubiquitin C-terminal hydrolase
UV-vis	Ultraviolet-visible
YUH1	Yeast ubiquitin hydrolase

The twenty common amino acids are abbreviated according to the standard three-letter or one-letter abbreviations.
Common abbreviations for nucleotides and nucleosides have been used.

Table of contents

Part I

Characterization and crystallization trials of ancient ferredoxins and *Pyrococcus furiosus* ferredoxin

Chapter one

1 Introduction

The first part of this dissertation primarily concerns reconstruction and characterization of an iron-sulfur protein, which is believed to be more than 3.5 billion years old. The protein in question has been named the Origin of Iron-Sulfur Proteins (Ori-ISP) and it is believed to be the evolutionary origin of ferredoxins (a class of iron-sulfur proteins). Reconstruction and preliminary characterization of a Last Common Ancestor (LCA) ferredoxin and crystallization trials of the [4Fe-4S] ferredoxin from *Pyrococcus furiosus* are also presented. The goal of this project was characterization and working towards crystallizing these proteins for crystal structure determination.

Based on phylogenetic comparisons of ferredoxins, the sequence of the LCA ferredoxin was previously derived by B. K. Davis (2002). Based on a theory of amino acid and genetic code evolution (Davis, 1999), LCA ferredoxin was traced even further back and the sequence of a pro-ferredoxin consisting of 23 amino acids, Ori-ISP, was deduced by Davis (2002). An understanding of Ori-ISP and its chemical and biophysical properties might aid in defining criteria for the origin of biological life on Earth. Ori-ISP and LCA ferredoxin can further be viewed as model systems for studying interactions between the iron-sulfur cluster and the protein moiety.

I worked on Ori-ISP during my M.Sc. studies (Christophersen, 2004) and developed a purification scheme for the sulfite-protected apo-protein and a procedure for performing *in vitro* iron-sulfur cluster reconstitution into the fusion protein Ub-Ori-ISP (Ub for ubiquitin). This Ph.D. thesis presents further characterization of Ori-ISP.

1.1 Outline of part I

In contemporary life, iron-sulfur proteins constitute one of the most widespread classes of metalloproteins. They are found in all living organisms and are believed to have evolved during the very early stages of evolution. Iron-sulfur proteins show remarkable structural and functional diversity. Their biological functions range from electron transfer, redox and non-redox catalysis to iron- and superoxide-sensing and gene regulation. Chapter 2 gives a brief overview on iron-sulfur proteins; emphasis is put on the ferredoxins since the characterised proteins belong to this class of iron-sulfur proteins.

How life evolved has puzzled mankind for many years, and an outline of the prevalent theories on this subject is presented in chapter 3. The main focus is put on a theory formulated by B. K. Davis (1999). This theory assumes an autotrophic origin of life which originated on mineral surfaces situated at hydrothermal vents. It suggests that coded amino acids originated from components of the central metabolism and that they co-evolved with the genetic code. How the sequences of LCA ferredoxin and Ori-ISP were derived is also presented.

An introduction to the experimental work is presented in chapter 4. Amongst others the anaerobic techniques employed are described. Chapters 5 to 11 constitute an outline of the experimental work and the obtained results. Chapter 5 presents the optimized schemes for expressing Ub-Ori-ISP in *E. coli* cells and purification of the sulfite-protected protein.
Chapter 6 presents cluster re-constitution into Ub-Ori-ISP and high performance liquid chromatographic purification of the reaction products using Ar/N_2 flushing for obtaining anaerobic conditions. The stabilities of the purified products were tested under various conditions. Electron paramagnetic resonance and Mössbauer spectroscopic measurements were performed on some of the obtained products. In chapter 7 anaerobic chambers were used for preparation and purification of holo-Ub-Ori-ISP. Products obtained were analysed by ultraviolet-visible (UV-vis) spectrophotometry and analytical gel filtration. A method for obtaining hydrolytic cleavage of the Ub-Ori-ISP fusion protein and cluster reconstitution into Ori-ISP followed by purification is presented. Chapter 8 presents crystallization trials of holo-Ub-Ori-ISP.
Ori-ISP as well as LCA ferredoxin were expressed and purified as ubiquitin fusions; Ub-Ori-ISP and Ub-LCA ferredoxin. A ubiquitin hydrolase is necessary for cleavage of these fusions. Chapter 9 presents expression and purification of yeast ubiquitin hydrolase (YUH1). Cleavage experiments on Ub-Ori-ISP and Ub-LCA ferredoxin are presented.
Chapter 10 presents expression, purification and initial characterization of Ub-LCA and LCA ferredoxin. Chapter 11 presents crystallization trials of the [4Fe-4S] ferredoxin from *Pyrococcus furiosus*.
Finally, Chapter 12 offers an overall conclusion of the experimental work and gives an outlook discussing further projects based on the results presented in this thesis.

Chapter two

2 Iron-sulfur proteins

2.1 Introduction to iron-sulfur proteins

Iron-sulfur (FeS) proteins constitute a large class of metalloproteins that occur in all life forms ranging from archaea and bacteria to plants and animals. FeS proteins were not discovered until the early 1960s and today hundreds of proteins and enzymes belonging to this class of proteins have been characterized, the number is still steadily rising (Beinert, 2000; Bian and Cowan, 1999; Imsande, 1999; Johnson,1998; Sticht and Rösch, 1998).

The presence of FeS proteins in ancient organisms and the versatility of the metal center in biological as well as inorganic systems have led to the hypothesis that these proteins evolved during the very early stages of evolution (Imsande, 1999; Meyer, 2008; Rao and Holm, 2004; Sticht and Rösch, 1998).

All FeS proteins contain one or more iron ions, and all (except the rubredoxins) also contain inorganic acid-labile sulfide. The iron ions and the inorganic sulfide constitute the iron-sulfur cluster (FeS cluster), where the iron ions are most frequently coordinated to the protein moiety via cysteinyl sulfurs (Bentrop *et al.*, 2001; Bertini *et al.*, 2007). However, histidine and aspartate have also been observed as naturally occurring ligands (Moulis *et al.*, 1996).

The most common types of FeS clusters are shown in figure 2.1. With only a few exceptions, each of the iron ions in these clusters is surrounded by four sulfur ions in a distorted tetrahedral geometry. This small coordination number is enforced by steric requirements of the large sulfur ions, and an important consequence of this is that the iron ions exclusively occur in a high-spin configuration (Bertini *et al.*, 2007; Cotton *et al.*, 1995).

FeS proteins show remarkable structural and functional diversity. Nature has not only created the cluster types shown in figure 2.1 but more sophisticated variants such as the hybrid cluster (Cooper *et al.*, 2000) and the P-cluster and MoFe cofactor of nitrogenase (Einsle *et al.*, 2002) have been characterized. The majority of the FeS clusters are involved in electron transfer, and they play an essential role in photosynthesis, cell respiration and nitrogen fixation. However, FeS clusters have many other functions. They have been recognized as sites for redox and non-redox catalysis and all-structural functions have also been documented. FeS proteins can serve as sensors of iron, dioxygen and superoxide, and they have also been found to play a role in gene regulation (Beinert *et al.*, 1997; Bentrop *et al.*, 2001; Bian and Cowan, 1999; Flint and Allen, 1996).

FeS proteins have characteristic brown colours due to ligand-to-metal-charge-transfer (LMCT) (Cotton *et al.*, 1995) and many of them are sensitive towards dioxygen. Electron paramagnetic resonance (EPR) spectroscopy and Mössbauer spectroscopy are methods that have been extensively used in inferring FeS cluster compositions and electron localization/

delocalization patterns (Mouesca and Lamotte 1998; Noodleman, 1995; Schünemann and Winkler, 1999).
This chapter focuses on the ferredoxins (a class of FeS proteins) since the proteins analysed in chapters 6 to 11 belong to this class. These small and very acidic proteins exclusively function in electron transport.

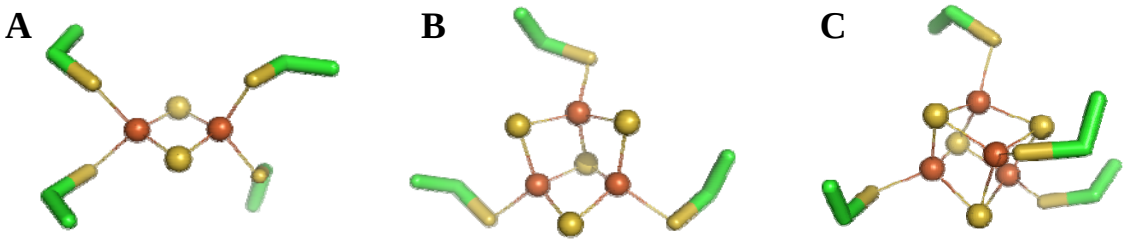


Figure 2.1 The most common cluster types. Iron and inorganic sulfide are shown as bonded spheres in orange and yellow, respectively. Cysteine ligands are shown as sticks; sulfur is yellow while carbon is green. A) The [2Fe-2S] cluster exemplified by the [2Fe-2S] ferredoxin isolated from *Anabaena* species strain PCC 7120 (PDB ID: 1FXA, Rypniewski *et al.*, 1991). B) The [3Fe-4S] cluster of ferredoxin II from *Desulfovibrio gigas* (PDB ID: 1FXD, Kissinger *et al.*, 1991). C) The structure of the [4Fe-4S] cluster exemplified by *Bacillus thermoproteolyticus* [4Fe-4S] ferredoxin (PDB ID: 1IQZ, Fukuyama *et al.*, 2002). The figure was prepared with PyMOL (DeLano W. L., 2002).

2.2 [2Fe-2S] ferredoxins

Ferredoxins containing a binuclear FeS cluster (includes two iron and two inorganic sulfide ions) have molecular weights from 11 to 14 kDa. The cluster-geometry can be described as two tetrahedral FeS_4 units sharing an edge (figure 2.1 A) (Bertini *et al.*, 2007; Zanetti *et al.*, 2001).
The [2Fe-2S] ferredoxins can be subdivided in three groups: the plant-type, the adrenodoxin-type and the *Clostridium*-type (Zanetti *et al.*, 2001). The plant-type ferredoxins are found in plants, algae and cyanobacteria, where they serve as terminal electron acceptors to photosystem I and as electron carriers in other metabolic pathways (nitrogen fixation and sulfite reduction). They are monomers and the cluster binding motif is Cys-x_4-Cys-x_2-Cys-x_n-Cys (Bentrop *et al.*, 2001; Zanetti *et al.*, 2001). The adrenodoxin-type ferredoxins occur in oxygenase systems and they serve as electron donors to cytochrome P450-monooxygenases or to bacterial hydroxylation systems. They are monomers but contain an interaction domain in the structure, which is not found in the plant-type ferredoxins. The cluster binding motif is Cys-$x_{5/3}$-Cys-x_2-Cys-x_n-Cys (Zanetti *et al.*, 2001). The *Clostridium*-type ferredoxins are found only in bacteria, and they are homo-dimers with one [2Fe-2S] cluster per monomer. The distribution of the coordinating cysteine residues in the amino acid sequence differs for the members of this group (Zanetti *et al.*, 2001). The structure of the plant- and adrenodoxin type [2Fe-2S] ferredoxins is characterized by a four-stranded anti-parallel β-sheet covered by an α-helix. This core is surrounded by two short β-strands, an α-helix and a short C-terminal helical turn. The FeS-cluster is located near the surface and is protected by a long loop harbouring three of the coordinating cysteines (Meyer, 2008; Zanetti *et al.*, 2001).
The cluster can exist in two oxidation states, $[2Fe\text{-}2S]^{1+/2+}$, and the reduction potential lies in the range -400 to -100 mV vs. SHE (standard hydrogen electrode). Proteins containing the

oxidized form often display characteristic UV-vis absorptions at about 330, 420, 460 and 550 nm (Zanetti *et al.*, 2001). Molar absorptivities of $[2Fe\text{-}2S]^{2+}$ ferredoxins are approximately 10 $mM^{-1}\cdot cm^{-1}$ at 420 nm (Zanetti *et al.*, 2001). Several examples of UV-vis spectra of proteins containing [2Fe-2S] clusters can be found in the literature (e.g. Chatelet and Meyer, 1999; Martínez-Espinosa *et al.*, 2003, Pan *et al.*, 2003) (these spectra are used for comparisons in chapters 6 and 7).

The $[2Fe\text{-}2S]^{2+}$ cluster contains two Fe^{3+} and the spin for each iron is S = 5/2. These spins are antiparallel and antiferromagnetically coupled with a resulting spin of S = 0, rendering the oxidized state EPR silent (Beinert, 2000; Beinert *et al.*,1997). The colour of the protein fades upon reduction. Mössbauer spectroscopy has shown that the "extra" electron in the $[2Fe\text{-}2S]^{1+}$ cluster is localized on one of the two iron ions. This means that the cluster contains a valence-localized Fe^{2+}-site and a valence-localized Fe^{3+}-site. The spins of the two sites are S = 2 and S = 5/2 respectively. The local spins couple antiferromagnetically to give a resultant spin of S = ½. This form of the protein is recognizable by EPR (Beinert, 2000; Beinert *et al.*,1997; Mouesca and Lamotte 1998).

2.3 Ferredoxins containing cuboidal iron-sulfur clusters

Ferredoxins containing cuboidal clusters are found in various bacteria. This group includes ferredoxins containing one or two [4Fe-4S] clusters, a single [3Fe-4S] cluster or ferredoxins containing one [3Fe-4S] and one [4Fe-4S] cluster (Sticht and Rösch, 1998). The two cluster types are shown in figure 2.1 B and C.

2.3.1 Monocluster ferredoxins

Monocluster ferredoxins contain either a [3Fe-4S] or a [4Fe-4S] cluster in the active site. These proteins often include about 55 to 60 amino acids. The overall fold is conserved and they contain two α-helices and two anti-parallel β-sheets. The placement of the coordinating cysteine residues in the amino acid sequence is also conserved; Cys-x_2-Cys-x_2-Cys-x_n-Cys-Pro. The [4Fe-4S] cluster coordinates to all four cysteines while the [3Fe-4S] cluster does not coordinate to the second cysteine. Most of these ferredoxins also contain two extra cysteines which constitute a disulfide bridge (Fukuyama, 2001; Moura *et al.*, 2001; Sticht and Rösch, 1998).

As a result of the conserved fold the two cluster types are interconvertible under certain conditions. An example is ferredoxins from *Desulfovibrio gigas* (*D. gigas*). Ferredoxin I and ferredoxin II have the same amino acid sequence, but they contain a [4Fe-4S] and a [3Fe-4S] cluster, respectively (Fukuyama, 2001; Moura *et al.*, 2001). Non-cysteine coordination to cuboidal clusters has also been observed in several proteins. An example is ferredoxin from *Pyrococcus furiosus* (*P. furiosus*) (section 2.5) where the second cysteine in the binding sequence has been exchanged with an aspartic acid residue. In this case, the carboxylate group coordinates to the FeS cluster. Cluster interconversion between the [4Fe-4S] and the [3Fe-4S] cluster also occurs (Brereton *et al.*, 1998).

Both cluster types can exist in two oxidation states, $[4Fe\text{-}4S]^{1+/2+}$ and $[3Fe\text{-}4S]^{0/1+}$. The reduction potential falls in the range -650 to -250 mV vs. SHE for the [4Fe-4S] cluster, and -145 to -100 mV vs. SHE for the [3Fe-4S] cluster (Fukuyama, 2001; Moura *et al.*, 2001;

Sticht and Rösch, 1998). The [4Fe-4S] ferredoxins show a characteristic broad UV-vis absorption band at ca. 390 nm (Fukuyama, 2001) while the [3Fe-4S] proteins show an absorption band at about 410 nm (Brereton *et al.*, 1998; Moura *et al.*, 2001). These proteins have molar absorptivities of ca. 16 to 17 $mM^{-1} \cdot cm^{-1}$ at their respective absorption maxima (Brereton *et al.*, 1998; Fukuyama, 2001; Moura *et al.*, 2001). The UV-vis absorption spectra of the *P. furiosus* ferredoxin containing a $[4Fe\text{-}4S]^{2+}$ or a $[3Fe\text{-}4S]^{1+}$ cluster are shown in figure 2.2.

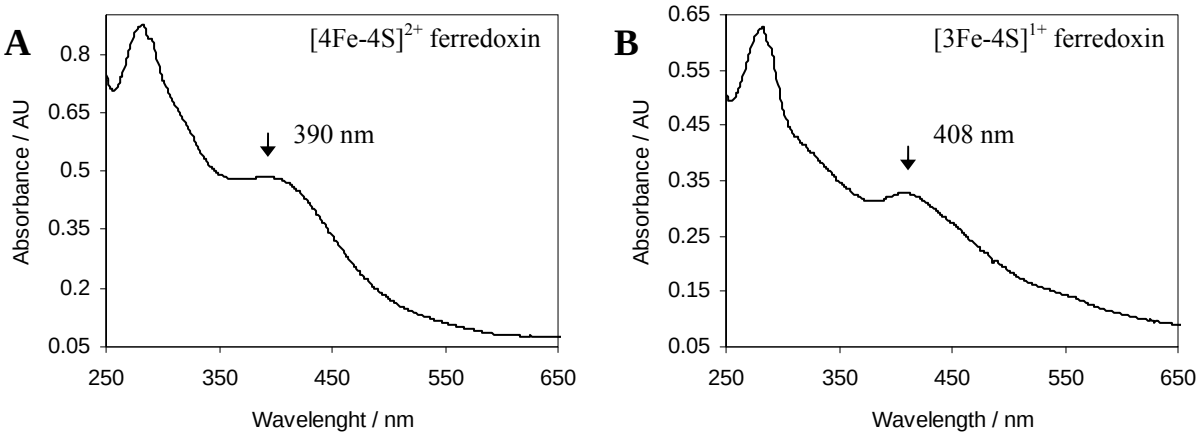


Figure 2.2 UV-vis absorption spectra of oxidized ferredoxins from *P. furiosus*. A) A $[4Fe\text{-}4S]^{2+}$ cluster is present in the ferredoxin. An absorption maximum at 390 nm is evident. B) A $[3Fe\text{-}4S]^{1+}$ cluster is present in the ferredoxin. An absorption maximum at 408 nm is observed.

Mössbauer spectroscopy has shown that the $[3Fe\text{-}4S]^{1+}$ cluster contains three equivalent Fe^{3+}. Each iron has a formal spin of S = 5/2, but the spins couple antiferromagnetically to form an S = ½ and the cluster is EPR active. The $[3Fe\text{-}4S]^{0}$ cluster formally contains one Fe^{2+} and two Fe^{3+}. Mössbauer spectroscopy has shown that the cluster contains one Fe^{3+} and two $Fe^{2.5+}$, which means that the "extra" electron is delocalized between two of the three iron ions. The formal spin of the Fe^{3+} and the delocalized $Fe^{2.5+}Fe^{2.5+}$ pair is S = 5/2 and S = 9/2, respectively. These spins couple antiferromagnetically to form an S = 2 state (Beinert, 2000; Beinert *et al.*,1997; Moura *et al.*, 2001).

The $[4Fe\text{-}4S]^{2+}$ cluster formally contains two Fe^{2+} and two Fe^{3+}. Mössbauer spectroscopy has shown that two mixed valence pairs ($2 \times Fe^{2.5+}Fe^{2.5+}$) are formed, each pair having a spin of S = 9/2. These spins are then antiferromagnetically coupled resulting in an EPR silent cluster. The $[4Fe\text{-}4S]^{1+}$ cluster formally contains three Fe^{2+} and one Fe^{3+}. In this case a mixed-valence pair ($Fe^{2.5+}Fe^{2.5+}$) and a ferrous pair ($Fe^{2+}Fe^{2+}$) are formed which are ferromagnetically coupled to spins of S = 9/2 and S = 4, respectively. The mixed-valence pair and the ferrous pair are then antiferromagnetically coupled resulting in a spin of S = ½ (Beinert, 2000; Beinert *et al.*,1997).

2.3.2 Dicluster ferredoxins

Dicluster ferredoxins contain a core structure characterized by a two-fold symmetrical arrangement of the two clusters in a $(\beta\alpha\beta)_2$ fold (Bentrop *et al.*, 2001; Meyer, 2008).

Dicluster ferredoxins containing two $[4Fe\text{-}4S]^{1+/2+}$ clusters primarily occur in anaerobic bacteria, especially in clostridial species. The primary sequence of these proteins contains about 55 to 60 amino acid residues. In the primary sequence the motif, Cys-x_2-Cys-x_2-Cys-x_3-Cys-Pro, occurs twice, but these cysteines do not constitute the cluster binding motif. In ferredoxin from *Clostridium acidiurici* (*C. acidiurici*), the N-terminal cluster (center I) is coordinated to Cys-8, Cys-11, Cys-14 and Cys-47, while the C-terminal cluster (center II) is coordinated to Cys-18, Cys-37, Cys-40 and Cys-43. Therefore, center I and center II have the binding motifs Cys-x_2-Cys-x_2-Cys-x_n-Cys and Cys-x_n-Cys-x_2-Cys-x_2-Cys, respectively. The reduction potential for center I and center II is ca. -400 mV and -660 mV vs. SHE, respectively (Sieker and Adman, 2001).
Ferredoxins containing one $[3Fe\text{-}4S]^{0/1+}$ and one $[4Fe\text{-}4S]^{1+/2+}$ cluster are thermo- and air-stable and can be handled aerobically for extended periods of time. In some of these proteins the $[3Fe\text{-}4S]^{0}$ cluster can be converted to the $[4Fe\text{-}4S]^{2+}$ cluster by insertion of Fe^{2+} *in vitro*. The primary sequence of these proteins contains about 60 to 110 amino acid residues. The placement of the [3Fe-4S]-coordinating cysteine residues in the amino acid sequence is Cys-$x_{5/7}$-Cys-x_n-Cys, while the placement of the [4Fe-4S]-coordinating cysteine residues is Cys-x_n-Cys-x_2-Cys-x_2-Cys. The reduction potentials are in the range -150 to -450 mV vs. SHE for the $[3Fe\text{-}4S]^{0/1+}$ cluster and -400 to -650 mV for the $[4Fe\text{-}4S]^{1+/2+}$ cluster (Stout, 2001).

2.3.3 Evolution of ferredoxins containing cuboidal clusters

A very large number of ferredoxins have been characterized, and the structural and sequential evidence gathered support the hypothesis that all ferredoxins containing cuboidal clusters originate from a common ancestor (Fukuyama, 2001; Meyer, 2008) (section 3.3).
The most widely accepted model assumes that all extant ferredoxins containing cuboidal clusters originate from an ancestral dicluster (2 × [4Fe-4S]) ferredoxin, which in turn had originated from gene duplication of a sequence encoding a short peptide. Extant dicluster ferredoxins then originated from insertion of C- or N-terminal extensions and Zn binding segments. The dicluster ferredoxins containing 55 residues constitute the core structure of this group (see section 2.3.2) and are believed to be very closely related to the ancestral ferredoxin. Extant monocluster ferredoxins originated from mutations of the cysteines coordinating the second cluster; in some cases a disulfide bridge was the result and in other cases the cysteines were all lost. Insertions of short peptide segments have also occurred (Fukuyama, 2001; Meyer, 2008; Sticht and Rösch, 1998). A crude sketch of the evolutionary events is presented in figure 2.3.

2.4 Core structure analysis of ferredoxins containing cuboidal clusters

To investigate structure-function properties and the properties of the residues in the close vicinity of the FeS-cluster, peptide-based model systems have been applied.
The Ori-ISP and LCA ferredoxin characterized in this project constitute one such model system in relation to evolution of ferredoxins containing cuboidal clusters. Cluster reconstitution into Ori-ISP is examined in chapters 6 and 7.
Cluster reconstitutions into peptide-based models have previously been accomplished and two such model systems are described in this section.

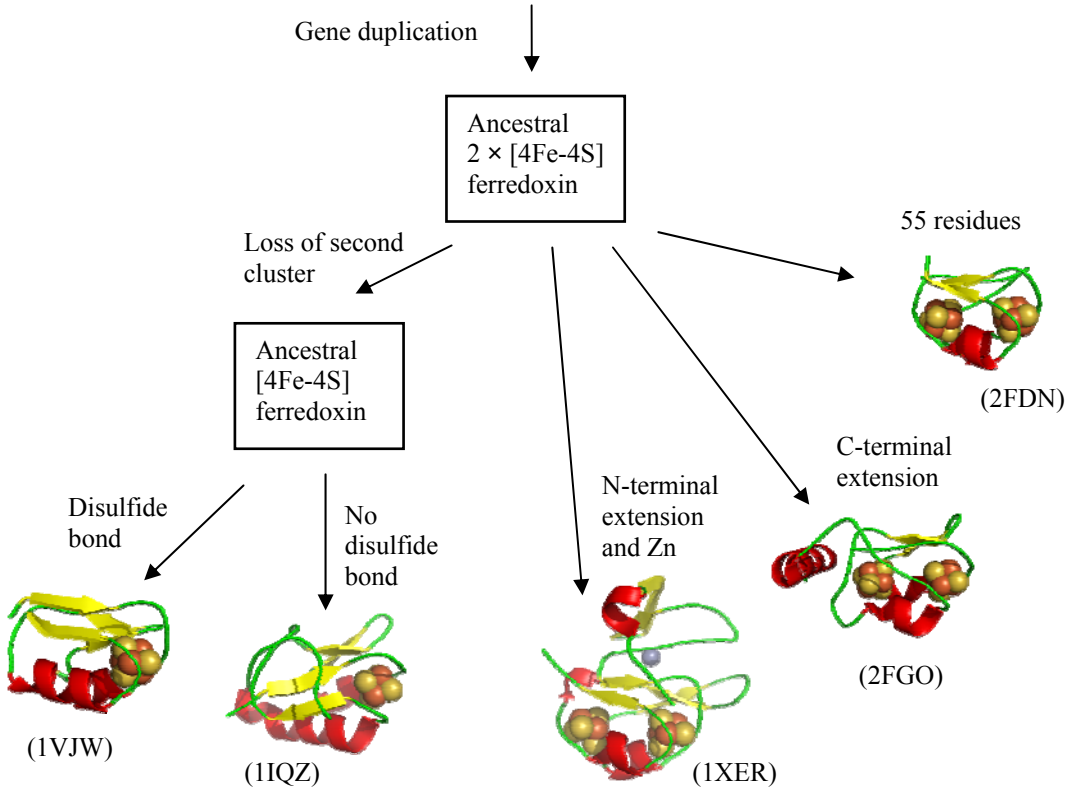


Figure 2.3 Crude sketch of evolution of ferredoxins containing cuboidal clusters. The figure was modified from Fukuyama (2001) and Meyer (2008). Ferredoxin structures were obtained from: *Thermotoga maritima* (PDB ID: 1VJW, Macedo-Ribeiro *et al.*, 1996), *Bacillus thermoproteolyticus* (PDB ID: 1IQZ, Fukuyama *et al.*, 2002), *Sulfolobus tokodaii* (PDB ID: 1XER, Fujii *et al.*, 1996), *Pseudomonas aeruginosa* (PDB ID: 2FGO, Giastas *et al.*, 2006) and *C. acidiurici* (PDB ID: 2FDN, Dauter *et al.*, 1998). Structures are shown in ribbon and α-helices and β-sheets are shown in red and yellow, respectively. Metal centers are shown as spheres; iron is orange, sulfur is yellow and zinc is grey. The structures were prepared with PyMOL (DeLano W. L., 2002).

2.4.1 *Desulfovibrio gigas* mini-ferredoxin

Sow *et al.* (1996) designed a mini-ferredoxin based on the structure and sequence of *D. gigas* ferredoxin II (58 amino acid residues). The model consisted of 31 residues and was made up of residues 3 to 17 and 43 to 56 joined via two residues (Thr and Lys) obtained from a 6-residue loop in trypsin. It was found that the cluster reconstituted mini-ferredoxin had roughly retained the properties from the native ferredoxin, but that it was less stable, suggesting that the entire peptide chain is involved in stabilizing the holo-protein (Sow *et al.*, 1996).

2.4.2 Ferredoxin maquettes

A prototype ferredoxin maquette (figure 2.4) containing 16 amino acid residues based on a consensus sequence of bacterial [4Fe-4S] ferredoxins was generated to probe the ligand and non-coordinating residue requirements for [4Fe-4S] cluster binding (Gibney *et al.*, 1996; Mulholland *et al.*, 1998). Upon cluster reconstitution the ferredoxin maquettte was found to be monomeric and to contain a [4Fe-4S] cluster. From mutational analysis it was found that three cysteines were necessary for formation of a monomeric species containing a [4Fe-4S]

cluster. Further it was found that the fourth cysteine was of particular importance for cluster binding (shown in pink in figure 2.4). Mutating all residues, except the cysteines, into glycines resulted in practically no formation of holo-protein, which shows that not only are the cysteines important, but the "core" sequence is of equal importance (Gibney *et al.*, 1996; Mulholland *et al.*, 1998). The minimum number of residues required to yield [4Fe-4S] cluster reconstitution was found to be the seven residues in the stretch including the three central cysteines (underlined in figure 2.4) (Mulholland *et al.*, 1999). Requirements for the non-cysteine residues in this seven residue peptide were analysed. It was found that isoleucine at the second or at the sixth position (while the remaining non-cysteine residues were mutated to glycine) resulted in cluster formation. Sequence analysis of 510 native [4Fe-4S] proteins revealed that Ile was the most frequent residue at the second position of the seven residue peptide while Gly was most often encountered at positions three and five while the sixth position was rather tolerant of all 20 amino acids, Ala being the most frequent (Mulholland *et al.*, 1999).

NH_2-KL**C**EGG**C**IA**C**GA**C**GGW-$CONH_2$

Figure 2.4 Sequence of ferredoxin maquette. Cysteines are shown in bold. A central seven residue peptide sequence constituting minimum requirements for cluster formation is underlined. The fourth cysteine in the sequence is of particular importance and shown in pink.

2.5 *Pyrococcus furiosus* ferredoxin

The hyperthermophilic archaeon *P. furiosus* lives under extreme conditions and has a growth optimum at 100ºC. The organism was originally isolated from geothermally heated marine volcanic sediments (Fiala and Stetter, 1986). It is a strict anaerobic heterotroph, which grows on carbohydrates and utilizes a modified Embden-Meyerhoff pathway for glycolysis (Kengen *et al.*, 1996).

A single [4Fe-4S] ferredoxin has been purified from *P. furiosus* (Aono *et al.*, 1989) and it was found to act as an electron carrier protein in the glycolysis. It functions as electron acceptor for glyceraldehyde-3-phosphate oxidoreductase and pyruvate oxidoreductase and as electron donor for a transmembrane multi-subunit hydrogenase complex (Silva *et al.*, 2000). The [4Fe-4S] ferredoxin consists of 66 residues and has a molecular weigh of 7.5 kDa including the FeS cluster (Conover *et al.*, 1990). As for the common ferredoxins the cluster exists in two oxidation states, $[4Fe\text{-}4S]^{1+/2+}$. This ferredoxin is highly thermostable and remains unaffected after 12 hours of incubation at 95 ºC (Aono *et al.*, 1989). However, the [4Fe-4S] cluster can be converted to a [3Fe-4S] cluster upon oxidation e.g. by oxygen (Conover *et al.*, 1990). The cluster binding motif differs from common [4Fe-4S] ferredoxins (see section 2.3.1) since the second Cys is replaced by Asp, which has been shown to coordinate the cluster (Calzolai *et al.*, 1995). Upon oxidation to the [3Fe-4S] cluster it is the iron coordinated by Asp which is lost (Conover *et al.*, 1990). UV-vis spectra of both forms are shown in figure 2.2.

The high stability of the ferredoxin along with the lability of the iron coordinated by Asp has rendered the ferredoxin an outstanding model system for studies of heterometallic clusters [M3Fe-4S], where M most often is an element of the first transition series (Conover *et al.*, 1990b; Fawcett *et al.*, 1998; Finnegan *et al.*, 1995; Fu *et al.*, 1994; Staples *et al.*, 1997).

The protein contains a disulfide bridge which, like the [4Fe-4S] cluster, is redox active. This yields four redox states for the protein (Gorst *et al.*, 1995). The disulfide bond is further found in two conformations; a left-handed and a right-handed spiral conformation (Sham *et al.*, 2002). The ferredoxin has been reported to be both a monomer (Conover *et al.*, 1990) and a dimer (Hasan *et al.*, 2002), this is elaborated and discussed in chapter 11. The crystal structure of the ferredoxin containing a [3Fe-4S] cluster has been solved (figure 2.5 A and B), which revealed a β-sheet interaction between two monomers supporting the suggestion that the ferredoxin could in fact be a functional dimer (Nielsen *et al.*, 2004). The overall fold is that of the family (section 2.3.1). The crystal structure of the D14C [4Fe-4S] variant has also been solved (Johannesen *et al.*, in preparation). The overall structure is similar to the structure of the native protein except for the difference in clusters (figure 2.5 C). The mutant ferredoxin further contained two molecules in the asymmetric unit; one generating a dimer when imposing crystallographic symmetry and the other generating a monomer. Crystallization of the native [4Fe-4S] ferredoxin is the subject of chapter 11. Solving the structure of the [4Fe-4S] native form will provide a detailed knowledge on Asp coordination and on structural and geometric differences of cluster and peptide chain enforced by the Cys/Asp-14 ligand.

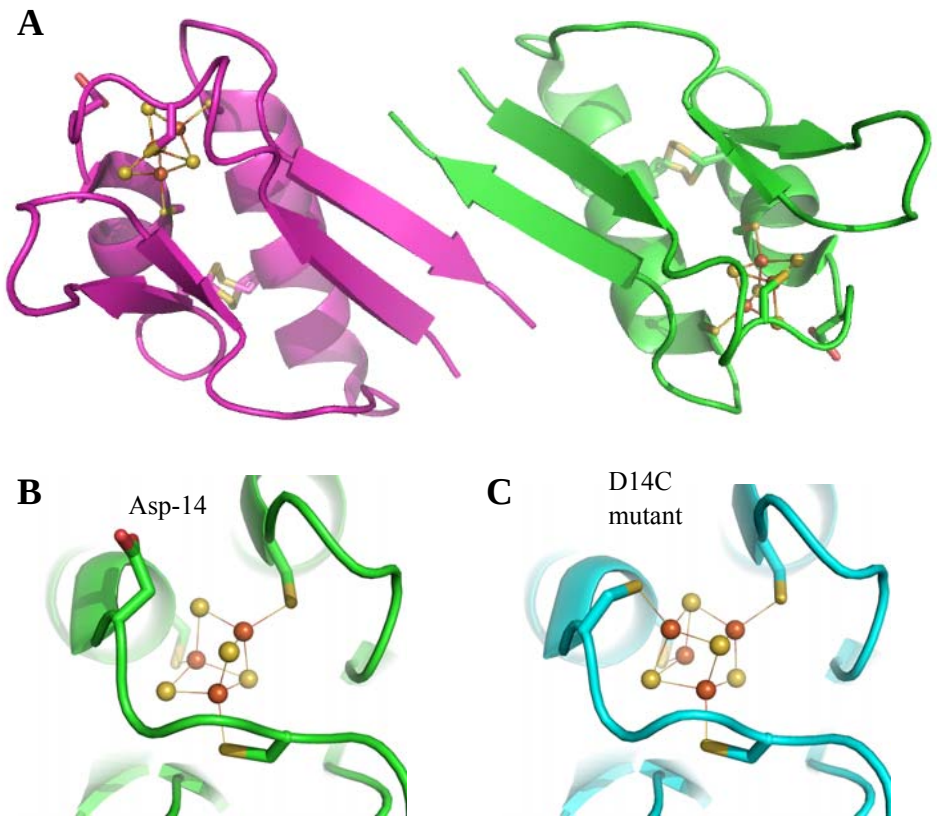


Figure 2.5 Crystal structures of *P. furiosus* ferredoxin. A) Structure of the [3Fe-4S] ferredoxin (PDB ID: 1SJ1, Nielsen *et al.*, 2004). The protein is a dimer in the asymmetric unit. The two monomers are connected via a β-sheet interaction. The monomers are shown in ribbon view in pink and green. The five cysteines present in the protein are shown as sticks and the FeS cluster of each monomer is shown as bonded spheres; iron is orange while sulfur is yellow. Asp-14, which coordinates the fourth iron in the [4Fe-4S] form, is also shown. Two conformations of the disulfide bond are evident. B) Close-up of the cluster in A. C) Close-up (same point of view as in B) of the [4Fe-4S] cluster of the D14C mutant (PDB ID: 2Z8Q, Johannessen *et al.*, in preparation). The protein is shown in cyan. The figure was prepared with PyMOL (DeLano W.L., 2002).

Chapter three

3 On the origin of life, evolution of the genetic code and ancient ferredoxins

How did life evolve? This question has always concerned the human race. How the chemosphere of early Earth was transformed to a biosphere is the key to understand the origin of life and to search for extraterrestrial life or life signatures.

Several theories and ideas on this subject have been put forward and extensive amounts of literature can be found on this subject. In this section the two most prevalent theories are presented; the prebiotic soup model and the autotrophic origin of live. A theory concerning co-evolution of amino acids and the genetic code (Davis, 1999) is also presented. Based on this latter theory and a phylogenetic analysis of ferredoxins the sequences of a Last Common Ancestor (LCA) ferredoxin and a pro-ferredoxin (Ori-ISP) have been deduced, and Ori-ISP have been dated back billions of years (Davis, 2002).

3.1 The origin of life

3.1.1 Early Earth

Radioactive dating indicates that Earth was formed about 4.6 billion years ago. Due to impacts of objects large enough to exterminate life on Earth today, the surface was very hot (500-1000ºC) for about half a billion years. According to geochemical estimations, Earth had cooled to about 100ºC, 3.9 billion years ago, giving rise to the formation of $H_2O(l)$, which is essential for the formation of life. Alternating evaporation/condensation processes must have led to forceful atmospheric phenomena such as electrical discharges and storms. The Earth was constantly irradiated by ultraviolet light since the ozone layer did not exist at that time and volcanic eruptions were frequent (Campbell and Farrel, 2003; Nisbet and Sleep, 2001; Orgel, 1998; Rode, 1999; Sleep *et al.*, 1989).

The atmosphere on early Earth was very different from the one we know today and it has probably gone through several stages before reaching its current composition. It is generally believed that the atmosphere was composed of NH_3, H_2S, CO, CO_2, CH_4, N_2, H_2, $H_2O(g)$ and $H_2O(l)$. However there is no agreement on the relative amounts of these compounds (Campbell and Farrel, 2003). The atmosphere of early Earth has been postulated to be everything from strongly reducing to slightly oxidizing (Berry, 2002; Edwards, 1998; Kasting, 1993; Orgel 1998).

The oldest known cell-like structures are 3.5 billion years old fossils located in minerals. This means that only a few hundred million years were available for the evolution of mono-cellular structures from simple inorganic compounds (Furnes *et al.*, 2004; van Zuilen *et al.*, 2002).

3.1.2 The prebiotic soup model

In the 1920s, A. I. Oparin (1938) and J. B. S. Haldane (1929) proposed that lightning could cause the molecules of the primordial reducing atmosphere to react and form simple organic compounds. These reactions were experimentally demonstrated by S. Miller in 1953 (Miller, 1953). He simulated the effects of lightning in the primordial atmosphere, subjecting a mixture of $NH_3(g)$, $CH_4(g)$, $H_2(g)$ and $H_2O(l)$ to electric discharges in a refluxing system for about a week. The resulting solution contained simple organic compounds like formaldehyde and hydrogen cyanide and several naturally occurring amino acids. These experiments were repeated under varying conditions (different compositions of the reducing "atmosphere", and ultraviolet radiation) leading to other molecules like nucleic bases and sugars (Campbell and Farrel, 2003; Orgel, 1998; Voet and Voet, 2003). These experiments were widely accepted as a heterotrophic model for synthesis of the bio-monomers needed for the creation of proteins and RNA. According to this model, the oceans slowly attained the consistency of a thin primordial soup of bio-monomers, which then somehow would react to create polypeptides and polynucleotides (Orgel, 1998; Voet and Voet, 2003).

A variant of the heterotrophic model assumes that the organic molecules of the primordial soup were delivered to earth by meteorites that bombarded Earth after it was formed (Orgel, 1998; Rode, 1999).

3.1.3 An autotrophic origin of life

An autotrophic (self-sustaining) origin of life would not be dependent on global conditions, but on local. Supporters of an autotrophic origin of life believe that life began on mineral surfaces. These surfaces were able to arrange primitive catalysts and substrates in simple arrays, and via a mineral-linked energy transduction mechanism, the first metabolic pathways were created (Edwards, 1998). Iron sulfide minerals, especially pyrite (FeS_2), have been considered as possible sites for primordial processes. Pyrite has a positively charged surface which is capable of binding not only inorganic anions, but also organic molecules containing anionic groups e.g. carboxylate and thiolate (Wächtershäuser, 1992).

G. Wächtershaüser (1990 and 1992) has proposed that life began chemoautotrophically on pyrite surfaces situated at hydrothermal vents and he has suggested an archaic reductive citrate cycle to be the origin of metabolism and bio-molecules. He suggested that the energy for reduction of CO_2 (CO_2 being the precursor molecule of bio-molecules) and for driving this cycle was the oxidative formation of FeS_2 from FeS and H_2S/HS^-. According to Wächtershaüser's theory, complex metabolic cycles self-organized on mineral surfaces and the products did not escape the surface. The reactions suggested by Wächtershäuser may be thermodynamically possible, but the thermodynamic calculations have been exposed to some critique (Kalapos, 2007) and the reactions have never been experimentally proved. Studies on CO_2 fixation and reaction on mineral surfaces have been carried out. The outcome of these studies was, that the reactions are highly unlikely due to large overpotentials involved in the fixation of CO_2, which means that further energy was necessary to carry out these reactions (Tributsch *et al.*, 2003).

Wächtershaüser has modified his original ideas to concern CO instead of CO_2 (Wächtershäuser, 2006). It was found that co-precipitated NiS and FeS could convert CO and CH_3SH into the thioester CH_3COSCH_3, which then hydrolysed to CH_3COOH. If the solution

was added a catalytic amount of selenium, CH_3SH and CH_3COOH were formed from CO and H_2S alone. NiS, FeS, CO and H_2S are found in the exhaust of hydrothermal vents. Therefore the above mentioned reactions can be considered to have participated in a chemoautotrophic origin of life (Huber and Wächtershäuser, 1997). Experiments have also shown that amino acids can form peptide bonds by activation with CO in the presence of NiS and FeS, and that a CO driven degradation of the peptides simultaneously proceeds, resulting in a peptide cycle. The ability of a primordial organism to recycle its components is necessary, otherwise functionless peptides would become a sink for valuable amino acids (Huber and Wächtershäuser, 1998; Huber *et al.*, 2003). The formations of α-hydroxy and α-amino-acids by CO fixation on catalytic transition metal precipitates have also been shown (Huber and Wächtershauser, 2006).
A variation of the pyrite theory is proposed by M. R. Edwards (1998) who has suggested that life could have originated on pyrite but that the origin was photoautotrophic. Studies of the photo-electrochemical behaviour of pyrite show that this mineral has photoelectric properties. Light absorbed by this mineral generates photocurrents and photo-potentials across the pyrite-liquid interface. If pyrite is put in contact with an $Fe^{2+/3+}$ electrolyte, it can convert light into photo-electrochemical energy (Tributsch, 2003). Edwards suggests that photoelectrons could have been used to reduce compounds in ancestral pathways (Edwards, 1998).

3.1.4 Self-replicating systems

Living systems/organisms have the ability to self-replicate. However, knowledge on the evolution of a self-replicating system is very limited (Orgel, 1998).
A discovery which has had deep implications for this discussion is that RNA is capable of catalyzing its own formation. This has led many scientists to believe that life originated from an "RNA world" in which RNA molecules functioned both as genetic material and as catalysts (de Lucreazia *et al.*, 2007). This system somehow evolved to a point when it was able to encode the synthesis of proteins, which are more effective catalysts. And at an even later stage, DNA took over as the genetic material (Campbell and Farrell, 2003). This hypothesis does not explain how the nucleotides of RNA originated and how they were able to polymerize. Polymerization could have proceeded on clay minerals. It has been found that montmorillonite, which is positively charged clay that is believed to have been plentiful on early Earth, is able to catalyze the formation of short RNA molecules (Orgel, 1998). It has also been proposed, that RNA was not the first genetic material, but that a simpler replicating system preceded RNA (Orgel, 1998). It has further been suggested that life appeared when different systems, some involved in metabolic cycles and others involved in replication, combined and entered into some sort of symbiotic process (Campbell and Farrell, 2003).

3.1.5 Membranes and cells

Lipid membranes are characteristic of extant life and they play a crucial role in energy metabolism of cells. How these cell boundaries arose and what their constituents were, is unknown (Berry, 2002). It has been proposed that cell membranes emerged very early in evolution. Several researchers think that the formation of macromolecules would have been assisted or even dependent on the compartmentalization that membranes would provide (Berry, 2002).

The main constituent of extant biomembranes is phospholipids, but the spontaneous formation of these under primordial conditions is not likely. Organic matter from meteorites have been found to form vesicles in water, which indicate that other molecules may have been suitable for the formation of these membranes. One theory says that the origin of the first cells was related to formation of cell membranes from terpenoids (Berry, 2002).
Minerals could have functioned as primordial cells. Hydrothermally formed iron sulfide chimneys (hydrothermal vents) have been found to be "three-dimensional cells" and not just two-dimensional surfaces. These minerals contain inner compartments which are incompletely separated from one another by membrane like sheaths of iron sulfide precipitate. One could imagine that life began inside these structures and that reactions could have been catalyzed on the iron sulfide surfaces (Martin and Russel, 2003).

3.2 Amino acid synthesis and evolution of the genetic code

B. K. Davis, who is a supporter of Wächtershaüsers theory, has proposed that primordial amino acid synthesis took place on mineral surfaces and that the genetic code evolved simultaneously (Davis, 1999). In this section the main aspects of Davis' theory are presented.

3.2.1 Primordial amino acid synthesis

According to Davis' theory, coded amino acids originate from branch reactions at various points in central metabolism. The components from which the amino acids originate include sugar derivatives and small carboxylated molecules. Under the anaerobic conditions that were found on early Earth, these compounds constituted the reductive citrate cycle (RCC), the reductive pentose phosphate cycle (RPC) and the central trunk (CT) that links these two cycles (figure 3.1). The RCC and RPC are autocatalytic CO_2 fixing cycles and they are believed to have taken place on mineral surfaces. The energy necessary for the reactions was obtained from a coupling of the endergonic reduction of CO_2 to the exergonic formation of pyrite. These reactions are based on ideas proposed by Wächtershäuser (1990 and 1992) and the correctness of these is questionable (section 3.1.3).
The RCC, RPC and CT contributed to the formation of coded amino acids (Davis, 1999). The number of amino acids originating from each system was used to identify which part of central metabolism that was first linked to protein synthesis and formation of the genetic code. The more a system contributes, the earlier in evolution it was involved in amino acid synthesis. The number of reactions in the biosynthetic pathway needed to synthesize an amino acid gives a hint to when that particular amino acid was formed. If an amino acid requires a large number of reaction steps it was formed at a late stage in evolution. Based on the above assumptions, the RCC was seen as the first cycle linked to protein synthesis (elaborated below). When counting the path length for the formation of an amino acid, no addition was made to the path length for the distance between components in the same cycle because these cycles predate protein synthesis. RCC was the first cycle and therefore it forms the basis of counting. The RPC is a distinct cycle and it is linked to the RCC via the CT, consequently two steps are added to amino acids synthesized from RPC components and one step is added to amino acids synthesized from the CT (figure 3.1) (Davis, 1999).

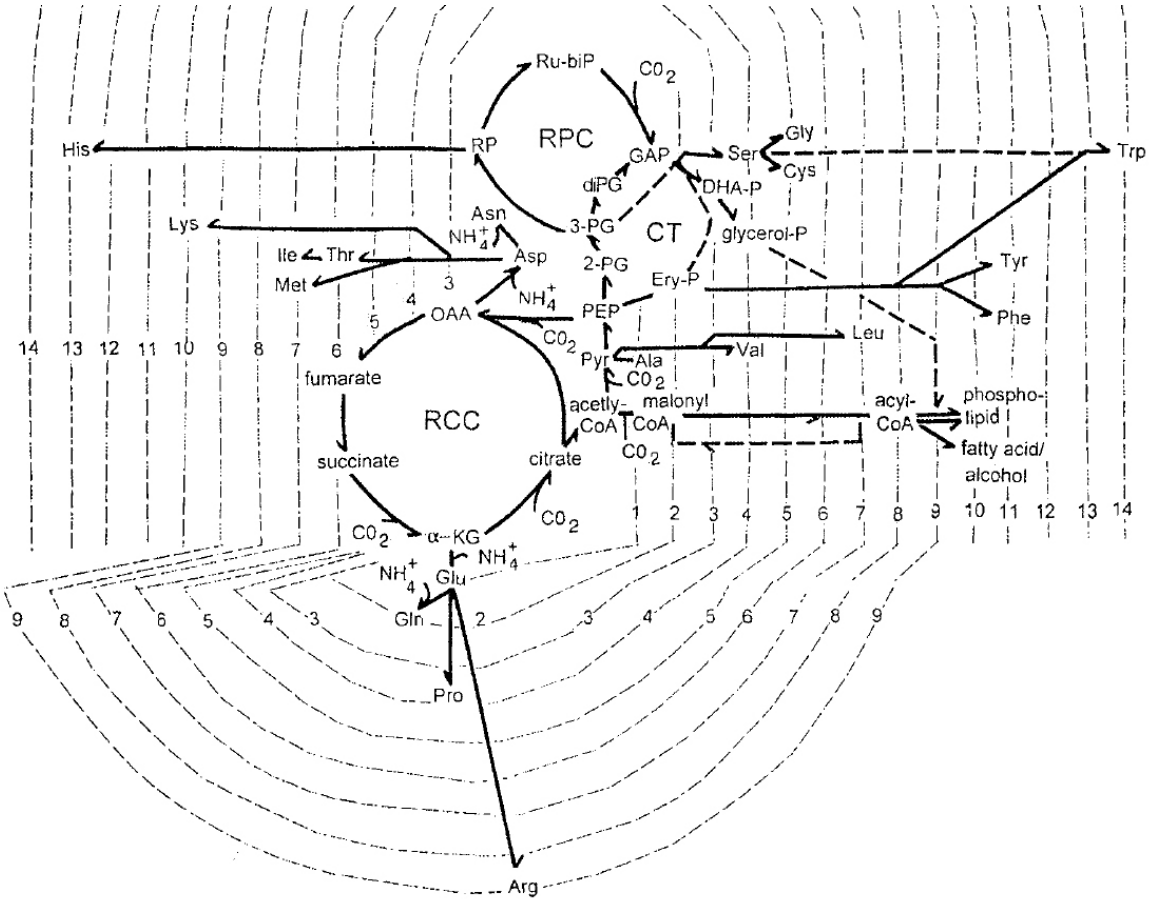


Figure 3.1 Biosynthetic pathways for amino acids. Coded amino acids originate from components of central metabolism; RCC, RPC and CT. Reprinted from Davis (1999) with permission from Elsevier Limited. The reader is referred to the original literature for details and abbreviations concerning these cycles.

Twenty α-amino acids (including selenocysteine, Sec) and one imino acid constitute the proteins of archeons and eukaryotes. Bacteria contain an additional amino acid; N-formylmethionine (f-Met). All 21 amino acids occur in all life forms, because they were incorporated into proteins before divergence from the Last Common Ancestor (LCA).

Table 3.1: Each amino acid originates from a component of central metabolism. The amino acids form six families and each family has a different precursor. The CT and RPC share Phe, Tyr and Trp but in counting the stage of addition to the genetic code, they are credited to the RPC (Davis, 1999).

Metabolic cycle	RCC			CT	CT/RPC	RPC
Precursor molecule	Oxalo acetate	α-keto-gluterate	Pyruvate	Phospho-glycerate	Phosphoenolpyruvate + ribose-5-phosphate	Ribose -5-phosphate
Amino acid	Asp Asn Thr Ile Met f-Met Lys	Glu Gln Pro Arg	Ala Val Leu	Ser Gly Cys Sec	Phe Tyr Trp	His

The amino acids form six families, and each family has a different precursor in central metabolism (table 3.1). 13 common amino acids and f-Met, originate from components of the

RCC. Six common amino acids and Sec originate from two components of the CT, and four amino acids originate from two components of the RPC (table 3.1). As the source of the highest number of amino acid combined with the shortest path lengths of amino acid synthesis, the RCC is considered to have had the longest association with protein synthesis. Hence, the RPC has had the shortest association with protein synthesis (Davis, 1999).
As stated above, the time of entry of each amino acid into the code is estimated from the number of reactions in its biosynthetic pathway. This number of reactions is therefore equal to the stage where a particular amino acid entered the genetic code, presented in table 3.2.
No amino acids were added to the genetic code at stage 3, 8 and 12. If one compares the path length of amino acid synthesis and the properties of the side chain, a pattern emerges (table 3.2). The first amino acids to enter the code were acidic or hydrophilic (stage 1 and 2); their side chains being carboxylate or amide groups. The next amino acids to enter the code (stage 4-7) display side chains of increasing hydrophobicity. And the last amino acids to enter the code (stage 9-14) display basic or aromatic side chains (Davis, 1999).
Path length analysis has also been applied to the synthesis of small amphipathic molecules, fatty acids, and phospholipids. Their formation is placed about 10 reactions (stage 10) from the RCC, giving rise to the formation of cells (Davis, 1999).

Table 3.2: The stage in evolution when each amino acid was added to the code. The colours indicate: **acidic**, **hydrophilic**, **hydrophobic**, **basic** and **aromatic** amino acids.

Stage	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Amino acid	**Asp Glu**	**Asn Gln**		**Ala Pro Ser Val**	**Cys Gly**	**Thr**	**Leu Ile Met**		**Arg**	**Lys**	**Phe Tyr**		**His**	**Trp**

3.2.2 Evolution of the genetic code

According to Davis (1999), the genetic code and amino acid synthesis co-evolved and this evolution was divided into three stages. Biosynthetically related amino acids often have related codons and a detailed examination of these relationships assisted in reconstruction of codon evolution. Davis has based codon assignment on the biosynthetic pathways of amino acid synthesis and tRNA phylogeny in extreme thermophiles and related archaean species. Consequently, the first part of the theory (section 3.2.1) supports the second part of the theory and vice versa (the main conclusions are described here) (Davis, 1999).
The genetic code evolved during several millions of years from including only a few codons to include all 64 codons. In the first stage of code evolution four sets of codons (a total of 16) constituted the genetic code. These codons coded for stage 1 and 2 amino acids (table 3.2). As only one fourth of all codons were assigned, mutations to unreadable codons often occurred. The second stage of code evolution comprised amino acids synthesized from stage 4 to 7. During this stage, the size of the code grew rapidly to include all 64 codons, and the number of coded amino acids increased from 4 to 14. Expansion of the code decreased the risk of mutation to an unassigned codon and chain termination (Davis, 1999). In the first two stages residues included in the genetic code captured unassigned codons. In the third stage (the post-expansion phase), incoming amino acids (synthesized from stage 9 to 14) needed to capture already assigned codons. These amino acids were basic or aromatic residues (table 3.2).

Capture of an already assigned codon required that the advantages contributed by an incoming amino acid exceed the cost of replacing the outgoing amino acid. The outgoing amino acid was already present in proteins, and changing this amino acid could change or destroy the function of a particular protein. Codon-reassignments that took place in the third stage led to the formation of the basic code, presently existing in animal mitochondria (Davis, 1999). Changes that affected single codons were further made in the basic code and these completed the formation of the standard code (the code of extant organisms). The standard code is not "frozen", and variant codes have formed after the divergence of archaea, bacteria and eukarya. This means that the genetic code continues to evolve (Davis, 1999).
The evolution of the genetic code and formation of living organisms took place within a few hundred million years. According to Davis, the code evolved in about 20 million years, which is almost instantaneous on a geological time scale. This result is based on tRNA divergence rates and error rates in RNA replication.

3.3 Ancient ferredoxins

This section provides an outline of how Davis (2002) deduced the sequences of Ori-ISP and LCA ferredoxin. Based on phylogenetic studies on ferredoxins, the sequence of LCA ferredoxin was constructed. The LCA ferredoxin was traced further back to a stage 5 protein that contained 23 amino acid residues; Ori-ISP (Davis, 2002).
From a completed study on ten different proteins and enzymes from more than 800 phylogenetic distinct sources, which included proteins like ferredoxins, reverse transcriptases, topoisomerases and ribonucleotide reductases, it was found that ferredoxins could be traced the furthest back in time. As a consequence of this study, Ori-ISP is believed to be the oldest known protein (it had the earliest residue profile).

3.3.1 Phylogenetic analysis of ferredoxin

To obtain sequences of ferredoxins, a representative of this class of proteins was chosen: *Clostridium pasteurianum* (*C. pasteurianum*) ferredoxin. It contains 55 amino acid residues and two [4Fe-4S] clusters (Davis, 2002).
A sequence homology search was performed, and 82 sequences were obtained. 60 sequences came from bacteria, while the remaining sequences came from archaea. The sequence identity with residues in the aligned reference sequence was between 29 and 89 percent. Based on these sequences, the phylogenetic trees of archaea and bacteria were constructed using the Neighbour-Joining method (Davis, 2002).

3.3.2 The Last Common Ancestor ferredoxin

Based on the phylogenetic analyses, ancestral sequences for ferredoxins from archaea and bacteria were constructed; each sequence contained 55 amino acid residues (figure 3.2). 29 of the 55 sites in the ancestral ferredoxin from archaea were resolved to a single amino acid. Doublets were found at four sites, and 22 sites were resolved to three or more amino acids. The ancestral ferredoxin from bacteria was resolved to 32, 10 and 13 sites with one, two and three or more amino acid residues, respectively. A closer look at the sequences revealed that each sequence consisted of two almost identical 23-residue segments (residue 1-23 and 30-

52), which indicated that a duplication of "a first" primary structure took place (section 2.3.3). These two segments were separated by six residues (24-29) and each sequence had a C-terminal "tail" consisting of three residues (53-55) (Davis, 2002). The sequences depicted in figure 3.2 were even further resolved when taking the 23-residue repeat into account. The resolved sequences are shown in figure 3.3.

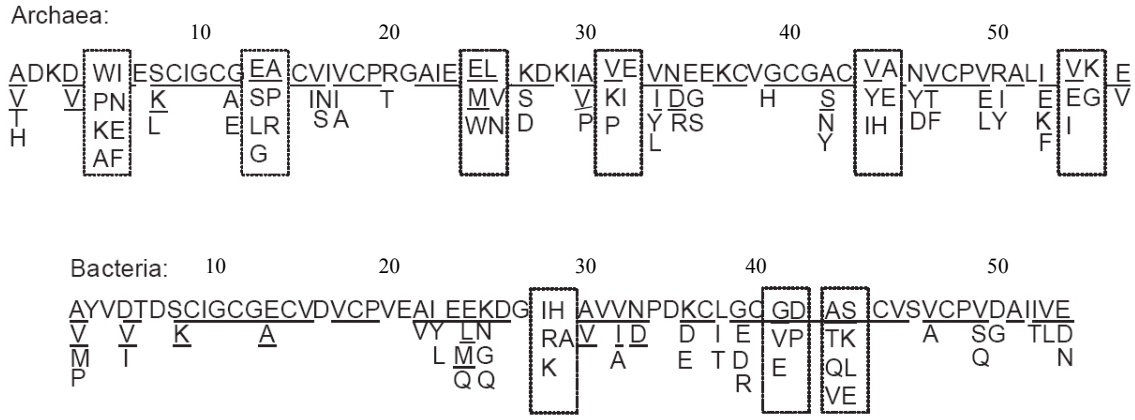


Figure 3.2 Ancestral sequences for ferredoxins from archaea and bacteria. Underlined amino acids are conserved between the two domains. Each box indicates one amino acid resolved to five or more residues (the sequences are not aligned). Reprinted from Davis (2002) with permission from Elsevier Limited.

Because a stage 10 amino acid (Lys) was conserved at one site in the 23-residue repeat, the duplication that must have occurred, could not have happened prior to stage 10 in evolution of the genetic code. Figure 3.3 shows how the two ancestral ferredoxin sequences arose from the sequence of a Last Common Ancestor ferredoxin (LCA ferredoxin). The amino acid sequence of the LCA ferredoxin was determined from the sequence of the two ancestral ferredoxins using a minimum of nucleotide mutations. Seven single and two double-base changes were needed to link the LCA ferredoxin and bacterial ferredoxin, and three single and four double-base changes were needed to link the LCA and archaean ferredoxin. In both cases a total of eleven base changes were necessary, placing bacteria and archaea equally far from the LCA. The majority of these mutations took place in the N-terminal of the internal repeat (Davis, 2002). Due to the high similarity with extant dicluster ferredoxins, LCA ferredoxin most likely share their fold and properties (section 2.3.2).
Chapter 10 presents expression, initial purification and characterization of LCA ferredoxin.

3.3.3 Ori-ISP

The LCA ferredoxin was traced even further back by Davis (2002). By removing the central 6-residue insert and the 3-residue C-terminal "tail", a protein containing 46 amino acids was obtained. This protein contained two identical 23-residue segments. Prior to duplication, the protein contained these 23 amino acids. This was a stage 10 protein (figure 3.3). Evidence of an earlier protein was found in the sequence, because early amino acids were dominant. Amino acids synthesized at stage 5 or earlier, accounted for 87 % of the residues in the stage 10 protein, but in extant ferredoxin this percentage was found to be only 73. Three amino

acids in the stage 10 protein were post-stage 5 amino acids; Ile (stage 7) at sites 9 and 23 and Lys (stage 10) at site 7 (Davis, 2002).
To reconstruct a stage 5 protein, these three amino acids had to be replaced by stage 5 or earlier amino acids. The codons specifying Lys (AAR, R = purines) can be changed to Glu codons (GAR) following a first base A:G substitution. In the ancestral sequence for bacteria, Glu occurred at site 36 (equivalent to site 7). Consequently Glu was placed at site 7 in the stage 5 protein. The codons specifying Ile (AUY and AUA, Y = pyrimidines) can be changed to Val codons (GUY and GUA) following a first base A:G substitution. In the ancestral sequence for archaea, Val occured at site 38 (equivalent to site 9). Consequently Val was placed at site 9 in the stage 5 protein. Val was also placed at site 23; it did however not occur in ancestral sequences at site 23 or 52, but in their descendants. These substitutions created the stage 5 protein, Ori-ISP (figure 3.3) (Davis, 2002).

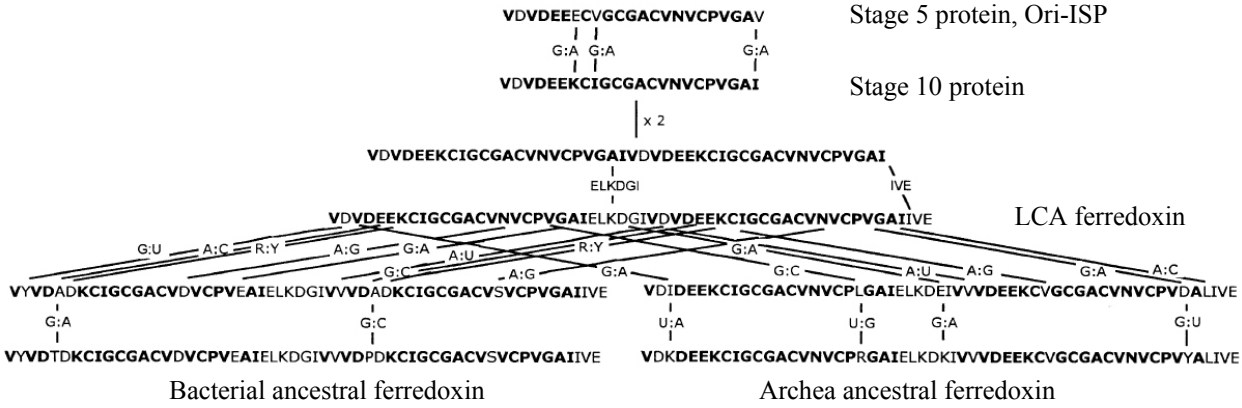


Figure 3.3 Evolution of ferredoxins in prokaryotes. The oldest known protein, Ori-ISP, was formed at stage 5 in the evolution of the genetic code. Three G to A transitions in the gene coding for Ori-ISP, led to the formation of a stage 10 protein. The gene was then somehow duplicated. A central 6-residue segment and a C-terminal 3-residue segment were inserted creating a LCA ferredoxin. After formation of the LCA, divergence of species occurred. Seven single and two double-base changes formed ancestral bacteria ferredoxin and three single and four double-base changes formed ancestral archaea ferredoxin. Re-printed from Davis (2002) with permission (Elsevier Limited).

3.3.4 Properties of Ori-ISP

The sequence of Ori-ISP is shown in figure 3.4 and it consists of 23 amino acids (the molecular weight is 2266 Da). The protein contains eight different amino acids; Val, Asp, Glu, Cys, Gly, Ala, Asn and Pro. Ori-ISP has a highly acidic N-terminal while the C-terminal contains polar and non-polar residues indicating that these have had different functions. The C-terminal contains four cysteines at site 8, 11, 14 and 18, which might have coordinated a [4Fe-4S] cluster that could have functioned as an electron transport center. The iron and sulfide needed for cluster formation could have originated from pyrite/hydrothermal vents. The N-terminal region could have anchored Ori-ISP to (positively charged) surfaces.
Since Ori-ISP is a stage 5 protein, it originated prior to the formation of cells, which according to Davis did not take place until stage 10 in evolution. Geological findings conclude that cellular organisms were present 3.5 billion years ago (section 3.1.1), which mean that Ori-ISP is at least this old.

VDVDEEE**C**VG**C**GA**C**VNV**C**PVGAV

Figure 3.4 The sequence of Ori-ISP. Red indicates acidic amino acids and cysteines are highlighted in bold.

The proposed sequence of Ori-ISP provides support for the chemoautotrophic model for the origin of life, and the early amino acid residue profile found in the LCA ferredoxin supports Davis' theory on evolution of the genetic code (Davis, 2002).
Chapters 6 to 8 concerns cluster reconstitution, purification and characterization of Ori-ISP.

3.3.5 Ori-ISP models

Model structures for Ori-ISP have been predicted by Kasper P. Jensen (Post doc at Department of Chemistry, DTU) and they are presented here with his permission.
C. pasteurianum ferredoxin (see section 3.3.1) was used to generate the models. The first 23 residues from the sequence of this ferredoxin align with Ori-ISP to a sequence identity of 43 % (the four cysteines are conserved) and the structure (PDB ID: 1CLF, Bertini *et al.*, 1995) was used as a template for generating two models. Upon creation of the models, residues of this stretch were mutated to the corresponding Ori-ISP residues (Thompson, 2004), hydrogens were added and the structure was geometry optimized with the UFF force field (Rappe, 1992). Two possibilities for cluster coordination were considered as constraints during formation of the models; an open and a closed form. In the open form only the first three cysteines in the sequence were forced to coordinate a [4Fe-4S] cluster (this coordination mode derives directly from the template). In the closed form the fourth cysteine was forced to coordinate the unsaturated iron ion. The calculated models are shown in figure 3.5. These models show great differences due to the different cluster coordination modes. It might be envisaged that the two forms are both present in solution and that an equilibrium exists. The open form provides the possibility of dimerization and this is further discussed in chapter 7.

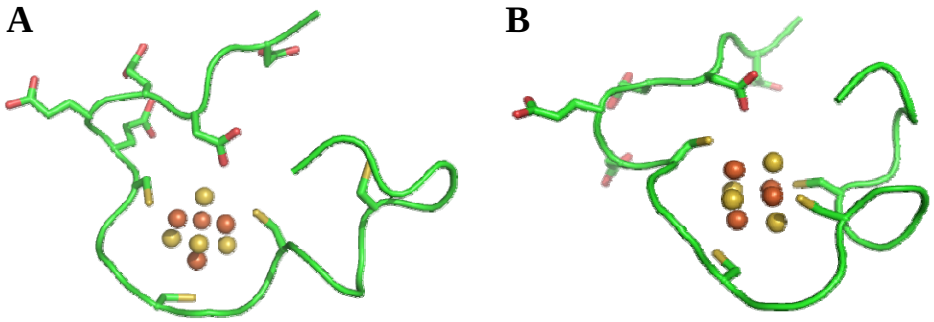


Figure 3.5 Models of Ori-ISP. A) The open conformation. The first three cysteines in the sequence coordinate the [4Fe-4S] cluster. B) The closed conformation. All four cysteines coordinate the [4Fe-4S] cluster. The models are shown in ribbon view. The side chains of the four cysteines and the five acidic residues are shown as sticks while the clusters are shown as spheres. Carbon, sulfur, oxygen and iron are shown in green, yellow, red and orange, respectively. The figure was prepared with PyMOL (DeLano W. L., 2002).

Chapter four

4 Introduction to the experimental work

This chapter presents an overview of some of the experimental techniques and instrumentation used throughout the project. Further, the *Escherichia coli* (*E. coli*) strains used for expressing the proteins and enzymes of interest are introduced.

4.1 Inert gas flushing techniques

Two different approaches to obtain anaerobic conditions were applied: constant Ar/N_2 flushing or anaerobic chambers (section 4.2).

One approach to remove oxygen from purification buffers and vessels was to keep a constant flow of Ar or N_2 to flush out O_2. This was accomplished using the system shown in figure 4.1 A. Glass vessels were fitted with Suba seal rubber septa. Teflon tubings were inserted through the rubber septa via stainless steel tubes. Ar or N_2 was used to deoxygenate solutions or vessels. Ar/N_2 was bubbled into buffer solutions using a moderate gas flow for one hour (per 100 ml) prior to application. Glass vessels used for reactions were flushed prior to use. Syringes were purged with N_2, or Ar several times prior to use. In the case of protein solutions, the gas stream was kept above the surface, because direct flushing will cause the proteins to denature. Columns were equilibrated and run using deoxygenated buffers. When high performance liquid chromatography (HPLC) was employed buffers were kept in closed vessels and a constant slight over-pressure or constant flushing using Ar or N_2 was applied during purification (figure 4.1 B).

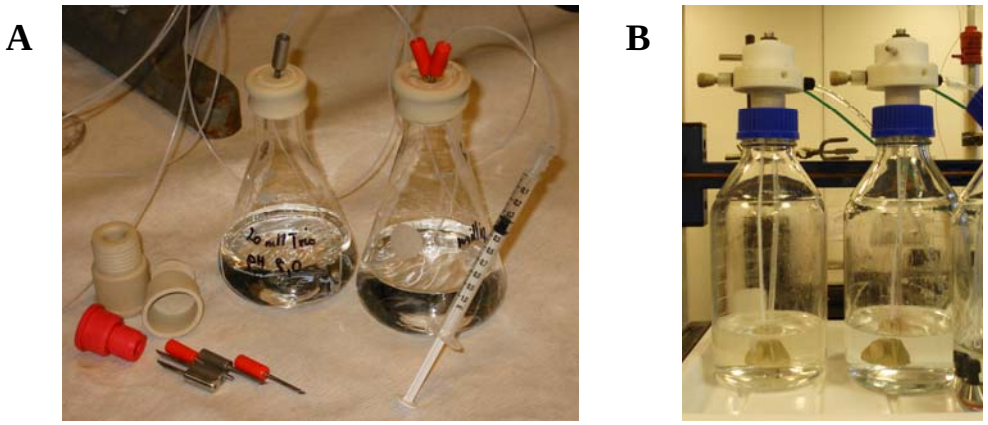


Figure 4.1 The system used for obtaining oxygen-free conditions through Ar/N_2 flushing. A) Glass vessels were fitted with Suba seal rubber septa and flushing was performed via Teflon tubings. B) During HPLC purifications buffers were kept in closed vessels and a slight overpressure or constant flushing was applied.

4.2 Anaerobic chambers

During my 10-weeks stay at Institute of Structural Biology (IBS, Grenoble, France), I had the opportunity to perform protein purification and protein crystallization experiments inside anaerobic chambers. These chambers along with working procedures and techniques used at IBS are described in section 4.2.1. Upon returning home an anaerobic chamber was equipped to perform protein purification and protein crystallization based on the observations I had made in France. This chamber is described in section 4.2.2.

4.2.1 Anaerobic chambers at Institute of Structural Biology

Two anaerobic chambers (both fabricated by Jacomex) were available. One chamber was employed for protein purification while the other was utilized for protein crystallization (figures 4.2 and 4.3). The working gas consisted of N_2 which contained ca. 3 % H_2. Chambers were supplied with a steady flow of gas to ensure a constant slight overpressure of < 5 mbar (pressure sensors were installed). Excess gas was removed by an oil safety valve. Both chambers were equipped with Cu-catalysts for removal of oxygen ($Cu + ½O_2 \rightarrow CuO$). Constant circulation of the gas through the catalyst was performed to ensure continuous removal of oxygen. Regeneration of the catalyst was performed with H_2 ($H_2 + CuO \rightarrow H_2O + Cu$). The composition of the working gas allowed continuous regeneration of the catalysts at room temperature. Gas was supplied via copper tubing from large tanks situated outside the building. Both chambers were equipped with an oxygen sensor and a refrigerating unit for maintaining the temperature at 20 ± 1ºC. A small and a large vacuum antechamber was

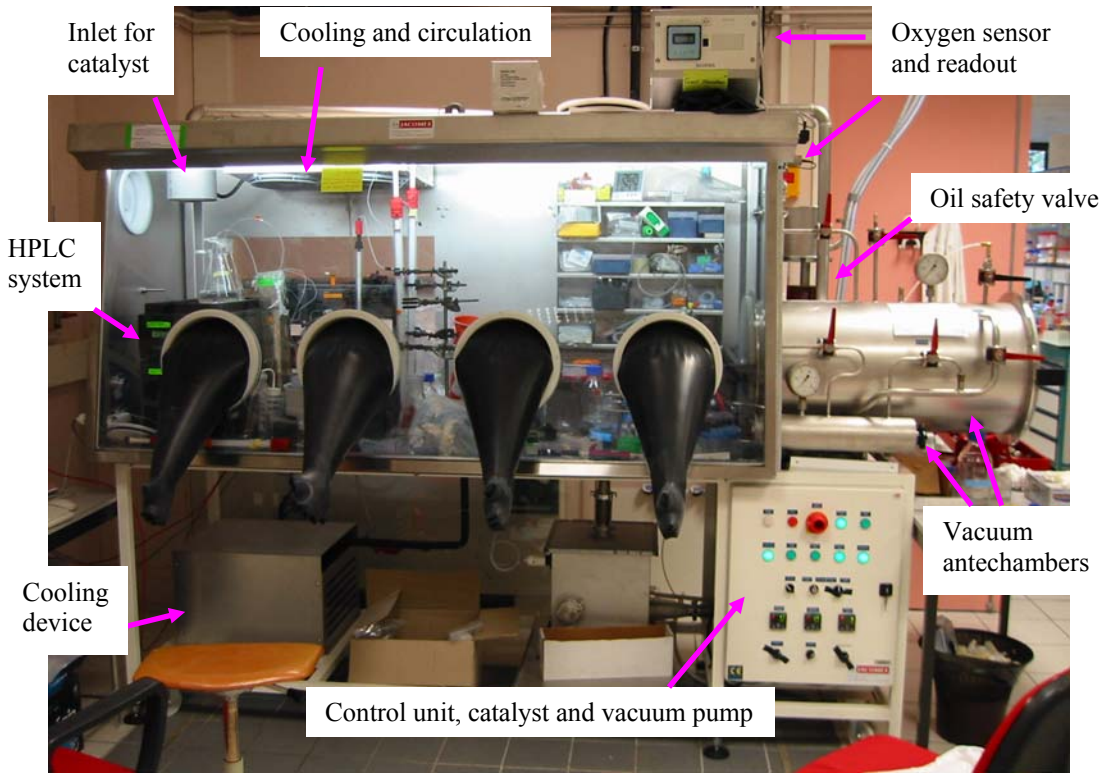


Figure 4.2 Anaerobic chamber used for protein purification (picture presented with permission from IBS).

embedded in the side of each chamber and connected to a vacuum pump. The chambers were equipped with common laboratory equipment (e.g. magnetic stirrers and small table centrifuges) and equipment necessary for performing protein purification and crystallization.

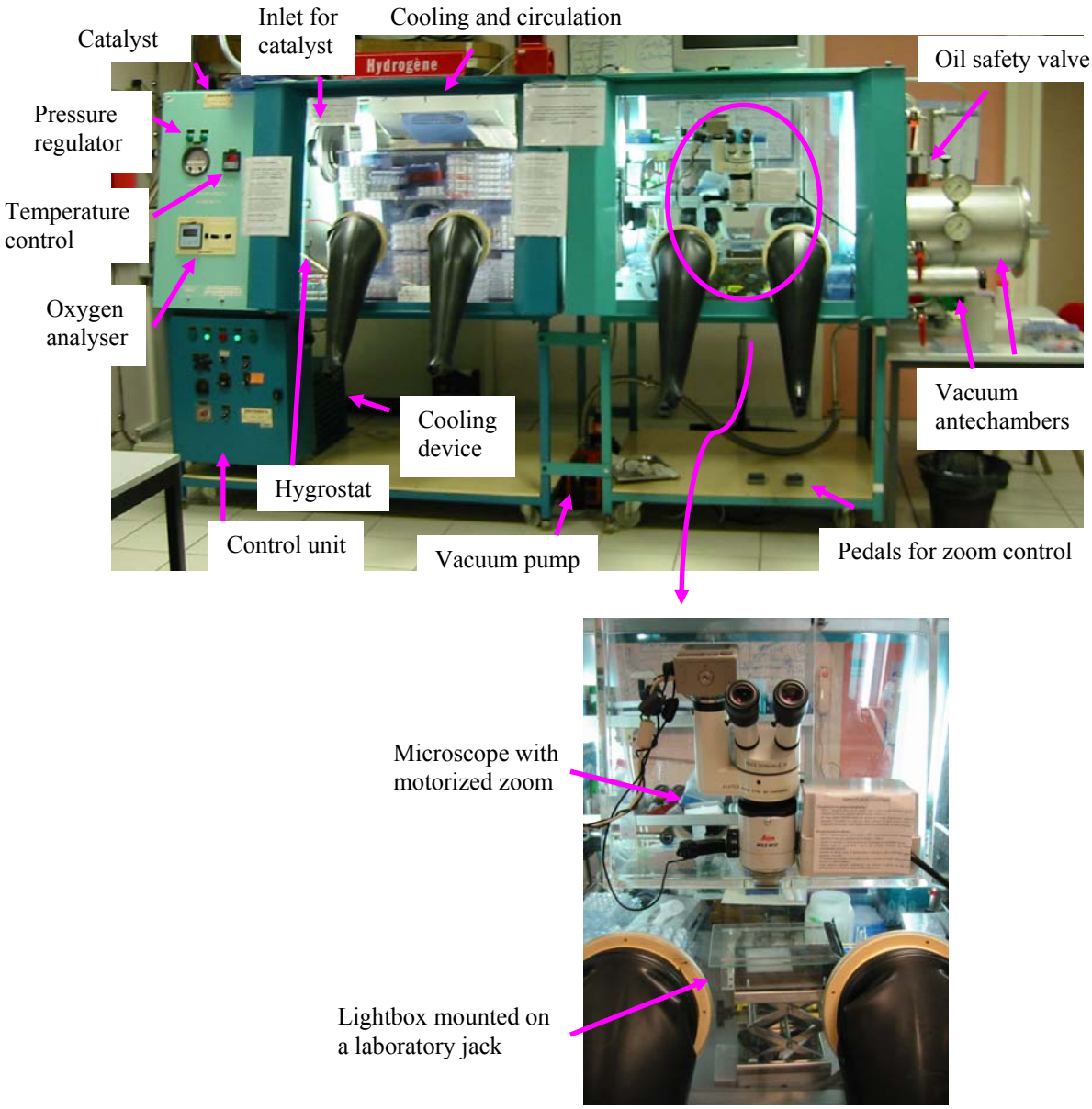


Figure 4.3 Anaerobic chamber used for protein crystallization (pictures presented with permission from IBS).

4.2.1.1 The anaerobic chamber for protein purification

The anaerobic chamber used for protein purification (figure 4.2) had an approximate volume of 1.5 m^3. 2 pairs of gloves were mounted on the front plate. One glove was mounted on the back panel of the chamber. The small and large vacuum antechambers had volumes of 3 L and 100 L, respectively. The oxygen concentration measured varied from 6 to 8 ppm. The chamber was equipped with an HPLC system (ÄKTAFPLC^{TM}), and a small refrigerator unit (approximately 25 × 25 × 30 cm). Valves for obtaining vacuum for degassing buffers

(connected to a water aspirator on the outside of the chamber) and compressed gas for ultra filtration (connected via cobber tubing to the inlet gas) were embedded in the back panel.

4.2.1.2 The anaerobic chamber for protein crystallization

The anaerobic chamber used for protein crystallization was constructed of two side-connected units each having a volume of 0.85 m^3 (figure 4.3). Each unit was equipped with one pair of gloves on the front and on the back plates (total of four pairs). The small and large vacuum antechambers had volumes of 2 L and 36 L, respectively. The oxygen concentration measured was below 2 ppm. A humidity control system consisting of a hygrostat connected to a water heating bath had been installed to maintain a relative humidity of at least 40 %. The chamber was equipped with an alarm system; if an oxygen concentration above 10 ppm was measured an audible alarm would be activated and oxygen concentrations above 100 ppm would trigger automatic flushing of the box with nitrogen. A microscope (figure 4.3) consisting of an internal part and an external part was mounted in the box. The internal part consisted of a light-box mounted underneath a transparent plate and placed on top of laboratory jack; this was used for focus adjustment. The external part consisted of a microscope (Leica, Wild M3Z) mounted in one of the front plates. This was equipped with motorized zoom control operated by two foot pedals, enabling full control of the microscope when working inside the chamber.

4.2.1.3 Procedures and equipment

As a precaution the vacuum antechambers were kept evacuated at all times when not in use.

Importing equipment and liquids into the chambers

All equipment and solutions required inside the chambers had to enter through the vacuum antechambers. Liquids had to enter via tightly closed flasks or containers to avoid evaporation. Containers had to be almost completely filled with liquid (leaving only as little air as possible inside the container) to avoid letting high amounts of O_2 inside. For each antechamber two rounds of evacuation had been found to be adequate to ensure depletion of O_2 prior to importing objects. One round of evacuation contained three steps; (i) complete evacuation of the chamber, (ii) when full vacuum was obtained the chamber was left evacuated for a certain period of time (explained below), (iii) the chamber was then filled with the working gas. When only importing a few objects (e.g. a box of pipette tips or a few flasks of buffers) the evacuation period, step (ii), should be 1 min when using the small antechamber and 5 min when using the large antechamber. When importing larger quantities, especially plastic ware (oxygen dissolves in plastic), then the evacuation period in one of the evacuation rounds should be at least two hours and preferably over night. When importing objects in this way no change in the oxygen concentration inside the chambers was measured.

Deoxygenating buffers and crystallization solutions

Buffers and crystallization solutions were not deoxygenated prior to entering the chambers. All solutions were deoxygenated by flushing with the inside gas using fish tank pumps. This was done by flushing through disposable pipette tips or serological pipettes connected to the pump via rubber tubing. Buffers used for HPLC purifications were further degassed by

evacuation (see section 4.2.1.1). Deoxygenation of buffers or solutions caused an unavoidable increase in the oxygen concentration (less than 2 ppm) depending on the volume of solution and the volume of air imported in the containers.

Crystallization set up
Crystallization set up was difficult when working with the large size gloves and therefore the sitting drop vapour diffusion method using crystal tape for sealing or the hanging drop vapour diffusion method using crystallization plates with screw caps were most often used.

Anaerobic flash cooling of crystals
To ensure that proteins crystallized under anaerobic conditions would not be damaged by mounting and flash cooling under aerobic conditions, a system for flash cooling crystals under anaerobic conditions had been developed. The rationale along with the system is described by Vernède and Fontecilla-Camps (1999). Here, a short introduction is presented along with an introduction to the updated equipment (unpublished).
The system was based on flash cooling of crystals in liquid propane instead of the commonly used liquid nitrogen. The cryogenic agent had to be brought inside the anaerobic chamber via the small vacuum antechamber. Propane (T_m = 85.5 K, T_b = 231.1 K) can be solidified by liquid nitrogen (T_b = 77.4 K) and then imported into the anaerobic chamber.
Gaseous propane was solidified in an aluminium container cooled in a nitrogen bath (figure 4.4 A). The container was then transferred to a polystyrene insulator (figure 4.4 B and C) and a cap-holder was placed on top. The equipment was then imported into the chamber (getting the propane through the small vacuum antechamber took about 3 min). When using the equipment described here (figure 4.4) and by Vernède and Fontecilla-Camps (1999), the propane had liquefied after approximately 4 min and after approximately 7 min the temperature had risen to 130 K. Only a very short time frame is therefore adequate for transferring a crystal from the crystallization drop or a soaking solution to the liquid propane. When the propane had liquefied, the collected crystal was plunged into the liquid propane and immediately the piece of equipment was taken out of the chamber. The cap could then be transferred to liquid nitrogen for storage of the crystal mounted in the loop. The system shown in figure 4.4 was adapted to magnetic caps and wands from Hampton Research (CrystalWand Magnetic and CrystalCap Magnetic) while the system described by Vernède and Fontecilla-Camps (1999) was adapted to cryoloops and caps made in-house.

4.2.2 Preparing an anaerobic chamber for protein purification and crystallization

The anaerobic chamber available in our own laboratories was a LabStar 50 (MBraun) having an approximate volume of 0.8 m^3. The chamber was equipped with a single H_2O/O_2 purifier unit. O_2 and H_2O sensors were not connected to the system. The chamber was equipped with a vacuum antechamber (approximate 74 L) connected to a vacuum pump. The working gas consisted of pure nitrogen supplied via copper tubing from 50 L high pressure gas flasks. A slight overpressure of < 5 mbar was applied and controlled via electro-pneumatic valves connected to the working gas inlet and the vacuum pump. Constant circulation of the gas through the purifier unit was performed to ensure continuous removal of oxygen. The purifier

unit could be regenerated through a regenerating cycle including heating and exposure to reformer gas (10 % H_2 and 90 % N_2).

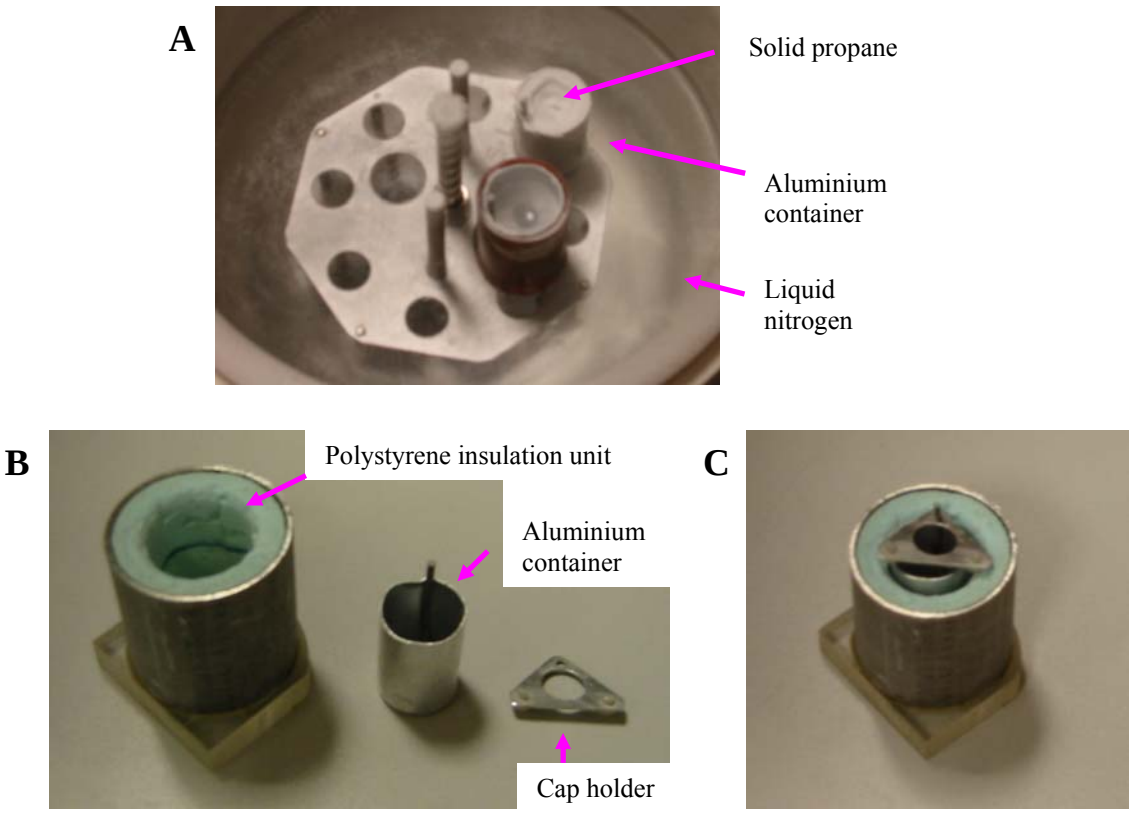


Figure 4.4 Equipment for anaerobic flash cooling of crystals. A) Propane was solidified in an aluminium container (ø 20 mm) immersed in a liquid nitrogen bath. B) The container for transferring solid propane into the anaerobic chamber and flash cooling crystals consists of three parts; a polystyrene insulator unit placed in a stainless steel cylinder (ø 40 mm), the aluminium container for solid/liquid propane and a cap holder. C) The assembled parts shown in panel B. The aluminium container is placed in the insulator and the cap holder is mounted on top (pictures presented with permission from IBS).

The chamber was equipped with an inside power supply so that several electrical instruments could be operated inside the chamber at the same time. The power cable had been cast in epoxy in an inside/outside connector device. To allow outside operation of instruments and data storage by computers, USB-cables were cast in epoxy in an inside/outside connecter device. The chamber was further equipped with an HPLC system (ÄKTAprime^{TM} plus), a spectrophotometer (NanoDrop ND-1000, NanoDrop Technologies), a magnetic stirrer and a small table centrifuge. No temperature control unit was connected to the chamber and it was found that during operation of the system the temperature would vary from approximately 22 to 32ºC primarily depending on the use of the electrical instruments (due to heat generation) and the temperature inside the laboratory (no air-conditioning system was installed).

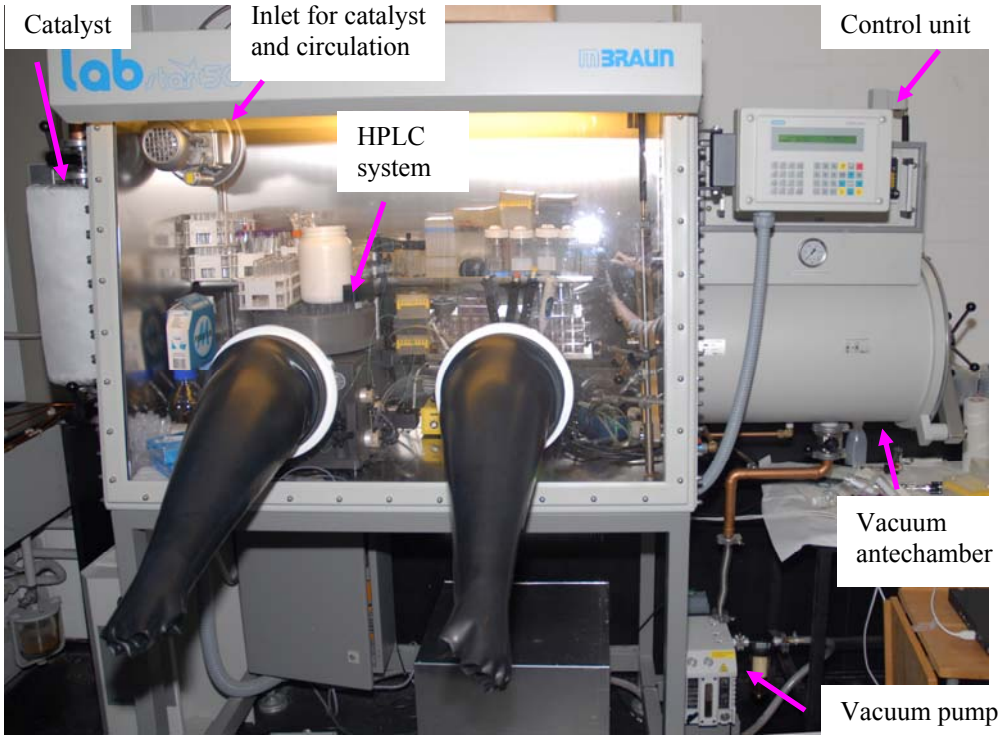


Figure 4.5 Anaerobic chamber used for protein purification and crystallization.

During operation of the anaerobic chamber, procedures for importing equipment and liquids into the chamber and deoxygenating buffers and crystallization solutions were adopted from those learned during my stay at Institute of Structural Biology (section 4.2.1.3). Deoxygenated solutions could not be degassed and the HPLC system was therefore operated with a flow restrictor (to avoid evolution of gas in the columns). Ultrafiltration was accomplished using a Vivacell 20 or 70 pressure cells (Sartorius Stedim), which could be sealed inside the chamber and then exported out of the chamber. Pressure could be applied to the cell under anaerobic conditions using pure nitrogen, and the cell could be imported though the vacuum antechamber after or during ultrafiltration. Since no humidity control system was present, solutions were kept in tightly closed containers when not in use.
During crystallization trials, set up could easily be performed inside the chamber but the crystallization drops could not be examined since no microscope was embedded in the front panel. It was therefore necessary to export plates from the chamber and then examine the drops under a microscope outside the chamber. When using screw cap plates (EasyXtal Tool plates from Qiagen) it was found that the plates could be re-imported into the chamber through the vacuum antechamber without damaging the drops. Since the plates were only kept outside the chamber for a short period of time, I believe that oxygen penetration is minimal. This offers a cheap alternative to embedding a microscope in the front plate. However, anaerobic seeding and flash cooling of crystals are rendered impossible.

4.3 Chromatographic media and HPLC systems

The pre-packed columns and column materials (table 4.1) used for protein purification throughout this project were purchased from Amersham Biosciences (GE Healthcare).

Table 4.1 Pre-packed columns and column materials

Pre-packed columns [A]	Column material	Specifications
	Source 30Q	Anion-exchange
Resource Q 6 ml	Source 15Q	Anion-exchange
Resource Q 1 ml	Source 15Q	Anion-exchange
HiLoad 16/10 Q Sepharose High Performance	Q Sepharose High Performance	Anion exchange
HiPrep 16/10 ANX Sepharose Fast Flow	ANX Sepharose Fast Flow	Anion exchange
	Q Sepharose Fast Flow	Anion-exchange
HiLoad 16/60 Superdex 75 prep grade	Superdex 75	High resolution fractionation
31/13 Superdex 75 HR	Superdex 75	High resolution fractionation
30/10 Superdex 75 HR	Superdex 75	High resolution fractionation
PD-10 Desalting	Sephadex G-25 Medium	Group separation
	Sephadex G-10	Group separation
HisTrap HP 5 ml	Ni Sepharose High Performance	Immobilized metal ion affinity

A: The dimensions, x/y, of some of the columns are given, x is the inner diameter of the column in mm and y is the height of the column in cm.

Purification experiments were either performed by gravity (for group separation and when a 15/2 Q Sepharose Fast Flow column was employed, chapters 6 to 8) or by HPLC. Three HPLC systems were employed, all purchased from Amersham Biosciences (GE Healthcare): ÄKTAFPLC™, ÄKTAprime™ plus and ÄKTA™ purifier 100.

4.4 Iron-sulfur cluster types inferred from ultraviolet-visible spectrophotometry

Throughout chapters 6 to 11, UV-vis spectrophotometry was used to aid in detecting the presence of FeS clusters in the given proteins and inferring the type of cluster present. UV-vis spectrophotometry is insufficient in itself since it measures the total absorbance of a given solution, and it cannot stand alone in inferring the composition of a sample (Skoog *et al.*, 1996).

If a sample contains a component with two or more different cluster types or several components with different cluster types, the different cluster types cannot be inferred from the UV-vis spectrum. If however a purified species with a measured UV-vis spectrum is identified to be of a certain composition using other spectroscopic methods (e.g. EPR and Mössbauer spectroscopy), the UV-vis spectrum can be used as a fast method to establish whether another sample contains the same purified species.

UV-vis spectra recorded of different purified FeS proteins known to contain the same cluster type may differ slightly since the absorbance is a consequence of LMCT (Cotton *et al.*, 1995) and the difference in energy levels are affected by the geometry of the cluster and ligands (Cotton *et al.*, 1995). Therefore, if a recorded spectrum of a given purified FeS protein is very

similar to a spectrum of a known FeS protein, it indicates but not verifies that your protein contains the same cluster type.
UV-vis spectrophotometry was also used for concentration determinations applying Beers law, when molar absorptivities are known or calculated at a certain wavelength.
UV-vis spectra were recorded either on a HP8453 (Hewlett Packard) diode-array spectrophotometer or on a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies). Unless otherwise stated, UV-vis spectra were recorded by the HP8453 diode-array spectrophotometer.

4.5 Ubiquitin fusion technology

Expression of recombinant proteins fused to ubiquitin (Ub) is a commonly used approach to obtain high yield expression and increased solubility of both peptides and large protein. Further, it can be combined with *in vivo* as well as *in vitro* cleavage (section 4.6) of the fusion tag (Butt *et al.*, 1989; Gohara *et al.*, 1999; Pilon *et al.*, 1997; Varshavsky, 2000).
Ub is a highly conserved protein consisting of 76 amino acids, and it is found in eukaryotic cells (Vijay-Kumar *et al.*, 1987), where it functions in the pathway for targeting proteins for degradation by the proteasome (Wilkinson, 2000). In the fusion proteins considered in this project, yeast (*Saccharomyces ceravisiae*) Ub (Wilkinson *et al.*, 1986) was used. Yeast Ub has a molecular weight of 8557 Da. It contains eleven basic residues, eleven acidic residues and one aromatic residue and has a calculated pI of 6.6 (calculated using the ExPASy Proteomics Server, Bjellqvist *et al.*, 1993 and 1994).

4.6 Ubiquitin hydrolases

The Ub fusion system was used for production of Ori-ISP and LCA ferredoxin. To fully study the properties of these proteins it is necessary to cleave off the Ub-fusion. This can be accomplished by the aid of ubiquitin hydrolases.
Ubiquitin hydrolases comprise a large family of enzymes that are able to hydrolyse Ub fusions having the general sequence $Ub^{1\text{-}72}$-Leu^{73}-Arg^{74}-Gly^{75}-Gly^{75}-X, where X is the protein or peptide fused to Ub. The enzymes cleave specifically at the peptide bond following the Arg-Gly-Gly sequence. Most of these enzymes are cysteine hydrolases that hydrolyse their substrates via nucleophilic attack by a Cys-SH/His-imidazole catalytic diad (Case *et al.*, 2006).
Yeast ubiquitin hydrolase (YUH1) belongs to the UCH (ubiquitin C-terminal hydrolase) family, which is comprised of low molecular weight enzymes. The UCHs hydrolyse short peptides or small thioester-, ester- and amide-linked adducts fused at the C-terminal of ubiquitin. YUH1 consist of 236 amino acid residues and has a molecular weight of 26.4 kDa (Case *et al.*, 2006; Miller *et al.*, 1989; Sakamoto *et al.*, 1999). The Ub/YUH1 system has previously been used to prepare short peptides in *E. coli* (Khono *et al.*, 1998; Miller *et al.*, 1989). YUH1 was used to perform *in vitro* cleavage of Ub-Ori-ISP and Ub-LCA ferredoxin.
The ubiquitin specific proteases (UBPs) are larger than the UCHs and vary in molecular mass from 50 to 300 kDa. The UBPs hydrolyse Ub-fusions when Ub is linked to proteins or other Ub molecules (Case *et al.*, 2006; Sakamoto *et al.*, 1999). UBP1 from yeast consists of 809 amino acid residues and has an approximate molecular weight of 93 kDa (Tobias and

Varshavsky, 1991). UBP1 was co-expressed with Ub-LCA ferredoxin to produce LCA ferredoxin in *E. coli* cells.

4.7 Strains used for expression of recombinant proteins and enzymes

The particular expression systems used for expression of both native and fusion proteins were based on the pET and pACYC vectors and the *E. coli* BL21(DE3) expression strain. This expression system employs the T7 RNA polymerase gene under the control of the *lac*UV5 promoter/operator and is therefore inducible by isopropyl-*β*-D-thiogalactopyranoside (IPTG) (Gohara *et al.*, 1999; Studier *et al.*, 1990; Sweet 2003). Five different strains expressing the proteins and enzymes of interest have been constructed by other members of our research group (table 4.2).

Table 4.2 Strains expressing proteins and enzymes

Strain number	Vector construct	The expressed protein(s)	*E. coli* strain	Selection marker
HC496	pET26b-Ub-Ori-ISP	Fusion protein, Ub-Ori-ISP	BL21 (DE3)	30 Kan [A]
HC298	pET3a-PfFd	Native, *P. furiosus* [4Fe-4S] ferredoxin	BL21 (DE3)	100 Amp [B]
HC1269	pET26b-Ub-LCA	Fusion protein, Ub-LCA ferredoxin	BL21 (DE3)	30 Kan [A]
HC1419	pET26b-Ub-LCA/ pACYC184-UBP1	Fusion protein, Ub-LCA ferredoxin Native, UBP1	BL21 (DE3)	30 Kan [A] 20 Cam [C]
HC1434	pET26-YUH1-His_6	His-tagged, YUH1-His_6	BL21 (DE3)	30 Kan [A]

A: 30 µg kanamycin sulfate pr millilitre of culture. B: 100 µg ampicillin pr millilitre of culture.
C: 20 µg chloroamphenicol pr millilitre of culture.

4.8 Chemicals and SDS-PAGE

All chemicals applied were of analytical grade. Water used throughout the experiments was 18.2 MΩcm QPAK Milli-Q (Millipore), abbreviated milli-q.

Samples that were to be analysed by sodium dodecyl sulfate polyacrylaminde gel electrophoresis (SDS-PAGE) were collected throughout the experiments and stored in Eppendorf tubes at -20ºC until further use. Precast gels from Bio-RAD (16,5% Tris-Tricine/Peptide) were used in combination with polypeptide SDS-PAGE molecular weight standards. SDS-PAGE was performed according to the instructions from the manufacturer of the gel.

Chapter five

5 Purification of sulfonated Ub-Ori-ISP

This chapter presents the optimized procedures for cultivation of cells expressing the fusion protein, Ub-Ori-ISP, and purification of the sulfonated apo-protein. The originally developed procedures are described in my M.Sc. thesis (Christophersen, 2004). Crystallization of the sulfonated apo-protein is also presented.

Ori-ISP is expressed as a fusion protein (Ub-Ori-ISP) due to the small size (23 residues) and the lack of aromatic and basic residues, which renders it difficult to visualize by SDS-PAGE and UV-vis spectrophotometric analysis. Since Ub contains a tyrosine and basic residues (section 4.5), Ub-Ori-ISP (10806 Da) is detectable by SDS-PAGE (stained with Coomassie blue G-250) and UV-vis spectrophotometry (Compton and Jones, 1985; Pace *et al.*, 1995).

It was previously attempted to purify holo-Ub-Ori-ISP produced *in vivo* however, this was not possible due to breakdown of the holo-protein. Therefore a different method was applied. The first step was to sulfonate the free cysteines of the apo-protein and purify the sulfonated protein (11126 Da) (Christophersen, 2004). Sulfonation of the free cysteines is performed to prevent formation of inter- and intra-molecular disulfide bonds and oxygenated species during purification. In the sulfonation reaction, the cysteinyl thiolates are bound to sulfite (- CH_2 - S - SO_3^-) rendering the protein stable for handling under aerobic conditions. Sulfonation of cysteinyl thiolates and di-sulfide bonds can be achieved by treating the protein with 2-nitro-5-(thiosulfo)-benzoate (NTSB), a denaturing agent (e.g. urea) and excess sulfite (Sow, 1997; Creighton, 1989). In the second step, the holo-protein was then obtained from the sulfonated apo-protein by de-protecting the sulfonated cysteines using a reducing agent and then performing *in vitro* iron-sulfur cluster reconstitution (Christophersen, 2004) (chapters 6 and 7).

5.1 Modifications to previously used protocols

The original method for cultivation of Ub-Ori-ISP producing cells gave a very low yield (ca. 5 mg/L cell culture) and this was traced to the cultivation protocol. It was found that the yield of protein was highly dependent on the optical density at 600 nm (OD_{600}) at the time of induction with IPTG and the length of the induction period (results not shown). The optimal conditions were found to be an OD_{600} between 1.2 and 1.5 at the time of induction and an induction period of 2 hours.

In the protocol for sulfonation and purification of apo-Ub-Ori-ISP, changes to the original protocols were also applied. An ultrafiltration step was introduced after the sulfonation reaction. Optimization of this step was performed by Laboratory Technician student, Ditte Skibstrup, under my supervision (results not shown). The first HPLC purification step was modified from including a one-step gradient (in the original procedure) to a two-step gradient.

5.2 Experimental

5.2.1 Cultivation of cells expressing Ub-Ori-ISP

This procedure is for cultivating 12 times 650 ml cell cultures. Cells from the HC496 glycerol stock were plated on an agar plate containing LB/30 Kan (37 g/L Luria Broth agar, 30 µg/ml kanamycin sulphate) and incubated overnight (16 hours) at 37ºC. Two single colonies were used to each inoculate 50 ml TB/30 Kan (45 g/L Terrific Broth, 0.4 % v/v glycerol, 30 µg/ml kanamycin sulphate) in 300 ml triple baffled shake flasks and incubated at 37ºC with shaking (250 rpm) until an OD_{600} of 0.6 to 1.0 was obtained. The pre-cultures were stored at 4ºC overnight. The cells of the two pre-cultures were then sedimented (centrifuged at 4ºC and 2300 g for 10 min) and supernatants discarded. The two cell portions were each re-suspended in 50 ml fresh TB/30 Kan. 6.5 ml portions, which were used to inoculate the large scale cultures, were transferred to each of twelve 2 L triple baffled shake flasks containing 650 ml TB/30 Kan. The cultures were incubated at 30ºC with shaking (250 rpm) until the OD_{600} was between 1.2 and 1.5. Then protein expression was induced by addition of IPTG (200 mM stock solution) to a final concentration of 0.1 mM, and the incubation was continued with shaking (250 rpm) for 2 hours at 30ºC. The cells were harvested by centrifugation at 3000 g and 4ºC for 15 min. The supernatants were discarded and each of the 12 portions of cells were re-suspended in 25 ml cold (4ºC) 20 mM Tris/HCl pH 8.0 (Tris(hydroxymethyl)-aminoethane, pH adjusted with 1 M HCl) and transferred to cold 50 ml polypropylene tubes. The cells were sedimented by centrifugation (3000 g, 4ºC, 15 min) and the supernatants discarded. The cells were kept at -80ºC for storage.

5.2.2 Sulfonation and purification of Ub-Ori-ISP

An overview of sulfonation and purification is presented in figure 5.1.

5.2.2.1 Sulfonation of Ub-Ori-ISP

Six portions of harvested cells (cells from 3.9 litres of culture, total volume of cells, ca. 45 ml) were thawed at 4ºC and re-suspended in cold sulfite-containing buffer (11.3 g of Na_2SO_3 dissolved in 165 ml 20 mM Tris/HCl pH 8.0). The cells were lysed by sonication (Satorius Labsonic P at 80 % amplitude) in three intervals of 30 seconds. The solutions were kept on ice at all times and allowed to cool between each sonication burst. The crude lysate was centrifuged at 24800 g for 40 minutes at 4ºC. The yellowish brown supernatant was decanted, and the pellet discarded. 36 g urea was dissolved in the supernatant (gently stirring) after which 40 ml 400 mM NTSB was added and the volume was adjusted to ca. 300 ml using 20 mM Tris/HCl pH 8.0 (final concentrations of sulfite, urea and NTSB: 0.3 M, 2 M and 0.05 M, respectively). The resulting orange solution was stirred gently for 3.5 hours at room temperature, and stirring was continued overnight at 4ºC. The now yellow solution was centrifuged at 24800 g for 40 minutes at 4ºC, the yellow supernatant decanted, and the pellet discarded. The supernatant was diluted ca. 3.5 times using 20 mM Tris/HCl pH 8.0 and the volume reduced (at 4ºC) to below 80 ml using the Vivaflow 50 filtering system (Sartorius Stedim) equipped with six 10.000 MWCO (molecular weight cut off) membranes. Dilution followed by reduction of the volume was continued (three times) until the conductivity of the

solution was below 5 mS/cm. The protein extract solution was centrifuged (24800 g, 40 min, 4ºC) and the supernatant kept for purification of sulfonated Ub-Ori-ISP.

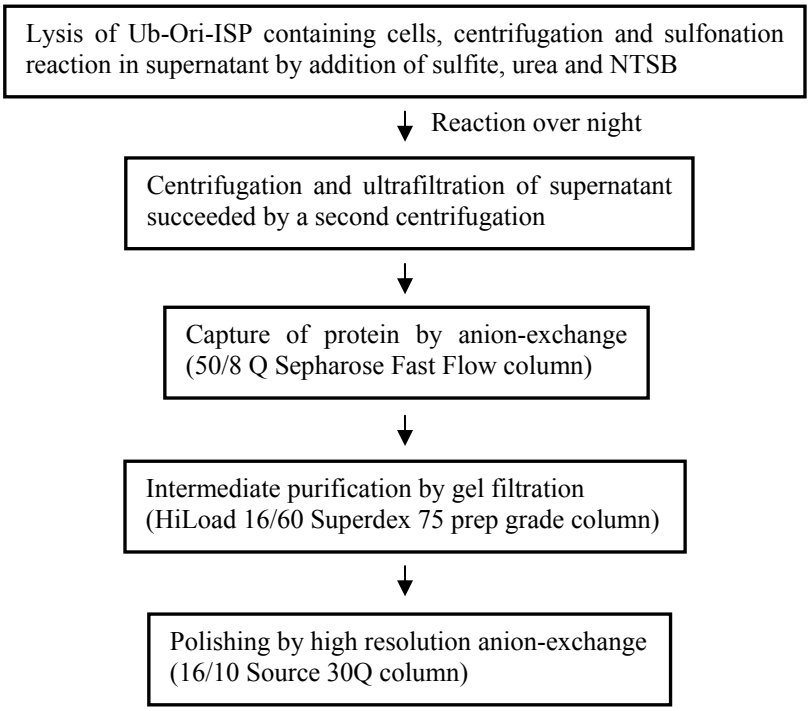


Figure 5.1 Flowchart of sulfonation, ultrafiltration and the subsequent three-step purification of Ub-Ori-ISP.

5.2.2.2 Purification of sulfonated Ub-Ori-ISP

The protein extract containing the sulfonated Ub-Ori-ISP (corresponding to protein from 12 portions of cells, section 5.2.1, and twice the amount sulfonated in section 5.2.2.1) was loaded onto a 50/8 Q Sepharose Fast Flow column equilibrated with 20 mM Tris/HCl pH 8.0. The protein was purified by HPLC (ÄKTA^{TM} purifier 100) and eluted (linear flow rate: 1.3 cm/min) using a stepwise gradient, 0.22 M NaCl in 2 CV (column volume) and 0.45 M NaCl in 4 CV. Fractions were collected and the total volume reduced to ca. 20 ml. The buffer was changed to 20 mM Tris/HCl, 0.15 M NaCl pH 8.0 by ultrafiltration using a stirred Amicon cell with a YM3 membrane. Ub-Ori-ISP was further purified on a HiLoad 16/60 Superdex 75 prep grade column equilibrated with 20 mM Tris/HCl pH 8.0, 0.15 M NaCl. The protein sample was filtered (0.45 µm) then loaded onto the column (a maximum of 4 ml was loaded for each run), and fractions were collected upon elution with the same buffer (linear flow rate: 0.2 cm/min). Ub-Ori-ISP was further purified on a 16/10 Source 30Q column equilibrated with 20 mM Tris/HCl pH 8.0, 0.05 M NaCl. Fractions collected from the gel filtration step were filtered (0.45 µm) and diluted three times with 20 mM Tris/HCl pH 8.0 before loading onto the column. The protein was eluted (linear flow rate: 1.2 cm/min) using a linear gradient, 0.15 M to 0.4 M NaCl in 10 CV. The volume of the collected fractions was reduced to ca. 15 ml and the buffer was changed to 0.1 M NH_4HCO_3 by ultrafiltration using a stirred Amicon cell with a YM3 membrane (Millipore). The amount of protein was determined by UV-vis spectrophotometry (using a molar absorptivity of 1490 $M^{-1}\cdot cm^{-1}$ at

280 nm, Pace *et al.*, 1995) and the solution was then freeze-dried. The freeze-dried protein was weighed and the percentage of remaining salt was determined. The protein was kept at -20ºC for storage. The purity of the protein was tested by SDS-PAGE.

5.2.3 Analytical gel filtration of sulfonated and purified Ub-Ori-ISP

Freeze-dried, sulfonated Ub-Ori-ISP was dissolved in 20 mM Tris/HCl, 0.15 M NaCl pH 8.0 to a final concentration of 5 mg/ml and analysed by analytical gel filtration. The sample was filtered (0.45 µm) and 250 µl was loaded onto a 20 mM Tris/HCl, 0.15 M NaCl pH 8.0 equilibrated 31/13 Superdex 75 HR column. Protein was eluted using the same buffer with a linear flow rate of 0.3 ml/min.

5.2.4 Crystallization of sulfonated Ub-Ori-ISP

Freeze-dried, sulfonated Ub-Ori-ISP was dissolved in 20 mM Tris/HCl, pH 8.0 and washed three times with the same buffer (using a stirred Amicon cell with a YM3 membrane). The solution was up-concentrated, filtered (0.22 µm) and the concentration was determined to be 17.9 mg/ml. The protein was kept at -20 ºC for storage.

Initial screening for crystallization conditions were performed using The Solubility Tool Kit (Molecular Dimensions). The hanging drop vapour diffusion method was applied and the drop (2 µl of protein solution mixed with 2 µl of reservoir solution) was equilibrated over 500 µl of reservoir solution at room temperature.

Micro crystals appeared when using ammonium citrate pH 9 as precipitant. Conditions for crystal growth were optimized by testing precipitant concentration versus pH and also testing addition of small volumes of PEG (polyethylene glycol) 4000 (experiments are not shown). Diffracting properties of some of the obtained crystals were tested at MaxLab (Lund, Sweden) beamline 9-11-3 and at European Synchrotron Radition Facility (ESRF) (Grenoble, France) beamline ID14-3.

5.3 Results and discussion

5.3.1 Cultivation of cells, sulfonation reaction and purification

The yield of cells (section 5.2.1) was about 5 to 10 ml cells per 650 ml of culture. The cells had a light brownish colour. Re-suspension of cells in buffer followed by cell lysis and centrifugation (section 5.2.2.1) yielded a light yellowish brown supernatant. Upon addition of the yellow NTSB, the supernatant turned bright orange indicating progress of sulfonation reaction. The orange colour is attributed to a bi-product (2-nitro-5-thiobenzoate) of the reaction (Sow, 1997 and Creighton, 1989). During ultrafiltration precipitation was observed which is probably due to the removal of the denaturing agent (urea), causing some protein to precipitate. Only protein from six portions of cells could be sulfonated in one batch, since ultrafiltration of a batch of this size took about 8 hours and it was not possible to leave the filtration to proceed overnight unattended.

The precipitation was removed by centrifugation prior to purifying the protein. Two batches of sulfonated protein (corresponding to protein from 12 portions of cells, section 5.2.1) was loaded onto the 50/8 Q Sepharose Fast Flow column and Ub-Ori-ISP was captured as

described (section 5.2.2.2) (figure 5.2). Sulfonated Ub-Ori-ISP was further purified by gel filtration (on a 16/60 Superdex 75 prep grade column) (figure 5.3). 5 to 6 runs loading a maximum of 4 ml in each run were necessary for purifying protein corresponding to 12 portions. Otherwise the column would be overloaded and the viscosity of the sample too high for optimal separation. The protein was finally purified by high-resolution anion-exchange in 6 runs on a 16/10 Source 30Q column (figure 5.4). The identity of the purified product as sulfonated Ub-Ori-ISP has previously been verified by mass spectrometry (Christophersen, 2004). The purity of the product was tested by SDS-PAGE (figure 5.5 lane 1) and two bands

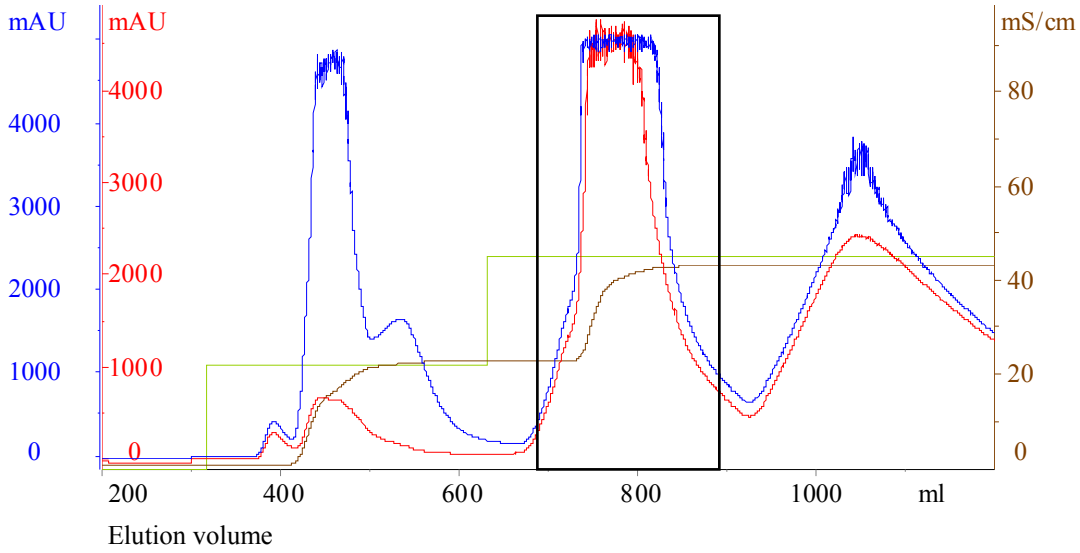


Figure 5.2 Chromatogram for purification of sulfonated Ub-Ori-ISP on 50/8 Q Sepharose Fast flow column. The collected fraction is indicated with a black box. The blue and the red lines indicate the absorbances at 236 nm and 280 nm, respectively. The brown and the green lines indicate the conductivity and the NaCl gradient, respectively. The concentration of NaCl can be read from the conductivity axis employing the unit 10^{-2} M.

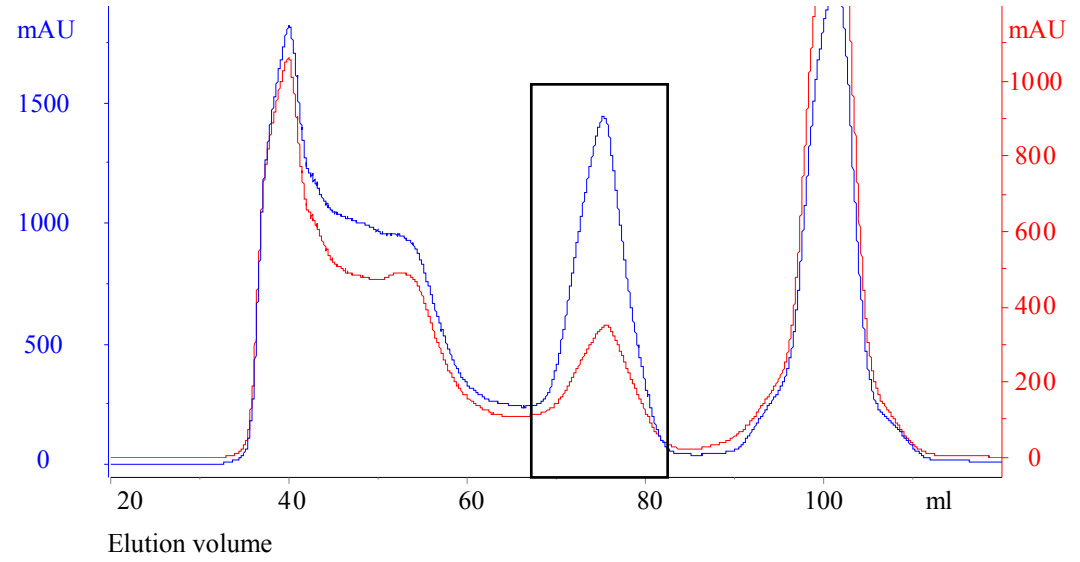


Figure 5.3 Chromatogram for purification of sulfonated Ub-Ori-ISP on 16/60 Superdex 75 prep grade column. The collected fraction is indicated with a black box. The blue and the red lines indicate the absorbances at 236 nm and 280 nm, respectively.

were apparent. The dominating band is known to be Ub-Ori-ISP (Christophersen, 2004) and the minor band is believed to be non-denatured, dimerized Ub-Ori-ISP. When re-purifying the protein by gel filtration, no secondary peak was observed from the chromatogram and SDS-PAGE yielded the same result (results not shown).
Including the ultrafiltration step in the purification strategy (see section 5.2.2.2) has had a great impact on the individual purification steps as compared to previously (Christophersen, 2004). When purifying protein from 12 portions of cells, the number of runs on the 50/8 Q Sepharose Fast flow column have been reduced from 4 to 1 and the number of runs on the 16/10 Source 30Q column have been reduced from 10 to 6. This has decreased buffer consumption and the time spent purifying the protein has been reduced from 9 to 7 working days. The yield of the Ub-Ori-ISP has also improved markedly from about 5 mg/L cell culture to 30 to 40 mg/L cell culture containing 5 to 30 % salt.

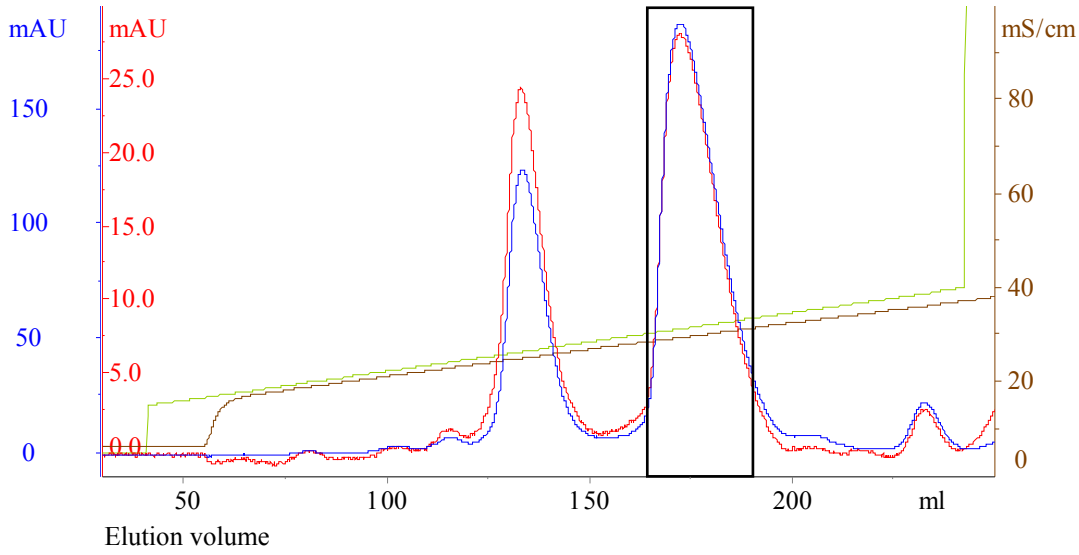


Figure 5.4 Chromatogram for purification of sulfonated Ub-Ori-ISP on 16/10 Source 30Q column. The collected fraction is indicated with a black box. The blue and the red lines indicate the absorbances at 236 nm and 280 nm, respectively. The brown and the green lines indicate the conductivity and the NaCl gradient, respectively. The concentration of NaCl can be read from the conductivity axis employing the unit 10^{-2} M.

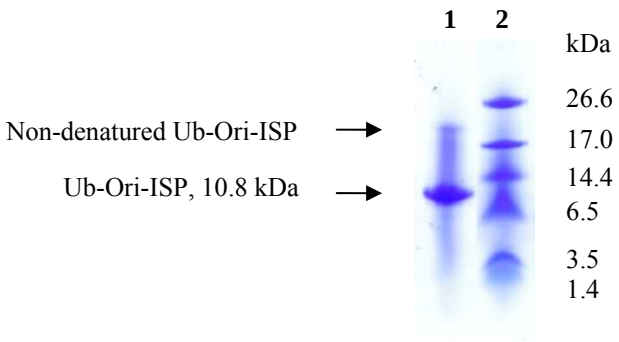


Figure 5.5 Purity of freeze-dried Ub-Ori-ISP tested by SDS-PAGE. Lane 1: 6 µl of an 8 mg/ml solution containing freeze-dried Ub-Ori-ISP dissolved in 20 mM Tris/HCl pH 8. Lane 2: Marker (1 µl).

5.3.2 Analytical gel filtration of sulfonated and purified Ub-Ori-ISP

The chromatogram from the analytical gel filtration (section 5.2.3) is shown in figure 5.6. One major peak eluted at 11.6 ml. In front of this peak a very small shoulder (ca. 10.5 ml) was observed. A minor peak at 12.9 ml was also observed. Since the SDS-PAGE results (figure 5.5) gave essentially only one band, the shoulder and the minor peak are also believed to be sulfonated Ub-Ori-ISP. The dominating peak at 11.6 ml is probably a dimer form, while the minor peak at 12.9 ml is a monomeric form. Protein found in the shoulder was attributed to polymeric forms. Chapter 7 presents thorough discussion of the presence of monomers and dimers. It is interesting that Ub-Ori-ISP seems to be present as a dimer in solution since no record has been found in the literature of wild-type ubiquitin acting as a dimer. Ubiquitin participates in the pathway for targeting proteins for degradation by the proteasome. Proteins are marked with chains of ubiquitin and then recognised by the proteasome. In these chains ubiquitin monomers are bound together by covalent bonds and are therefore not dimers, trimers or polymers in the traditional sense (Jackson, 2006; Fang and Weissman, 2004; Wilkinson, 2000). However, upon analysis of fragments of human ubiquitin it was found that a fragment encompassing residues 1 to 51 was able to dimerise (Bolton *et al.*, 2001). It is therefore likely that the fusion of Ub to Ori-ISP induces dimerisation. This could be due to structural changes of ubiquitin or the presence of Ori-ISP.

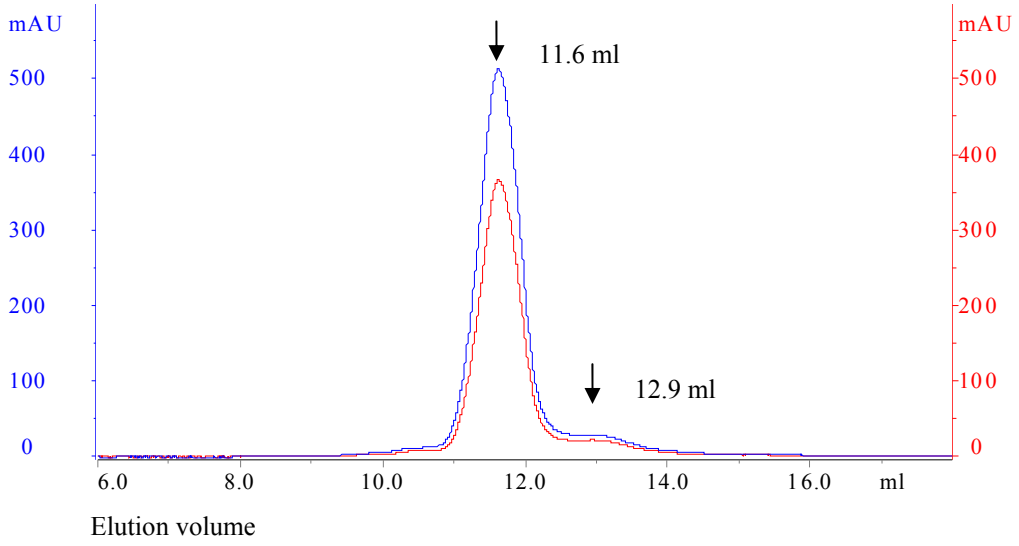


Figure 5.6 Chromatogram for analytical gel filtration obtained when analysing sulfonated and purified Ub-Ori-ISP. The blue and the red lines indicate the absorbance at 236 nm and 280 nm, respectively.

5.3.3 Crystallization of sulfonated Ub-Ori-ISP

To gain knowledge on the crystallization properties of Ub-Ori-ISP, crystallization of the sulfonated apo-protein was attempted. It was hoped that conditions for crystallization of the apo-protein could be modified and used to obtain crystals of the holo-protein.

Crystallization trials were carried out and conditions optimized for growth of regular, three-dimensional single crystals (section 5.2.4). Figure 5.7 shows some of the obtained crystals. Cubic shaped crystals were formed in tri-ammonium citrate in the concentration range 1.4 M

to 1.5 M and in the pH range 6.5 to 8.5. When PEG 4000 was added (0.5 % to 1.0 %) larger crystals were obtained. Diffraction properties of some of the crystals were tested. At MaxLab only diffraction from the largest crystals could be measured. However, these were extremely poor (diffraction limits about 15 Å). At the ESRF, diffraction to ca. 7 Å was obtained. Annealing the crystals had no effect on the diffracting properties. No further attempts were made to optimize the diffraction properties. The poor diffraction properties are believed to be due to disorder in the fusion protein either originating from flexibility in the Ori-ISP part of the molecule or at the junction of the two proteins. It is hoped that formation of the holo-protein will inhibit possible flexibility, and hereby produce better diffracting crystals. Another possibility is to cleave off the ubiquitin part of the molecule and then crystallize holo-Ori-ISP (discussed further in chapter 8).

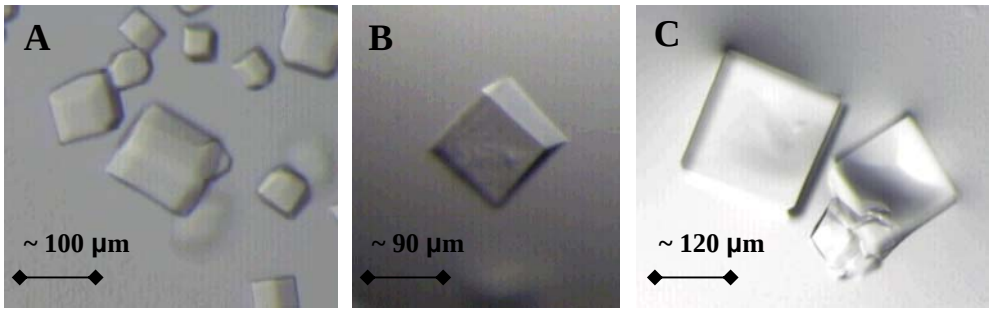


Figure 5.7 Pictures showing some of the obtained crystals. A) Crystals were grown from 1.4 M tri-ammonium citrate pH 7.8, 1 % PEG 4000. B) Crystals were grown from 1.5 M tri-ammonium citrate pH 6.5. C) Crystals were grown from 1.4 M tri-ammonium citrate pH 7.8, 0.75 % PEG 4000.

5.4 Conclusion

The cultivation procedure of cells expressing Ub-Ori-ISP and the purification procedure for sulfonated Ub-Ori-ISP have been optimized. This has resulted in an increased yield of purified protein from ca. 5 mg/L cell culture to 30 to 40 mg/L and a reduction of time consumption from 9 to 7 days in the purification procedure. Analytical gel filtration has shown the protein to be a dimer in solution. The sulfonated, purified protein was crystallized; however, the obtained crystals diffracted poorly and this was believed to be due to flexibility of the Ori-ISP part of the fusion protein.

Chapter six

6 Iron-sulfur cluster reconstitution of Ub-Ori-ISP followed by purification

This chapter provides a presentation of the experimental work and results obtained when performing *in vitro* iron-sulfur cluster reconstitution of Ub-Ori-ISP followed by purification using Ar/N_2 flushing (section 4.1) for obtaining anaerobic conditions.
In vitro FeS cluster reconstitution was first accomplished in 1966 by Malkin and Rabinowitz, and since then, their method has been modified and optimized. Much literature is found on this subject, and all methods in which the FeS cluster reconstruction progresses *in situ* have in common the presence of iron ions, sulfide and thiols (reducing agent) as necessary components for the reconstitution to occur (Coghlan and Vickery, 1991; Hong and Rabinowitz, 1970; Mulholland *et al.*, 1998; Ollagnier *et al.*, 1999; Pierrel *et al.*, 2003).
The procedure used for cluster reconstitution in Ub-Ori-ISP has been developed and optimized by the writer (Christophersen, 2004) and was based on previous procedures for cluster reconstitution in ferredoxins (Sow, 1997). In this procedure $FeCl_3$ was used as the iron source for cluster reconstitution. After reconstitution the protein solution was desalted on a PD-10 column and purified on a 15/2 Q Sepharose Fast Flow column. Further purification of the product was attempted on a 1 ml Resource Q column, but total breakdown of the holo-protein was observed (Christophersen, 2004).
In this chapter, the outcome of performing cluster reconstitution using different iron sources followed by purification either by gravity or by HPLC is described. Further, purification was performed with or without glycerol in the purification buffers when purification was performed by HPLC. The stability of the obtained product has been tested under various conditions.

6.1 Experimental

6.1.1 Cluster reconstitution reaction and desalting

13.6 mg freeze-dried sulfite-protected Ub-Ori-ISP (chapter 5) was weighed out in a round bottomed flask and dissolved in 1.9 ml 100 mM Tris/HCl pH 8.5. 200 µl 0.12 M dithiothreitol (DTT) solution (20-fold molar excess, DTT was dissolved in 100 mM Tris/HCl pH 8.5) was added to the protein solution followed by incubation for 1.5 hours with gentle stirring. Then, 200 µl freshly prepared 50 mM iron solution (8-fold molar excess) was added to the protein solution. The iron source was either $NH_4Fe(III)(SO_4)_2 \cdot 12H_2O$, $(NH_4)_2Fe(II)(SO_4)_2 \cdot 6H_2O$ or $FeCl_3$ dissolved in milli-q water or metallic iron dissolved in sulphuric acid (section 6.1.1.1). Right after addition of the iron solution, 200 µl freshly prepared 50 mM $Na_2S \cdot 9H_2O$ solution

(8-fold molar excess, dissolved in milli-q water) was added drop wise over a period of approximately two minutes. The resultant solution was incubated for one hour with gentle stirring, after which desalting on a PD-10 column (by gravity) into either 20 mM Tris/HCl, 0.15 M NaCl pH 8.0 or 20 mM Tris/HCl, 0.15 M NaCl, 20 % glycerol pH 8.0 was carried out. The protein was further purified either by gravity on a 15/2 Q Sepharose Fast Flow column (section 6.1.2) or by HPLC purification on a 10/5 Source 30Q column (section 6.1.3). Cluster reconstitution, desalting and purification were carried out at room temperature while all obtained products were kept on ice. Figure 6.1 presents a flow chart of the cluster reconstitution and purification.

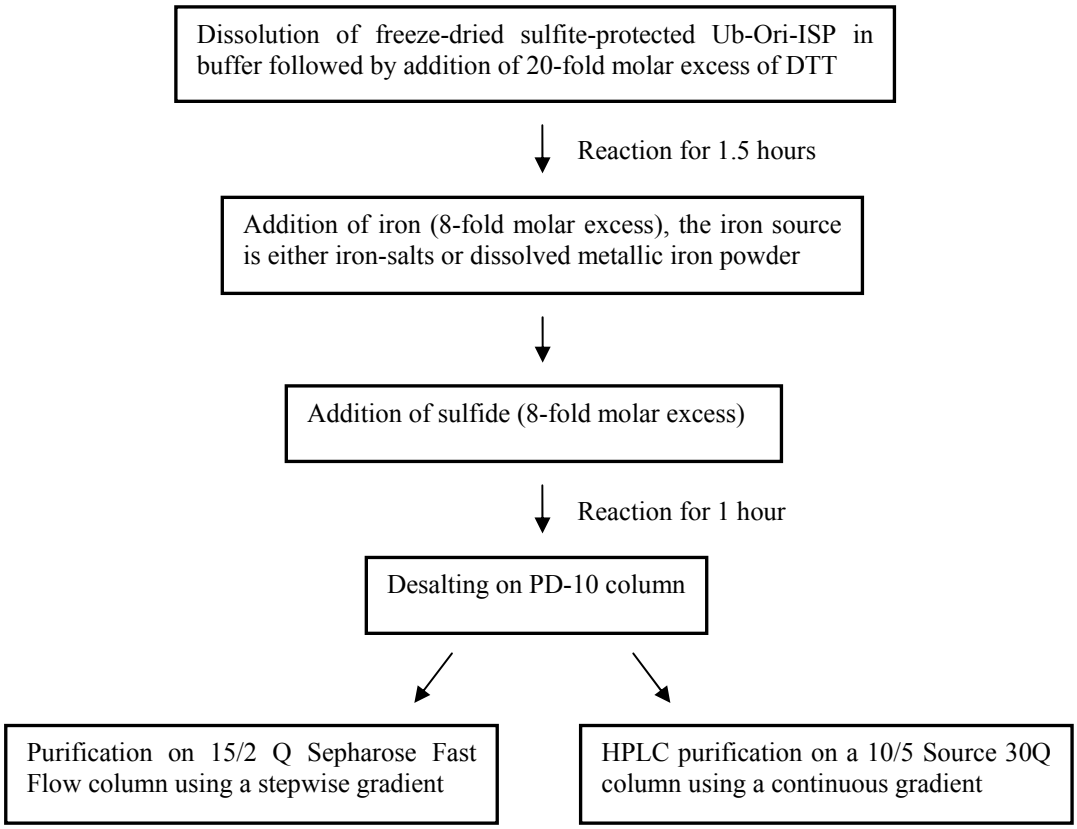


Figure 6.1 Flowchart of cluster reconstitution reaction into Ub-Ori-ISP and subsequent desalting and purification of reaction products.

6.1.1.1 Dissolution of iron powder in sulphuric acid

Dissolution of metallic iron powder was carried out as described (Jørgensen, 2003). 685 µl 1.4 M H_2SO_4 was added to 13.4 mg of metallic iron powder (H^+ is in 8-fold molar excess) and incubated for 18 hours at 65ºC. This was not carried out under anaerobic conditions. NaOH was added to the solution (5-fold molar excess compared to iron, from 1 M stock solution) to avoid overloading the buffer capacity of the protein solution (section 6.1.1). Then, the iron concentration was adjusted to 50 mM with milli-q water to obtain the wanted concentration (see section 6.1.1). After dilution the pH was < 1. An iron solution prepared as stated was determined (Nørgaard, 2006) to contain ca. 80 % Fe^{2+} and 20 % Fe^{3+}. The iron powder was either ^{57}Fe or common iron powder (containing natural isotope distribution).

6.1.2 Purification on a 15/2 Q Sepharose Fast Flow column

Protein desalted into 20 mM Tris/HCl, 0.15 M NaCl pH 8.0 (section 6.1.1) was diluted with 20 mM Tris/HCl pH 8.0 to a final NaCl concentration of 0.1 M and loaded onto a 15/2 Q Sepharose Fast Flow column (by gravity) equilibrated with 20 mM Tris/HCl, 0.1 M NaCl pH 8.0. After washing with 2 CV 20 mM Tris/HCl, 0.1 M NaCl pH 8.0, the protein was eluted with 20 mM Tris/HCl, 0.5 M NaCl pH 8.0. Fractions were collected and UV-vis spectra recorded.

6.1.2.1 Experiments with product obtained from purification by gravity

Several experiments were carried out to test the stability of the product obtained from the Q Sepharose Fast Flow column and to identify stabilising conditions. The effects of temperature, dilution, buffer, NaCl, glycerol, citrate and reducing agents were tested. Some of the experiments are outlined below. In all experiments $(NH_4)_2Fe(II)(SO_4)_2 \cdot 6H_2O$ was used as the iron source for the cluster reconstitution reaction and protein eluted in the main fraction from the anion-exchange was used. UV-vis spectra were recorded throughout the experiments. All stock solutions were freshly prepared.

1. A sample was diluted 2.5 times with the elution buffer and left undisturbed in the cuvette in air and at room temperature for 3 hours.
2. A sample was diluted 12.5 times with 20 mM Tris/HCl pH 8.0 (final salt concentration of 0.04 M) and kept at room temperature for 3 hours.
3. A sample was diluted 12.5 times with milli-q water (final buffer concentration of 1.6 mM and final salt concentration of 0.04 M), kept at room temperature for 3 hours.
4. A sample was diluted 12.5 times with elution buffer and kept at room temperature for 3 hours.
5. A sample was diluted 12.5 times with elution buffer and kept at 37ºC for 3 hours.
6. A sample was diluted 12.5 times with 20 mM Tris/HCl pH 8.0 (final salt concentration of 0.04 M) and tri-sodiumcitrate ($C_6H_5Na_3O_7 \cdot 2H_2O$, 100 mM stock solution) was added to a final concentration of 2.7 mM (100 times molar excess). The sample was kept at room temperature for 16 hours.
7. Same as experiment 2 however, the sample was kept at 0ºC for 48 hours.
8. Same as experiment 3 however, the sample was kept at 0ºC for 48 hours.
9. Same as experiment 4 however, the sample was kept at 0ºC for 48 hours.
10. A sample was diluted 12.5 times with 20 mM Tris/HCl, 5 % glycerol pH 8.0 (final salt and glycerol concentrations: 0.04 M and 4.6 %, respectively), kept at 0ºC for 24 hours.
11. A sample was diluted 12.5 times with 20 mM Tris/HCl, 10 % glycerol pH 8.0 (final salt and glycerol concentrations: 0.04 M and 9 %, respectively), kept at 0ºC for 48 hours.
12. A sample was diluted 12.5 times with 20 mM Tris/HCl, 20 % glycerol pH 8.0 (final salt and glycerol concentrations: 0.04 M and 18 %, respectively), kept at 0ºC for 48 hours.
13. A sample was diluted 12.5 times with 20 mM Tris/HCl, 30 % glycerol pH 8.0 (final salt and glycerol concentrations: 0.04 M and 28 %, respectively), kept at 0ºC for 48 hours.
14. A sample was diluted 12.5 times with 20 mM Tris/HCl, 40 % glycerol pH 8.0 (final salt and glycerol concentrations: 0.04 M and 37 %, respectively), kept at 0ºC for 48 hours.

15. An undiluted sample was frozen at -20ºC for 14 hours. The sample was sealed under anaerobic conditions and then transferred to the freezer. The sample was removed from the freezer, and then thawed under argon. UV-vis spectra were recorded before and after freezing.
16. A sample was diluted 2 times with 20 mM Tris/HCl, 40 % glycerol pH 8.0 (to a final salt concentration of 0.25 M, and a final glycerol concentration of 20 %). It was frozen at -20ºC for 24 hours and then thawed as described above. UV-vis spectra were recorded before and after freezing.
17. An undiluted sample was left at 0ºC for 72 hours.
18. An undiluted sample was left at room temperature for 72 hours.
19. A sample diluted 10 times with elution buffer was kept at 0ºC for 72 hours.
20. 50 times molar excess $Na_2S_2O_4$ (0.5 M stock solution, final concentration of 10 mM) was added to an undiluted sample and left at 0ºC for 72 hours.
21. 50 times molar excess $Na_2S_2O_4$ (0.5 M stock solution, final concentration of 1mM) was added to a sample diluted 10 times with elution buffer and left at 0ºC for 72 hours.
22. 20 times molar excess DTT (0.5 M stock solution, final concentration of 0.4 mM) was added to a sample diluted 10 times with elution buffer and left at 0ºC for 72 hours.

6.1.3 HPLC purification on a 10/5 Source 30Q column

If the protein was desalted into buffer containing 20 % glycerol (section 6.1.1), buffers used for further purification also contained 20 % glycerol. This is indicated in the following description with ± 20 % glycerol. Purification was performed an ÄKTA^{TM} purifier 100 HPLC system capable of measuring absorption at three wavelengths.

Protein exchanged into 20 mM Tris/HCl, 0.15 M NaCl pH 8.0 (± 20 % glycerol) was diluted with 20 mM Tris/HCl pH 8.0 (± 20 % glycerol) to a final NaCl concentration of 0.1 M and loaded onto a 10/5 Source 30Q column equilibrated with 20 mM Tris/HCl, 0.1 M NaCl pH 8.0 (± 20 % glycerol). After washing with 2 CV 20 mM Tris/HCl, 0.1 M NaCl pH 8.0 (± 20 % glycerol), the protein was eluted using a linear gradient, 0.10 M to 0.45 M NaCl in 20 CV and a linear flow rate of 2.6 cm/min. Fractions were collected and UV-vis spectra recorded.

It was tested if the obtained product could be re-purified. Product obtained when using $(NH_4)_2Fe(II)(SO_4)_2{\cdot}6H_2O$ as an iron source and purification buffers containing glycerol was diluted to ca. 0.1 M NaCl using 20 mM Tris/HCl, 20 % glycerol pH 8.0 and purified as described above.

6.1.3.1 Experiments with product obtained from purification by HPLC

Experiments were carried out on holo-protein obtained when cluster reconstitution reaction was performed using $FeCl_3$ as iron source and in the absence of glycerol in the purifications buffers. UV-vis spectra were recorded throughout the experiments and some of the experiments are outlined below.

23. To a sample containing undiluted protein, iron ($(NH_4)_2Fe(II)(SO_4)_2{\cdot}6H_2O$, 10 mM stock solution, final concentration of 0.2 mM) and cysteine (cysteine$\cdot H_2O\cdot$HCl, 0.05 M stock solution adjusted to pH 7.5 with 1 M NaOH, final concentration of 0.9 mM) were added

in 12 and 50 times molar excess, respectively. The sample was left at room temperature for 35 minutes.

24. Control sample for experiment 1. Same conditions as in experiment 1 but with no protein in the sample. Instead an equal volume of 20 mM Tris/HCl pH 8.0, 0.5 M NaCl was used.
25. An undiluted sample was frozen at -20ºC for two weeks and then thawed as described in section 6.1.2.1. UV-vis spectra were recorded before and after freezing.

6.1.4 Preparation of holo-protein for EPR and Mössbauer spectroscopic analyses

This part of the project was carried out in collaboration with M.Sc. student, Hanne Nørgaard. The procedure for cluster reconstitution is described in section 6.1.1 using 82 mg of protein (6 times 13.6 mg) dissolved in the reaction buffer to a concentration of 7.2 mg/ml. As previously, DTT was added in 20-fold molar excess (0.3 M stock solution) and iron and sulfide were both added in 8-fold molar excess (100 mM stock solutions). Dissolved ^{57}Fe (section 6.1.1.1) was used as the iron source and the product from the cluster reconstitution reaction was purified in buffers containing 20 % glycerol. Anion-exchange was performed on a 16/5 Source 30Q column under the same conditions as described in section 6.1.3. The main fraction was split into three portions. The first portion was concentrated without addition of any reducing agent while a 25-fold and a 100-fold excess of $Na_2S_2O_4$ (200 mM stock solution) were added to the second and third portions, respectively, prior to concentration of the samples. Samples were concentrated using a Vivacell 70 cell equipped with a 5000 molecular weight cut off (MWCO) membrane. From each of the three solutions, aliquots were taken for EPR and Mössbauer spectroscopic measurements. Samples for Mössbauer spectroscopic measurements were transferred to polymethylmethacrylate sample holders and then frozen in liquid nitrogen. Samples for EPR spectroscopic measurements were transferred to EPR tubes and frozen by immersing the tubes into isopentane cooled in a dry ice/ethanol bath. The EPR tubes were transferred to liquid nitrogen and stored herein.

The active iron species in Mössbauer spectroscopy is the isotope ^{57}Fe which has a nuclear spin of I = ½. When using ^{57}Fe in EPR samples, the nuclear spin causes broadening of the measured signal (Nørgaard, 2006). EPR and Mössbauer measurements were performed and analysed by Hanne Nørgaard (Nørgaard, 2006).

6.2 Results

6.2.1 Cluster reconstitution and desalting

In the cluster reconstitution procedure (section 6.1.1), sulfite-protected Ub-Ori-ISP was first de-protected with DTT. If Fe^{3+} was used as the iron source, the solution turned dark reddish brown upon addition (figure 6.2 A). If Fe^{2+} was used as the iron source, the solution turned yellowish brown upon addition. Independent on the iron source, the colour of the solution changed to dark blackish brown upon addition of sulfide (figure 6.2 B).

In the desalting step, the column turned brown upon loading of the protein (figure 6.2 C) solution and a dark brown fraction was collected upon elution. After elution the column

sometimes appeared almost clean (almost no coloured species seen on the column) while at other times the column was dark grey. Also, sometimes a broad pink band was seen on the column, and on the top of the column, a green precipitate remained (figure 6.2 D).

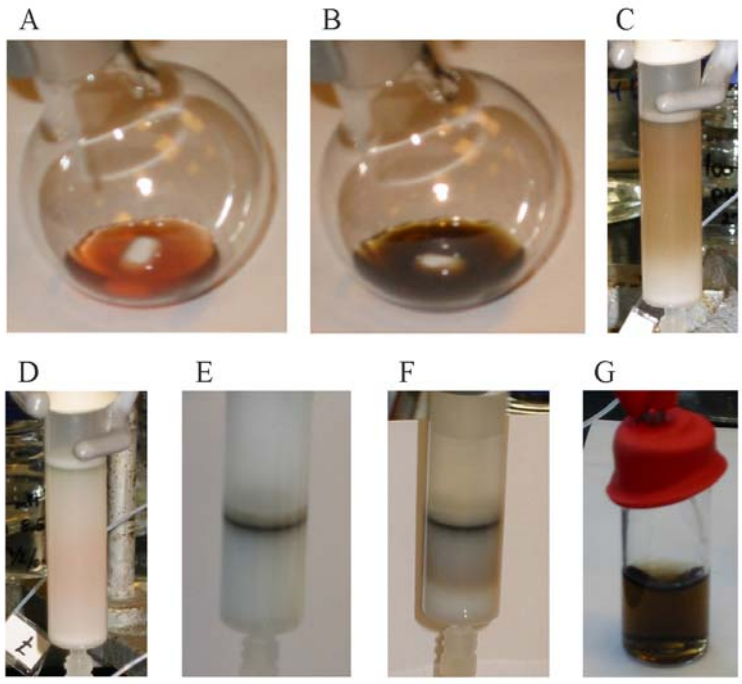


Figure 6.2 Reconstitution and purification of holo-Ub-Ori-ISP. A) Upon addition of $FeCl_3$ to the solution of sulfite de-protected Ub-Ori-ISP, the solution turned reddish brown. B) When sulfide was added, the solution turned dark blackish brown. C) Upon loading the solution onto the PD-10 column, a broad brown band appeared. D) After elution, a broad pink band was sometimes observed on the column, and a green precipitate remained on top. E) When the desalted solution was loaded to the Q Sepharose Fast Flow column, a narrow blackish brown band appeared on the column. F) Upon elution, a dark brown band moved down the column. G) The brown band was collected in one main fraction.

6.2.2 Purification on Q Sepharose Fast Flow column

When loading the protein sample onto the Q Sepharose Fast Flow column a dark narrow blackish brown band appeared at the top of the column (figure 6.2 E). Upon elution a dark broad band moved down the column (figure 6.2 F) and was collected in a primary fraction (dark brown, figure 6.2 G) and a few secondary fractions (tail end of the band). The dark brown colour is consistent with the colour of iron-sulfur proteins (section 2.1). Only the primary fraction was used for further experiments. After elution a black band remained at the top of the column and the column had a grey colour, which could be removed when the column was washed with 1 M HCl.

Typical UV-vis spectra of the primary fractions obtained when using ferrous or ferric iron sources are shown in figure 6.3. In both spectra an absorption maximum at ca. 280 nm, a shoulder at ca. 310 nm and a shoulder or an absorption maximum at ca. 390 nm can be seen.

The yield of the reaction under different conditions was calculated based on an assumed molar absorptivity of 17 $mM^{-1} \cdot cm^{-1}$ at 390 nm (Brereton *et al.*, 1999; Conover *et al.*, 1990).

The yield of the reaction was found in the primary fraction to vary from 30 % to 50 % and in the secondary fractions to vary from 2 % to 10 %. A total yield of ca. 40 % to 55 % was obtained.

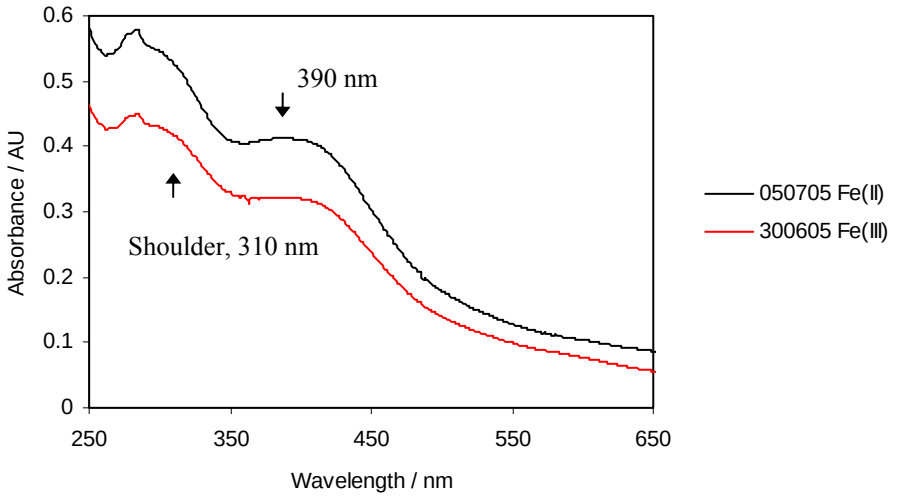


Figure 6.3 UV-vis spectra of reconstituted Ub-Ori-ISP. Spectra are those of main fractions (diluted 12.5 times with elution buffer) obtained from purification on Q Sepharose Fast Flow column. They were obtained when using $(NH_4)_2Fe(II)(SO_4)_2 \cdot 6H_2O$ or $FeCl_3$ as iron source and shown in black and red, respectively.

6.2.2.1 Experiments with product obtained from purification by gravity

Short term stability (experiments 1 to 5)

Experiments 1 to 5 (section 6.1.2.1) explored the short term stability of the obtained product at different conditions. In figure 6.4 A the spectrum of a sample exposed to air at room temperature (experiment 1) was recorded over a period of three hours. Within the first 80 minutes the shoulder at 310 nm was unchanged while a minor flattening of the shoulder at 390 nm was observed. A marked increase of the absorption above 450 nm was observed. After 110 minutes almost all characteristics had disappeared and after 180 minutes they were all gone. Figure 6.4 B shows the spectrum of a sample diluted in 20 mM Tris/HCl pH 8.0 and kept under anaerobic conditions at room temperature (experiment 2) over a period of three hours. Unlike in the previous experiment, the shoulders at 310 nm and 390 nm were preserved. However, a slight decrease in absorbance was observed throughout the spectrum. The results of experiments 3 to 5 are not shown since they resembled the results obtained from experiment 2.

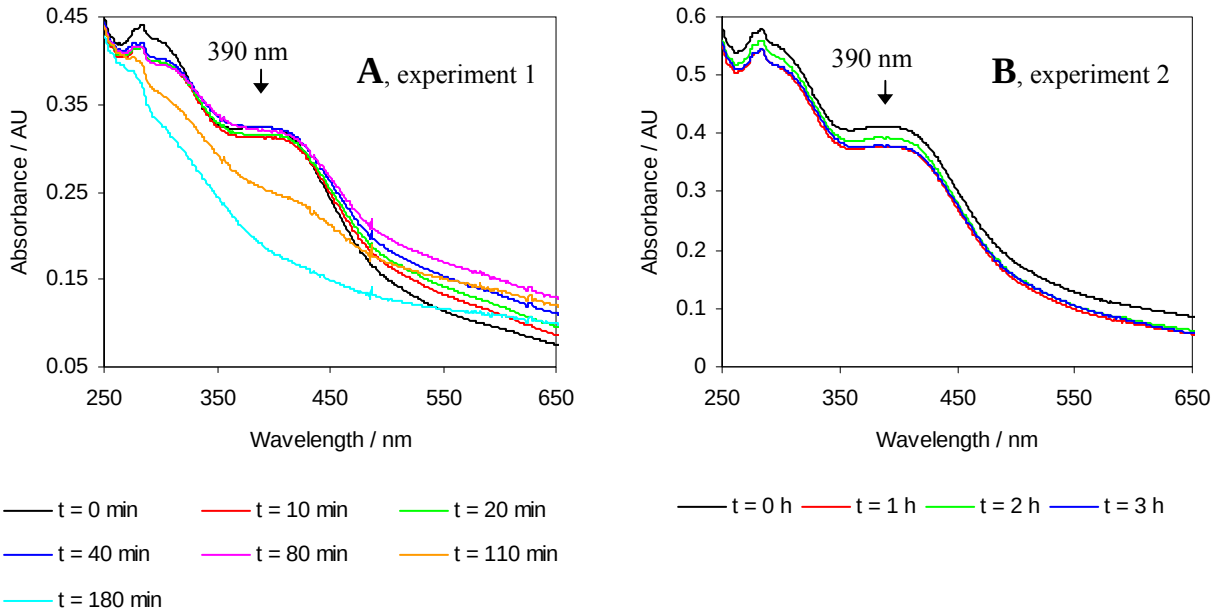


Figure 6.4 Short term stability of holo-Ub-Ori-ISP obtained from purification on Q Sepharose Fast Flow column. Results from experiment 1 and 2 (section 6.1.2.1) are shown in A and B, respectively. UV-vis spectra were recorded over a period of three hours as indicated in both panels. A) A sample diluted with elution buffer was exposed to air. B) A sample diluted with elution buffer was kept under anaerobic conditions.

Effect of citrate exposure (experiment 6)

Experiment 6 (section 6.1.2.1) explored the stability of the holo-protein in the presence of citrate. From the UV-vis spectra (figure 6.5) it is seen that the shoulder at 310 nm turned into a distinct absorption maximum within three hours, and after 16 hours it became the dominating feature of the spectrum with a maximum at ca. 330 nm. At the same time the absorption maximum at 390 nm shifted to higher wavelengths and decreased in intensity. After 16 hours an absorption maximum at about 420 nm and a shoulder at about 465 nm were observed.

Stability upon dilution and effect of glycerol (experiments 7 to 14)

Experiments 7 to 14 (section 6.1.2.1, results shown in figure 6.6) explored the stability of the holo-protein at 0ºC when diluted with either milli-q water, 20 mM Tris/HCl pH 8.0, elution buffer or 20 mM Tris/HCl pH 8.0 containing different concentrations of glycerol (5 % - 40 %) over a period of 48 hours (experiment 10 only 24 hours). The general trend in all the experiments was that the absorbance decreased over time and that the absorption maximum at 390 nm was reduced to a shoulder. In experiment 14 (40 % glycerol in dilution buffer, figure 6.6 H) the distinctive features disappeared after 24 hours and in experiment 7 (dilution with 20 mM Tris/HCl pH 8.0, figure 6.6 A) the features were reduced to very faint shoulders. Experiment 12 (20 % glycerol in dilution buffer, figure 6.6 F) showed the slowest decrease in absorbance.

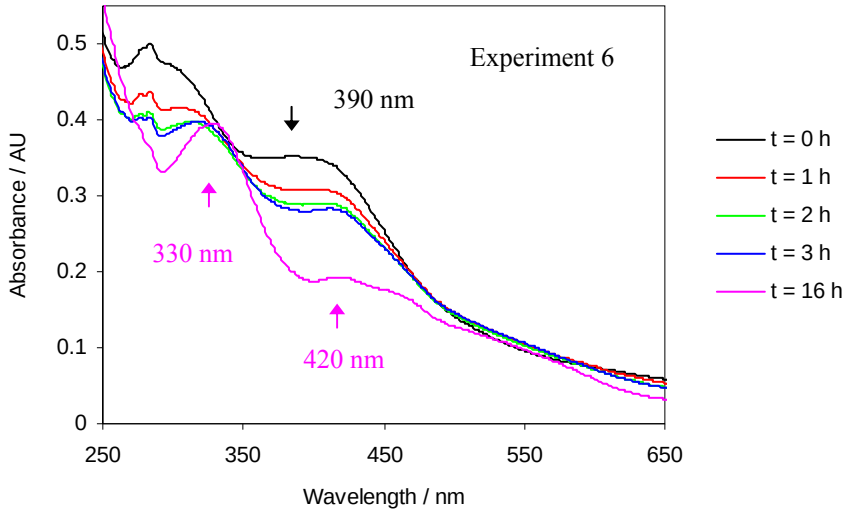
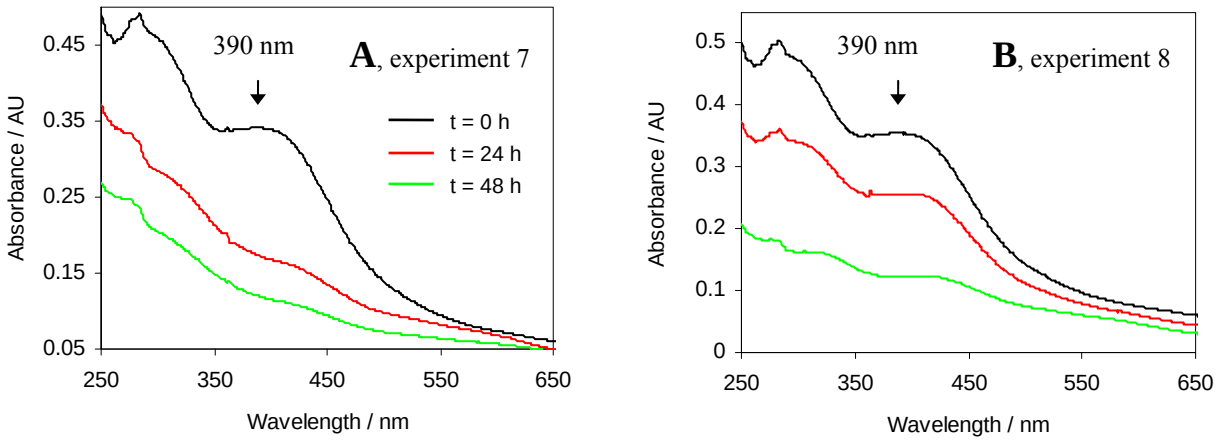


Figure 6.5 Results of experiment 6 (section 6.1.2.1). Tri-sodium citrate was added to a diluted sample obtained from purification on Q Sepharose Fast Flow column and UV-vis spectra were recorded throughout a time period of 16 hours.

Stability upon freezing and thawing (experiment 15 and 16)

Experiment 15 and 16 (section 6.1.2.1, results shown in figure 6.7) explored the stability of the holo-protein upon freezing followed by thawing. Figure 6.7 A shows that when an undiluted sample was frozen for 14 hours and then thawed a decrease in absorption was observed below 550 nm. Also, the absorption maximum at 390 nm was reduced to a flat shoulder. Figure 6.7 B shows that freezing a sample in 20 % glycerol caused a slight shift in the shoulder at 310 nm to about 325 nm. The maximum at 390 nm shifted to a shoulder at 410 nm while the absorbance above 450 nm increased.



*(**Figure 6.6**: figure legend found on the next page)*

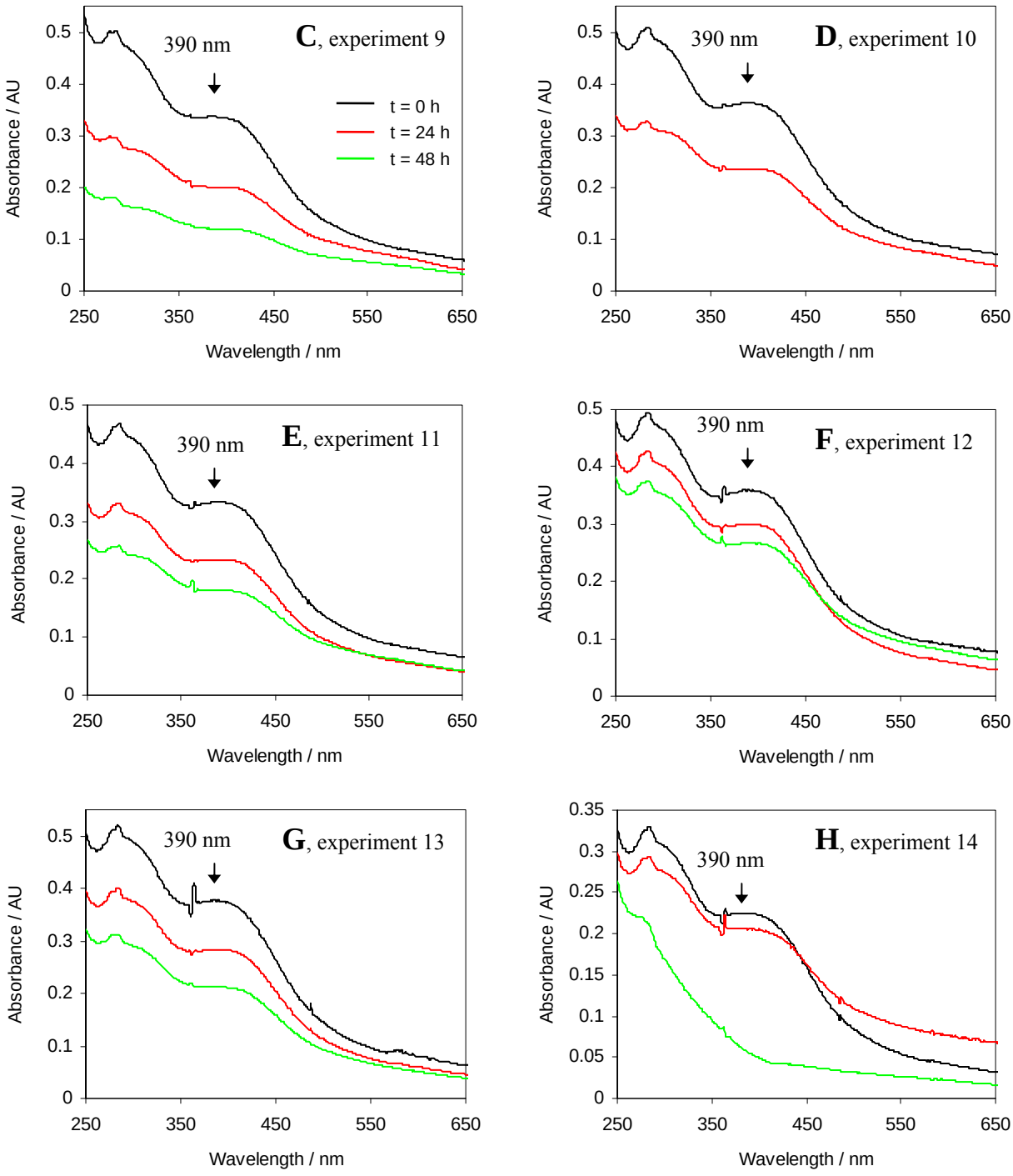


Figure 6.6 Stability of diluted samples kept at 0ºC from purification on Q Sepharose Fast Flow column. The results of experiment 7 to 14 (section 6.1.2.1) are shown in panels A to H. Black spectra were recorded just after dilution, red spectra were recorded after 24 hours while green spectra were recorded after 48 hours. The feature at 360 nm is an instrument artefact. A) Experiment 7; protein solution was diluted with 20 mM Tris/HCl pH 8.0. B) Experiment 8; protein solution was diluted with milli-q water. C) Experiment 9; protein solution was diluted with elution buffer. D) Experiment 10; protein solution was diluted with 20 mM Tris/HCl, 5 % glycerol pH 8.0. E) Experiment 11; protein solution was diluted with 20 mM Tris/HCl, 10 % glycerol pH 8.0. F) Experiment 12; protein solution was diluted with 20 mM Tris/HCl, 20 % glycerol pH 8.0. G) Experiment 13; protein solution was diluted with 20 mM Tris/HCl, 30 % glycerol pH 8.0. H) Experiment 14; protein solution was diluted with 20 mM Tris/HCl, 40 % glycerol pH 8.0.

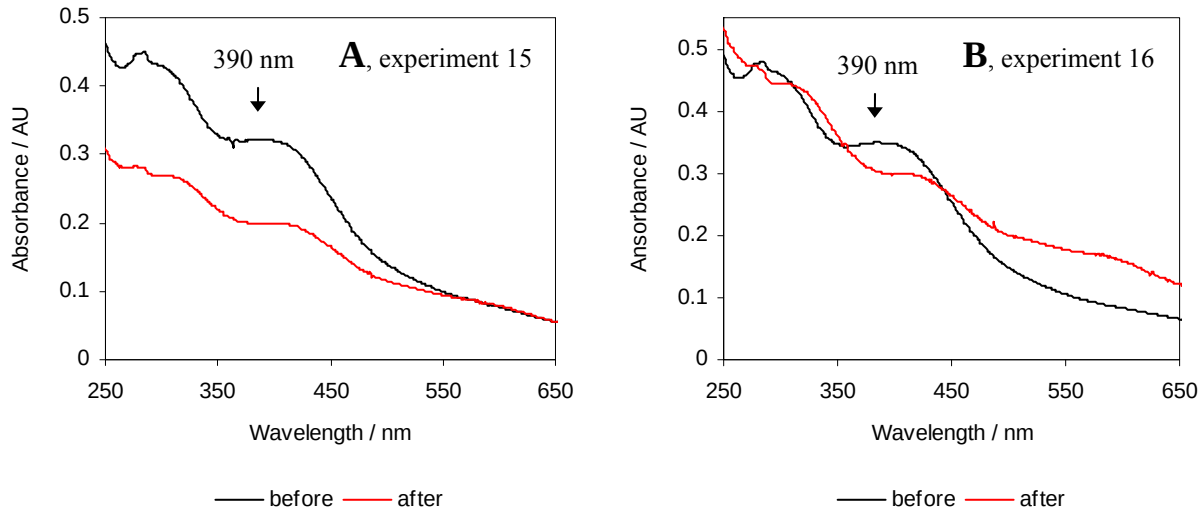


Figure 6.7 Stability of the product obtained from purification on Q Sepharose Fast Flow column upon freezing under different conditions. The results from experiments 15 and 16 (see section 6.1.2.1 for details) are shown in panels A and B, respectively. The black spectra were recorded prior to freezing the samples while the red spectra were recorded after thawing the samples. In experiment 15 an aliquot was diluted 12.5 times prior to recording a spectrum, while in experiment 16 the aliquot was diluted 5 times.

Effects of temperature and dilution (experiments 17 to 19)

Experiments 17 to 19 (section 6.1.2.1, results shown in figure 6.8) explored the stability of undiluted samples left at 0ºC and at room temperature, respectively, versus a diluted sample left at 0ºC. In all experiments a flattening of the peaks was observed, the absorption maximum at 390 nm was reduced to a faint shoulder. In the experiment (19) leaving a diluted sample on ice the characteristics had almost disappeared after 48 hours, which was also the case when an undiluted sample was left at room temperature (experiment 18). In the experiment with an undiluted sample left at 0ºC (experiment 17) the characteristics were still present after 72 hours.

Effects of reducing agents (experiments 20 to 22)

Experiment 20 to 22 (section 6.1.2.1, results shown in figure 6.9) explored the effect of a reducing agent on the holo-protein. When dithionite was added to a diluted or undiluted sample (figure 6.9 A and B), the characteristic features in the spectra disappeared just after addition and then returned. However, the absorption at 390 nm was a shoulder. This effect was more pronounced in the undiluted sample (experiment 20), where the characteristics were still present after 72 hours while they had disappeared for the diluted sample. In the undiluted sample an absorption maximum observed at about 315 nm can be assigned to the dithionite radical (Mayhew, 1978). Upon addition of DTT to a diluted sample, all characteristics disappeared and they did not return (figure 6.9 C).

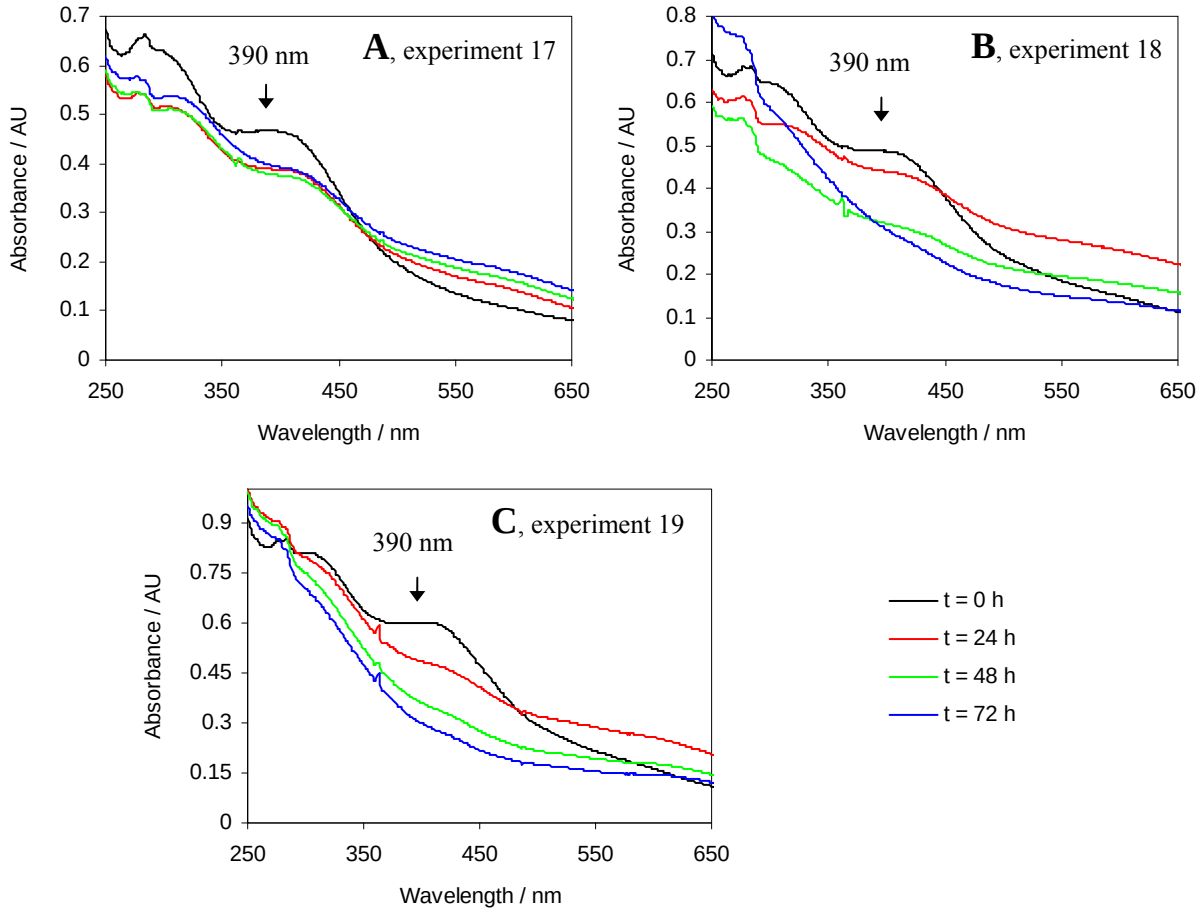


Figure 6.8 Stability of undiluted samples left at 0ºC (A) and at room temperature (B), respectively, versus a diluted sample left at 0ºC (C). Results of experiments 17 to 19 are shown in panels A to C. UV-vis spectra were recorded over a period of 72 hours as indicated in the figure. In experiment 17 and 18 an aliquot was taken out and diluted 10 times prior to recording a spectrum.

6.2.3 HPLC purification on a 10/5 Source 30Q column

Observations made during HPLC protein purification were comparable to those made using the Q Sepharose Fast Flow column (see section 6.2.2). The chromatograms obtained using different iron sources and purification buffers with or without glycerol are presented in figure 6.10. In all chromatograms a main peak is observed followed by tailing. However, the size and resolution of the main peak compared to the tailing and minor peaks differ markedly.

$FeCl_3$, $NH_4Fe(III)(SO_4)_2 \cdot 12H_2O$ and $(NH_4)_2Fe(II)(SO_4)_2 \cdot 6H_2O$ were tested as iron sources without glycerol in the purification buffers. Chromatograms for the purifications are shown in figure 6.10 A to C and the main peaks eluted from 19.5 mS/cm to 21.0 mS/cm. UV-vis spectra corresponding to main peak fractions are presented in figure 6.11 A to C. Two absorption maxima at about 320 nm and 420 nm, respectively, are present in the three spectra; the exact positions differ in each spectrum. Also, a slight shoulder at about 460 nm is apparent.

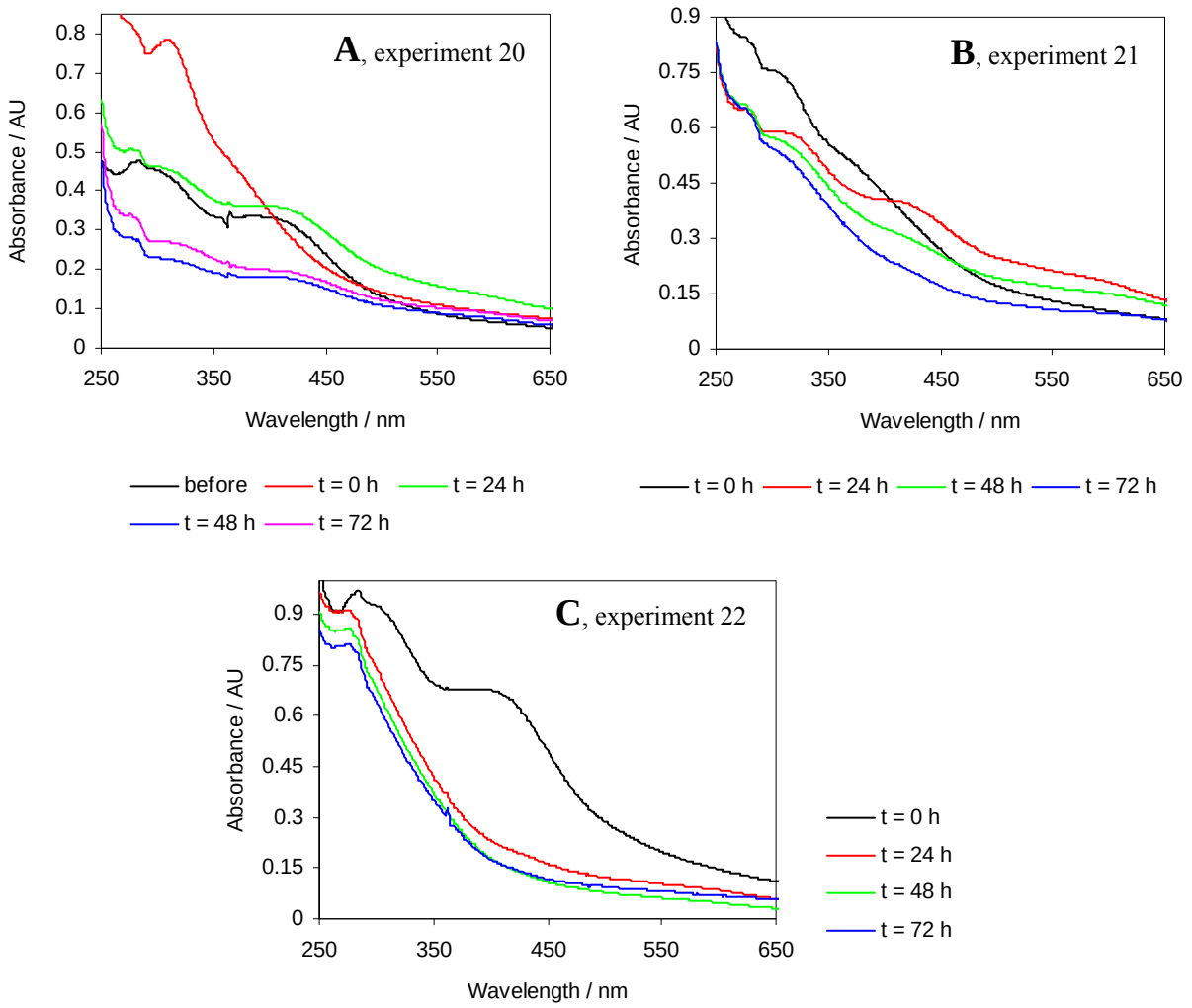
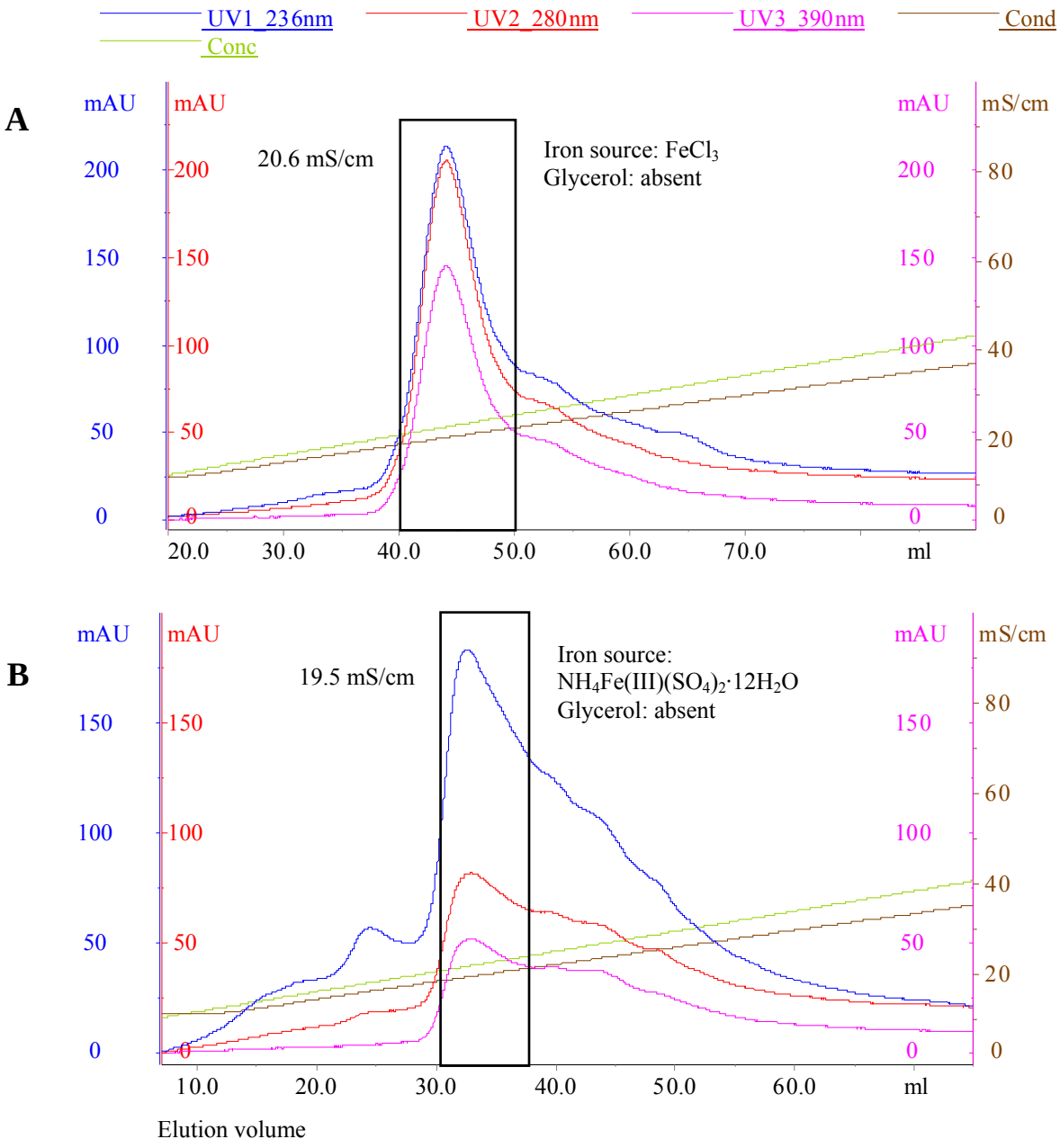
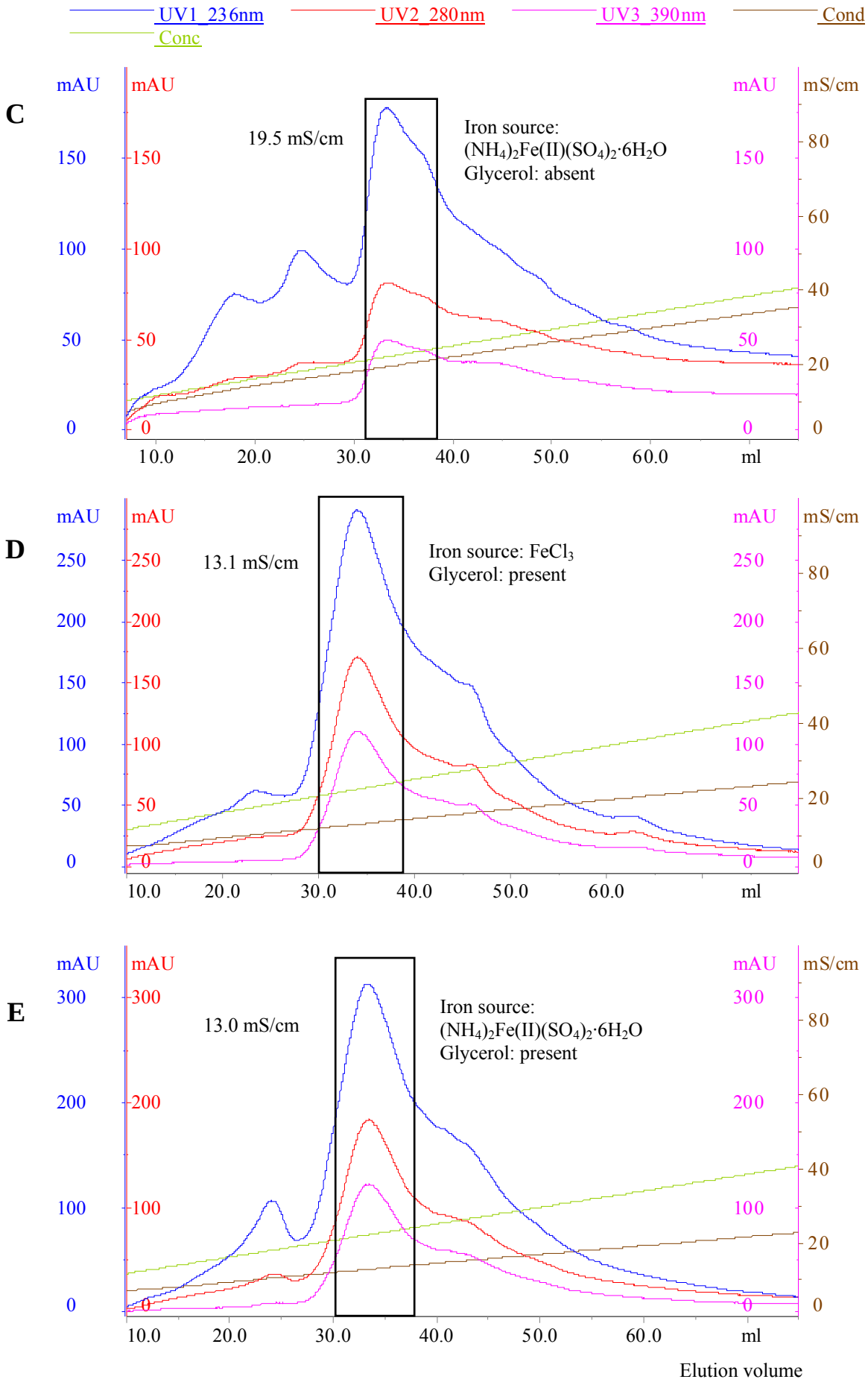


Figure 6.9 Reduction of samples from purification on Q Sepharose Fast Flow column. The results from experiments 20 to 22 (see section 6.1.2.1 for details) are shown in panels A to C. A) Reduction of an undiluted sample with dithionite. Prior to recording a spectrum an aliquot was diluted 10 times. B) Reduction of a diluted sample with dithionite. C) Reduction of a diluted sample with DTT.

$FeCl_3$, $(NH_4)_2Fe(II)(SO_4)_2 \cdot 6H_2O$, ^{57}Fe powder and common iron powder were tested as iron sources with glycerol in the purification buffers. Chromatograms for the purifications are shown in figure 6.10 D to G and the main peaks eluted from 13.0 to 14.0 mS/cm (glycerol reduces the conductivity of a solution). When iron-powder was used the UV-vis spectra of main fractions (figure 6.11 F and G) resembled those obtained from purification on the Q Sepharose Fast Flow column (sections 6.1.2 and 6.2.2 and figure 6.3). A maximum at 390 nm was present when ^{57}Fe-powder was used and a shoulder at about 400 nm was observed when common Fe-powder was used. The UV-vis spectra obtained when glycerol was present in purification buffers and $FeCl_3$ or $(NH_4)_2Fe(II)(SO_4)_2 \cdot 6H_2O$ was used (figure 6.11 D and E), resembled those obtained when no glycerol was used (figure 6.11 A to C).



*(**Figure 6.10**: figure legend found on page 56)*



(***Figure 6.10**: figure legend found on the next page)*

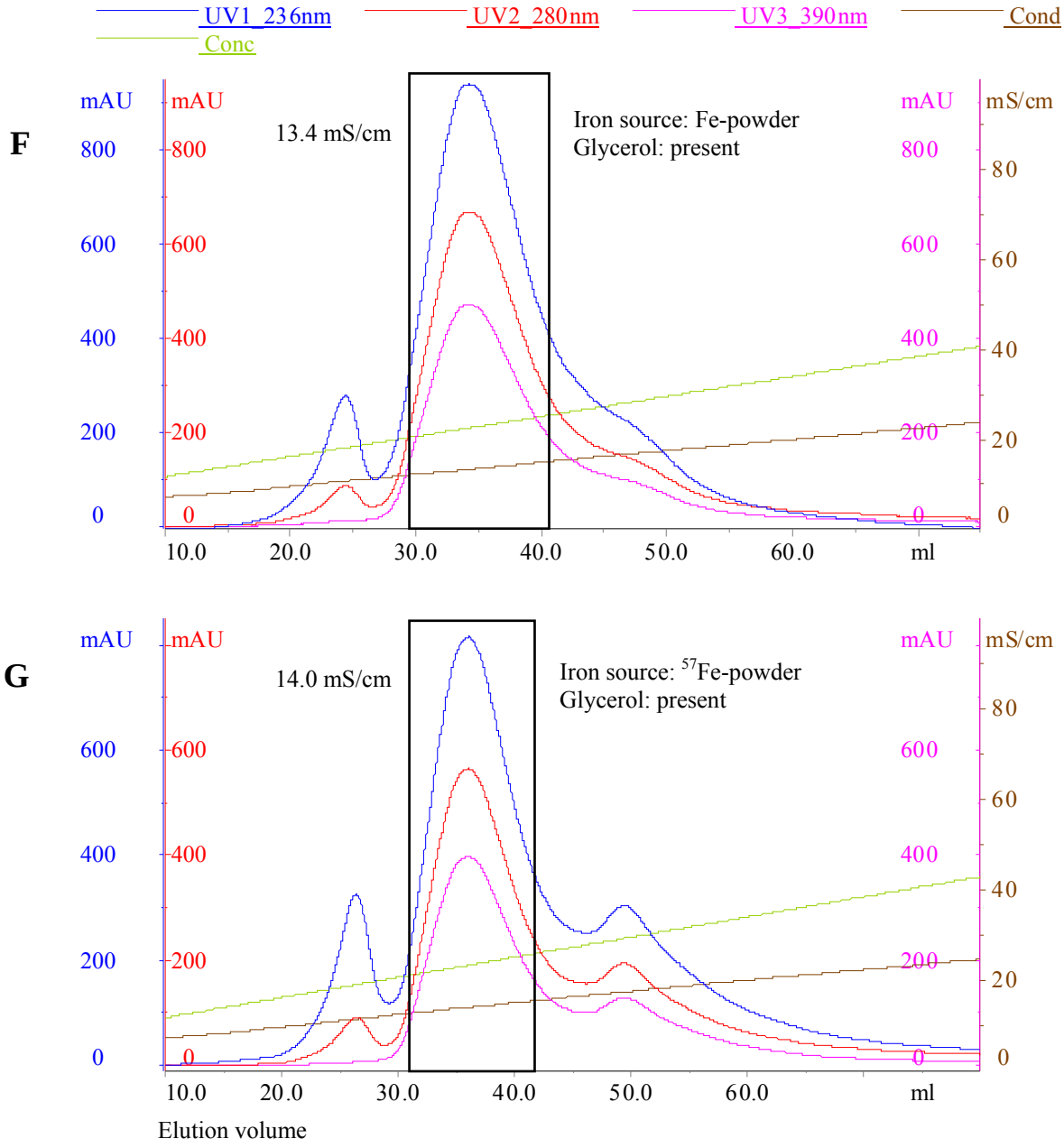


Figure 6.10 Chromatograms for purification on a 10/5 Source 30Q column of holo-protein obtained under different conditions. In each chromatogram the collected fraction is indicated with a black box and the conductivity at maximum absorption at 280 nm is also presented. The blue, red and pink lines indicate the absorbances at 236 nm, 280 nm, and 390 nm, respectively. The brown and the green lines indicate the conductivity and the NaCl gradient, respectively. The concentration of NaCl can be read from the conductivity axis employing the unit 10^{-2} M.

A) $FeCl_3$ was used as iron source and no glycerol was added to the purification buffers.

B) $NH_4Fe(III)(SO_4)_2 \cdot 12H_2O$ was used as iron source and no glycerol was added to purification buffers.

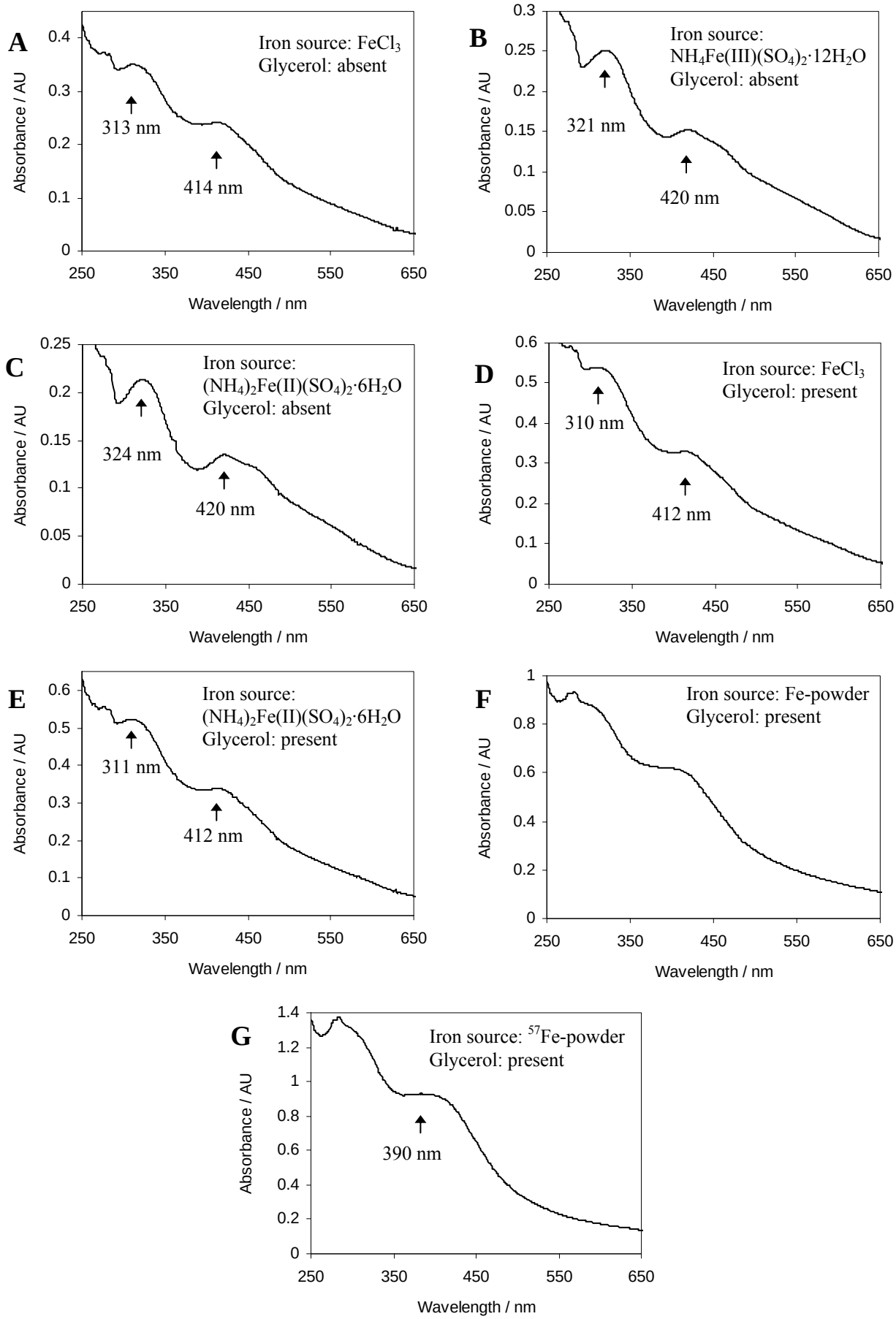
C) $(NH_4)_2Fe(II)(SO_4)_2 \cdot 6H_2O$ was used as iron source and no glycerol was added to purification buffers.

D) $FeCl_3$ was used as iron source and glycerol was added to the purification buffers.

E) $(NH_4)_2Fe(II)(SO_4)_2 \cdot 6H_2O$ was used as iron source and glycerol was added to purification buffers.

F) Dissolved Fe-powder was used as iron source and glycerol was added to purification buffers. Twice the amount of protein used in A to E was purified.

G) Dissolved ^{57}Fe-powder was used as iron source and glycerol was added to purification buffers. Same amount of protein as in F was purified.



*(**Figure 6.11**: figure legend found on the next page)*

Figure 6.11 UV-vis spectra of main fractions collected from purification of holo-protein on 10/5 Source 30Q column (see figure 6.10). In each spectrum absorption maxima are indicated. Numbering of the spectra corresponds to the numbering of the chromatograms in figure 6.10 from which the analysed fractions originated. In F, the fraction was diluted two times prior to recording the spectrum
A) $FeCl_3$ was used as iron source and no glycerol was added to the purification buffers.
B) $NH_4Fe(III)(SO_4)_2 \cdot 12H_2O$ was used as iron source and no glycerol was added to purification buffers.
C) $(NH_4)_2Fe(II)(SO_4)_2 \cdot 6H_2O$ was used as iron source and no glycerol was added to purification buffers.
D) $FeCl_3$ was used as iron source and glycerol was added to the purification buffers.
E) $(NH_4)_2Fe(II)(SO_4)_2 \cdot 6H_2O$ was used as iron source and glycerol was added to purification buffers.
F) Dissolved Fe-powder was used as iron source and glycerol was added to purification buffers. Twice the amount of protein used in A to E was purified.
G) Dissolved ^{57}Fe-powder was used as iron source and glycerol was added to purification buffers. Same amount of protein as in F was purified.

The chromatogram obtained from re-purification of one of the main-peak fractions (section 6.1.3) is shown in figure 6.12. Only this one re-purification was performed. One main peak was observed followed by tailing and the absorption was very low compared to the original chromatogram (from which the re-purified fraction was obtained) (figure 6.10 E). UV-vis spectra were recorded however, only absorption at about 280 nm was apparent from the spectra (results not shown) even though the chromatogram indicated weak absorption at 390 nm.

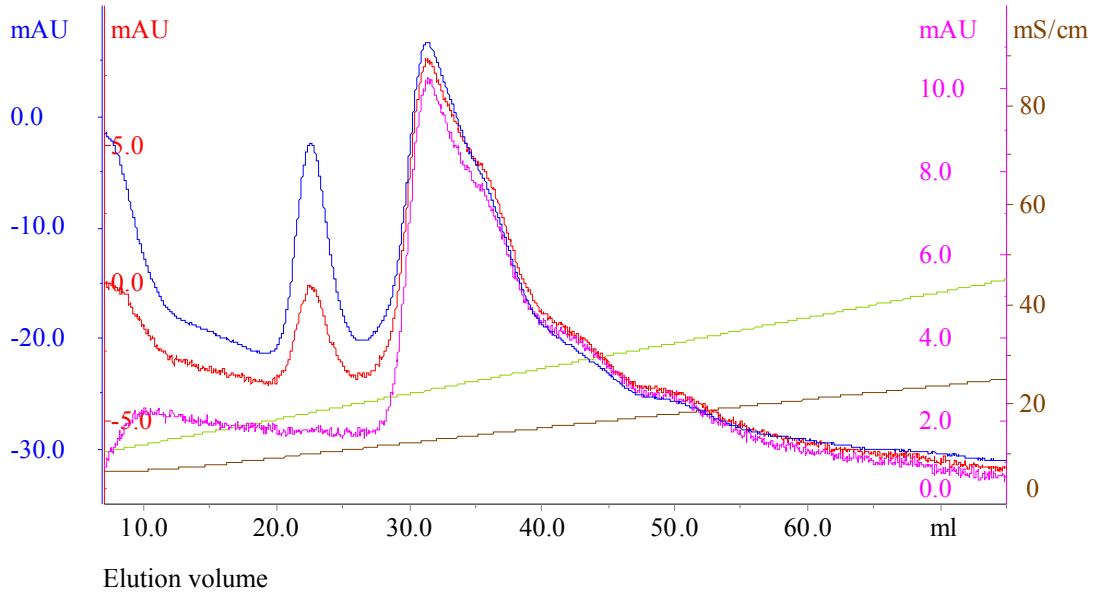


Figure 6.12 Chromatogram for re-purification on a 10/5 Source 30 Q column of purified holo-protein (figure 6.10 E) obtained using $(NH_4)_2Fe(II)(SO_4)_2 \cdot 6H_2O$ as the iron source and glycerol in purification buffers. The blue, red and pink lines indicate the absorbances at 236 nm, 280 nm, and 390 nm, respectively. The brown and the green lines indicate the conductivity and the NaCl gradient, respectively. The concentration of NaCl can be read from the conductivity axis employing the unit 10^{-2} M.

6.2.3.1 Experiments with product obtained from purification by HPLC

The results obtained from experiment 23 and 24 (section 6.1.3.1) are shown in figure 6.13. Addition of cysteine and Fe^{2+} to the protein solution (experiment 23), caused a shift in the absorption maximum at 414 nm to 385 nm, while the maximum at 313 nm disappeared. The absorption maximum at 280 nm remained but the absorbance increased. The absorbance above ca. 400 nm decreased. In the control experiment (24), upon mixing Fe^{2+} and cysteine, slight shoulders at ca. 290 nm and 350 nm as well as a broad peak at ca. 500 nm were observed. However, the absorbance was very low compared to when holo-protein was present.

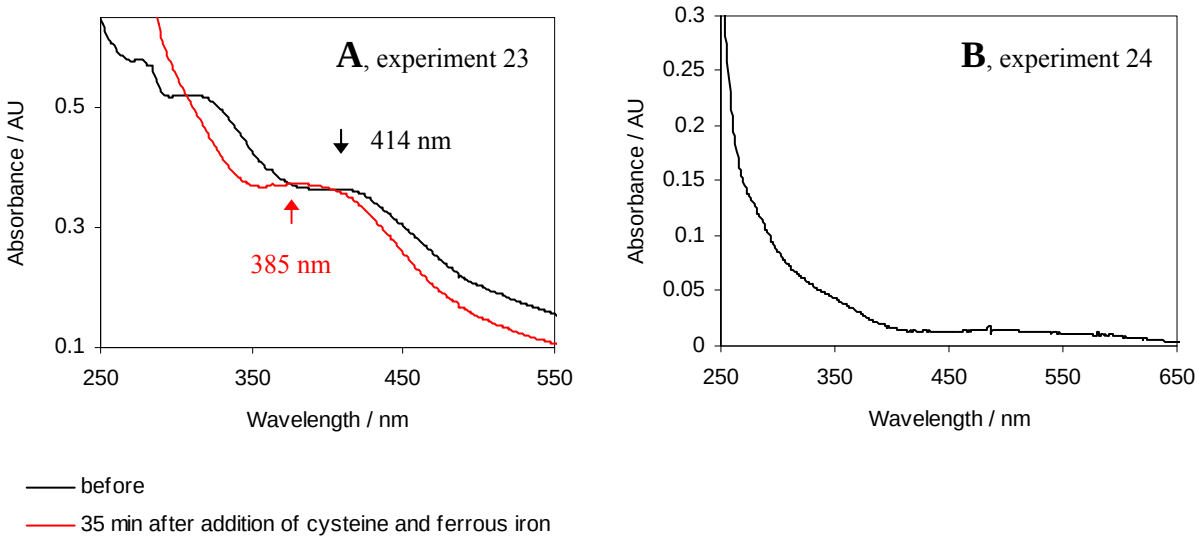


Figure 6.13 Effect of addition of cysteine and ferrous iron to purified holo-protein. The results of experiment 23 and 24 is shown in A and B, respectively. A) UV-vis spectra recorded before and 35 minutes after addition of cysteine and ferrous iron to purified holo-protein. B) UV-vis spectrum obtained from control experiment.

The result of experiment 25 is shown in figure 6.14. Upon freezing the protein and thawing two weeks after, the absorption maximum at 414 nm disappeared and the maximum at 313 nm had been reduced to a shoulder. The absorbance below ca. 400 nm had increased markedly.

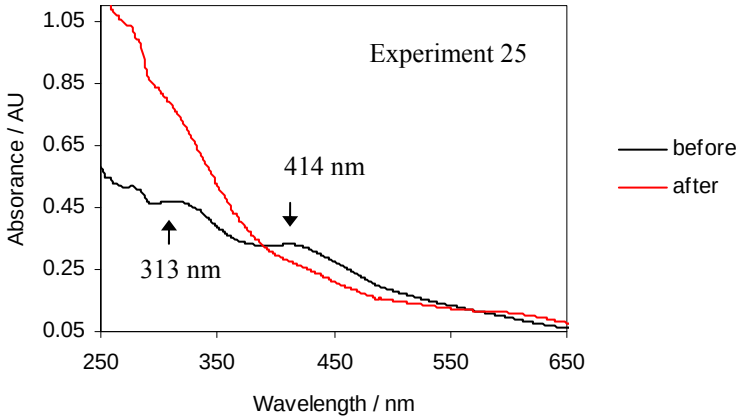


Figure 6.14 Effect of freezing a purified protein sample for two weeks. The result of experiment 25 is shown in the figure. A UV-vis spectrum was recorded before and after freezing the sample.

6.2.4 EPR and Mössbauer analyses of holo-protein

The chromatogram obtained from purification of sample to be used for EPR and Mössbauer spectroscopic measurements (see section 6.1.4) is shown in figure 6.15. One main peak was observed followed by a little tailing and poorly resolved minor peaks. In front of the main peak a minor peak was present with tailing in front of the peak. This peak did not show absorption at 390 nm. The UV-vis spectrum recorded of the main peak is shown in figure 6.16 A. Absorption maxima were observed at ca. 280 nm and at 390 nm and a shoulder was present at ca. 310 nm. UV-vis spectra were also recorded of the minor peaks following the main peak (not shown) and the spectra very closely resembled that of the main fraction. The yield of the reaction and purification was ca. 45 % (calculated for the main peak as described in section 6.2.2).

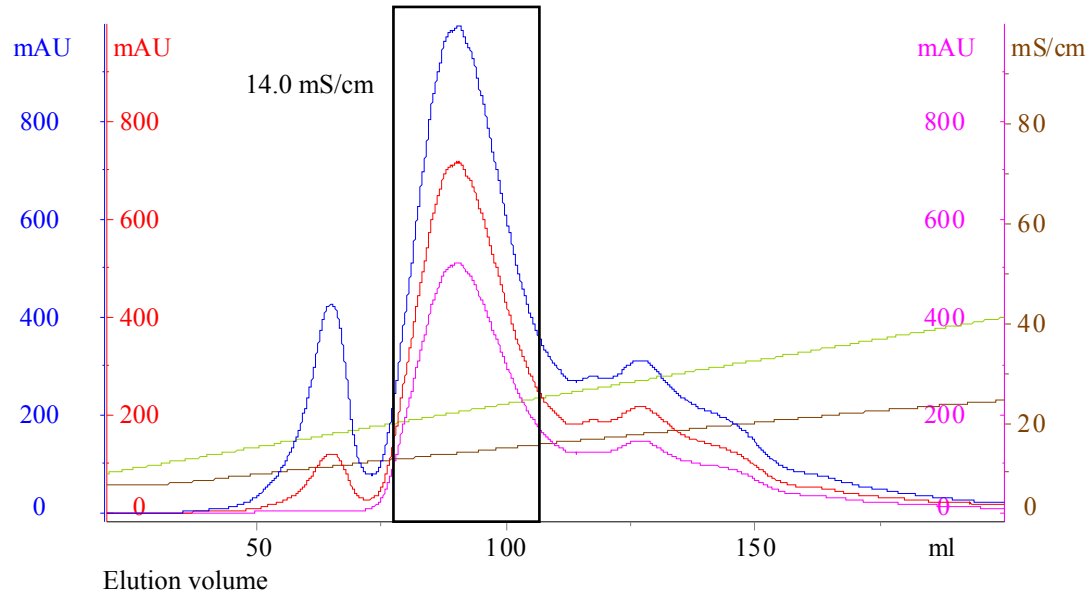


Figure 6.15 Chromatogram for purification of holo-protein for preparation of EPR and Mössbauer samples. In the chromatogram the collected fraction is indicated with a black box and the conductivity at maximum absorption at 280 nm is also presented. The blue, red and pink lines indicate the absorbances at 236 nm, 280 nm, and 390 nm, respectively. The brown and the green lines indicate the conductivity and the NaCl gradient, respectively. The concentration of NaCl can be read from the conductivity axis employing the unit 10^{-2} M.

From the main peak three samples were prepared for EPR and Mössbauer spectroscopic measurements (section 6.1.4 and table 6.1). The UV-vis spectra of the samples are shown in figure 6.16 B. From the spectra of sample 1 and 2 it is seen that the absorption maximum found previously at 390 nm (figure 6.16 A) had been reduced to a shoulder and the absorbance above 450 nm had increased. The spectrum of sample 3 on the other hand resembled the spectrum of the main fraction (figure 6.16 A).

The EPR spectra are not shown since the recorded signals were very weak and could not be convincingly assigned to any compounds. The Mössbauer results obtained from sample 1 and 2 are not shown since they could not be unambiguously interpreted. The Mössbauer spectrum of sample 3 obtained at 6 K is shown in figure 6.17 along with simulated spectra of

adventitious Fe^{2+} and a $[4Fe\text{-}4S]^{2+}$ cluster and the sum of the fits. These results indicate that 95.3 % of the ^{57}Fe was found as $[4Fe\text{-}4S]^{2+}$ clusters while 4.7 % was found as adventitiously bound Fe^{2+}. A detailed interpretation of the EPR and Mössbauer results has been presented elsewhere (Nørgaard, 2006).

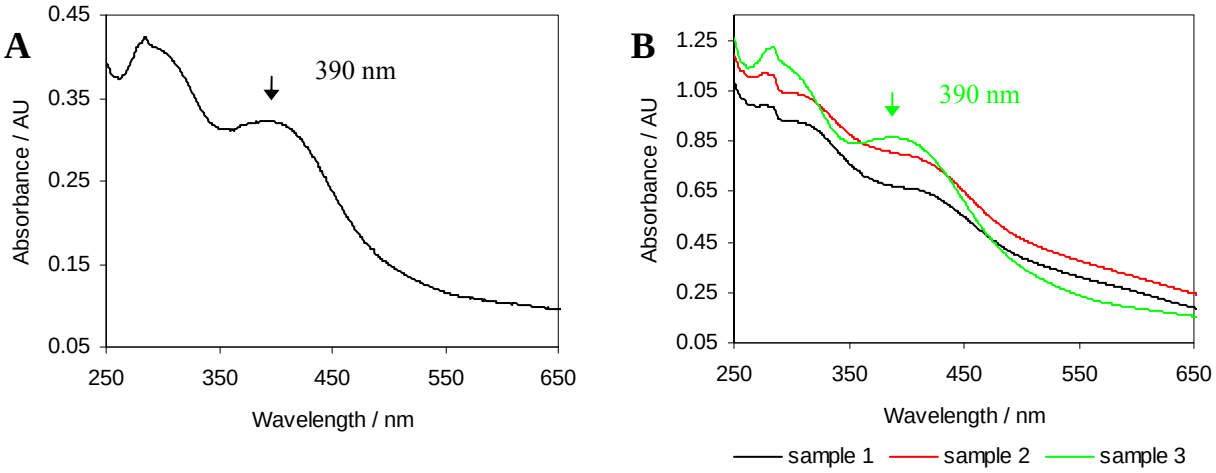


Figure 6.16 UV-vis spectra of purified holo-protein and samples prepared for EPR and Mössbauer spectroscopic measurements. A) UV-vis spectrum of the main fraction obtained upon purification. The fraction was diluted five times prior to recording the spectrum. B) UV-vis spectra of the EPR and Mössbauer samples (see table 6.1). The samples were diluted 10 times prior to recording the spectra.

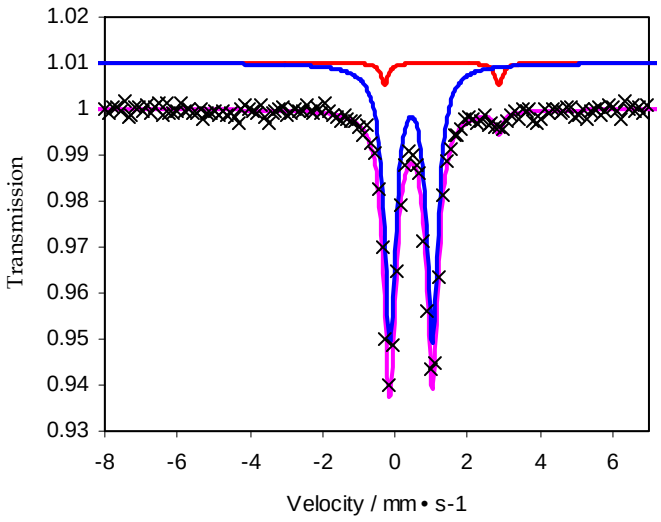


Figure 6.17 Mössbauer spectrum of sample 3. The dots correspond to the measured spectrum. The red and blue curves are fits of adventitious bound Fe(II), 4.7%, and simulated $[4Fe\text{-}4S]^{2+}$ cluster, 95.3%, respectively. These are offset 0.01 on the y-axis. The pink curve is the sum of fits. The figure was prepared by Hanne Nørgaard and printed with permission.

Table 6.1 Overview of EPR and Mössbauer samples

Sample	Protein concentration* / mM	Molar excess dithionite added prior to concentration
1	0.39	-
2	0.47	25
3	0.51	100

* The concentration was determined by measuring the absorbance at 390 nm and assuming a molar absorbtivity at this wavelength of 17 $mM^{-1} \cdot cm^{-1}$

6.3 Discussion

In all experiments a brown coloured protein solution was obtained when performing cluster reconstitution reaction follow by purification, which indicates the formation of FeS-clusters. A general trend was observed in all experiments, namely that the intensity of the brown colour faded and turned into a blackish brown colour over time. This is probably due to cluster breakdown and formation of iron sulfide. However, no visible precipitation was formed.

When performing desalting after cluster reconstitution the appearance of the column differed from experiment to experiment (section 6.2.1). The pink band and the green precipitation are probably due to formation of Fe-DTT complexes, while the grey species are from precipitations of iron sulfide. No relation could be established between the state of the column after elution and the iron source.

6.3.1 Holo-protein obtained form purification by gravity

After purification on the Q Sepharose Fast Flow column a black band remained on the column and the column material had turned grey. These grey or black species can be assigned to some form iron sulfide. However, one would expect these species to have been removed during the desalting step unless they are part of some polymer material or "bound" to the protein. The massive colouring of the column material may also to a minor extent be due to breakdown of the holo-protein during purification because of exposure to residual oxygen (further discussed in chapter 7).

The UV-vis spectra of the purified products (figure 6.3) clearly resemble spectra of [4Fe-4S] ferredoxins (chapter 2 figure 2.2) since a maximum or a very distinct shoulder is present at 390 nm. The absorption at 280 nm can be assigned to a tyrosine residue found in the ubiquitin part of the Ub-Ori-ISP protein (section 4.5). However, disulfide bonds also absorbs in this part of the UV-vis spectrum (Pace *et al.*, 1995).

The overall yield of the reaction varied from ca. 40 % to ca. 55 % independent on iron source. This might be explained by varying degree of oxygen exposure. The method used for keeping anaerobic conditions (Ar/N_2 flushing, section 4.1) is not 100 % efficient in keeping out oxygen and the concentration of oxygen in the reaction flask etc. probably varied from experiment to experiment.

6.3.1.1 Experiments with product obtained from purification by gravity

Short term stability (experiments 1 to 5)
The obtained product is very sensitive towards dioxygen and when exposed to air, the spectrum of a diluted sample lost all characteristic features within a period of three hours (experiment 1, section 6.1.2.1 and figure 6.4 A). The dilution and temperature can be ruled out as being major contributing factor, since samples diluted with milli-q water, 20 mM Tris/HCl pH 8.0 or elution buffer, kept at the same temperature or at 37ºC over the same time period under anaerobic conditions, remained stable (experiment 2-5, section 6.1.2.1 and figure 6.4 B). The breakdown progress is interesting (see figure 6.4 A) since an increase in absorption above 450 nm is seen before the shoulder at 390 nm disappears. No other cluster types could be distinguished from the UV-vis spectra during breakdown. One could envisage cluster breakdown taking place in a series of oxidation reactions e.g. the oxidation of the $[4Fe\text{-}4S]^{2+}$ cluster to a $[3Fe\text{-}4S]^{1+}$ cluster and its subsequent conversion to the $[2Fe\text{-}2S]^{2+}$ cluster. This may still happen, but undetectable by UV-vis spectrophotometry. The stability of the obtained product at least for a short period of time (at least three hours) in various solutions and in the temperature interval 0ºC to 37ºC is a necessity for analysing the composition and properties of the obtained product.

Effect of citrate exposure (experiment 6)
The sulfonated apo-protein was crystallized using a citrate salt (section 5.2.4) as precipitant and it was hoped that these conditions could be applied and further optimized for crystallization of the holo-protein. Therefore, the effect of adding citrate to holo-protein was investigated (experiment 6, section 6.1.2.1). As seen in figure 6.5 a major change of the UV-vis spectrum was observed (see section 6.2.2.1 for a description). The resultant spectrum resembles spectra of the $[3Fe\text{-}4S]^{1+}$ or $[2Fe\text{-}2S]^{2+}$ ferredoxins (see sections 2.2 and 2.3), showing absorption maxima at 330 nm, 420 nm and a shoulder at ca. 465 nm. The $[3Fe\text{-}4S]^{1+}$ ferredoxin from *P. furiosus* has an absorption maximum at 408 nm (figure 2.2 B) while the $[2Fe\text{-}2S]^{2+}$ ferredoxin from *A. vinelandii* has absorption maxima at 330 nm, 420 nm, 460 nm and 550 nm (Chatelet and Meyer, 1999). The observed absorption maxima fits closely to the $[2Fe\text{-}2S]^{2+}$-ferredoxin while the UV-vis spectrum above 400 nm resembles more closely the UV-vis spectrum of a $[3Fe\text{-}4S]^{1+}$ ferredoxin. Citrate may act as an iron scavenger and here it appears to facilitate cluster conversion from a $[4Fe\text{-}4S]^{2+}$ cluster to possibly a $[2Fe\text{-}2S]^{2+}$ or a $[3Fe\text{-}4S]^{1+}$ cluster or a mixture of clusters. Therefore, citrate salts should not be applied as precipitants for holo-protein during crystallization experiments since it induces cluster breakdown.

Stability upon dilution and effect of glycerol (experiments 7 to 14)
The stability of the product at 0ºC diluted with different solutions (milli-q water, 20 mM Tris/HCl pH 8.0, elution buffer or 20 mM Tris/HCl pH 8.0 containing 5 to 40 % glycerol) was investigated over a period of 48 hours (experiment 7 to 14, section 6.1.2.1 and figure 6.6). Glycerol was added since it is commonly used to keep proteins and enzymes stable in solution or during freezing. In all cases, a drop in absorbance was seen. When 20 mM Tris/HCl, 20 % glycerol pH 8.0 was used to dilute the sample, the least drop in absorbance was observed

(figure 6.6 F) and the maximum at 390 nm had not been reduced to a shoulder or completely disappeared as in the other cases. Based on the results of experiments 7 to 14 it can be concluded that adding 20 % glycerol to the holo-protein has a stabilizing effect. It was therefore decided to add 20 % glycerol to the purification buffer in some of the HPLC purification procedures (see section 6.1.3). It is interesting to note that the sample diluted with milli-q water appears to be more stable than the sample diluted with 20 mM Tris/HCl pH 8.0. One would expect the protein to be more stable in a buffered solution unless the buffer itself takes part in protein breakdown. However, this rapid breakdown is not seen in the remaining UV-vis spectra (figure 6.6, except when 40 % glycerol was added). The only obvious explanation is that this sample was somehow exposed to a higher degree of oxygen.

Stability upon freezing and thawing (experiment 15 and 16)
The stability of the holo-protein upon freezing was investigated (experiment 15 and 16, section 6.1.2.1 and figure 6.7). The results indicate massive cluster degradation. It can therefore be concluded that even though glycerol appear to have a stabilizing effect on the protein in solution (see above) it does not have a similar effect upon freezing the protein hence freezing of the holo-protein should be avoided. The UV-vis spectra recorded after freezing of the samples in the two experiments do have some differences (each spectrum is described in section 6.2.2.1). When no glycerol was added (experiment 15) the absorbance below 550 nm decreased and when glycerol was added the absorbance above 450 nm increased. However, the spectra are similar and the difference in absorbance after thawing might be due to different absorption properties of the breakdown products.

Effects of temperature and dilution (experiments 17 to 19)
The results obtained from experiment 17 to 19 (section 6.1.2.1 and figure 6.8) clearly indicate that an undiluted sample left at 0ºC is more stable than a diluted sample left at 0ºC or an undiluted sample left at room temperature. It is expected that degradation will take place at a higher rate when the temperature is high. The diluted solution and the solution left at room temperature were transferred to new vessels prior to starting the experiments hereby possibly introducing small amounts of oxygen in the process, while the undiluted solution left at 0ºC was left in the original vessel. This might have affected the outcome of the experiments, which could explain why the diluted sample degraded much faster than the undiluted (both left at 0ºC).

Effects of reducing agents (experiments 20 to 22)
Addition of dithionite to a diluted or undiluted sample (experiment 20 and 21, section 6.1.1.2 and figure 6.9 A and B) immediately reduced the holo-protein as seen from the recorded UV-vis spectra since the feature at 390 nm disappeared. After 24 hours the feature at 390 nm returned indicating the holo-protein had been re-oxidized. The feature was however present as a shoulder indicating slight degradation. Degradation within 24 hours was also observed in the previous experiments. Right after addition of DTT to a diluted sample (experiment 22, section 6.1.1.2 and figure 6.9 C), the feature found at 390 nm was still present, which indicates that the holo-protein had not been reduced. DTT may not be powerful enough for reduction of holo-Ub-Ori-ISP or the reaction is very slow. After 24 hours however, all

features associated with the presence of holo-protein had disappeared and they did not return. DTT is able to form complexes with iron and one could therefore envisage DTT acting as an iron scavenger. When citrate was added to holo-protein, the UV-vis spectrum changed and resembled more those of [3Fe-4S] or [2Fe-2S] ferredoxins. However, here total degradation was observed.

6.3.2 HPLC purification on a 10/5 Source 30Q column

As observed previously (section 6.3.1) a black/grey band assigned to some form of iron sulfide remained on the column after purification.

6.3.2.1 The obtained products

The chromatograms obtained upon purification of holo-protein on the Source 30Q column (section 6.1.3 and figure 6.10) show great differences (section 6.2.3).

$(NH_4)_2Fe(II)(SO_4)_2{\cdot}6H_2O$, $FeCl_3$ and $NH_4Fe(III)(SO_4)_2{\cdot}12H_2O$ as iron sources

In the absence of glycerol in purification buffers the best resolved main peak was found using $FeCl_3$ as the iron source. When using $NH_4Fe(III)(SO_4)_2{\cdot}12H_2O$ or $(NH_4)_2Fe(II)(SO_4)_2{\cdot}6H_2O$ the main peak was very poorly resolved (figure 6.10 A to C). It appears that the differences are independent on the oxidation state of iron but dependent on counter-ions since poorly resolved peaks were obtained using iron compounds containing sulphate and ammonium. Sulphate was however present when using iron powder as the iron source (see below), hence ruled out as the cause. The difference in elution patterns could arise from different species present. However, upon comparing the UV-vis spectra of the main fractions (figure 6.11 A to C) they appear very similar, only a slight difference in the exact position of the absorption maxima is evident. These spectra are comparable to the spectrum obtained when treating (what appears to be) a $[4Fe\text{-}4S]^{2+}$-cluster containing protein with citrate (experiment 6 section 6.1.2.1 and figure 6.5). It appears that the main fraction contains holo-protein containing oxidized forms of [2Fe-2S] or [3Fe-4S] or a mixture (argument found in section 6.3.1.1). However, these two cluster types would not be expected to elute simultaneously due to their overall charge difference of 1.

When glycerol was present in purification buffers and $FeCl_3$ was used as the iron source, the obtained chromatogram showed a less resolved main peak compared to when glycerol was absent (comparison of figure 6.10 A and D). The opposite was the case when $(NH_4)_2Fe(II)(SO_4)_2{\cdot}6H_2O$ was used as the iron source since glycerol in purification buffers provided a better resolved main peak (comparison of figure 6.10 C and E). The effect of glycerol is therefore inconclusive. UV-vis spectra of the main fractions are comparable in the presence and absence of glycerol (compare panels in figure 6.11 A to E). Components of the main fractions are discussed above.

Fe-powder as iron source

It was decided to use glycerol in purification buffers when using iron powder as the iron source since the dissolved iron contained about 80 % Fe^{2+} (section 6.1.1.1), and glycerol appeared to improve the chromatogram when Fe^{2+} was used as the iron source (see above). When iron powder was used as iron source the best results were observed (figure 6.10 F and

G) since the resolution of the main peak had markedly improved compared to the previous experiments (compare panels F and G to panels A to E in figure 6.10). The UV-vis spectra of the main fractions shown in figure 6.11 F and G resemble those of $[4Fe\text{-}4S]^{2+}$ ferredoxins. When ^{57}Fe was used, the closest resemblance was observed (absorption maximum at 390 nm). Iron powder was dissolved using sulphuric acid and consequently sulphate was present during cluster reconstitution. However, no ammonium ions were present (which was the case for the other Fe(II) source). When iron powder was used, twice the amount of protein was purified, and it may be a combination of the iron source and the amount of protein loaded onto the column or either of the two, which are responsible for the improved resolution and the difference in the obtained UV-vis spectra. This is further discussed below.

Cluster types and degradation

Since the product obtained when purifying the reconstituted protein on the Source 30Q column appeared to contain a $[4Fe\text{-}4S]^{2+}$ cluster when using iron powder as iron source (see above), it can be envisaged that the majority of the observed products from HPLC purification were in fact breakdown products of a [4Fe-4S] specie. $[2Fe\text{-}2S]^{2+}$ and $[3Fe\text{-}4S]^{1+}$ contain only Fe^{3+} contrary to $[4Fe\text{-}4S]^{2+}$ and have been observed to be oxidative breakdown products of [4Fe-4S] clusters (Crack *et al.*, 2004; Khoroshilova *et al.*, 1997; Leal *et al.*, 2004; Liu *et al.*, 2000; Tilley *et al.*, 2001). The time the holo-protein spent on the column was longer during HPLC purification (compared to purification by gravity) and it may therefore have been exposed to higher amounts of oxygen which leads to oxidative breakdown. The products obtained from purification on the Q Sepharose Fast Flow column did in fact appear to be a $[4Fe\text{-}4S]^{2+}$ protein independent on the iron source (section 6.3.1). Buffer exiting the HPLC when anaerobic conditions were applied was shown to contain about 0.5 mg/L dissolved oxygen (results not shown). It can be anticipated that the concentration of oxygen might vary depending on the conditions of the system. If this is the case, the difference in oxygen concentration might be responsible for the differences observed in the chromatograms and UV-vis spectra (figure 6.10 and 6.11) and not the iron source or glycerol. When loading e.g. a double amount of protein onto the column, it is likely that a smaller percentage of oxidative breakdown will occur and it is therefore more likely to get a better resolution in the chromatogram and a UV-vis spectrum of a [4Fe-4S]-cluster containing protein.

The results obtained when using glycerol in purification buffers indicate that holo-protein containing a $[4Fe\text{-}4S]^{2+}$ cluster eluted at a slightly higher conductivity than what appeared to be a $[2Fe\text{-}2S]^{2+}$ or a $[3Fe\text{-}4S]^{1+}$ protein (figures 6.10 and 6.11). Since the $[4Fe\text{-}4S]^{2+}$ cluster and $[2Fe\text{-}2S]^{2+}$ cluster have the same net charge holo-protein containing these clusters are expected to elute at approximately the same conductivity, while protein containing a $[3Fe\text{-}4S]^{1+}$ cluster is expected to elute at a higher conductivity (anion-exchange) due to the less positive charge ($|\Delta| = 1$). The fact that only a slight variation in the elution point of the main fraction was observed favours the $[2Fe\text{-}2S]^{2+}$ cluster as the product of degradation. Another possibility could be formation of monomers, dimers etc. since the elution would be expected to be dependent on the degree of polymerization.

As discussed above, differences in the position of the elution point (measured conductivity) of the main peak is evident which may indicate the presence of different species. However in

most cases the shape of the main peak is symmetrical, which indicates that only one species is present in the fraction corresponding to this particular peak, or that this peak contains several species with the same surface charge distribution under the given conditions. The tailing in front of or after the main peak and poorly resolved peaks indicate either breakdown of holo-protein during purification, presence of multiple species or slightly differing conformations of species. Since cluster reconstitution was performed using purified protein, all peaks can be assigned to different forms of the protein. Different conformations are imaginable since the cluster formation was performed under unnatural conditions compared to the "cluster formation machinery" found in living organisms (Frazzon and Dean, 2003; Lill and Mühlenhoff, 2005).
However, most likely the major differences seen in the chromatograms and the poorly resolved peaks are due to breakdown caused by oxygen.
The yields of the reactions were not calculated since in most of the cases it was difficult to assign the obtained products to a certain cluster species and different species have different molar absorptivities (see chapter 2).

6.3.2.2 Re-purification of the obtained product

When comparing the chromatograms shown in figure 6.10 E and 6.12 it is clear that re-purification of holo-protein found in the main peak (figure 6.10 E) was not possible under the given conditions (section 6.1.3). When attempting to re-purify the protein, very weak absorbance was observed from the chromatogram (figure 6.12) and tailing and several peaks were observed indicating breakdown of the holo-protein. UV-vis spectra of fractions corresponding to the peaks revealed absence of holo-protein. Degradation is most likely caused by oxygen as discussed in section 6.3.2.1. Re-purification was only performed this one time and it was performed on a poorly resolved main peak. However, it is still believed that breakdown would also have been observed had re-purification been carried out on well resolved main peaks. Therefore it appears that re-purification with an acceptable yield is not possible under the given conditions.

6.3.2.3 Experiments with product obtained from purification by HPLC

When adding cysteine and Fe^{2+} to the obtained product (experiment 23, section 6.1.3.1 and figure 6.13) the UV-vis spectrum was seen to undergo a striking change (described in section 6.2.3.1) from what looks like a $[2Fe\text{-}2S]^{2+}/[3Fe\text{-}4S]^{1+}$ protein to a $[4Fe\text{-}4S]^{2+}$ protein with an absorption maximum at 385 nm. From the control experiment, no indication of an absorption maximum in this region was evident, indicating the change had happened to the cluster of the protein. It is however very likely that this newly formed cluster is coordinated by the "external" cysteine and not cysteine from the protein. The cluster re-arrangement could also be achieved using cysteine alone but when DTT and iron was used no change was observed (results not shown), supporting the above stated assumption. It was attempted to desalt the protein solution (results not shown), but the UV-vis spectrum of the product was comparable to the spectrum obtained prior to addition of Fe^{2+} and cysteine indicating breakdown of the obtained product. Cluster re-arrangement is not an uncommon phenomenon. In biotin synthase from *E. coli* two $[2Fe\text{-}2S]^{2+}$ clusters can be transformed into a $[4Fe\text{-}4S]^{2+}$ cluster by reduction (Duin *et al.*, 1997).

When freezing the obtained product (experiment 25, section 6.1.3.1 and figure 6.14) degradation of the holo-protein was observed as previously seen (section 6.3.1.1).

6.3.3 EPR and Mössbauer analyses on holo-protein

The holo-protein obtained from HPLC purification for EPR and Mössbauer spectroscopic analyses (sections 6.1.4 and 6.2.4) very closely resembled $[4Fe\text{-}4S]^{2+}$ ferredoxins since absorption at 390 nm was observed from the UV-vis spectrum (figure 6.16 A).
Three samples were prepared for measurements (see section 6.1.4 and table 6.1). The samples to which no dithionite and 25-fold molar excess of dithionite was added prior to up-concentration (sample 1 and 2, respectively) did appear to be partly broken down as evident from the UV-vis spectra (figure 6.16 B) since only a shoulder and no absorption maximum was present at 390 nm. The sample to which 100-fold molar excess of dithionite was added prior to concentration looked like the spectrum recorded prior to addition of dithionite and up-concentration. It is very likely that that the holo-protein found in this sample contained a $[4Fe\text{-}4S]^{2+}$ cluster in the active site. It appears that adding 25-molar excess of dithionite did not reduce the protein or keep it from breakdown during concentration. This is probably due to oxygen exposure. When 100-fold molar excess of dithionite was added, the protein was kept from breakdown. However, the amount of dithionite was apparently not enough to reduce the protein (or keep it reduced). Extra dithionite should probably have been added after up-concentration of the protein to keep it reduced since it has been shown that the protein can be reduced by dithionite (section 6.3.1.1).
The EPR results of the three samples showed only very weak signals which could not be assigned to any compounds, only a very small amount of adventitious ferric iron could be assigned in all spectra. This however indicates that no paramagnetic species were present in any of the samples hence excluding the presence of $[3Fe\text{-}4S]^{1+}$, $[4Fe\text{-}4S]^{1+}$, $[4Fe\text{-}4S]^{3+}$ and $[2Fe\text{-}2S]^{1+}$ (chapter 2).
The Mössbauer results of samples 1 and 2 could not be unambiguously interpreted since they appeared to contain multiple cluster types which could not be simulated to fit the spectra. The Mössbauer results of sample 3 on the other hand could be fitted to contain 4.7 % of the iron as adventitiously bound ferrous iron and 95.3 % as present in $[4Fe\text{-}4S]^{2+}$ clusters.
EPR and Mössbauer results were interpreted by Hanne Nørgaard and thoroughly discussed elsewhere (Nørgaard, 2006).
The results very strongly suggest that holo-protein present in sample 3 contained a $[4Fe\text{-}4S]^{2+}$ cluster in the active site (see figure 6.16 B for the corresponding UV-vis spectrum with an absorption maximum at 390 nm and a shoulder at about 310 nm). The results indicate that holo-protein obtained when purifying the protein on the Q Sepharose Fast Flow column most likely contained a $[4Fe\text{-}4S]^{2+}$ cluster in the active site (figure 6.3). The products obtained when using iron powder as iron source and performing HPLC purification were also dominated by a $[4Fe\text{-}4S]^{2+}$ cluster in the active site (figure 6.10 F and G).

6.3.4 Fusion protein

The protein was purified as a fusion with ubiquitin since it provides a means of identifying the protein throughout purification (chapter 5). It is however possible that the presence of ubiquitin interferes with the formation of holo-protein and affects the properties of the holo-

protein. Therefore, the fusion should be removed when investigating the intrinsic properties of holo-Ori-ISP.

6.4 Conclusion

From UV-vis spectrophotometric analyses, reconstituted protein purified on a Q Sepharose Fast Flow column appear to contain a $[4Fe\text{-}4S]^{2+}$ cluster in the active site irrespective of the iron source used. It was not possible to purify the reconstituted protein by HPLC unless the ion-exchange by gravity was omitted. Depending on the conditions of cluster reconstitution and purification, different cluster types ($[2Fe\text{-}2S]^{2+}$ / $[3Fe\text{-}4S]^{1+}$ or $[4Fe\text{-}4S]^{2+}$) appeared to be present in the active site of the purified holo-protein.

The holo-protein is very sensitive towards oxygen and it is only stable for a short while (less than 24 hours) under anaerobic conditions and 0ºC before breakdown was apparent. To preserve the protein it should be left undisturbed after purification (undiluted) and at 0ºC under anaerobic conditions. Freezing and thawing of the holo-protein resulted in degradation. Re-purification of the obtained product was not possible due to degradation. Glycerol appeared to have a stabilizing effect on the holo-protein, but the results were not conclusive.

Reduction of holo-protein was achieved by dithionite, however, DTT appeared to break down the holo-protein. Cluster conversions from a [4Fe-4S] cluster to a [2Fe-2S] / [3Fe-4S] cluster and vice versa appeared to take place by addition of citrate and cysteine, respectively.

Mössbauer spectroscopy identified a $[4Fe\text{-}4S]^{2+}$ cluster to be the absolute dominating species in one sample. Since this sample had a UV-vis spectrum with an absorption maximum at 390 nm and a shoulder at about 310 nm, it was concluded that other samples having comparable UV-vis spectra contained a $[4Fe\text{-}4S]^{2+}$ cluster as the dominant cluster type in the active site.

Since oxygen appears to be a parameter which is difficult to exclude and control using Ar/N_2 flushing for obtaining anaerobic conditions, the results obtained in this chapter are to some degree flawed. Residual oxygen has to be controlled since the holo-protein is extremely sensitive towards it, hence continual work on this protein should be carried out under strictly controlled conditions in an anaerobic chamber.

Chapter seven

7 Reconstitution and purification of Ub-Ori-ISP and Ori-ISP in anaerobic chambers

The results obtained in chapter 6 strongly indicate that further work on holo-Ub-Ori-ISP should be carried out in a strictly controlled environment depleted of oxygen.

This chapter describes experimental work and results obtained when performing cluster reconstitution and purification in anaerobic chambers. The experiments described were carried out during a 10-week stay at IBS (Grenoble, France) and in our own laboratories. The anaerobic chamber used for the experiments in France contained 6-8 ppm oxygen, while the concentration in the chamber used in our own laboratories was not determined. Further information on the anaerobic chambers is found in chapter 4. All experiments were carried out at room temperature, and samples were left at room temperature unless otherwise stated.

Three experiments were performed using $FeCl_3$ as the iron source for cluster reconstitution into Ub-Ori-ISP. Two identical experiments were performed in France to test reproducibility. A third experiment was performed in Denmark to test reproducibility in a different anaerobic chamber with an unknown oxygen concentration prior to performing cluster reconstitution into Ori-ISP in experiment 4. One experiment (experiment 5) was performed using $(NH_4)_2Fe(II)(SO_4)_2 \cdot 6H_2O$ as the iron source for cluster reconstitution into Ub-Ori-ISP. This experiment was performed in France.

In each of the five experiments described in this chapter, parallel anion-exchange purifications were carried out on a 15/2 Q Sepharose Fast Flow column and a 10/5 Source 30Q column, respectively. This was done to test differences in the obtained products under the same conditions.

7.1 Experimental

HPLC purifications were carried out on an ÄKTAFPLC™ system in France and on an ÄKTAprime™ plus system in Denmark. Both systems were capable of measuring the absorbance at 280 nm. No glycerol was added to any of the purification buffers due to the inconclusive results in chapter 6. All experimental work was carried out inside the anaerobic chamber except UV-vis spectrophotometric recordings on the diode array spectrophotometer.

7.1.1 Cluster reconstitution and purification using $FeCl_3$ as the iron source

7.1.1.1 Cluster reconstitution into Ub-Ori-ISP and purification

Experiment 1 and 2 (performed in France)
Experiment 1 and 2 were identical and in each experiment cluster reconstitution was performed as described in section 6.1.1. However, 34.0 mg protein was used and dissolved in reaction buffer to a concentration of 7.4 mg/ml. DTT was added in 20-fold molar excess (0.3 M stock solution) and iron and sulfide were both added in 8-fold molar excess (0.12 M stock solutions). After cluster reconstitution the protein solution was divided into two portions and each was desalted into 20 mM Tris/HCl, 0.15 M NaCl pH 8.0 as described in section 6.1.1. Holo-protein from both solutions was further purified by anion-exchange. One portion was purified on a 15/2 Q Sepharose Fast Flow column (by gravity) as described in section 6.1.2. The other solution was purified on a 10/5 Source 30Q column (by HPLC) as described in section 6.1.3, but with the gradient extended to 0.6 M NaCl (28.6 CV). Fractions collected from each of the two anion-exchange purifications were analysed by UV-vis spectrophotometry and by analytical gel filtration on a 20 mM Tris/HCl, 0.15 M NaCl pH 8.0 equilibrated 10/30 Superdex 75 HR column (by HPLC). 50 µl samples were loaded onto the column and eluted using the same buffer (linear flow rate: 0.6 cm/min). Fractions were collected and UV-vis spectra recorded. Figure 7.1 presents a flow chart of the cluster reconstitution and purification.

Experiment 3 (performed in Denmark)
Experiment 1 (see above) was repeated. However, 27.2 mg protein dissolved in reaction buffer to a concentration of 6.0 mg/ml was used. During the purification of the obtained protein on the Q Sepharose Fast Flow column (section 6.1.2) a second elution step was introduced, namely elution with 20 mM Tris/HCl, 1.0 M NaCl. Analytical gel filtration was performed on a 13/31 Superdex 75 HR column with 500 µl aliquots of sample (linear flow rate: 0.3 cm/min). Some fractions were re-analysed on the analytical gel filtration column.

7.1.1.2 Cluster reconstitution into Ori-ISP and purification

One experiment was carried out performing cluster reconstitution into Ori-ISP (experiment 4) using the optimized conditions for Ub-Ori-ISP (section 6.1.1). However, the reaction time for de-protection was raised from 1.5 to 22 hours (section 9.4.2). The rationale was to de-protect the cysteines and perform simultaneous proteolytic cleavage (by the enzyme YUH1) of the Ub-fusion. Since the cleavage was to take place in the reaction mixture the amount of YUH1 necessary for cleavage under the given conditions was optimized (section 9.3.3). After cleavage and de-protection of the cysteines cluster reconstitution was to be carried out leaving YUH1 and ubiquitin in the reaction mixture. During anion-exchange YUH1 and ubiquitin could be separated from holo-Ori-ISP due to differences in total charge and surface charge distributions.
27.2 mg freeze-dried Ub-Ori-ISP was dissolved in 4.3 ml 100 mM Tris/HCl pH 8.5. 200 µl 0.25 M DTT solution (20-fold molar excess) was added to the protein solution followed by

189 µl of YUH1 solution (17.7 mg/ml, the molar ratio Ub-Ori-ISP:YUH1 was 20:1). The solution was incubated for 22 hours with gentle stirring. Then, 200 µl freshly prepared 0.10 M $FeCl_3$ solution (8-fold molar excess) was added to the protein solution followed by 200 µl freshly prepared 0.10 M $Na_2S{\cdot}9H_2O$ solution (8-fold molar excess). The resultant solution was incubated for one hour with gentle stirring (see section 6.1.1 for further details on cluster reconstitution). The solution was split into two portions and each was further purified as described in section 7.1.1.1. However, desalting was carried out on two 15/8 Sephadex G-10 columns (by gravity). When purifying the first portion on the Q Sepharose Fast Flow column (section 7.1.1.1 and 6.1.2) a second elution step was included, namely elution with 20 mM Tris/HCl, 1.5 M NaCl. When purifying the second portion on the 10/5 Source 30Q column (section 7.1.1.1 and 6.1.3), the gradient was extended to 1.0 M NaCl (43 CV). Fractions collected from each of the two anion-exchange purifications were further separately analysed by analytical gel-filtration on a 20 mM Tris/HCl, 0.15 M NaCl pH 8.0 equilibrated 13/31 Superdex 75 HR column (by HPLC). 500 µl samples were loaded onto the column and eluted using the same buffer (linear flow rate: 0.3 cm/min). This analysis was carried out right after anion-exchange and again after a period of 24 hours. Fractions were collected and UV-vis spectra recorded. An overview of the cluster reconstitution and purification is presented in figure 7.2.

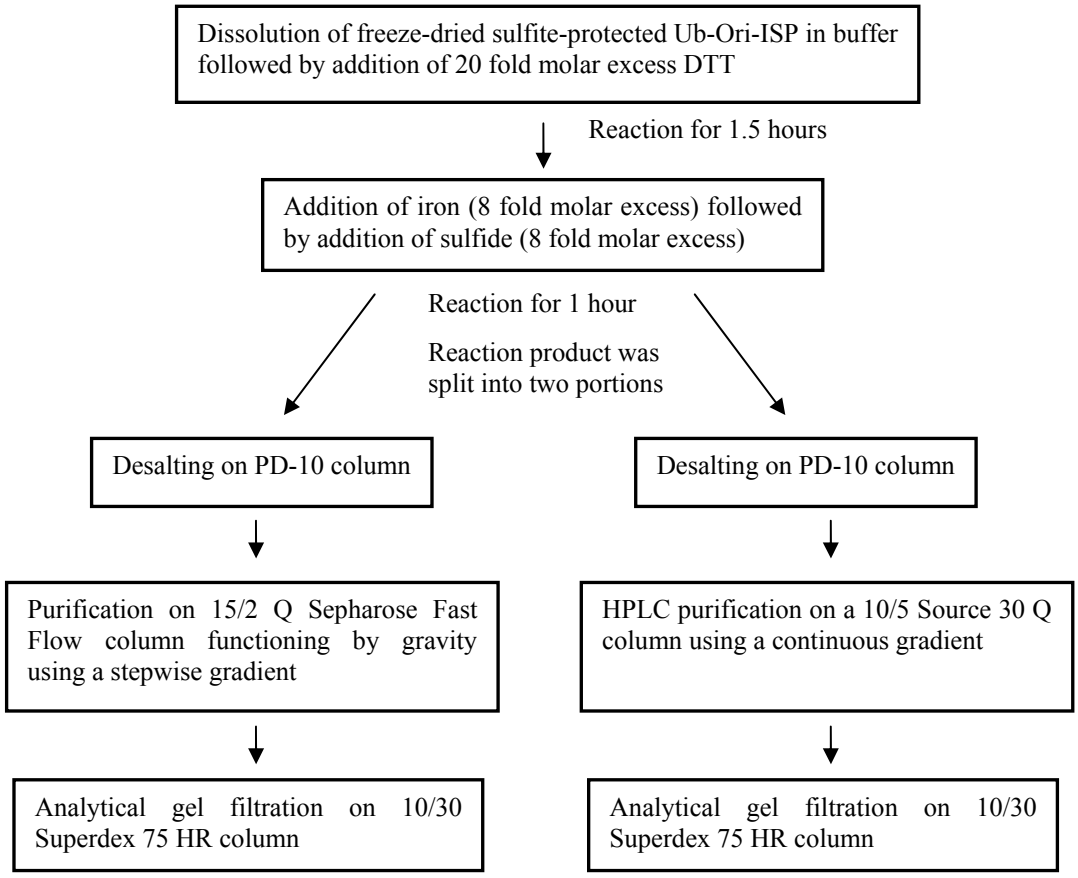


Figure 7.1 Flowchart of cluster reconstitution reaction into Ub-Ori-ISP followed by splitting of the solution into two portions and subsequent parallel desalting and purification of reaction products.

7.1.2 Cluster reconstitution into Ub-Ori-ISP and purification using Fe(II) as the iron source

Experiment 5 was performed in France and cluster reconstitution and purification was carried out as described in section 7.1.1.1 (experiment 1 and 2). However, $(NH_4)_2Fe(II)(SO_4)_2 \cdot 6H_2O$ was used as the iron source.

7.1.3 Stability of holo-protein purified by gravity

Holo-protein obtained from purification on the Q Sepharose Fast Flow column (section 7.1.1.1, experiment 1 and 2) was up-concentrated to approximately 70 mg/ml using a Microcon YM-3 centrifugal filter unit. The holo-protein was stored at 4ºC and the stability was measured by recording a UV-vis spectrum of the obtained product over a period of 49 days.

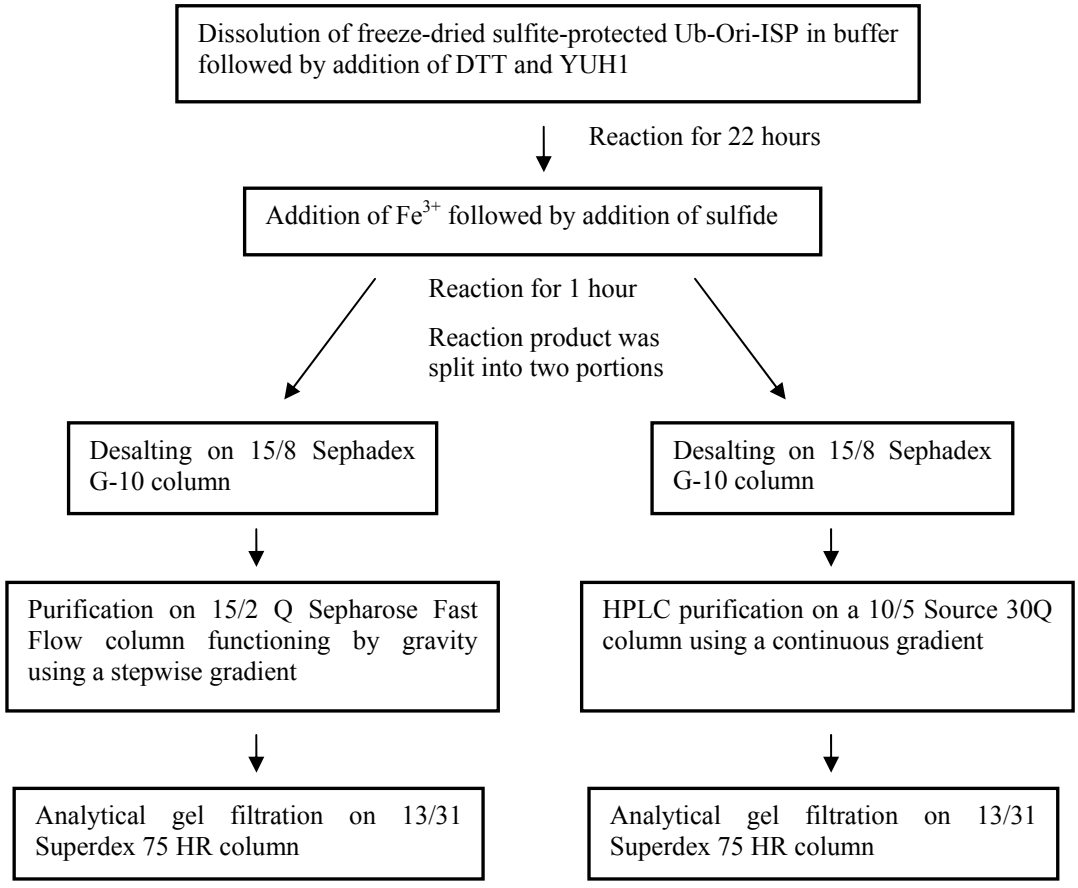


Figure 7.2 Flowchart of proteolytic cleavage of Ub-Ori-ISP, cluster reconstitution reaction into Ori-ISP followed by splitting of the solution into two portions and subsequent parallel desalting and purification of reaction products.

7.2 Results and discussion

7.2.1 General comments on cluster reconstitution and desalting

Observations made during cluster reconstitution and desalting in this chapter are comparable to those made previously (section 6.2.1). This indicates that oxygen was not the determining factor when different coloured by-products and differing amounts of by-products were observed. Other parameters such as the batch from which apo-protein originated, the speed of which iron and sulfide were added, the iron source, the temperature in the reaction mixture and the reaction time may be important parameters. Even though experiments were carried out in the exact same manner, there are still minor differences since they are not automated.

When using $(NH_4)_2Fe(II)(SO_4)_2 \cdot 6H_2O$ as the iron source inside the anaerobic chamber (section 7.1.2), the colour of the reaction mixture was not nearly as intense as using the same iron source outside the anaerobic chamber or Fe^{3+} in general (section 6.2.1). When adding Fe^{2+} to a reducing solution in a non-oxidizing environment it is difficult to imagine oxidized FeS-cluster forms being synthesized (further discussed in section 7.2.5). Reduced forms are more faintly coloured than the oxidized forms (chapter 2) and this is probably the reason for the faintly coloured solution.

7.2.2 General comments on anion-exchange and analytical gel filtration

When performing anion-exchange purifications the observations made were comparable to those observed previously (section 6.2.2 and 6.2.3). However, when eluting the holo-protein from the Q Sepharose Fast Flow column larger broadening of the band was observed. When eluting the protein from the Source 30Q column, two bands were observed. In the following description (sections 7.2.3 and 7.2.4) all fractions had a brown colour unless otherwise stated.

As observed in chapter 6, anion-exchange columns had a black band at the top and a greyish colour after elution of holo-protein which could be assigned to some form of iron sulfide (see section 6.3.1 for further discussion). This is expected since iron and sulfide were both added in molar excesses. In all five experiments (section 7.1.1 and 7.1.2) fractions collected from anion-exchange were analysed by analytical gel filtration. Two different analytical gel filtration columns were used.

When using the 10/30 Superdex 75 HR column (experiments 1, 2 and 5) main peaks were found to elute at two positions, namely ca. 10.9 ml and ca. 12.4 ml. When using the 13/31 Superdex 75 HR column in experiment 3, main peaks were also found at two positions, ca. 11.7 ml and 12.8 ml (comparable to those in experiments 1, 2 and 5). In experiment 4 the 13/31 Superdex 75 HR column was also used and main peaks at two positions were also found. However, these were found at higher volumes since the ubiquitin fusion was cleaved off. In each of the five experiments, the protein present in the main peaks eluting at the lowest volume are considered dimeric, while the protein present in the main peaks eluting at the highest volume are considered monomeric. These assumptions are substantiated in section 7.2.7.

7.2.3 Purification of holo-protein when $FeCl_3$ was used as the iron source

7.2.3.1 Purification of holo-Ub-Ori-ISP

Experiment 1

The chromatogram obtained from the HPLC purification on the Source 30Q column is presented in figure 7.3. A minor peak eluted with a maximum absorbance at 24.7 mS/cm (Source fraction 1) and a major peak eluted from 37 to 46 mS/cm (Source fraction 2). Small and poorly resolved peaks were also observed between these two peaks.

Source fraction 1 eluted at a somewhat higher conductivity than the main peak observed previously (figure 6.10 A to C) when performing purification outside the anaerobic chamber (chapter 6). The UV-vis spectrum of Source fraction 1 (figure 7.4 A black spectrum) reveals a $[4Fe\text{-}4S]^{2+}$ cluster to be present in the protein (absorption maximum at 390 nm, see section 6.3.3). Therefore, this peak corresponds to the previously observed main peaks (see sections 6.2.3 and 6.3.2). Source fraction 2, which was the major peak in the chromatogram, has not previously been observed. The UV-vis spectrum of this fraction (figure 7.4 A, red spectrum) resembles those of $[3Fe\text{-}4S]^{1+}$ ferredoxins (absorption maximum at 408 nm, figure 2.2 B). This is in agreement with electrostatic considerations since a $[3Fe\text{-}4S]^{1+}$ cluster containing protein would be expected to elute at a higher conductivity than a corresponding $[4Fe\text{-}4S]^{2+}$ cluster containing protein.

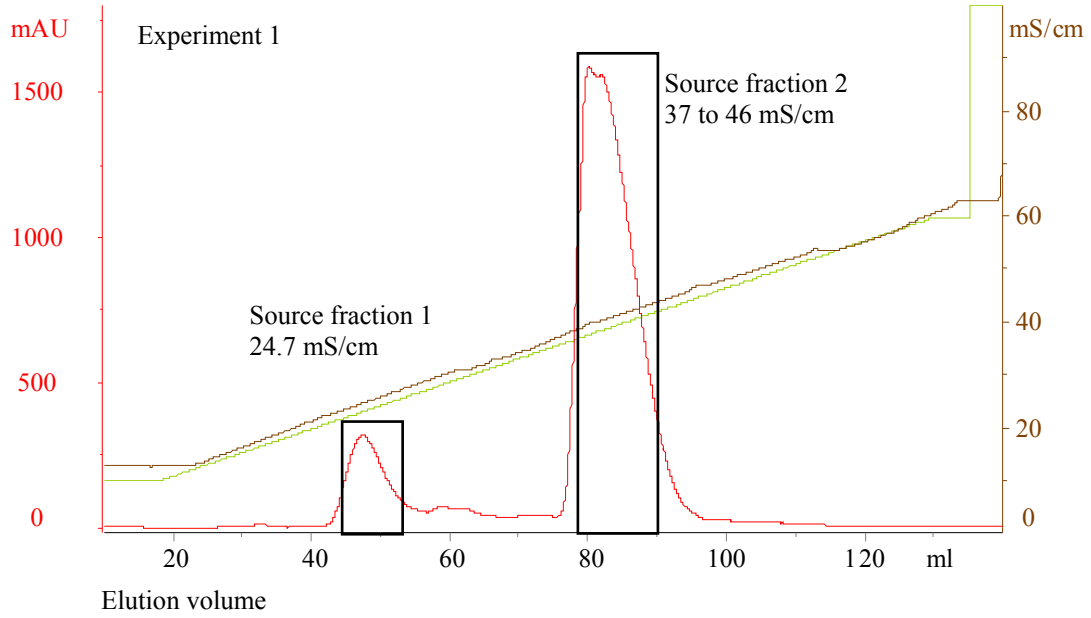


Figure 7.3 Chromatogram for purification of holo-Ub-Ori-ISP on the Source 30Q column in experiment 1, when using $FeCl_3$ as the iron source. The red line indicates the absorbance at 280 nm. The brown and the green lines indicate the conductivity and the NaCl gradient, respectively. The concentration of NaCl can be read from the conductivity axis employing the unit 10^{-2} M. Collected fractions are indicated with black boxes and the conductivity at maximum absorption is also presented.

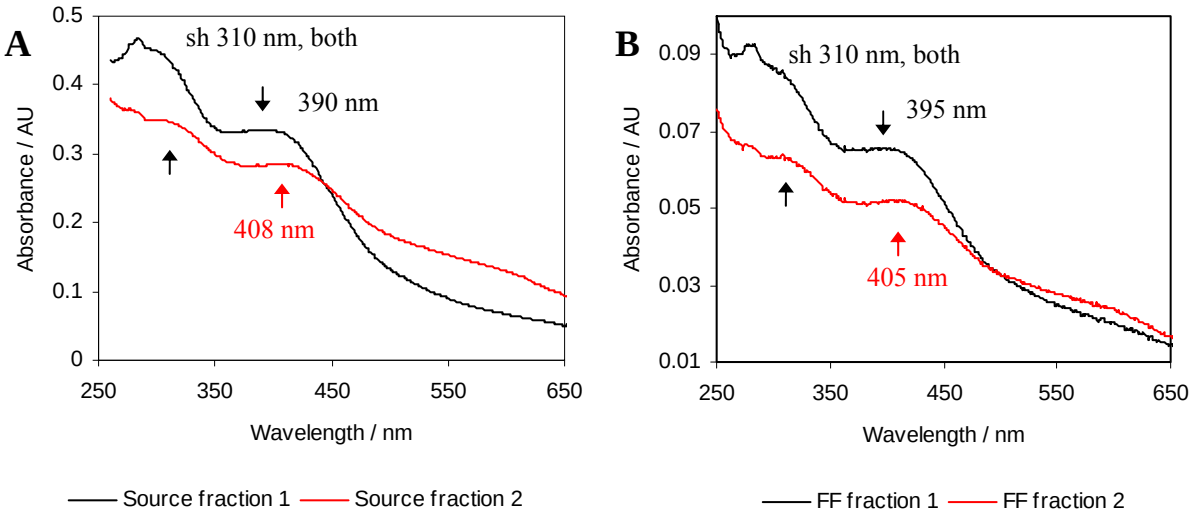


Figure 7.4 UV-vis spectra of main fractions from anion-exchange in experiment 1. A) Spectra obtained from fractions from purification on Source 30Q column. Source fraction 2 was diluted 5 times (with 20 mM Tris/HCl pH 8.0) prior to recording the spectrum. B) Spectra obtained from fractions from purification on Q Sepharose Fast Flow column. Both fractions were diluted 50 times (with 20 mM Tris/HCl pH 8.0) prior to recording the spectra. Maxima and shoulders (sh) are indicated with arrows.

An experiment where the holo-protein was purified on a Q Sepharose Fast Flow column (by gravity) and eluted in one step was performed. Two main fractions (FF fractions 1 and 2) were collected and from the UV-vis spectra shown in figure 7.4 B, FF fraction 1 is seen to be comparable to Source fraction 1 while FF fraction 2 is comparable to Source fraction 2.
Prior to analytical gel filtration, Source fractions 1 and 2 were concentrated ca. 20 and 5 times, respectively, and FF fractions 1 and 2 were both concentrated ca. 5 times (concentration was performed using a stirred Amicon cell equipped with a YM3 membrane). The four chromatograms obtained from the analytical gel filtrations on each of the four collected fractions are presented in figure 7.5. From the purifications of Source fractions 1 and 2 and FF fraction 2, one main peak along with one or two minor peak were observed while two similarly sized peaks were observed from purification of FF fraction 1. Minor peaks are difficult to observe in the figure due to the scale on the y-axis. The elution volume at maximum absorption of each of the peaks is presented in table 7.1 and UV-vis spectra corresponding to the main peaks are presented in figure 7.6 (absorption maxima and shoulders are indicated in the figure).
Source fraction 1, gave one main peak (12.5 ml) and a minor peak (10.9 ml) in the chromatogram (figure 7.5, red line). This strongly indicates that Source fraction 1 contains a monomer. The UV-vis spectrum of the main peak (figure 7.6 A, black spectrum) is slightly different from the spectrum recorded prior to the analysis (figure 7.4 A, black spectrum) since maxima at 308 nm and 403 nm are now observed. This spectrum indicates the presence of a $[3Fe\text{-}4S]^{1+}$ cluster or a mixture of clusters, which could have been formed due to oxidative breakdown of the $[4Fe\text{-}4S]^{2+}$ cluster either prior to or during the analysis (discussed in section 7.2.8). The chromatogram for Source fraction 2 (figure 7.5, green line), shows main peak at 10.9 ml indicating the presence of a dimer. However, a minor peak corresponding to a monomer (12.4 ml) is also present. The UV-vis spectrum of the main peak (figure 7.6 A, red

spectrum) resembles the spectrum recorded prior to performing the analysis (figure 7.4 A, red spectrum) however, no absorption maxima are evident.
In summary Source fraction 1 appears to be a monomer containing a $[4Fe\text{-}4S]^{2+}$ cluster while Source fraction 2 appears to be a dimer containing one or two $[3Fe\text{-}4S]^{1+}$ clusters or a combination of two cluster types.

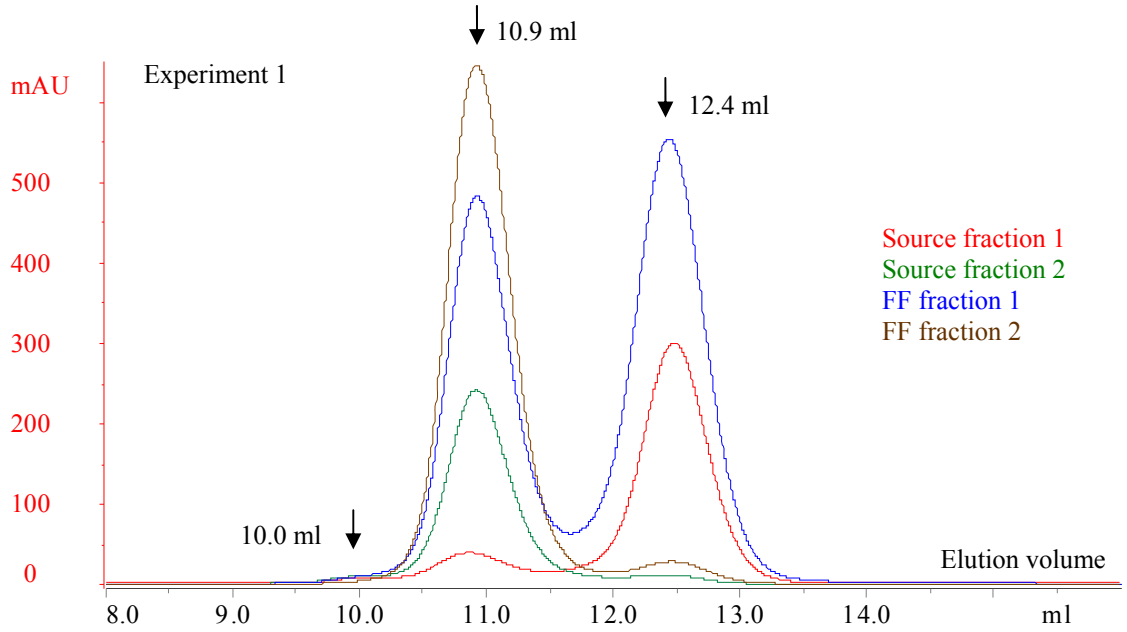


Figure 7.5 Chromatograms for analytical gel filtrations in experiment 1 obtained when analysing fractions collected from anion-exchange. Absorption was measured at 280 nm. Red, green, blue and brown lines indicate overlaid chromatograms for Source fraction 1, Source fraction 2, FF fraction 1 and FF fraction 2, respectively.

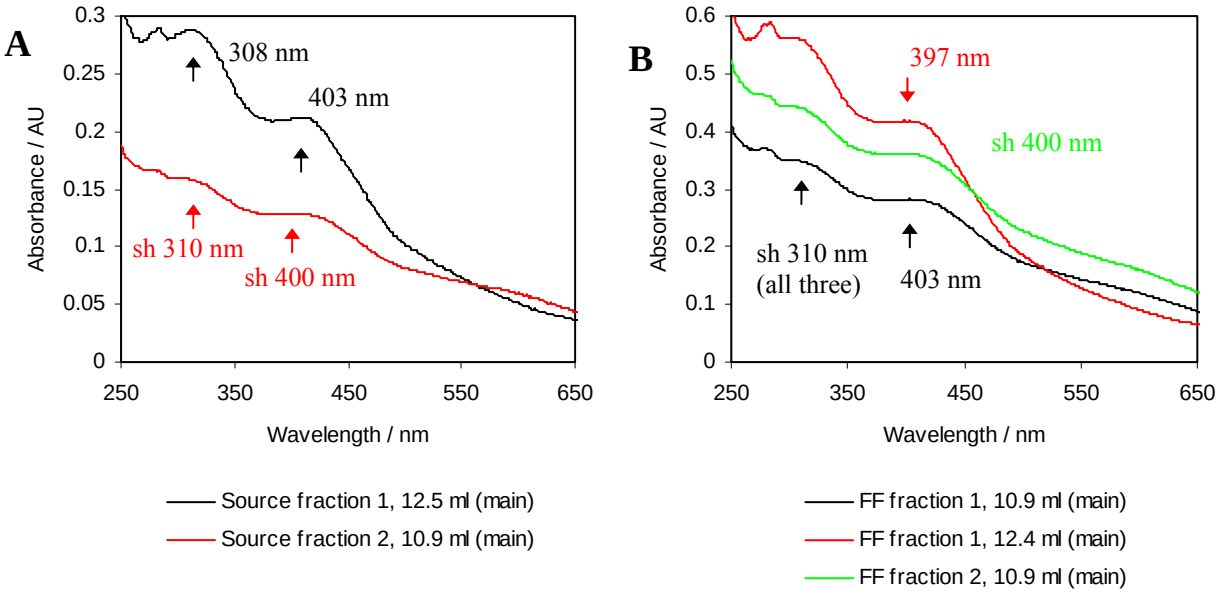


Figure 7.6 UV-vis spectra of fractions corresponding to peaks observed when performing analytical gel filtration in experiment 1. A) Source fractions 1 and 2. B) FF fractions 1 and 2. Maxima and shoulders (sh) are indicated.

The chromatogram of FF fraction 1 (figure 7.5, blue line), showed two equally sized main peaks (10.9 ml and 12.4 ml) indicating the presence of both a dimer and a monomer. The UV-vis spectrum of the monomer peak (figure 7.6 B, red spectrum) resembles a $[4Fe\text{-}4S]^{2+}$ ferredoxin while the UV-vis spectrum of the dimer (figure 7.6 B, black spectrum) resembles that of a $[3Fe\text{-}4S]^{1+}$ ferredoxin (see figure 2.2). FF fraction 2 (figure 7.5, brown line) gave one main peak at 10.9 ml indicating the presence of a dimer and a minor peak (12.5 ml) corresponding to a monomer. The UV-vis spectrum of the main peak (figure 7.6 B, green spectrum) resembles the spectrum recorded prior to performing the analysis (figure 7.4 B, red spectrum).
When eluting the holo-protein from the Q Sepharose Fast Flow column, all was eluted in one step and collected in two fractions. Therefore FF fraction 1 contains holo-protein corresponding to Source fractions 1 and 2 (monomer and dimer) while FF fraction 2 contains holo-protein corresponding to Source fraction 2 (dimer).

Table 7.1 Elution points of peaks observed in chromatograms obtained from analytical gel filtration.

Experiment	Fractions analysed by gel filtration	Point of elution presented in ml. Corresponding peaks from different chromatograms are presented in the same columns. The magnitude of each peak is presented in parenthesis (main, minor)				
1 [a]	Source fraction 1		10.9 (minor)	12.5 (main)		
	Source fraction 2	10.0 (minor)	10.9 (main)	12.4 (minor)		
	FF fraction 1		10.9 (main)	12.4 (main)		
	FF fraction 2		10.9 (main)	12.5 (minor)		
2 [a]	Source fraction 1	10.1 (minor)	10.9 (main)	12.4 (main)		
	Source fraction 2	10.0 (minor)	10.9 (main)	12.5 (minor)		
	FF fraction 1		10.9 (minor)	12.4 (main)		
5 [a]	Source fraction 1		10.9 (minor)	12.5 (main)		
	Source fraction 2		11.0 (minor)	12.4 (main)		
	FF fraction 1	10.1 (minor)	10.9 (minor)	12.2 (main)		
	FF fraction 2		shoulder	12.3 (main)		
3 [b]	Source fraction 1		shoulder	12.7 (main)		
	Source fraction 2	10.7 (minor)	11.6 (main)	12.8 (minor)		
	FF fraction 1		tailing	12.8 (main)		
	FF fraction 2		11.7 (main)	13.0 (minor)		
	Source fraction 1, 12.7 ml (main)	10.7 (minor)	11.4 (main)	12.9 (minor)		
	Source fraction 2, 11.8 ml (main)	Shoulder	11.5 (main)	12.8 (minor)		
4 [b]	Source fraction 1				shoulder	14.4 (main)
	Source fraction 1, 24 hours				13.1 (main)	14.4 (main)
	Source fraction 2				13.2 (main)	Shoulder
	FF fraction 1		11.3(YUH1)		shoulder	14.5 (main)
	FF fraction 2				13.3 (main)	shoulder

[a] The 10/30 Superdex 75 HR column was used
[b] The 13/31 Superdex 75 HR column was used

Experiment 2

Experiment 2 was a repetition of experiment 1 to test reproducibility of the results. The chromatogram obtained when purifying the reconstituted protein on the Source 30Q column (figure 7.7) contains two main peaks. The major peak eluted with a maximum absorption at 23.9 mS/cm (Source fraction 1) and a minor peak eluted at 41.8 mS/cm (Source fraction 2). Some small and poorly resolved peaks were observed between these two peaks. UV-vis spectra of the two fractions are shown in figure 7.8 A. The two main peaks are comparable to the peaks investigated in experiment 1 (Source fractions 1 and 2). However, here the first peak is the dominating peak. Even though experiment 1 and 2 were carried out in the same way, this major difference is seen in the chromatograms. The reasons for this discrepancy might be those discussed in section 7.2.1. As observed in experiment 1, Source fraction 1 (figure 7.8 A, black spectrum) contains a $[4Fe\text{-}4S]^{2+}$ protein (absorption maximum at 390 nm) while the UV-vis spectrum of Source fraction 2 (figure 7.8 A, red spectrum) resembles that of a $[3Fe\text{-}4S]^{1+}$ ferredoxin (weak absorption maximum at 402 nm) or a mixture of clusters.

An experiment purifying holo-protein on a Q Sepharose Fast Flow column (by gravity) was performed and the holo-protein was eluted in one step (section 7.1.1.1). One main fraction was collected, FF fraction 1, and from the UV-vis spectra shown in figure 7.8 B, it is seen to be comparable to Source fraction 1.

Prior to performing analytical gel filtration, Source fractions 1 and 2 were concentrated ca. 10 and 15 times, respectively, and FF fraction 1 was concentrated ca. 5 times (concentration was performed using a stirred Amicon cell equipped with a YM3 membrane). The three chromatograms obtained from the analytical gel filtrations are presented in figure 7.9. The elution volume at maximum absorption of each of the peaks is presented in table 7.1 and UV-vis spectra corresponding to the main peaks are presented in figure 7.10.

For Source fraction 1 (figure 7.9, red line), two main peaks were observed at 10.9 ml and 12.4 ml, respectively, along with a minor peak at 10.1 ml in the chromatogram. This indicates that Source fraction 1 contains monomer, dimer and a small amount of polymer, which was not the case in experiment 1 (see above). The multiple components might be due to oxidative breakdown prior to analysis or during analysis or it might be due to conversion between monomer and dimer (further discussed in section 7.2.8). UV-vis spectra of the main fractions reveal that the monomer fraction (12.4 ml) appears to contain a $[4Fe\text{-}4S]^{2+}$ protein (figure 7.10 A red spectrum, absorption maximum at 390 nm) while the cluster type for the dimer fraction is difficult to infer (10.9 ml, figure 7.10 A black spectrum). However, the UV-vis spectrum is comparable to that of Source fraction 2 (figure 7.8 A, red spectrum). Source fraction 2 (figure 7.9 green line) gave one main peak at 10.9 ml, indicating the presence of a dimer, and a minor peak (12.4 ml) indicating a monomer, as well as a peak at 10.0 ml indicating a polymer form. The UV-vis spectrum of the main peak (figure 7.10 A, blue spectrum) resembles the spectrum recorded prior to performing the analysis (figure 7.8 A, red spectrum). The UV-vis spectrum of the polymer form (10.0 ml, figure 7.10 A, green spectrum) differs slightly since it has a maximum at 306 nm and the holo-protein corresponding to this peak probably contains multiple cluster types.

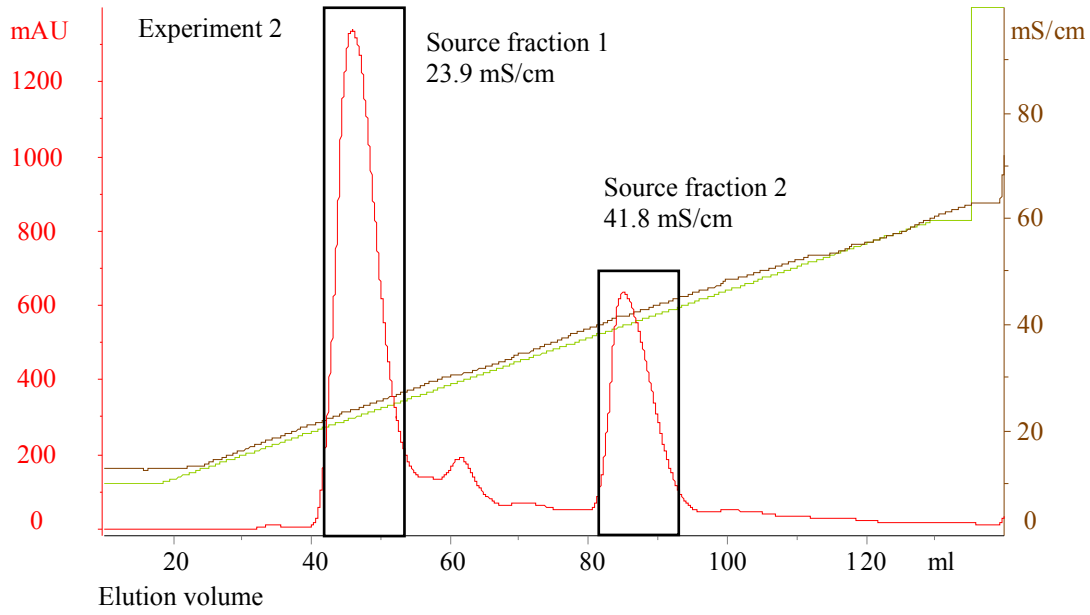


Figure 7.7 Chromatogram for purification of holo-Ub-Ori-ISP on the Source 30Q column in experiment 2, when using $FeCl_3$ as the iron source. The red line indicates the absorbance at 280 nm. The brown and the green lines indicate the conductivity and the NaCl gradient, respectively. The concentration of NaCl can be read from the conductivity axis employing the unit 10^{-2} M. Collected fractions are indicated with black boxes and the conductivity at maximum absorption is also presented.

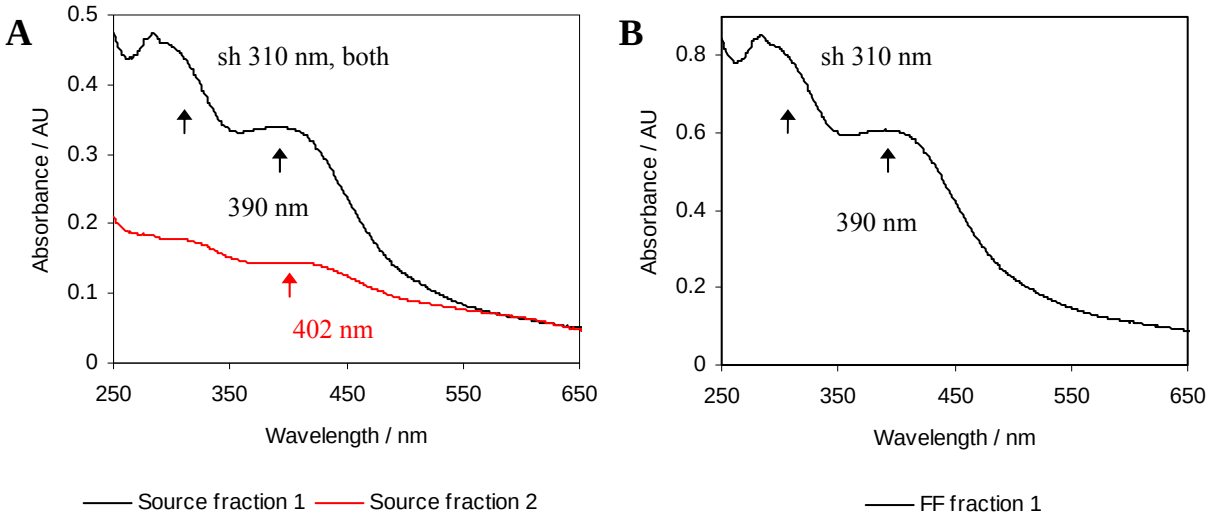


Figure 7.8 UV-vis spectra of main fractions from anion-exchange in experiment 2. A) Fractions from purification on Source 30Q column. Both fractions were diluted 5 times (with 20 mM Tris/HCl pH 8.0) prior to recording the spectra. B) Fraction collected when performing purification on Q Sepharose Fast Flow column. The fraction was diluted 20 times (with 20 mM Tris/HCl pH 8.0) prior to recording the spectrum. Maxima and shoulders (sh) are indicated.

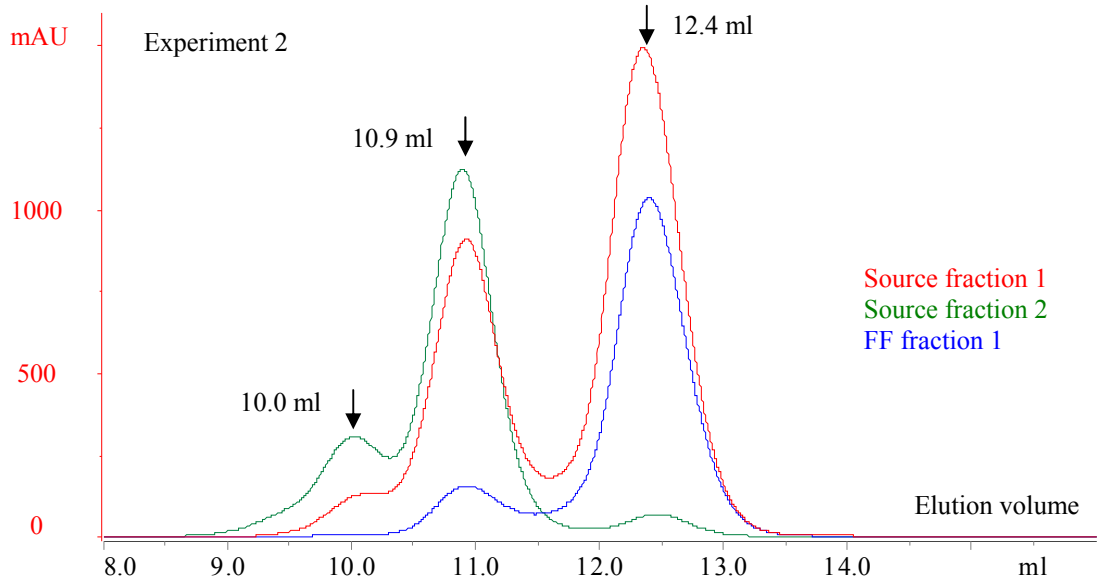


Figure 7.9 Chromatograms for analytical gel filtrations in experiment 2 obtained when analysing fractions collected from anion-exchange. Absorption was measured at 280 nm. Red, green and blue lines indicate overlaid chromatograms for Source fraction 1, Source fraction 2, and FF fraction, respectively.

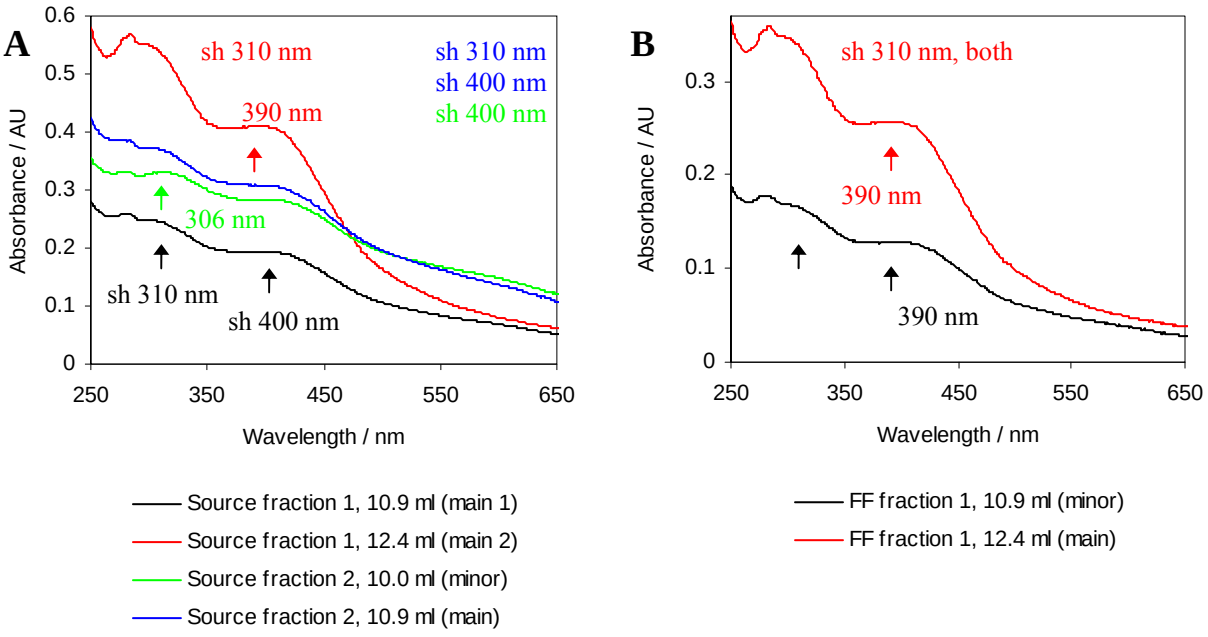


Figure 7.10 UV-vis spectra of fractions corresponding to peaks observed when performing analytical gel filtration in experiment 2. A) Source fractions 1 and 2. Fractions were three times diluted (with 20 mM Tris/HCl pH 8.0) prior to recording the spectra except for Source fraction 2, 10.0 ml (minor), which was undiluted. B) FF fractions 1. FF fraction 1, 12.4 ml (main) was three times diluted (with 20 mM Tris/HCl pH 8.0). Maxima and shoulders (sh) are indicated.

FF fraction 1, gave one main peak at 12.4 ml (figure 7.9, blue line) indicating the presence of a monomer. A minor peak was observed at 10.9 ml indicating the presence of a dimer. UV-vis spectra of both of these peaks (figure 7.10 B) closely resemble the spectrum recorded prior to performing the analysis (figure 7.8 B), indicating the presence of a $[4Fe\text{-}4S]^{2+}$ cluster (absorption maximum at 390 nm). When eluting the holo-protein from the Q Sepharose Fast Flow column, the holo-protein was eluted in one step and therefore FF fraction 1 contains holo-protein corresponding to Source fraction 1, and holo-protein corresponding to Source fraction 2 was probably eluted in minor fractions (tailing) that were not analysed. It is interesting to note that FF fraction 1 is more homogenous than Source fraction 1 (as judged by analytical gel filtration), which was not the case in experiment 1. This might be due to breakdown or conversion as discussed in section 7.2.8.

Experiment 3

Since experimental work was now carried out in an anaerobic chamber in our own laboratory and since no oxygen sensor was present, experiment 1 carried out in France was repeated to test reproducibility prior to performing cluster re-constitution into Ori-ISP.

Throughout this experiment the nano-drop UV-vis spectrophotometer (see section 4.4) was employed since this was present inside the anaerobic chamber. This resulted in poor spectra rendering it impossible to determine exact absorption maxima.

The chromatogram obtained when purifying the reconstituted protein on the Source 30Q column is presented in figure 7.11. Two main peaks eluted at 22.7 mS/cm (Source fraction 1) and 30.8 mS/cm (Source fraction 2), respectively. Some small and poorly resolved peaks were also observed between these two peaks. It is seen that the gradient was too short and only the first part of Source fraction 2 was collected. The UV-vis spectra recorded of Source fractions 1 and 2 are of quite poor quality (figure 7.12 A) which makes it difficult to distinguish any particular cluster types. The chromatogram is comparable to the chromatograms obtained in experiment 1 and 2 (figures 7.3 and 7.7) however, the two main fractions are found to elute at somewhat lower conductivities. This might be due to a higher temperature of the anaerobic chamber (see section 4.2). The size of the two main peaks are somewhat comparable (Source fraction 2 being the dominating peak), which was not the case previously (figures 7.3 and 7.7) since Source fraction 2 dominated in experiment 1 while Source fraction 1 dominated in experiment 2. The chromatogram obtained in this experiment can be regarded as being something in between the chromatograms obtained in experiment 1 and 2. Even though these three experiments were carried out in the same way, differences in the chromatograms are apparent. Oxygen can be ruled out as the main contributor and reasons for these differences might be those discussed in section 7.2.1.

An experiment where the holo-protein was purified on a Q Sepharose Fast Flow column (by gravity) was performed and the holo-protein was eluted in two steps. One fraction was eluted using 0.5 M NaCl, FF fraction 1, and one fraction was eluted using 1.0 M NaCl, FF fraction 2. The UV-vis spectra of these fractions are shown in figure 7.12 B. The quality of the spectra is quite poor. However, FF fractions 1 and 2 have absorption maxima at about 395 nm and 390 nm, respectively, indicating that both fractions contain a $[4Fe\text{-}4S]^{2+}$ protein.

Prior to performing analytical gel filtration, Source fractions 1 and 2 were both concentrated ca. 10 times while FF fractions 1 and 2 were not (concentration was performed using Vivacell

70 equipped with a 5000 Da MWCO membrane). The four chromatograms obtained from the analytical gel filtrations are presented in figure 7.13. From each chromatogram one main peak was apparent however, shoulders or minor peaks were also present as indicated in table 7.1. UV-vis spectra corresponding to the main peaks are not presented since they resemble those seen in figure 7.12 A. As seen previously (in experiment 1 and 2) holo-protein eluted at two positions, namely ca. 11.6 ml and ca. 12.8 ml (see table 7.1). These positions are different from what was previously observed (see table 7.1) since analytical gel filtration was performed on a different column. Elution at ca. 11.6 ml is corresponds to a dimeric compound while elution at ca. 12.8 ml is corresponds to a monomeric compound (discussed in section 7.2.7).

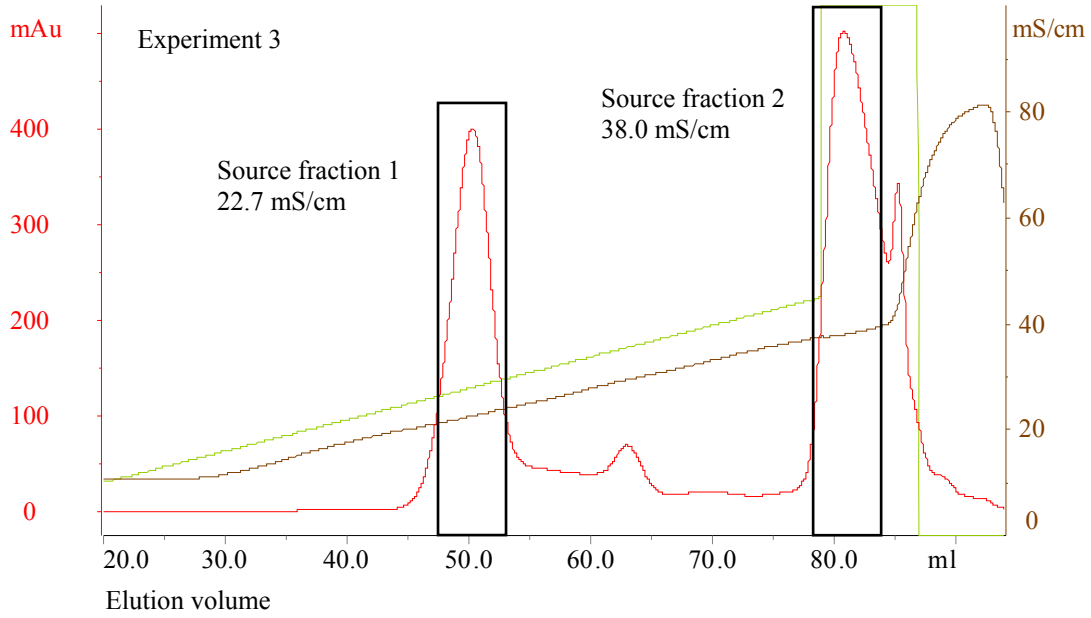


Figure 7.11 Chromatogram for purification of holo-Ub-Ori-ISP on the Source 30Q column in experiment 3, when using $FeCl_3$ as the iron source. The red line indicates the absorbance at 280 nm. The brown and the green lines indicate the conductivity and the NaCl gradient, respectively. The concentration of NaCl can be read from the conductivity axis employing the unit 10^{-2} M. Collected fractions are indicated with black boxes and the conductivity at maximum absorption is also presented.

Analytical gel filtrations on Source fraction 1 and FF fraction 1 are comparable and both gave one main peak (monomer) eluting at 12.7 ml and 12.8 ml, respectively (figure 7.13, red and blue lines). Analytical gel filtrations on Source fraction 2 and FF fraction 2 are also comparable and both gave one main peak (dimer) eluting at 11.6 ml and 11.7 ml, respectively (figure 7.13, green and brown lines). These observations are in agreement with those made in experiment 1 and 2.

The main peaks observed in the chromatograms when analysing Source fractions 1 and 2 (figure 7.13) by gel filtration were collected and re-purified by analytical gel filtration to test whether the species formed were stable or not. The result of re-purifying Source fraction 1, main peak (12.7 ml, monomer), is shown in figure 7.14 A. It is seen that the one peak observed previously has turned into three poorly resolved peaks (the dimer peak is now dominating) and broadening in the region 8 to 10 ml. This indicates breakdown, resulting in

formation of dimers and polymers. This might have taken place prior to analysis or during analysis (further discussed in section 7.2.8). When re-analysing Source fraction 2, main peak (11.7 ml), the resulting chromatogram (figure 7.14 B) shows formation of shoulders on both sides of the main peak which either indicates breakdown or slight conversion of dimer to monomer and polymer (further discussed in section 7.2.8).

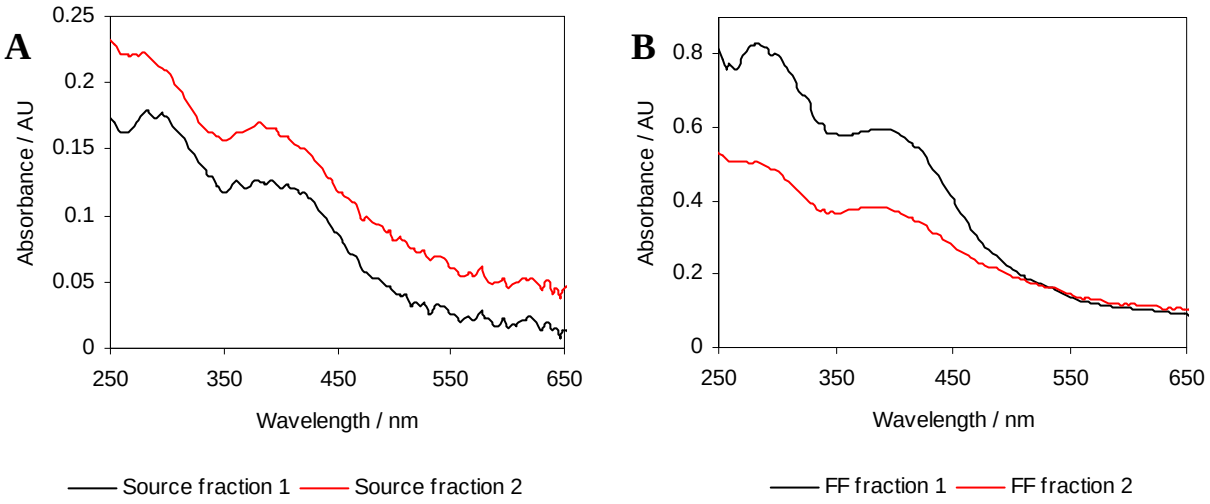


Figure 7.12 UV-vis spectra of main fractions from anion-exchange in experiment 3. A) Fractions from purification on Source 30Q column. The spectra are too poor to determine exact absorption maxima. B) Fractions collected when performing purification on Q Sepharose Fast Flow column. All fractions have shoulders at 310 nm while FF fraction 1 and FF fraction 2 have an absorption maximum at ca. 395 nm and 390 nm, respectively.

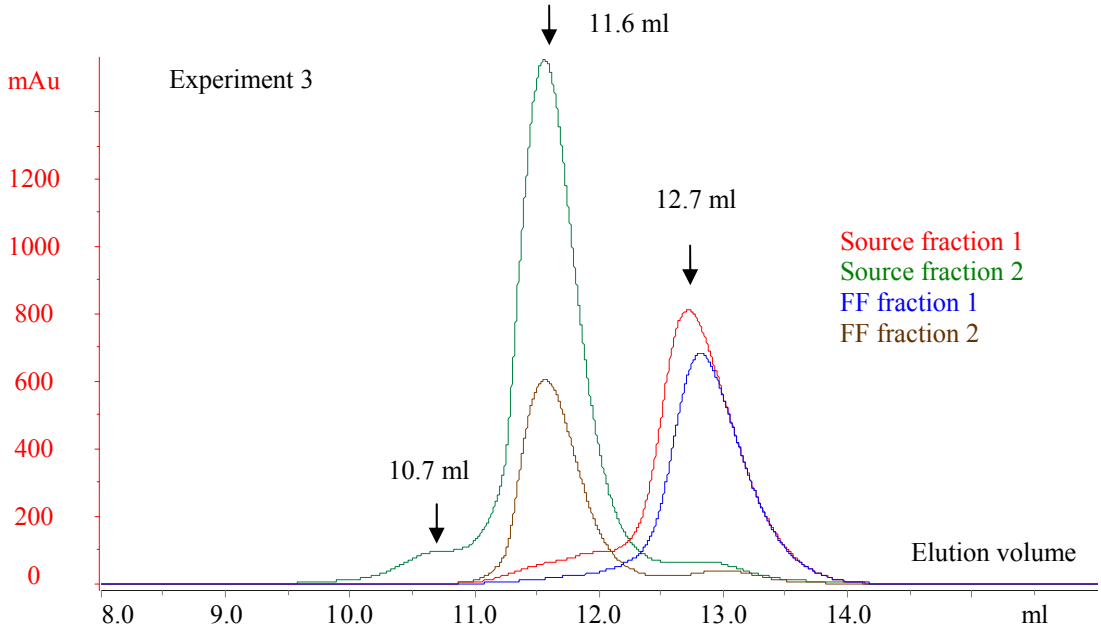


Figure 7.13 Chromatograms for analytical gel filtrations obtained when analysing fractions collected from anion-exchange in experiment 3. Absorption was measured at 280 nm. Red, green, blue and brown lines indicate overlaid chromatograms for Source fraction 1, Source fraction 2, FF fraction 1 and FF fraction 2, respectively.

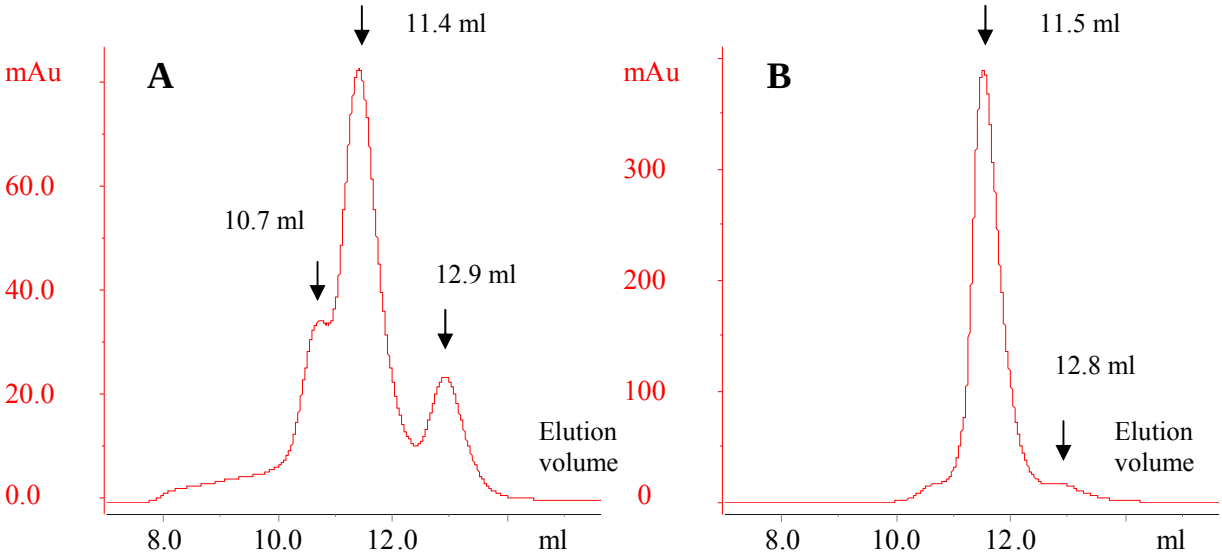


Figure 7.14 Chromatograms obtained when re-analysing main peak fractions obtained from analytical gel filtrations. A) Source fraction 1, 12.7 ml (main) obtained from analytical gel filtration was re-purified. B) Source fraction 2, 11.8 ml (main) obtained from analytical gel filtration was re-purified.

7.2.3.2 Purification of holo-Ori-ISP

After performing proteolytic cleavage on Ub-Ori-ISP using the protease YUH1 (see section 7.1.1.2, experiment 4) the composition of the solution was analysed by SDS-PAGE to test the extent of cleavage, which was found to be complete (results not shown). Cluster reconstitution was then performed into Ori-ISP. The observations made during cluster reconstitution and desalting were comparable to those made previously (section 7.2.1). Also when performing anion-exchange purifications the observations made were comparable to those observed previously (7.2.3.1). However, the holo-protein was found to elute at higher conductivities as would be expected.

The chromatogram obtained from the HPLC purification is presented in figure 7.15. A small peak (uncoloured) eluted at 29.0 mS/cm and it was assigned to YUH1 due to the position of the peak and UV-vis spectrophotometric analysis revealed only apo-protein (results not shown). Ubiquitin did not bind to the column (results not shown). A larger peak followed by tailing eluted at 36.5 mS/cm (Source fraction 1). The tailing could indicate slow breakdown during purification or presence of multiple species. The dominating peak eluted at 72.1 mS/cm (Source fraction 2). The chromatogram differs from those previously obtained (see section 7.2.3.1) since comparable coloured fractions elute at much higher conductivities in this experiment. The elution at higher conductivities is a result of a change in overall charge and charge distribution upon cleavage of the Ub-fusion. UV-vis spectra of Source fractions 1 and 2 recorded right after purification and after 24 hours are shown in figure 7.16 (absorption maxima and shoulders are outlined in the figure). These spectra reveal that Ub is no longer part of the molecules since no distinct absorption is seen at 280 nm however, what was previously seen to be a shoulder at 310 nm has now turned into a maximum in the same region. Source fraction 1 and Source fraction 2 are comparable to fractions obtained previously named the same way (section 7.2.3.1). The spectrum of Source fraction 1 shows that the fraction most likely contains a $[4Fe\text{-}4S]^{2+}$ protein (figure 7.16 A, shoulder at 390 nm)

while the spectrum of Source fraction 2 (figure 7.16 B, shoulder at 400 nm) resembles more closely the spectra previously suggested to contain $[3Fe\text{-}4S]^{1+}$ or different cluster types (section 7.2.3.1). After 24 hours the spectrum of Source fraction 1 changed (figure 7.16 A), the maximum at 305 nm shifted to 310 nm and flattening of the spectrum was evident. It appears that breakdown has happened. No marked changes were observed in Source fraction 2 within the same period (figure 7.16 B).

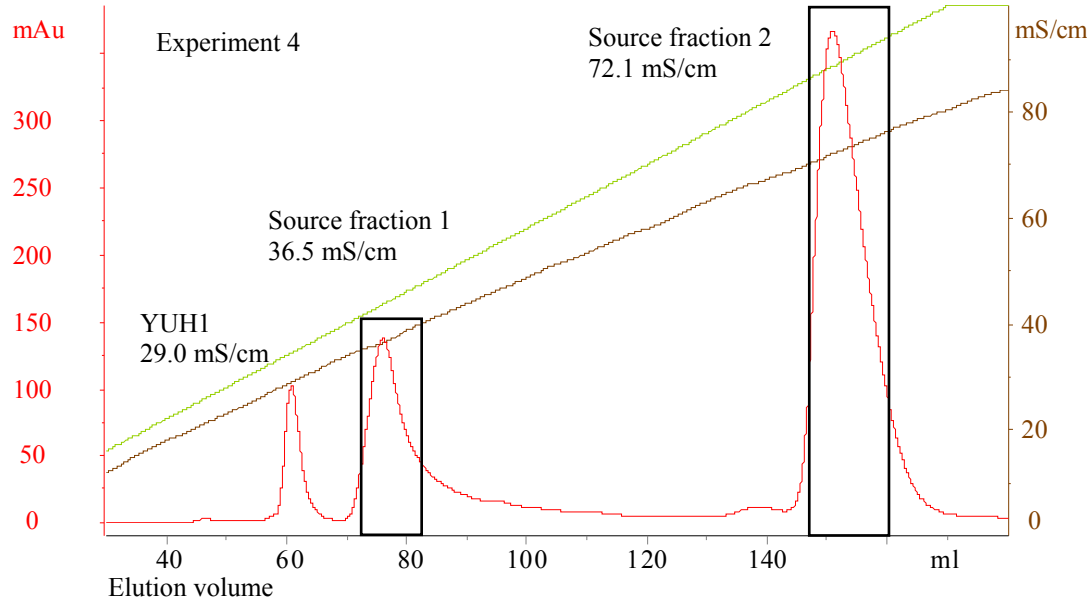


Figure 7.15 Chromatogram for purification of holo-Ori-ISP on the Source 30Q column in experiment 4. The red line indicates the absorbance at 280 nm. The brown and the green lines indicate the conductivity and the NaCl gradient, respectively. The concentration of NaCl can be read from the conductivity axis employing the unit 10^{-2} M. Collected fractions are indicated with black boxes and the conductivity at maximum absorption is also presented.

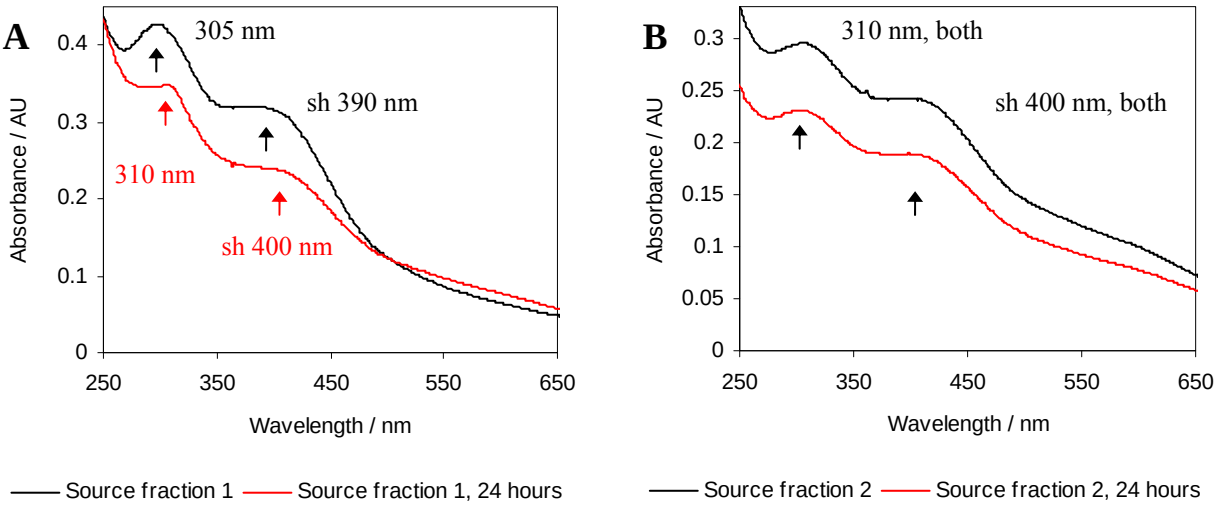


Figure 7.16 UV-vis spectra of fractions from anion-exchange on Source 30Q column in experiment 4. Spectra were recorded right after elution and 24 hours later. A) Source fraction 1. B) Source fraction 2, diluted ca. three times prior to recording the spectra. Maxima and shoulders (sh) are indicated.

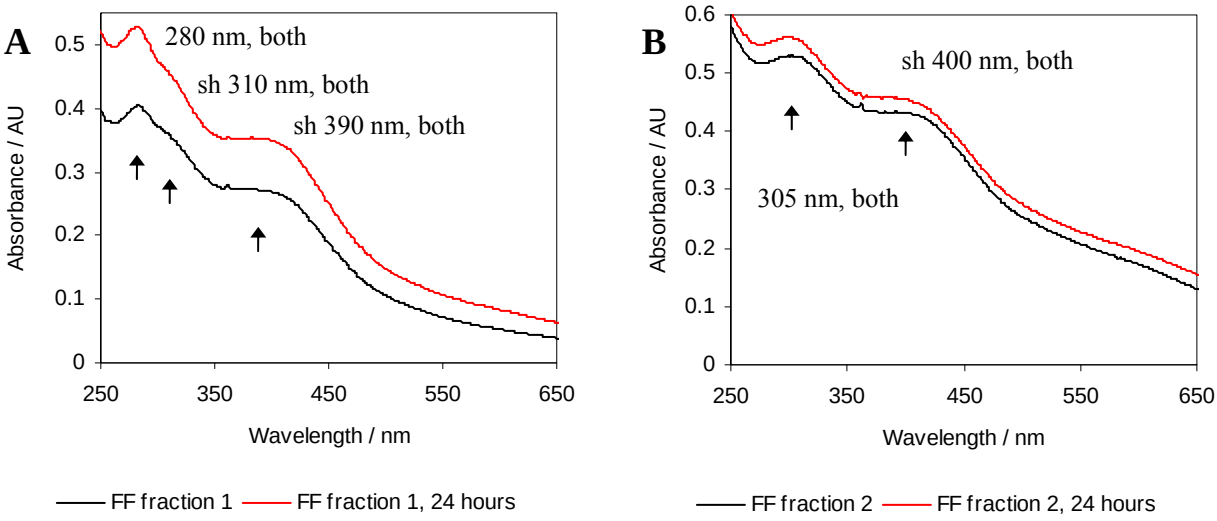


Figure 7.17 UV-vis spectra of fractions from anion-exchange on the Q Sepharose Fast Flow column in experiment 4. Spectra were recorded right after elution and 24 hours later. A) FF fraction 1. B, FF fraction 2. Fractions 1 and 2 were diluted ca. three and five times, respectively, prior to recording the spectra. Maxima and shoulders (sh) are indicated.

An experiment where the holo-protein was purified on a Q Sepharose Fast Flow column (by gravity) was performed and the holo-protein was eluted in two steps. One fraction was eluted using 0.5 M NaCl, FF fraction 1, and one fraction was eluted using 1.5 M NaCl, FF fraction 2. The UV-vis spectra of these fractions are shown in figure 7.17. FF fraction 1 is comparable to Source fraction 1 (figure 7.16 A, containing a $[4Fe\text{-}4S]^{2+}$ protein, shoulder at 390 nm). However, a maximum at ca. 280 nm shows that YUH1 is found in the fraction (this is because of the stepwise elution). FF fraction 2 is comparable to Source fraction 2 (figure 7.16 B, indicating the presence of $[3Fe\text{-}4S]^{1+}$ or multiple cluster types, shoulder at 400 nm). No changes were observed in the UV-vis spectra (figure 7.17) over a 24 hour period for both fractions.

It was tested whether the obtained holo-protein could be concentrated using a 3000 MWCO membrane. However, this was not possible since the holo-protein passed through.

Analytical gel filtration was performed on the four collected fractions right after preparation (figure 7.18) and again after 24 hours, and as seen previously (sections 7.2.1.2 and 7.3.2), holo-protein was found to elute at two positions, namely ca. 13.2 ml and ca. 14.4 ml (see table 7.1). These positions are different from those observed previously on the same column. This is because ubiquitin is no longer part of the molecules. Elution at ca. 13.2 ml corresponds to a dimeric protein while elution at ca. 14.4 ml is corresponds to a monomeric protein (discussed in section 7.2.7).

Analytical gel filtrations on Source fraction 1 and FF fraction 1 both gave main peak (monomer) eluting at 14.4 ml and 14.5 ml, respectively (figure 7.18, red and blue lines). A minor peak at 11.4 ml was also observed when analysing FF fraction 1, which is assigned to YUH1. Analytical gel filtrations on Source fraction 2 and FF fraction 2 also both resulted in one main peak (dimer) eluting at 13.2 ml and 13.3 ml, respectively (figure 7.18, green and brown lines). Shoulders were seen to be present for all the main peaks. The observations made

in this experiment are in agreement with those made previously when using $FeCl_3$ as the iron source (see section 7.2.3.1). The main peak observed when analysing FF fraction 2 is very broad and the top of the peak appears rounded. This could indicate multiple components with comparable sizes.
When analysing the four fractions again (by analytical gel filtration) 24 hours later, no changes were observed for Source fraction 2 and FF fractions 1 and 2. However, analysis of Source fraction 1 revealed two similarly sized main peaks in the chromatogram (figure 7.19) indicating the presence of monomer and dimer. Changes were already apparent from the UV-vis spectrum recorded prior to the analysis (see above and figure 7.16 A). This change is probably due to breakdown (discussed in section 7.2.8).
Figure 7.20 shows UV-vis spectra of the main peak fractions obtained by analytical gel filtration. The spectra of Source fraction 2 and FF fraction 2 main peaks (blue and cyan) are comparable to spectra obtained prior to the analysis (figures 7.16 B and 7.17 B). Spectra of Source fraction 1 and FF fraction 1 main peaks (black, red, green and pink) are comparable however, they differ from spectra recorded prior to the analysis (figures 7.16 A and 7.17 A). They now resemble more closely those of [2Fe-2S] or [3Fe-4S] ferredoxins (see section 6.3.1.1 for a discussion of a comparable spectrum, figure 6.5). The observed changes in the spectra are probably due to breakdown of the cluster of the holo-protein from one cluster type to another during the analyses (discussed in section 7.2.8). These results show that the dimeric protein is more stable than the monomeric protein. This could be attributed to the assumed overall lower oxidation state of the monomeric compound, which can easily be oxidized by trace amounts of oxygen (further discussed in section 7.2.8). Further, the dimeric protein might provide better protection of the cluster from the surroundings.

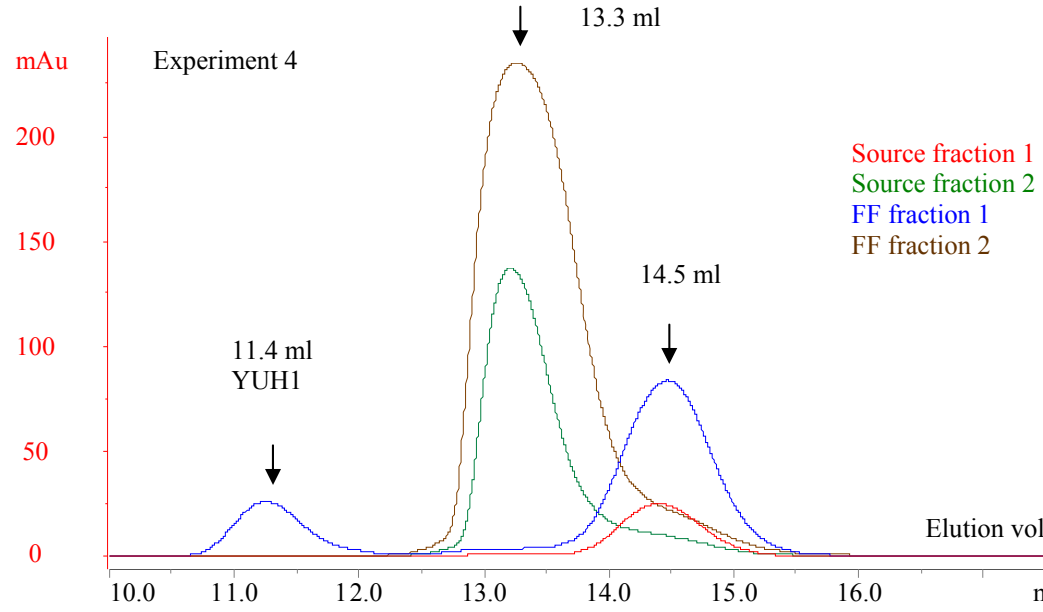


Figure 7.18 Chromatograms for analytical gel filtrations obtained when analysing fractions collected from anion-exchange in experiment 4. Absorption was measured at 280 nm. Red, green, blue and brown lines indicate overlaid chromatograms for Source fraction 1, Source fraction 2, FF fraction 1 and FF fraction 2, respectively.

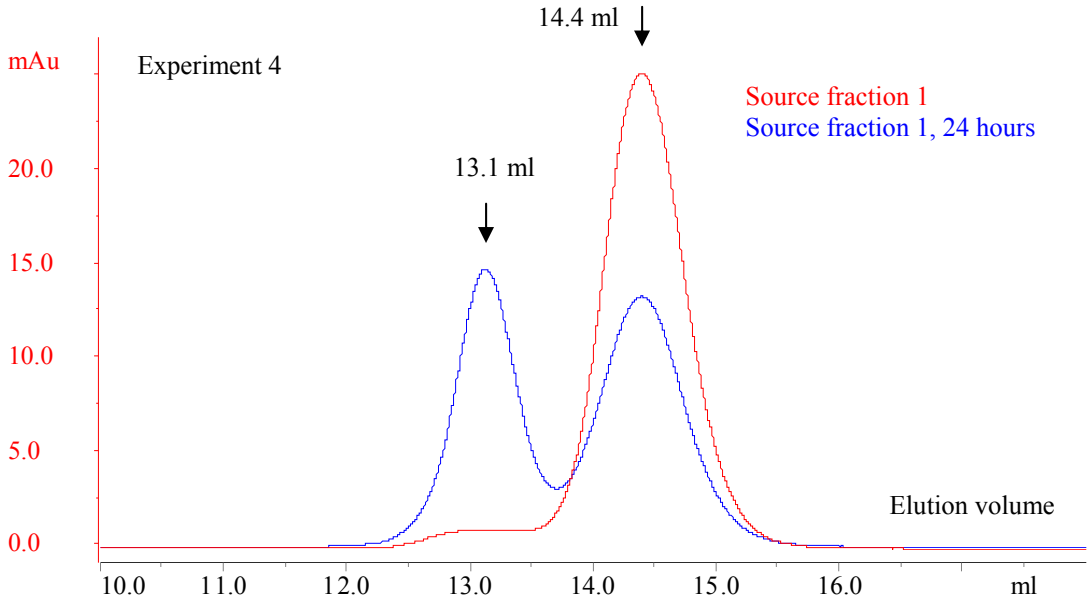


Figure 7.19 Chromatograms for analytical gel filtrations of Source fraction 1 collected from anion-exchange in experiment 4. The analysis was carried out just after obtaining the fraction 1 (red line) and 24 hours after obtaining the fraction (blue line). Absorption was measured at 280 nm.

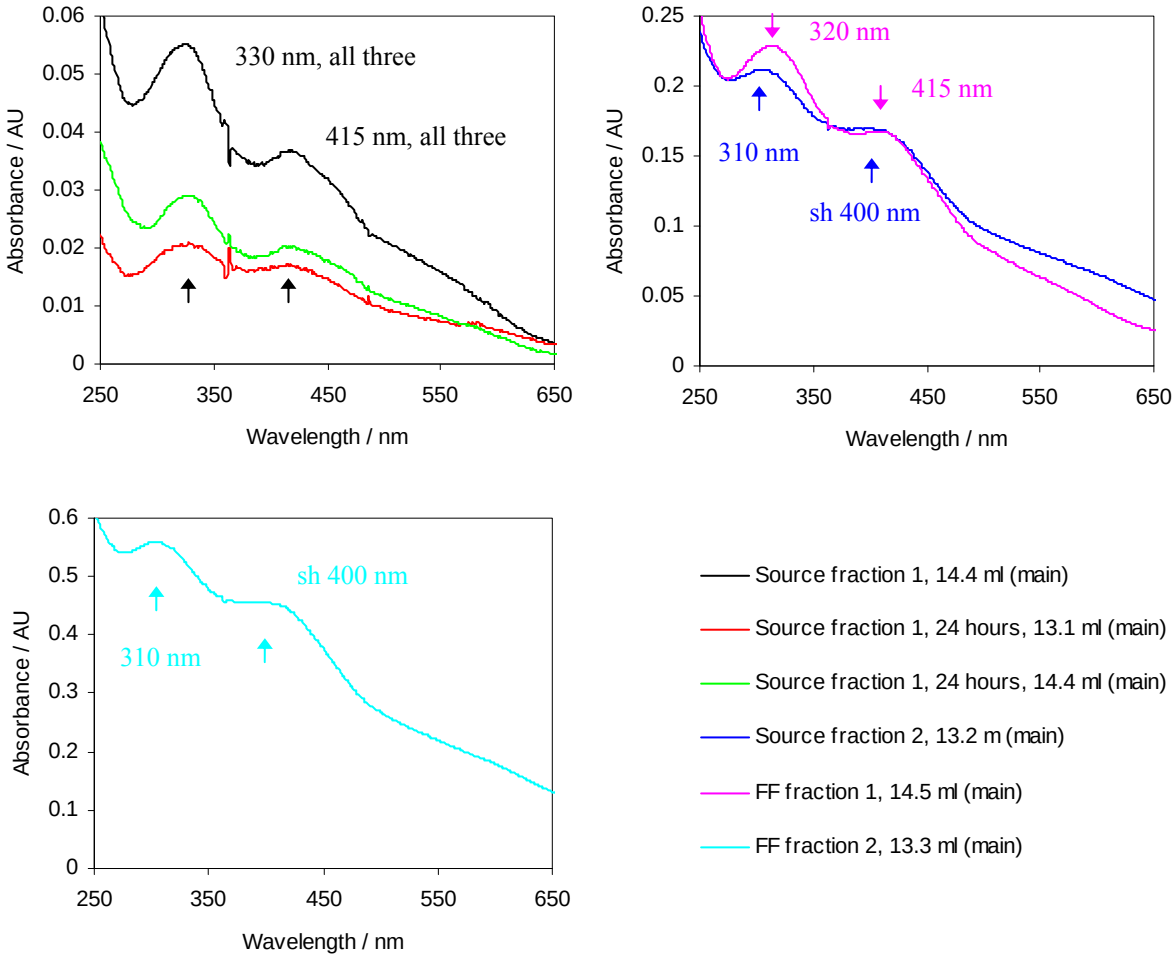


Figure 7.20 UV-vis spectra of fractions corresponding to main peaks observed when performing analytical gel filtrations in experiment 4. Maxima and shoulders (sh) are indicated.

7.2.4 Purification of holo-Ub-Ori-ISP using Fe(II) as the iron source

In experiment 5, $(NH_4)_2Fe(II)(SO_4)_2 \cdot 6H_2O$ was used as iron source and differences observed during cluster reconstitution are explained and discussed in sections 7.2.1 and 7.2.5. The chromatogram obtained when purifying the reconstituted protein on the Source 30Q column (figure 7.21) differs from those obtained in experiment 1 to 3 (figures 7.3, 7.7 and 7.11). A minor peak (uncoloured) eluted at 20.1 mS/cm and from UV-vis spectrophotometric analysis this peak was found to contain apo-protein (results not shown). A small poorly resolved peak eluted at 25.1 mS/cm (Source fraction 1) and a major peak eluted at 36.6 mS/cm (Source fraction 2). UV-vis spectra of the two fractions recorded right after elution and after 24 hours are shown in figure 7.22. Source fraction 1 is comparable to Source fraction 1's from the previous experiments (section 7.2.3) and the UV-vis spectrum (figure 7.22 A) shows that this fraction very likely contains $[4Fe\text{-}4S]^{2+}$ protein (shoulder at 390 nm). Source fraction 2 is not comparable to Source fraction 2's from the previous experiments due to the earlier elution. The obtained UV-vis spectra are also not comparable. Just after preparation no distinct cluster types can be inferred (figure 7.22 B), but after 24 hours the protein appears to contain a $[4Fe\text{-}4S]^{2+}$ cluster (shoulder at 390 nm, figure 7.22 B). This strongly suggests that Source fraction 2 contained reduced holo-protein (containing a $[4Fe\text{-}4S]^{1+}$cluster) right after elution.

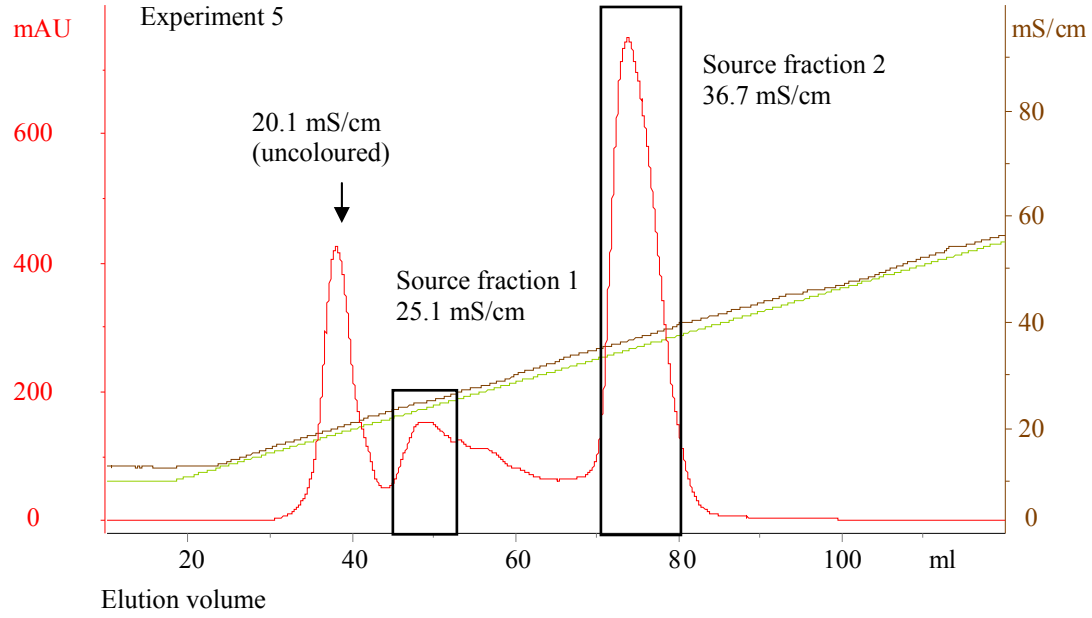


Figure 7.21 Chromatogram for purification of holo-Ub-Ori-ISP on the Source 30Q column in experiment 5, when using $(NH_4)_2Fe(II)(SO_4)_2 \cdot 6H_2O$ as the iron source. The red line indicates the absorbance at 280 nm. The brown and the green lines indicate the conductivity and the NaCl gradient, respectively. The concentration of NaCl can be read from the conductivity axis employing the unit 10^{-2} M. Collected fractions are indicated with black boxes and the conductivity at maximum absorption is also presented.

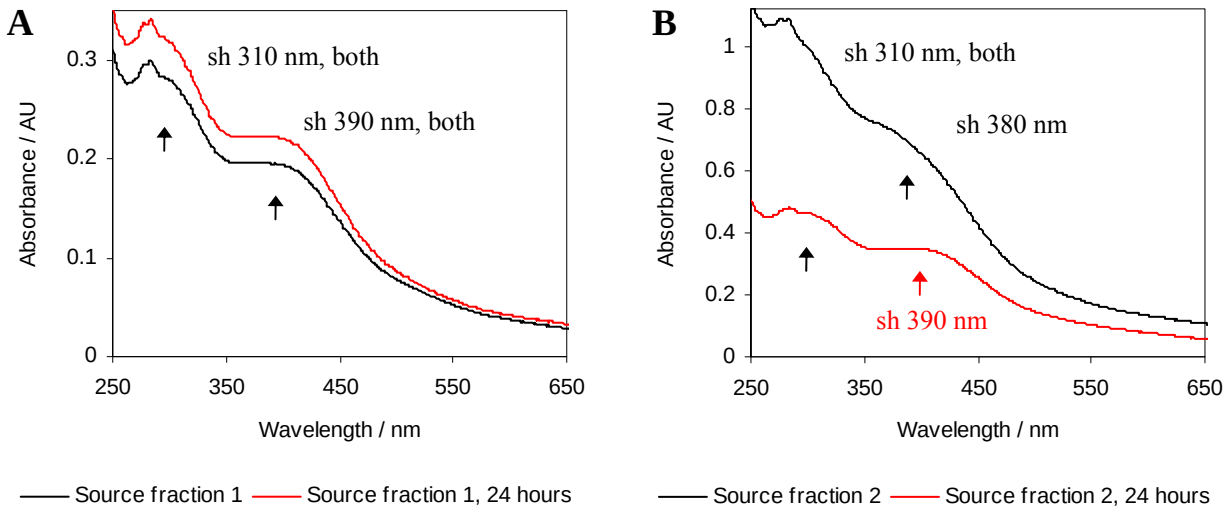


Figure 7.22 UV-vis spectra of fractions from anion-exchange on Source 30Q column in experiment 5. Spectra were recorded right after elution and 24 hours later. A Source fraction 1. B) Source fraction 2. When recording the spectrum after 24 hours the fraction was diluted two times. No distinct absorption maxima are present, but shoulders (sh) are presented in the figures.

An experiment where the holo-protein was purified on a Q Sepharose Fast Flow column (by gravity) was performed and the holo-protein was eluted in one step. Two main fractions were collected (FF fractions 1 and 2). The UV-vis spectrum of FF fraction 1 (figure 7.23 A) is dominated by apo-protein (absorption at 280 nm) and no distinct cluster types can be inferred. From the UV-vis spectrum of FF fraction 2 (figure 7.23 B), it is also difficult to infer any particular cluster types and leaving the sample for 24 hours at room temperature did not change the spectrum as was the case for Source fraction 2.

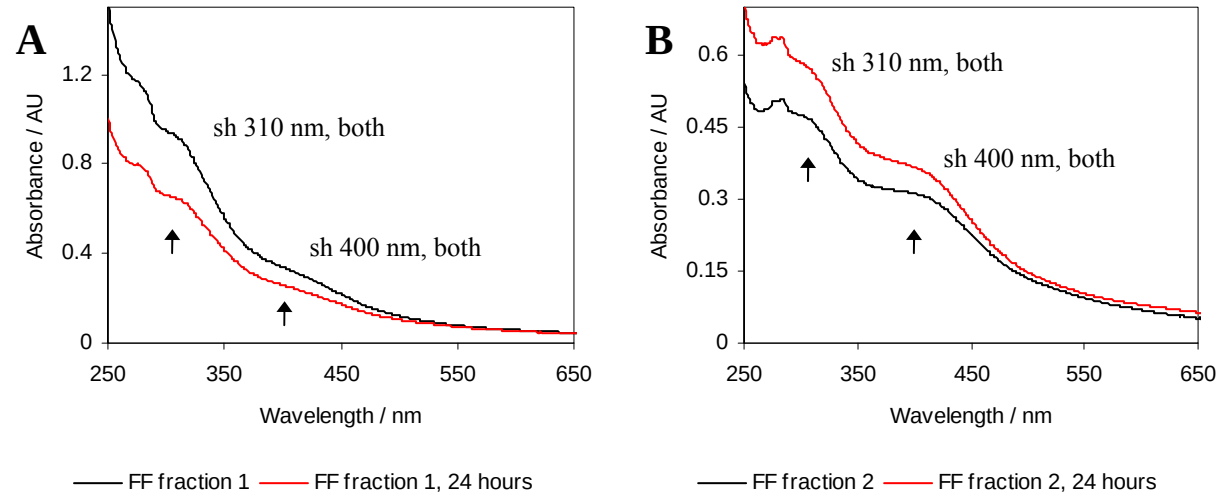


Figure 7.23 UV-vis spectra of fractions from anion-exchange on Q Sepharose Fast Flow column in experiment 5. Spectra were recorded right after elution and 24 hours later. All spectra were recorded on ca. ten times diluted samples. A FF fraction 1. B) FF fraction 2. No distinct absorption maxima are present but shoulders (sh) are presented in the figures.

Prior to performing analytical gel filtration, Source fractions 1 and 2 were concentrated ca. 20 and 10 times, respectively, and FF fractions 1 and 2 were concentrated ca. 10 and 5 times, respectively (concentration was performed using a stirred Amicon cell equipped with a YM3 membrane).
The four chromatograms obtained from the analytical gel filtrations are presented in figure 7.24. One main peak was observed in all four chromatograms (varying from 12.2 ml to 12.5 ml) indicating they were all dominated by monomeric species. However, tailing (FF fraction 2), minor resolved peaks (FF fraction 1) or very small peaks (Source fractions 1 and 2) were found to elute prior to the main peak. The elution volume at maximum absorption of each of the peaks is presented in table 7.1 and UV-vis spectra corresponding to the main peaks are presented in figure 7.25.

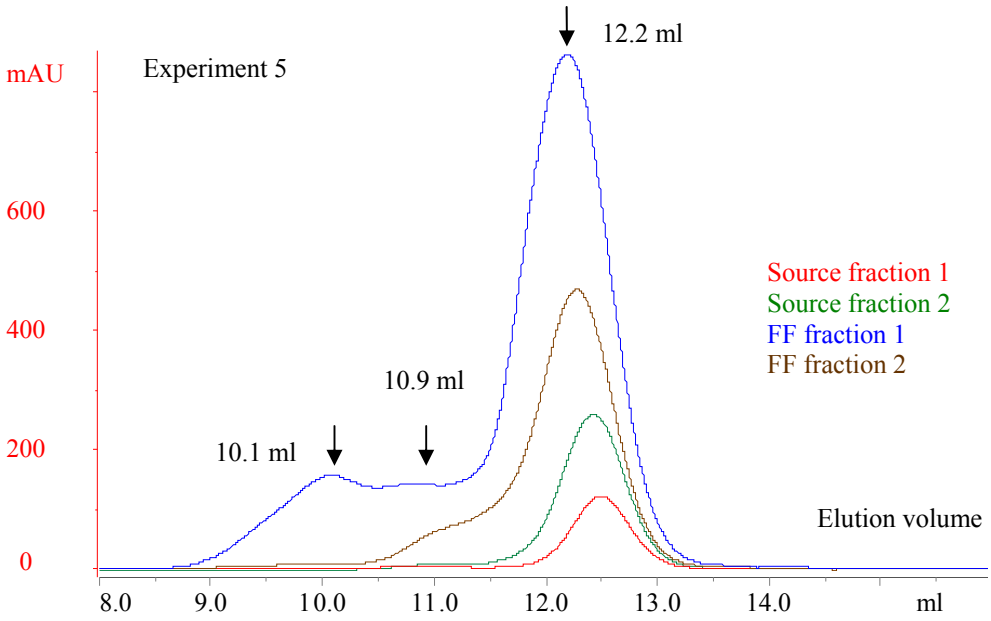


Figure 7.24 Chromatograms for analytical gel filtrations in experiment 5 obtained when analysing fractions collected from anion-exchange. Absorption was measured at 280 nm. Red, green, blue and brown lines indicate overlaid chromatograms for Source fraction 1, Source fraction 2, FF fraction 1 and FF fraction 2, respectively.

For Source fraction 1, the UV-vis spectrum of the main peak (figure 7.25, blue spectrum) is different compared to the spectrum recorded prior to the analysis (figure 7.22 A). It is difficult to infer any particular cluster type. However, the spectrum is comparable to those previously inferred to be dominated by $[3Fe\text{-}4S]^{1+}$ protein or multiple cluster types (see section 7.2.3). This probably indicates breakdown (discussed in section 7.2.8). Source fraction 2 gave a main fraction with the UV-vis spectrum resembling a $[4Fe\text{-}4S]^{2+}$ protein (figure 7.25, pink spectrum, maximum at 390 nm). For FF fraction 1, the main peak is seen to contain what appears to be a $[4Fe\text{-}4S]^{2+}$ protein (figure 7.25, red spectrum, shoulder at 390 nm), but apo-protein is still a dominating factor in the spectrum (absorption at 280 nm). The minor peak (10.1 ml) is dominated by apo-protein as seen from the UV-vis spectrum (figure 7.25, black spectrum, absorption at 280 nm). However, some coloured species were also apparent, but

cannot be resolved. The UV-vis spectrum of the main peak for FF fraction 2 after analysis (figure 7.25, green spectrum) is different compared to the spectrum recorded prior to the analysis (figure 7.23 B). A maximum is now apparent at 385 nm, strongly indicating the presence of a $[4Fe\text{-}4S]^{2+}$ protein. This probably means that FF fraction 2 did in fact contain reduced protein (containing a $[4Fe\text{-}4S]^{1+}$ cluster) just like Source fraction 2. FF fraction 1 and Source fraction 1 are comparable. However, FF fraction 1 contains the apo-protein eluting prior to Source fraction 1. FF fraction 2 and Source fraction 2 are comparable.

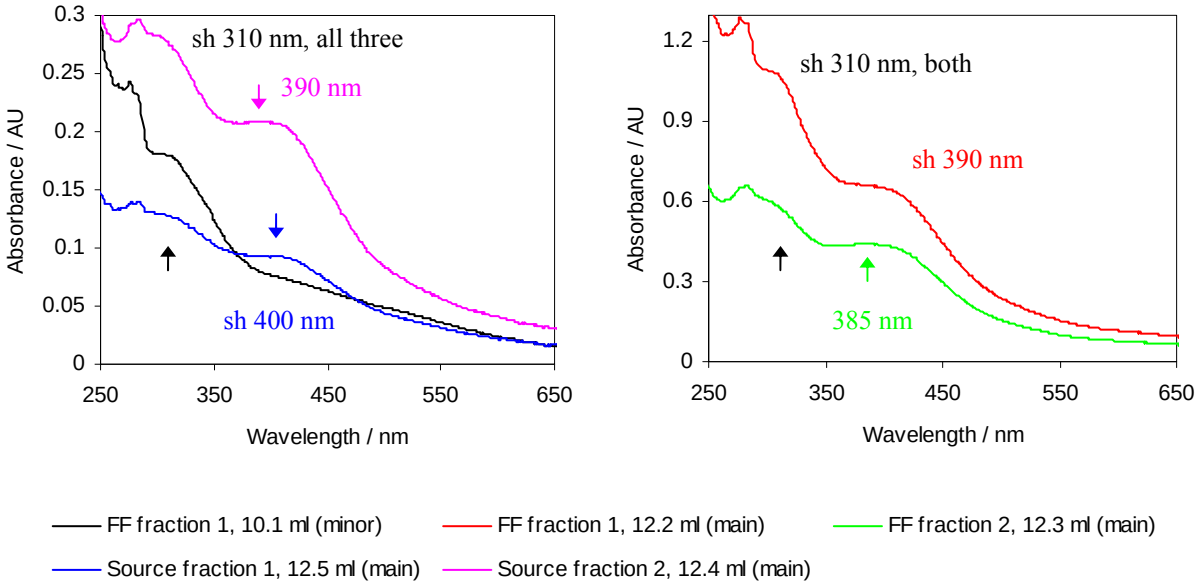


Figure 7.25 UV-vis spectra of fractions corresponding to peaks observed when performing analytical gel filtration in experiment 5. Maxima and shoulders (sh) are indicated.

7.2.5 Fe(II) versus Fe(III) as iron sources and reproducibility

As discussed in sections 7.2.3 and 7.2.4 a great difference in the obtained chromatograms and the products was observed depending on the iron source. When an Fe(III) source was used, the products appeared to contain oxidized cluster forms ($[4Fe\text{-}4S]^{2+}$ and $[3Fe\text{-}4S]^{1+}$), and dimeric and monomeric compounds were obtained. The majority of the iron remained in the high oxidation state. When an Fe(II) source was used under the reducing conditions, it is quite difficult to obtain Fe(III), which is necessary for formation of common cluster types (chapter 2). However, the oxygen concentration inside the anaerobic chamber was 6-8 ppm (see section 4.2) which could contribute to formation of Fe(III). The major component obtained when purifying the protein was in fact found to be reduced monomeric holo-protein ($[4Fe\text{-}4S]^{1+}$), which was slowly oxidized. Using Fe(II), some apo-protein was also observed during purification, and it might be necessary to increase the time of reaction to gain higher yield of holo-protein. A super reduced $[4Fe\text{-}4S]^{0}$ protein is highly unlikely to form under the given circumstances and has only been generated from clusters of higher oxidation states (Scott *et al.*, 2005).

Using $FeCl_3$ as the iron source, three experiments were carried out performing cluster reconstitution into Ub-Ori-ISP followed by purification (sections 7.1.1.1 and 7.2.3.1). There were some discrepancies in the sizes of comparable peaks in the chromatograms obtained

from anion-exchange. However, the components found in comparable peaks were believed to be identical based on UV-vis spectrophotometric analysis and analytical gel filtrations. The experiments are therefore to some degree reproducible. However, unknown parameters do affect the ratio of produced species (see section 7.2.1). When using $FeCl_3$ as the iron source, cluster reconstitution into Ori-ISP followed by purification was performed successfully. The experiment was repeated (results not shown) with a highly similar outcome showing that the experiment was reproducible. Only one experiment was performed using Fe(II) for performing cluster reconstitution into Ub-Ori-ISP.

7.2.6 Stability of holo-protein purified by gravity

The stability of holo-protein obtained when $FeCl_3$ was used as iron source and the obtained holo-protein was purified on the Q Sepharose Fast Flow column was investigated (see section 7.1.3) by recording UV-vis spectra over a period of 49 days. As seen from the UV-vis spectra in figure 7.26, the characteristics associated with a $[4Fe\text{-}4S]^{2+}$ protein are preserved throughout the measuring period (shoulder at 310 nm and maximum at 390 nm), indicating that the $[4Fe\text{-}4S]^{2+}$ protein is not degraded significantly. However, the absorption increases throughout the entire spectrum over time (a factor 2 within 49 days). One explanation could be formation of breakdown products increasing the background absorption. Another explanation could be up-concentration of the holo-protein due to evaporation. The sample had an approximate volume of 200 µl and was kept in an eppendorf tube at 4ºC. Since the sample was opened regularly evaporation could at least be part of the explanation.

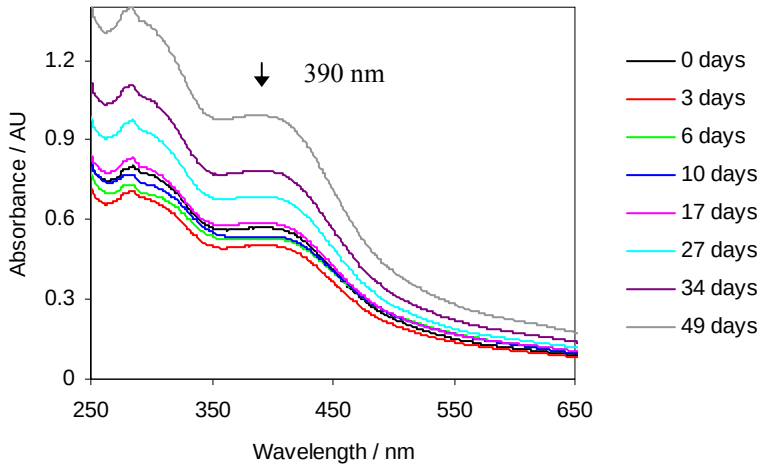


Figure 7.26 Stability of a concentrated sample of holo-protein obtained from purification on Q Sepharose fast Flow column. UV-vis spectra were recorded throughout a time period of 49 days. Prior to recording a spectrum an aliquot was diluted 200 times.

Even though the characteristic features do not disappear, there is no guarantee that conversion between monomer, dimer or polymers and slight breakdown has not happened (see discussion section 7.2.8). It is not known whether the sample is homogeneous consisting of either monomer or dimer to begin with, since it was not analysed by analytical gel filtration.

The stability of the obtained product has however greatly improved in comparison to the results obtained in chapter 6, where the features were only present for a day or two and massive breakdown was apparent.

7.2.7 Molecular weights of species obtained from analytical gel filtrations

When performing analytical gel filtration on holo-protein purified by anion-exchange, main peaks were found to elute at two distinct positions (high and low molecular weight) when $FeCl_3$ was used as an iron source (experiment 1 to 4). Using Fe(II) (experiment 5) produced only species of the lower molecular weight (see section 7.2.4 and table 7.1). The properties of the eluting compounds as monomers, dimers or polymers are discussed in this section.
A calibration curve had not been determined for the analytical gel filtration column used for the experiments carried out in France (experiment 1, 2 and 5, 10/30 Superdex 75 HR column). However, a calibration curve was obtained (appendix A.1) for the column used in our own laboratory (experiment 3 and 4, 13/31 Superdex 75 HR column). Using this curve and the elution points, approximate molecular weights were determined for the eluting species in experiment 3 (see table 7.2). A monomeric or dimeric holo-Ub-Ori-ISP containing one [4Fe-4S] cluster per monomer would have molecular weights of ca. 11.2 kDa and 22.3 kDa, respectively. This indicates that the species which eluted at 11.6 ml and 12.7 ml are most likely to be trimeric and dimeric, respectively. Since no compounds are found to elute at higher volumes than at about 12.7 ml and YUH1 (27.2 kDa) is found to elute at 11.4 ml (see section 7.2.3.2), it is more likely that the species which eluted were dimeric (ca. 11.6 ml) and monomeric (12.7 ml). Elution at lower volumes is considered to be of polymeric forms.

Table 7.2 Calculated molecular weights of species eluting from the 13/31 Superdex 75 HR column

Point of elution (ml)	10.7	11.6	12.7
Molecular weight (kDa)	44.2	31.4	20.0

Since experiment 3 (using the 13/31 Superdex 75 HR column) is a repetition of experiment 1 and 2 (using the 10/30 Superdex 75 HR column) the peaks observed using the 10/30 Superdex 75 HR column were also considered to be of monomeric and dimeric species.
Purified, sulfonated apo-protein has also been analysed on the analytical gel filtration column (section 5.3.2) and the main peak was found to elute at 11.6 ml (just like the dimeric holo-protein) suggesting that the apo-protein is found as a dimer. However, a minor peak is also present at 12.9 ml, suggesting that monomer is also present. Since apo-protein appears to be present as both dimeric and monomeric forms, it is possible that the dimer/monomer forms of the holo-protein might be an effect of the ubiquitin part of the molecule and not the Ori-ISP part. To establish whether this is the case, ubiquitin should be cleaved off. When performing analytical gel filtration on holo-Ori-ISP main peaks at two positions were also found to be present, which proves that two differently sized species were formed and that this happened without the ubiquitin being part of the molecules. This suggests that the formation of monomer/dimer is independent on ubiquitin and the fact that apo-Ub-Ori-ISP most likely is a dimer does not appear to affect formation of holo-protein as a dimer or monomer.
Massive conversion from monomer to dimer and not the other way around has been observed (sections 7.2.3 and 7.2.4). This has been observed with and without ubiquitin as part of the

molecule, which indicates that ubiquitin is not the main factor contributing to formation of dimers or polymers. It is much more likely that the protein takes these conformations as part of a breakdown mechanism (section 7.2.8). The protein originates from di-cluster ferredoxins (chapter 3) and it is very likely that it is this conformation it "tries" to adopt (see section 7.2.9).

7.2.8 Breakdown, cluster interconversion and monomer-dimer conversions

As observed from the gel filtration analyses throughout this chapter, the cluster in the holo-protein appear to either convert or break down to other cluster forms and monomeric holo-protein appear to be able to convert to dimers. It is possible that equilibrium between different forms exists. It is also very likely, that these transformations happen as a consequence of breakdown, since no dimers were observed to transform into monomers to any noteworthy extent.

Breakdown might happen as a series of steps where one cluster type is oxidized to another and so forth. Formation of dimers and polymers might also be a part of this mechanism. This is substantiated by the obtained results (see section 7.2.3 and 7.3.4) since the reduced [4Fe-4S] cluster was found as a monomer while the oxidized form was found primarily as a monomer with a small amount as dimer. At the same time what appeared to contain a $[3Fe\text{-}4S]^{1+}$ cluster (an oxidized form) or a mixture of clusters was found to be a dimer.

The most obvious reason for breakdown is oxygen, since no reducing agents were added to purification buffers and a small amount of oxygen was present in the anaerobic chamber (6-8 ppm in the chamber in France). When leaving a solution shielded from the surroundings this small amount of oxygen might only be slightly damaging. During purification on the other hand, the holo-protein is continuously exposed to fresh buffer containing small amounts of oxygen. This constant exposure might prove harmful to a very oxygen-sensitive protein. This could be why possible breakdown is seen during purification.

Due to the very small size of Ori-ISP (only 23 amino acid residues) the FeS cluster is exposed and the protein moiety provides less protection against solubilised destructive or oxidizing species than in the larger FeS-protein, rendering it quite unstable. The small size probably prevents extensive secondary structure and hydrogen bonding, leaving the chain which is wrapped around the cluster quite flexible.

During purification, interaction with the column material might also be detrimental. When performing anion-exchange it is imaginable that the interaction between the negatively charged protein and the positively charged groups of the media might be enough to rip the cluster and protein apart. This is however not the case when performing gel filtration since no interaction with the media should take place. If the holo-protein is to some extent unstable due to the small size (explained above) the stress of purification might still cause some breakdown.

7.2.9 Dimer conformations

Models of Ori-ISP as a monomer containing a [4Fe-4S] cluster have been made previously, and they are presented in section 3.3.5.

No models are available for holo-Ori-ISP as a dimer. One could however imagine two scenarios for the formation of a dimer containing two clusters; parallel and anti-parallel alignment of the two peptide chains (see figure 7.27).

The parallel alignment would have to contain a disulfide bond for binding the two monomers together since the three remaining cysteines from each monomer would be expected to be coordinated to each their cluster (see figure 7.27 A). In this scenario the most likely cluster type to be coordinated would be a [3Fe-4S] cluster since only three cysteines are available for coordination.

The anti-parallel alignment is a much more likely scenario. This is also the coordination mode that dicluster ferredoxins adopt (section 2.3.2). Tryptic cleaved apo-ferredoxin from *C. acidiurici* (holo-form contains two [4Fe-4S] clusters) was found to form a dimer identical to the native ferredoxin upon *in vitro* cluster reconstitution (evident from EPR spectroscopy, Orme-Johnson, 1973) indicating that the anti-parallel mode is the most likely to form. In the anti-parallel mode, the three first cysteines from monomer 1 and the fourth cysteinee from monomer 2 share in coordinating one cluster. At the same time the three first cysteines from monomer 2 and the fourth cysteine from monomer 1 share in coordinating the second cluster (see figure 7.27 B). This coordination mode would be able to accommodate each of the three common cluster types: [2Fe-2S], [3Fe-4S] and [4Fe-4S]. Accommodating a [3Fe-4S] cluster would mean that one of the cysteines in the sequence would not coordinate the cluster. In di-cluster ferredoxins this would be comparable to the second cysteine in the sequence (section 2.3.2).

Several scenarios for forming a dimeric protein containing only one cluster can be envisaged (not shown). These would have to contain free cysteines and/or disulfide bonds.

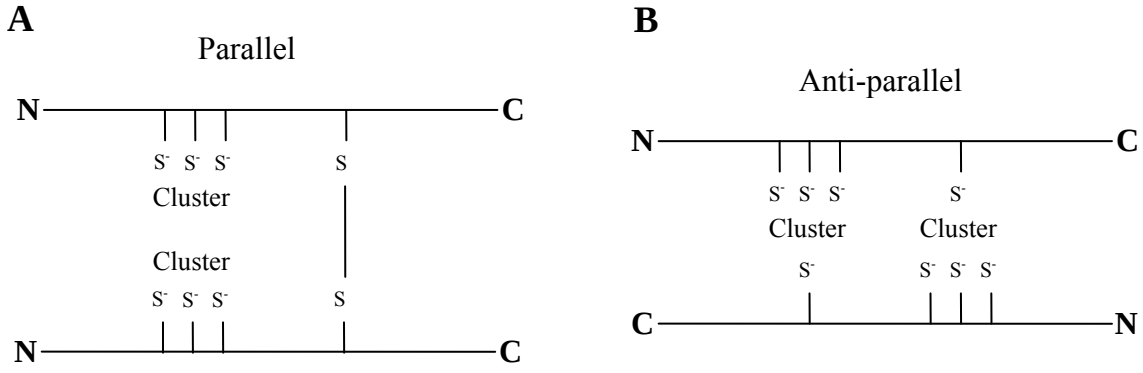


Figure 7.27 Possible dimer conformations of holo-Ori-ISP. A) Example of parallel coordination of two clusters. B) Example of anti-parallel coordination of two clusters.

A crude model for the anti-parallel scenario can be made by aligning two Ori-ISP peptides to the sequence of a dicluster ferredoxin with a known structure and then taking the structure and deleting all residues not superposed by Ori-ISP. Only looking at the backbone and

coordinating cysteines, yields a very rough picture of a possible coordination mode. This was done using the sequence and structure of *C. acidiurici* 2 × [4Fe-4S] ferredoxin (PDB ID: 2FDN, Dauter *et al.*, 1997). The structure was found by performing a blast search (National Center for Biotechnology Information, Altschul *et al.*, 1990) of the Ori-ISP sequence against the PDB and thereafter search for 2 × [4Fe-4S] ferredoxins. The sequence alignment based on cysteines (performed by hand) is shown in figure 7.28 and the modified structure is shown in figure 7.29. From the sequence alignment it is seen that the sequences have 47.8 % sequence identity (calculated based on the 46 aligned residues). From this crude model it can be seen that the two N-terminals point in opposite directions. This means that this model is also plausible when ubiquitin is fused to Ori-ISP since the ubiquitin C-terminal is fused to the Ori-ISP N-terminal.

```
AYVINEACISCGACEPECPVNAISSGDDRYVIDADTCIDCGACAGVCPVDAPVQA
VDVDEEECVGCGACVNVCPVGAV------VDVDEEECVGCGACVNVCPVGAV
```

Figure 7.28 Sequence alignments of two Ori-ISP peptides with the sequence of *C. acidiurici* ferredoxin. The alignment was made by hand using the repeated cluster coordination cysteine pattern. The cysteines are underlined and identical residues are coloured green.

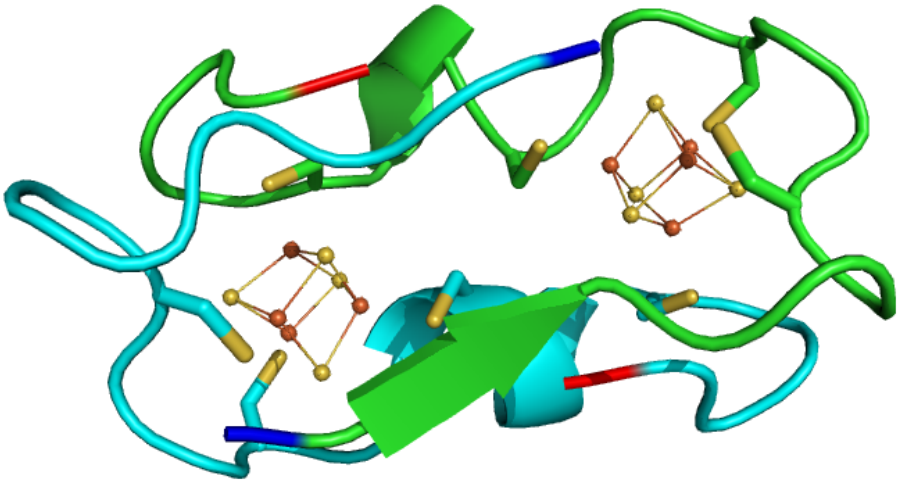


Figure 7.29 Crude model for the anti-parallel dicluster Ori-ISP dimeric protein. The model was generated by first aligning two Ori-ISP peptide sequences with the sequence of *C. acidiurici* 2 × [4Fe-4S] ferredoxin (see figure 7.28). Residues from the ferredoxin not aligned with the Ori-ISP sequences were then deleted in the structure of the ferredoxin (PDB ID: 2FDN, Dauter *et al.*, 1997). Here, residues 1 to 23 (corresponding to the first Ori-ISP monomer) are coloured green while residues 30 to 52 (corresponding to the second Ori-ISP monomer) are coloured cyan. N- and C-terminals of the two segments are coloured blue and red, respectively. The model is shown in ribbon view and only the side chains of the eight cysteines are shown. The two [4Fe-4S] clusters are shown in ball and sticks; sulphur is yellow while iron is orange. The figure was prepared with PyMOL (DeLano W. L., 2002).

7.3 Conclusion

Cluster reconstitution and purification of holo-Ub-Ori-ISP was successfully carried out in anaerobic chambers and better resolved peaks and less tailing was observed when performing HPLC anion-exchange inside anaerobic chambers compared to previously obtained results (chapter 6). Further, a marked increase in stability of the obtained products was observed. The characteristics of holo-protein purified by gravity were found to be preserved for at least 49 days as determined by UV-vis spectrophotometry.

The obtained chromatograms were only reproducible to some degree since great differences in the sizes of comparable peaks were observed. Oxygen was ruled out as the determining factor and other determining factors have been discussed. The products obtained during reconstitution and purification do appear reproducible and the general trend of purification is that when $FeCl_3$ was used as iron source, a monomeric $[4Fe\text{-}4S]^{2+}$ protein was eluted at low conductivities while a dimeric $[3Fe\text{-}4S]^{1+}$ protein or a protein containing two different clusters or multiple cluster types was found to elute at higher conductivities. When Fe(II) was used as the iron source, only monomeric compounds were obtained. $[4Fe\text{-}4S]^{2+}$ protein was found to elute at low conductivity, while the reduced form, containing a $[4Fe\text{-}4S]^{1+}$ cluster, was found to elute at a higher conductivity.

Breakdown, oxidation of holo-protein and monomer to dimer conversions were observed in some cases when performing analytical gel filtration experiments. The primary reasons for this are believed to be the presence of trace amounts of oxygen or interaction with the column material.

A preliminary experiment performing cluster reconstitution into Ori-ISP (the ubiquitin fusion was cleaved) using a slightly modified procedure for cluster reconstitution (using $FeCl_3$ as the iron source) was found to be successful. The obtained products were comparable to those obtained for Ub-Ori-ISP, namely a monomeric $[4Fe\text{-}4S]^{2+}$ protein and a dimeric protein containing $[3Fe\text{-}4S]^{1+}$ clusters or different cluster types.

Chapter eight

8 Initial crystallization trials of holo-Ub-Ori-ISP

This chapter describes initial crystallization trials of holo-Ub-Ori-ISP. The experiments were carried out in the anaerobic chamber described in section 4.2.1.2.

In the previous chapter, holo-protein eluting in Source fraction 1 and FF fraction 1 (experiment 1 and 2, section 7.2.3.1) was found to contain a $[4Fe\text{-}4S]^{2+}$ cluster. However, both fraction types could be highly homogeneous (containing mostly monomer) or in-homogeneous (a mixture of monomer and dimer). Even though a completely homogeneous sample could not be attained, it was decided to set up initial crystallization trials. $FeCl_3$ was used as iron source and holo-protein was purified by gravity, since this purification step is less time consuming than performing HPLC purification. Protein purified by gravity was found to be quite stable; the $[4Fe\text{-}4S]^{2+}$cluster features were still present after 49 days (section 7.2.6).

8.1 Experimental

Cluster reconstitution was performed as described (section 6.1.1) using 34.0 mg freeze-dried protein (section 7.1.1.1). The protein was desalted and purified by anion-exchange (sections 6.1.1 and 6.1.2). The obtained product was up-concentrated (Microcon YM-3 centrifugal filter) and stored at 4ºC. Using holo-protein concentrated to 75 mg/ml (concentration determination, section 6.2.2), initial screening was performed with Crystal Screen I and II (Hampton Research) and the pH Clear Suite (Qiagen). 96-well plates were set up with reservoir volumes of 100 µl and sitting drops of 1 µl protein solution and 1 µl reservoir solution. The vapour diffusion method was applied and plates were stored at 20ºC.

From a second batch of holo-protein (100 mg/ml) experiments were carried out to optimize promising conditions using the hanging drop vapour diffusion method (appendix A.2).

8.2 Results

After setting up Crystal Screen I and II approximately half of the drops had lost their colour within a few days. In most of the remaining drops the colour turned blackish brown within the same period of time. No precipitation or crystalline material was observed in any of the coloured drops and no general trends could be inferred from the results.

When setting up the pH Clear Suite, no precipitation was observed when using NaCl as precipitant. However, drops lost their colour at pH 4 and 5 which was also the case when PEG 6000, MPD (2-methyl-2,4-pentanediol) and ammonium sulphate (0.8 and 1.6 M) was used. At high concentrations (20 or 30 %) PEG 6000 resulted in phase separation above pH 6 which further turned into black precipitation. 2.4 M ammonium sulphate resulted in blackish brown precipitation which turned grey within one week. 3.2 M ammonium sulphate resulted in precipitation at all examined pH values (4 to 9); the lower the pH the more blackish was the

brown precipitation. The precipitation observed was fluffy looking and not appearing to have "sharp edges". 65 % MPD resulted in phase separation above pH 6.
Using an increased holo-protein concentration, trials to optimise the ammonium sulphate conditions (appendix A.2) only resulted in clear drops. Trials to optimise the MPD conditions resulted in phase separation when 60 % MPD was used and fluffy precipitate when salts were added. A home made version of The Solubility Tool Kit (originally manufactured by Molecular Dimensions) (appendix A.2) only resulted in clear drops.

8.3 Discussion

No crystalline precipitation or crystals were obtained when using commercially available screens. The protein was found to break down (loss of colour and formation of black iron sulfide species) under most of the analysed conditions and when exposed to solutions having a pH of 5 or below. The only leads to follow (however not particularly convincing) were fluffy precipitation when using ammonium sulphate as precipitant or phase separation using MPD as precipitant. Trials to optimize these conditions did not yield any usable results.
The sensitivity towards pH is expected since it is a general feature of FeS-proteins that the inorganic sulfide present in clusters is acid labile (section 2.1).
The protein concentration used for setting up the experiments was quite high (75 mg/ml and 100 mg/ml) and a too low protein concentration is therefore probably not the reason for the lack of coloured precipitation. One reason for not obtaining any usable results is breakdown of holo-protein when exposed to oxygen or precipitants inducing breakdown. The protein solution is probably not homogeneous (as discussed in chapter 7) and conversions between monomer and dimer forms or between different cluster types might also be problematic (discussed in section 7.2.8) when trying to crystallize the protein. If such conversions indeed do take place it is likely that they are dependent on protein concentration, salt concentration or other compounds present in solution and equilibria might shift as a consequence of the applied conditions for crystallization. Adventitiously bound iron (section 6.3.3) or iron sulfide (section 6.3.1) might also be a problem since these compounds are probably bound to the surface of the protein. These compounds may be difficult to remove or avoid since the FeS cluster was assembled in the protein by inorganic synthesis. It was previously seen that the sulfonated apo-protein could be crystallized using ammonium citrate (section 5.3.3). The [4Fe-4S] cluster present in holo-protein was found to convert to another cluster type (see section 6.3.1.1) when exposed to citrate, rendering the conditions found for crystallizing apo-protein unsuitable for crystallization of the holo-protein.
It may prove very difficult to crystallize the holo-protein. It could be advantageous to cleave off the Ub-fusion before further attempting to crystallize the holo-protein.

8.4 Conclusion

It was attempted to crystallize holo-Ub-Ori-ISP however, no crystals were obtained. This is believed to be due to either breakdown of holo-protein, an inhomogeneous sample or presence of adventitiously bound iron.

Chapter nine

9 Purification of YUH1 and proteolytic cleavage of ubiquitin fusions

This chapter describes the final procedures for cultivation of YUH1 expressing cells and purification of YUH1. The final optimization experiment on proteolytic cleavage of sulfonated Ub-Ori-ISP is also presented along with a preliminary experiment on Ub-LCA ferredoxin cleavage.

YUH1 (section 4.6) has previously been expressed and purified (Kohno *et al.*, 1998, Liu *et al.*, 1989; Miller *et al.*, 1989), however, it was decided to create new expression and purification protocols suitable for equipment present in our laboratories.

Optimization of the cultivation procedure was performed by Laboratory Assistant, Stephanie Boy and Miriam G. Pedersen (project student) under my supervision, and it was based on in-house procedures for cultivation of non-metalloenzymes. Initial purification trials were performed by Miriam G. Pedersen under my supervision.

9.1 Considerations prior to development of the purification procedure

It was decided to purify the enzyme at pH 7.0 since the enzyme contains a free cysteine in the active site. At this pH value the cysteine is most likely protonated (pKa 8.3) and less prone to oxidation by oxygen. Therefore thiol-containing reducing agents and EDTA were not added to purification buffers as described elsewhere (Kohno *et al.*, 1998; Liu *et al.*, 1989; Miller *et al.*, 1989). Ubiquitin hydrolases purified in the absence of thiol-containing reducing agents have been found to be inactive. They can be activated by addition of e.g. DTT, although full activity has not been obtained (Duerksen-Huges *et al.*, 1989). The enzyme should not be frozen during the purification procedure since repeated freeze-thaw cycles have been found to render ubiquitin hydrolases inactive. Activity loss has also been observed when the cleavage reaction proceeds at 25ºC compared to 37ºC (Duerksen-Huges *et al.*, 1989).

Even though the enzyme is expressed as His-tagged (YUH1-His_6) the strategy was, that purification should be attempted by omitting purification on a His-binding column. It is a general experience of mine, that His-binding column materials that have been in contact with protein/enzyme extract (obtained from cells lysis) are tedious to clean and regenerate, while this is easily accomplished with anion-exchange materials.

9.2 Considerations prior to performing proteolytic cleavage of Ub-Ori-ISP

As described in section 7.1.1.2 proteolytic cleavage of sulfonated Ub-Ori-ISP and de-protection of the sulfite-protected thiolate groups were intended to take place simultaneously. Since the de-protection reaction has already been optimized (Christophersen, 2004) the amount of YUH1 needed to perform proteolytic cleavage was tested under the following

conditions; 6 mg/ml Ub-Ori-ISP, 100 mM Tris/HCl pH 8.5, 10 mM DTT and room temperature.
No record of a pH optimum for the proteolytic cleavage of Ub-fusions by YUH1 has been found however cleavage reactions at pH 8.0 and 7.5 have been performed (Kohno *et al.*, 1998; Liu *et al.*, 1989). pH optima for other ubiquitin hydrolases between pH 7.5 and 8.5 have been reported (data obtained from the Braunschweig Enzyme Database, Barthelmes *et al.*, 2007).

9.3 Experimental

9.3.1 Cultivation of YUH1 expressing cells

This procedure is for cultivation of 6 times 650 ml cell cultures. Cells from the HC1434 glycerol stock were plated on an agar plate containing LB/30 Kan and incubated overnight (16 hours) at 30ºC. A single colony was used to inoculate 50 ml LB/30 Kan (25 g/L Luria Broth, 30 μg/ml kanamycin sulfate) poured into a 300 ml triple baffled shake flask and incubated at 37ºC with shaking (250 rpm) until an OD_{600} of 0.6 was obtained. The pre-culture was stored at 4ºC overnight. The cells were sedimented (centrifuged at 4ºC and 2300 g for 10 min), the supernatant discarded and cells were re-suspended in 50 ml fresh LB/30 Kan. 6.5 ml were transferred to each of six 2 L triple baffled shake flasks each containing 650 ml TB/30 Kan. The cultures were incubated at 30ºC with shaking (250 rpm) until the OD_{600} was 0.6. Then protein expression was induced by addition of IPTG (200 mM stock solution) to a final concentration of 0.1 mM, and the incubation was continued with shaking (250 rpm) overnight (16 hours) at 30ºC. The cells were harvested by centrifugation at 3000 g and 4ºC for 15 min. The supernatants were discarded and each of the six portions of cells were re-suspended in 25 ml cold (4ºC) 20 mM Bis-Tris propane/HCl pH 7.0 (1,3-bis(tris-(hydroxymethyl)methylamino)propane, pH adjusted with 1 M HCl) and transferred to 50 ml polypropylene tubes. The cells were sedimented by centrifugation (3000 g, 4ºC, 15min) and the supernatants discarded. The cells were kept at -80ºC for storage.

9.3.2 Purification of YUH1

2 portions of harvested cells (cells from 1.3 litres of culture, total volume of cells, ca. 8 ml) were thawed at 4ºC and re-suspended in cold (4ºC) 20 mM Bis-Tris propane/HCl pH 7.0 to a final volume of 80 ml. The cells were lysed by sonication (Satorius Labsonic P at 80 % amplitude) in three intervals of 30 seconds. The solutions were kept on ice at all times and allowed to cool between each sonication burst. The crude lysate was centrifuged at 24800 g for 40 minutes at 4ºC. The enzyme extract (supernatant) was decanted, and the pellet discarded.
The enzyme was then purified by HPLC (ÄKTA^{TM} purifier 100) in a three-step purification procedure (anion-exchange → gel filtration → anion-exchange). The enzyme was kept on ice or at 4ºC between each purification step, which were all performed at room temperature. The extract was loaded onto a 20 mM Bis-Tris propane/HCl pH 7.0 equilibrated HiPrep 16/10 ANX Sepharose Fast Flow column and eluted (linear flow rate: 0.8 cm/min) using a linear gradient, 0 M to 0.6 M NaCl in 15 CV, and fractions were collected. The buffer was changed to 20 mM Bis-Tris propane/HCl, 0.15 M NaCl pH 7.0 and the volume reduced to ca. 15 ml by

ultrafiltration using a stirred Amicon cell with a YM3 membrane. The enzyme sample was filtered (0.45 μm) and loaded onto a 20 mM Bis-Tris propane/HCl, 0.15 M NaCl pH 7.0 equilibrated HiLoad 16/60 Superdex 75 prep grade column (a maximum of 4 ml was loaded for each run) and fractions were collected upon elution with the same buffer (linear flow rate: 0.2 cm/min). The enzyme was finally purified on a 20 mM Bis-Tris propane/HCl, 0.05 M NaCl pH 7.0 equilibrated 16/10 Source 30Q column. Fractions collected from the gel filtration step were filtered (0.45 μm) and diluted three times with 20 mM Bis-Tris propane/HCl pH 7.0 before loading onto the column. The protein was eluted (linear flow rate: 1.2 cm/min) using a linear gradient, 0.15 M to 0.3 M NaCl in 10 CV, and fractions were collected. The buffer was changed to 20 mM Bis-Tris propane/HCl, pH 7.0 and the volume reduced to ca. 4 ml by ultrafiltration using a stirred Amicon cell with a YM3 membrane. The enzyme solution was diluted with glycerol to a final concentration of 50 % (w/v) for storage at -20ºC (the final buffer concentration was ca. 10 mM). The enzyme concentration and yield was determined by UV-visible spectrophotometric analysis (the molecular weight of the His_6-tagged enzyme was calculated to 27.2 kDa and the molar absorptivity was calculated to 28 $mM^{-1} \cdot cm^{-1}$, (ExPASy Proteomics Server; Gasteiger *et al.*, 2005)) and the purity was tested by SDS-PAGE.

9.3.3 Enzymatic cleavage of sulfonated Ub-Ori-ISP

The progress of enzymatic cleavage of sulfonated Ub-Ori-ISP under the given conditions (section 9.2) was tested at two different concentrations of YUH1. The molar ratios of Ub-Ori-ISP:YUH1 tested were 20:1 and 50:1. Two solutions were prepared and both solutions contained 6 mg/ml sulfonated Ub-Ori-ISP and 10 mM DTT dissolved in 100 mM Tris/HCl pH 8.5. The first solution furthermore contained 0.7 mg/ml YUH1, and the second solution contained 0.3 mg/ml YUH1. The cleavage experiments were performed at room temperature using Ar-flushing for obtaining anaerobic conditions (section 4.1) and progress was followed by taking out samples for SDS-PAGE analysis after 1, 2, 4 and 22 hours.

9.3.4 Enzymatic cleavage of Ub-LCA ferredoxin

A preliminary experiment on cleavage of Ub-LCA ferredoxin using YUH1 was performed. The progress of cleavage was tested at two different concentrations of YUH1. The molar ratios of Ub-LCA ferredoxin:YUH1 tested were 1:1 and 5:1. Two solutions were prepared, which contained 1.5 mg/ml purified Ub-LCA ferredoxin (chapter 10) and 4 mM DTT dissolved in 100 mM Tris/HCl pH 8.5 and 2.6 mg/ml or 0.6 mg/ml YUH1, respectively. The cleavage experiments were performed as described in section 9.3.3.

9.4 Results and discussion

9.4.1 Purification of YUH1

The enzyme was purified using a three-step purification procedure (anion-exchange → gel filtration → anion-exchange, section 9.3.2). Enzyme extract from two portions of cells was loaded onto the HiPrep16/10 ANX Sepharose Fast Flow column and purified by anion-exchange in the first step (figure 9.1). In the second step gel filtration (on a HiLoad 16/60

Superdex 75 prep grade column) was performed (figure 9.2). 4 runs with a maximum loading of 4 ml in each run were necessary for purification of protein from two portions of cells, otherwise the column was overloaded and the viscosity of the sample too high for optimal separation. The protein was finally purified by high-resolution anion-exchange on a 16/10 Source 30Q column (figure 9.3). Again, 4 runs were necessary to ensure optimal conditions.

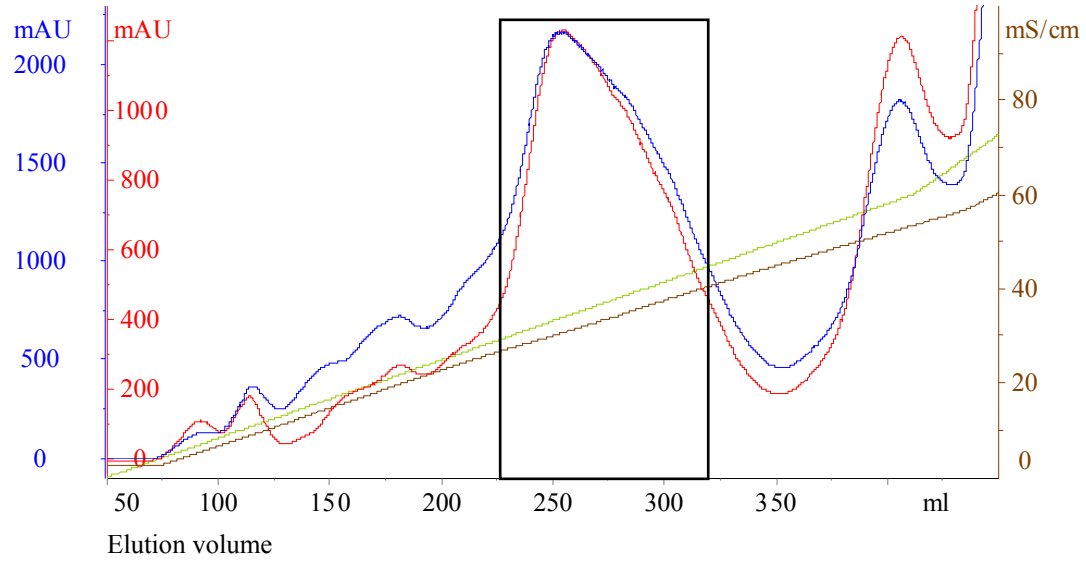


Figure 9.1 Chromatogram for purification of YUH1 on the HiPrep 16/10 ANX Sepharose Fast Flow column. The collected fraction is indicated with a black box. The blue and the red lines indicate the absorbance at 236 nm and 280 nm, respectively. The brown and the green lines indicate the conductivity and the NaCl gradient, respectively. The concentration of NaCl can be read from the conductivity axis employing the unit 10^{-2} M.

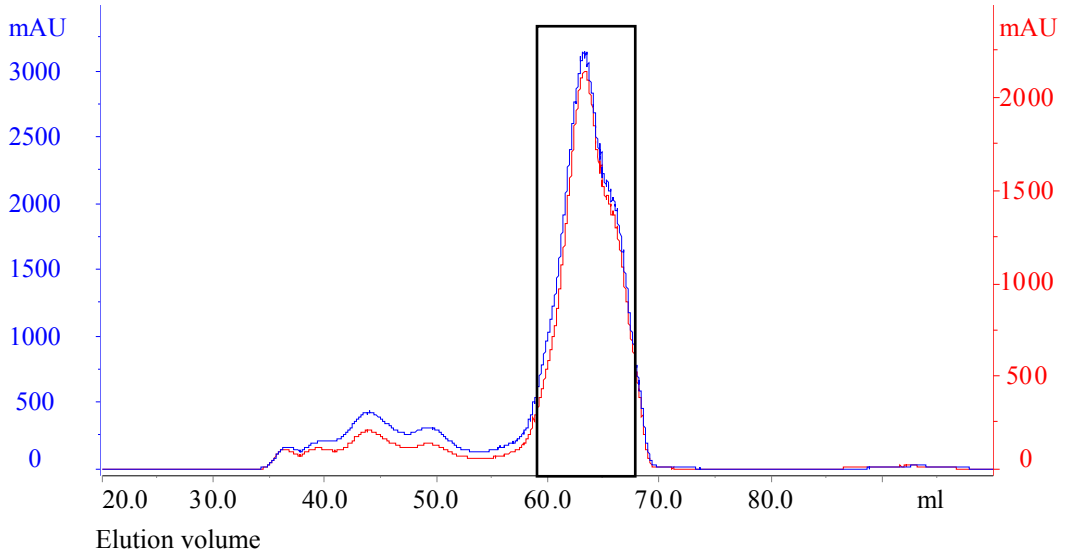


Figure 9.2 Chromatogram for purification of YUH1 on the HiLoad 16/60 Superdex 75 prep grade column. The collected fraction is indicated with a black box. The blue and the red lines indicate the absorbance at 236 nm and 280 nm, respectively.

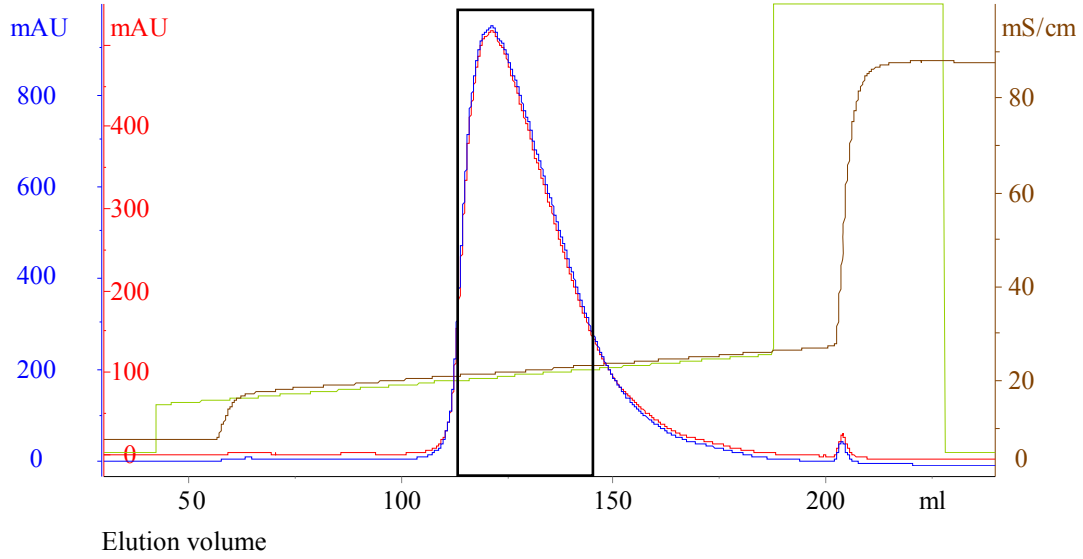


Figure 9.3 Chromatogram for purification of YUH1 on the 16/10 Source 30Q column. The collected fraction is indicated with a black box. The blue and the red lines indicate the absorbance at 236 nm and 280 nm, respectively. The brown and the green lines indicate the conductivity and the NaCl gradient, respectively. The concentration of NaCl can be read from the conductivity axis employing the unit 10^{-2} M.

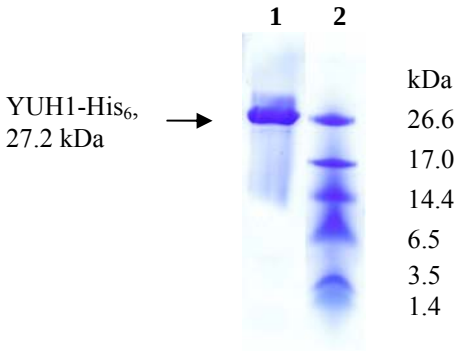


Figure 9.4 Purity of YUH1 tested by SDS-PAGE. Lane 1: Purified YUH1, 1 µl of a 17.6 mg/ml solution. Lane 2: Marker, 1 µl.

The purity of the product was tested by SDS-PAGE (figure 9.4) and only one band was observed, which indicates a highly pure product. The yield of enzyme was calculated to 150 mg/L cell culture, which is very high compared to yields obtained when purifying other proteins/enzymes in our laboratories (see chapters 5 and 10).

The optimized purification procedure is a “small scale” procedure that takes three working days. Bearing in mind the high yield, the procedure should not be scaled for larger columns or other column materials until the effects of freezing and long-term storage on enzyme activity and stability have been examined. The effect of adding a reducing agent during purification on enzyme activity and stability should also be investigated.

9.4.2 Enzymatic cleavage of sulfonated Ub-Ori-ISP

The SDS-PAGE results obtained from proteolytic cleavage of sulfonated Ub-Ori-ISP is shown in figure 9.5. Cleavage reaction was performed with 20:1 and with 50:1 molar ratio of Ub-Ori-ISP:YUH1. It was observed that a ratio of 20:1 resulted in complete cleavage after 22 hours (only the Ub band is visible, figure 9.5 lane 6). Cleavage was still incomplete after 4 hours (Ub-ori-ISP and Ub bands visible, figure 9.5 lane 5). A ratio of 50:1 gave incomplete cleavage (figure 9.5 lane 7 to 10). However, only a very weak Ub-Ori-ISP band was left after 22 hours (lane 10).

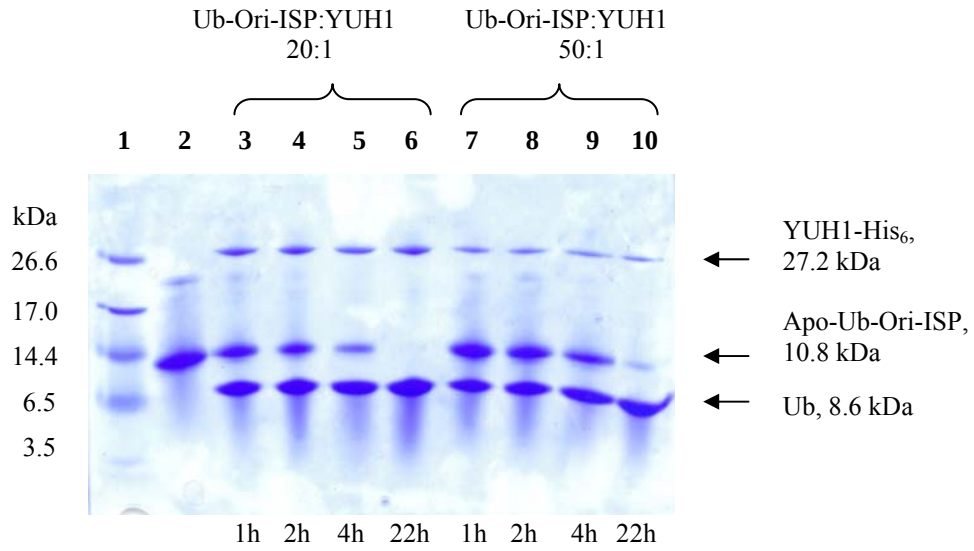


Figure 9.5 SDS-PAGE results showing the extent of Ub-Ori-ISP cleavage by YUH1. Lane 1: Marker, 1 µl. Lane 2: 4 µl of a solution containing 6 mg/ml un-cleaved sulfonated Ub-Ori-ISP. Lanes 3 to 6 show the extent of cleavage when the Ub-Ori-ISP:YUH1 molar ratio was 20:1. 4 µl samples were analysed 1, 2, 4 and 22 hours after the reaction was initiated. Lanes 7 to 10 show the extent of cleavage when the Ub-Ori-ISP:YUH1 molar ratio was 50:1. 4 µl samples were analysed 1, 2, 4 and 22 hours after the reaction was initiated.

Therefore a ratio of 20:1 and a cleavage and de-protection period of 22 hours should be used prior to performing cluster reconstitution into Ori-ISP (see section 7.1.1.2). In the cluster reconstitution procedure (chapters 6 and 7) the time period for de-protection of the sulfonated thiolate groups is 1.5 hours. Increasing this period might affect the outcome of cluster reconstitution but, the reaction is supposed to take place in a closed vessel under reducing conditions (DTT) inside an anaerobic chamber. Therefore it is unlikely that the free cysteines will become oxidised. It might however prove necessary to add extra DTT just prior cluster reconstitution. If a period of 1.5 hours is optimal for the de-protection reaction, then more YUH1 or a higher reaction temperature (see section 9.1) is necessary for the cleavage reaction. The effect of temperature on de-protection has not been investigated. The amount of YUH1 added should be kept as low as possible since it is not known how the presence of this enzyme affects the cluster reconstitution reaction. As YUH1 contains a His_6-tag it may act as an iron scavenger.

9.4.3 Enzymatic cleavage of Ub-LCA ferredoxin

Since a pH of 8.5 had proven successful in cleavage of Ub-Ori-ISP (section 9.3.3), the same pH was applied in the preliminary experiment of Ub-LCA ferredoxin cleavage (section 9.3.4). The DTT concentration was lowered from 10 mM to 4 mM in the reaction mixture since the effect of a high DTT concentration on the iron-sulfur clusters in LCA ferredoxin is not known. A lower protein concentration was also applied. Cleavage reaction was performed when the molar ratio of Ub-LCA ferredoxin:YUH1 was 1:1 and 5:1 and the SDS-PAGE results are presented in figure 9.6. The extend of cleavage was followed by the presence of Ub in the reaction mixture, since LCA ferredoxin is not visible from SDS-PAGE (chapter 10). With a ratio of 5:1 almost no cleavage was observed after 22 hours (only a very faint Ub band is visible, figure 9.6 A lane 3). When a ratio of 1:1 was used about 20 % was cleaved as judged by the intensity of the bands on the gel (Ub and Ub-LCA ferredoxin bands visible, figure 9.6 B, lanes 3 to 5). It appears that cleavage has stagnated after 2 hours since the intensity of comparable bands were about the same throughout the time period (2, 4 and 22 hours). Stagnation and incomplete cleavage of Ub-fusions using YUH1 have previously been

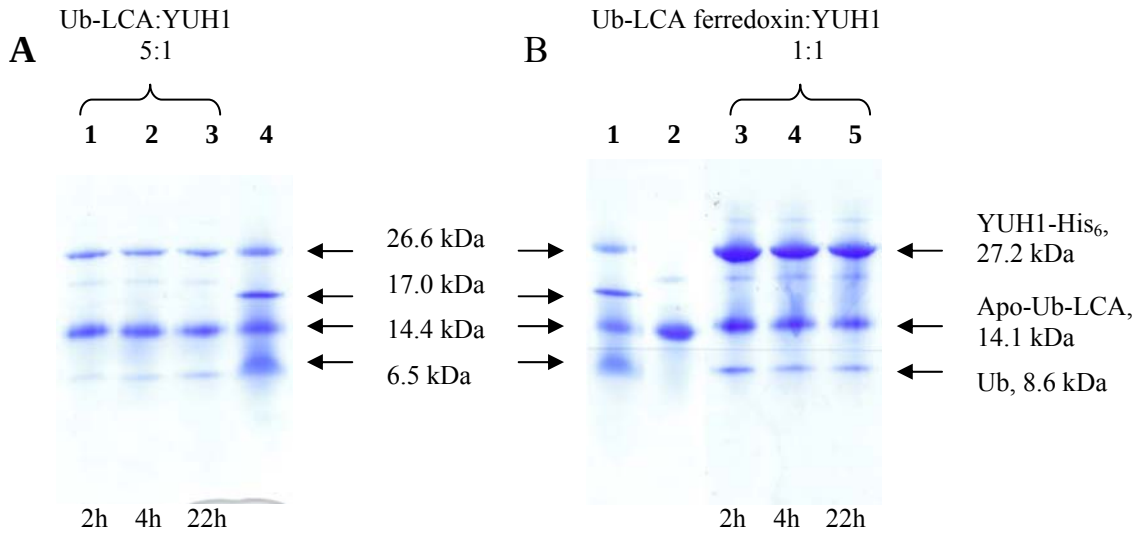


Figure 9.6 SDS-PAGE results showing the extent of Ub-LCA ferredoxin cleavage by YUH1. A) Shows the extent of cleavage when the Ub-LCA ferredoxin:YUH1 molar ratio was 5:1. Lanes 1 to 3: 4 µl samples were analysed 2, 4 and 22 hours after the reaction was initiated. Lane 4: Marker, 1 µl. B) Shows the extent of cleavage when the Ub-LCA ferredoxin:YUH1 molar ratio was 1:1. Lane 1: Marker, 1 µl. Lane 2: 4 µl of a solution containing 1.5 mg/ml un-cleaved Ub-LCA ferredoxin. Lanes 3 to 5: 4 µl samples were analysed 2, 4 and 22 hours after the reaction was initiated.

observed (Ub was fused to a peptide containing 38 amino acid residues) and was attributed to the presence of un-cleavable conformations of the Ub-fusion (Lui *et al.*, 1989). YUH1 has also been found to be nearly inactive with larger fusions such as Ub-β-galactosidase (Miller *et al.*, 1989). The high ratios (Ub-fusion:YUH1) used here have previously been necessary for complete cleavage of other Ub-fusions (Kohno *et al.*, 1998). Therefore, the low percentage of cleavage using a high ratio is not surprising. It is however interesting that Ub-Ori-ISP was almost completely cleaved after 22 hours using a ratio of 50:1 (section 9.4.2). The most

obvious reason for this difference is that the size of the protein fused to ubiquitin plays a major role in the progress of the reaction. If the fusion is too big, then the reaction is hindered. It may be necessary to employ a UBP (see section 4.6) to obtain complete cleavage. The concentrations of DTT, the Ub-fusion and YUH1 may also be important for the progress of the reaction. It can be anticipated that a higher percentage of cleavage may be accomplished by raising the reaction temperature to 37ºC, lowering the pH to e.g. 8.0 or 7.5 (see section 9.1) or raising the concentration of the reactants. The effect of raising the temperature on LCA ferredoxin stability is not known.

9.5 Conclusion

A procedure for expression of YUH1 in *E. coli* cells has been prepared, and an optimized three-step small scale procedure for purification of YUH1 has been achieved. The final yield of purified enzyme was ca. 150 mg/L cell culture.

Complete cleavage of sulfonated Ub-Ori-ISP using a Ub-Ori-ISP:YUH1 molar ration of 20:1 was accomplished in a time period of 22 hours under the conditions used.

Preliminary cleavage experiments were performed on Ub-LCA ferredoxin, however, the progress of cleavage was only about 20 % when a Ub-LCA ferredoxin:YUH1 molar ratio of 1:1 was applied under the conditions used and hence the procedure should be further optimized.

Chapter ten

10 Purification of Ub-LCA and LCA ferredoxin

This chapter presents the cultivation of *E. coli* cells expressing Ub-LCA or LCA ferredoxin and small scale purification procedures for both proteins. A large scale purification procedure for Ub-LCA ferredoxin and experiments concerning cleavage of Ub-LCA ferredoxin on a large scale and purification of the reaction products are also presented.

Two approaches for obtaining LCA ferredoxin are exploited in this chapter. One is to obtain Ub-LCA ferredoxin and then cleave off ubiquitin using the Ub-protease, YUH1. The other is to obtain LCA ferredoxin directly from cells co-expressing Ub-LCA ferredoxin and UBP1. The co-expression is hoped to yield LCA ferredoxin by proteolytic cleavage (ection 4.6). Strains expressing Ub-LCA ferredoxin were used since high yields are obtainable (section 4.5).

To perform characterization of LCA ferredoxin it is necessary at least at some point to obtain the ferredoxin without the Ub-fusion. Preliminary characterization can however be performed on the Ub-LCA ferredoxin since most methods for characterization of the protein and the FeS cluster(s) are assumed to be more or less independent on ubiquitin (UV-visible spectrophotometry, EPR, Mössbauer and electrochemical measurements). Since it was not possible to obtain suitable crystals on the apo-Ub-Ori-ISP fusion protein (section 5.3.3) it may also prove difficult to obtain diffracting crystals on the Ub-LCA ferredoxin fusion protein. Therefore a method to obtain LCA ferredoxin is sought.

The sequence of LCA ferredoxin was obtained from phylogenetic analyses of dicluster ferredoxins and it contains the cysteines necessary for coordination of two FeS clusters (see section 3.3.2). It is hoped that holo-LCA ferredoxin is obtainable directly from expression in *E. coli* cells (contrary to holo-Ub-Ori-ISP, chapter 5). Obtained holo-LCA and holo-Ub-LCA ferredoxin are further assumed to contain two FeS clusters.

10.1 Experimental

All purifications and cleavage experiments were performed anaerobically using either inert gas flushing or an anaerobic chamber (see sections 4.1 and 4.2.2).

10.1.1 Cultivation of cells

The cultivation protocol presented in this section was based on strategies previously developed in our laboratories for cultivation of cells expressing FeS proteins (Caspersen *et al.*, 2000). The two strains used are HC1269 (selection marker: 30 μg/ml Kan), which expresses Ub-LCA ferredoxin, and HC1419 (selection markers: 30 μg/ml Kan and 20 μg/ml Cam), which co-expresses Ub-LCA ferredoxin and UBP1 (see section 4.7).

The following procedure is for cultivation of 6 times 650 ml cell cultures using either HC1269 or HC1419. Cells from the glycerol stock were plated on an LB agar plate containing the appropriate selection marker and incubated overnight at 37ºC. Two single colonies were used to each inoculate 50 ml TB media containing the appropriate selection marker. Incubation with shaking (250 rpm) in 300 ml triple baffled shake flasks was performed at 37ºC until an OD_{600} of 0.6 to 1.0 was obtained. The pre-cultures were stored at 4ºC overnight. The cells of the two pre-cultures were then sedimented (centrifuged at 4ºC and 2300 g for 10 min) and supernatants discarded. The two cell portions were each re-suspended in 23 ml fresh TB media containing the appropriate selection marker. 6.5 ml portions were transferred to six 2 L triple baffled shake flasks each containing 650 ml TB media (and the appropriate selection marker). The cultures were incubated at 30ºC with shaking (250 rpm) until the OD_{600} was 2.4. Protein expression was induced by addition of IPTG (200 mM stock solution) to a final concentration of 0.1 mM, and the incubation was continued with shaking (250 rpm) for 2 hours at 30ºC. The cells were harvested by centrifugation at 3000 g and 4ºC for 15 min. The supernatants were discarded and each of the six portions of cells was re-suspended in 25 ml cold (4ºC) 20 mM Tris/HCl pH 8.0 and transferred to 50 ml polypropylene tubes. The cells were sedimented by centrifugation (3000 g, 4ºC, 15 min) and the supernatants discarded. The cells were kept at -80ºC for storage.

10.1.2 Small scale purification of Ub-LCA and LCA ferredoxin

Protein obtained from the two strains was purified by the following procedure using inert gas flushing for obtaining anaerobic conditions. Protein extract was prepared aerobically.
3 portions of harvested cells (cells from 1.95 litres of culture, total volume of cells was ca. 35 ml) were thawed at 4ºC and re-suspended in cold (4ºC) 20 mM Tris/HCl pH 8.0 to a final volume of 120 ml. $Na_2S_2O_4$ was added to a final concentration of 2 mM (1 M stock solution, freshly prepared). The cells were lysed by sonication (Satorius Labsonic P at 80 % amplitude) in three intervals of 30 seconds. The solutions were kept on ice at all times and allowed to cool between each sonication burst. The crude lysate was centrifuged at 24800 g for 40 minutes at 4ºC. The protein extract (supernatant) was decanted, and the pellet discarded.
The protein was then purified by HPLC (ÄKTA^{TM} purifier 100) in a three-step purification procedure (figure 10.1). The protein was frozen (-20ºC) between each of the three purification steps, which were performed at room temperature. The protein extract was diluted (ca. 3 times) with 20 mM Tris/HCl, 2 mM $Na_2S_2O_4$ pH 8.0 to a conductivity below 5 mS/cm and loaded onto a 20 mM Tris/HCl, 2 mM $Na_2S_2O_4$, 0.05 M NaCl pH 8.0 equilibrated HiLoad 16/10 Q Sepharose High Performance column (linear flow rate: 0.6 cm/min). Protein was eluted using a linear gradient, 0.1 M to 0.6 M NaCl in 12.5 CV, and fractions were collected. For each of the collected fractions the buffer was changed to 20 mM Tris/HCl, 2 mM $Na_2S_2O_4$, 0.15 M NaCl pH 8.0 and the volume reduced to ca. 4 ml by ultrafiltration using a stirred Amicon cell with a YM3 membrane. Each sample was filtered (0.45 µm) and loaded onto a 20 mM Tris/HCl, 2 mM $Na_2S_2O_4$, 0.15 M NaCl pH 8.0 equilibrated HiLoad 16/60 Superdex 75 prep grade column and fractions were collected upon elution with the same buffer (linear flow rate: 0.2 cm/min). Ub-LCA or LCA ferredoxin was finally purified on a 20 mM Tris/HCl, 2 mM $Na_2S_2O_4$, 0.05 M NaCl pH 8.0 equilibrated 16/10 Source 30Q column. Fractions collected from the gel filtration step were filtered (0.45 µm) and diluted

three times with 20 mM Tris/HCl, 2 mM $Na_2S_2O_4$ pH 8.0 before loading onto the column. The protein was eluted (linear flow rate: 1.2 cm/min) using a linear gradient; 0.17 M to 0.35 M NaCl in 12 CV (when purifying Ub-LCA ferredoxin) or 0.22 M to 0.40 M NaCl in 12 CV (when purifying LCA ferredoxin), and fractions were collected. The purity was tested by SDS-PAGE. The yield of protein was determined by UV-vis spectrophotometric analysis by assuming a molar absorptivity of 30.6 $mM^{-1} \cdot cm^{-1}$ at 390 nm (this value is valid for *C. acidiurici* 2 × [4Fe-4S] ferredoxin (Lode *et al.*, 1974)) and using molecular weights of holo-Ub-LCA and holo-LCA ferredoxin calculated to 14.8 kDa (apo-form, 14.1 kDa) and 6.3 kDa (apo-form, 5.6 kDa), respectively.

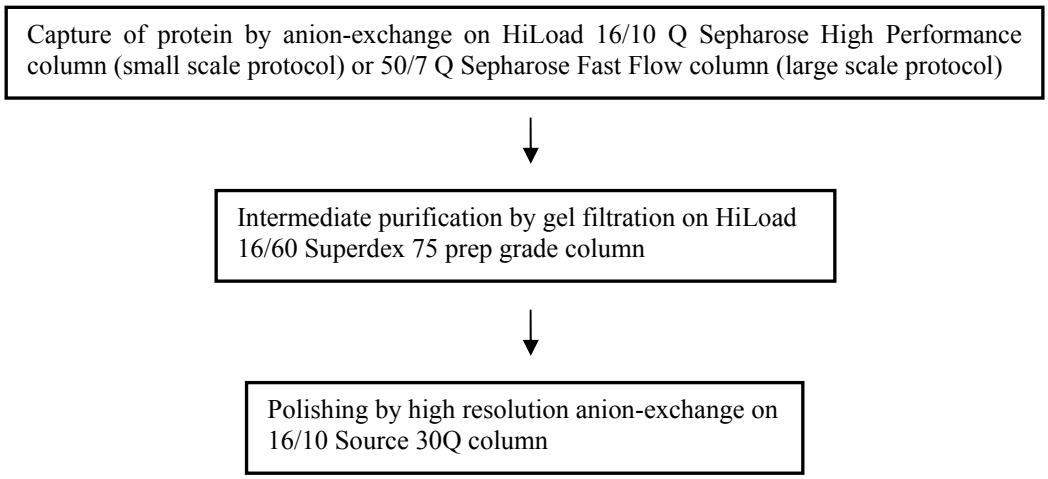


Figure 10.1 Flowchart of the three-step purification procedure of Ub-LCA or LCA ferredoxin.

10.1.3 Large scale purification of Ub-LCA ferredoxin

The small scale purification procedure (section 10.1.2) was scaled up to allow purification of protein from 12 portions of Ub-LCA ferredoxin producing cells (HC1269, from 7.8 litres of culture). Protein extract was prepared as described, but the first purification column was replaced by a 50/7 Q Sepharose Fast Flow column (figure 10.1). The diluted protein extract was loaded onto the 20 mM Tris/HCl, 2 mM $Na_2S_2O_4$, 0.05 M NaCl pH 8.0 equilibrated column (linear flow rate: 0.6 cm/min) and eluted using a linear gradient, 0.15 M to 0.45 M NaCl in 5 CV. The volume of the collected fraction was reduced to ca. 15 ml and the protein was further purified by gel filtration (a maximum of 4 ml was loaded for each run) and high resolution anion-exchange (section 10.1.2).

10.1.4 Proteolytic cleavage of Ub-LCA ferredoxin

Experiments described in this section were performed in an anaerobic chamber (section 4.2.2) and HPLC purifications were performed on an ÄKTAprime^{TM} plus system. Two cleavage experiments were performed.

In the first experiment 22.6 mg purified Ub-LCA ferredoxin (in 10 ml 20 mM Tris/HCl, 2 mM $Na_2S_2O_4$ pH 8.0) and 3.0 ml 17.6 mg/ml YUH1 (ca. 10 mM Bis-Tris propane/HCl pH 7.0, 50 % (w/v) glycerol) were mixed and the buffer was changed to 100 mM Tris/HCl pH 8.5 via ultrafiltration using a Vivacell 70 equipped with a 5000 MWCO membrane. The

volume was adjusted to 5.0 ml and DTT was added (500 mM stock solution) to a final concentration of 8 mM. The final Ub-LCA ferredoxin and YUH1 concentrations were 4.4 mg/ml and 13.2 mg/ml, respectively. The molar ratio of Ub-LCA ferredoxin:YUH1 was approximately 1:1. The solution was left with gentle stirring for 22 hours. The buffer was changed to 20 mM Tris/HCl, 20 mM imidazole, 0.5 M NaCl pH 7.5 using a Vivacell 20 equipped with a 3000 MWCO membrane and the volume adjusted to 10 ml. The purification was carried out in 5 runs. In each run 2 ml of the solution was loaded onto a 20 mM Tris/HCl, 20 mM imidazole, 0.5 M NaCl pH 7.5 equilibrated 5 ml HisTrap HP column (flow rate: 2 ml/min), and the flow through and wash out containing LCA ferredoxin and un-cleaved Ub-LCA ferredoxin was collected. YUH1 was eluted using 20 mM Tris/HCl, 0.5 M imidazole, 0.5 M NaCl pH 7.5 and discarded. LCA and Ub-LCA ferredoxin were separated by anion-exchange. The solution containing LCA and Ub-LCA ferredoxin was diluted five times with 20 mM Tris/HCl pH 8.0 and loaded onto a 20 mM Tris/HCl, 0.05 M NaCl pH 8.0 equilibrated 6 ml Resource Q column (flow rate: 4 ml/min). Proteins were eluted using a linear gradient, 0.17 M to 0.5 M NaCl in 22 CV, and fractions were collected.

In the second experiment 14 mg Ub-LCA ferredoxin (in 3.5 ml 20 mM Tris/HCl) and 3.0 ml 17.6 mg/ml YUH1 (ca. 10 mM Bis-Tris propane/HCl pH 7.0, 50 % (w/v) glycerol) were mixed and the buffer was changed to 20 mM Tris/HCl, 0.15 M NaCl pH 8.0 via ultrafiltration using a Vivacell 70 equipped with a 5000 MWCO membrane. The volume was adjusted to 2.5 ml and DTT was added (500 mM stock solution) to a final concentration of 20 mM. The final Ub-LCA ferredoxin and YUH1 concentrations were 5.6 mg/ml and 21.1 mg/ml, respectively. The molar ratio of Ub-LCA ferredoxin:YUH1 was approximately 1:2. The solution was left with gentle stirring for 22 hours. Purification was then performed as described above.

Purified LCA ferredoxin from the two experiments was pooled, concentrated to 14.3 mg/ml and the buffer was changed to 20 mM Tris/HCl pH 8.0 via ultrafiltration using a Vivacell 20 equipped with a 3000 MWCO membrane. The purity was tested by SDS-PAGE.

10.1.5 Solubility tests of LCA ferredoxin

To find the optimal concentration of LCA ferredoxin necessary for crystallization, The Solubility Tool Kit (Molecular Dimensions) was set up. This was performed in the anaerobic chamber (section 4.2.2) at room temperature. For each precipitant available, only the lowest and the highest concentrations were tested. When testing salts as precipitants, the pH was predetermined (pH 4.5 and 9.0) and when testing PEGs and MPD a pH of 5.5 was applied. The hanging drop vapour diffusion method was applied and each drop was equilibrated over 500 µl of reservoir solution. Drops were set up using 1 µl of protein solution (14.3 mg/ml in 20 mM Tris/HCl pH 8.0) and 1 µl of reservoir solution.

10.2 Results

10.2.1 Expression and small scale purifications of LCA and Ub-LCA ferredoxin

Independent on the strain (HC1269 or HC1419) used, the cells obtained from the cultivation (section 10.1.1) had a greenish brow colour and ca. 10 ml to 15 ml cell paste was obtained from 650 ml culture. The protein extract obtained from cell lysis had a yellowish brown colour. When performing SDS-PAGE on cells or cell extract (results not shown) no distinct bands were observed for either strain.

10.2.1.1 Purification of Ub-LCA ferredoxin

When the protein extract (obtained from HC1269) containing Ub-LCA ferredoxin was loaded onto the 16/10 Q Sepharose High Performance column a greenish brown band appeared close to the top of the column. Upon elution the brown band moved down the column and was collected in one fraction (figure 10.2). The maximum absorption of the peak corresponding to the brown band was observed at a conductivity of 22.9 mS/cm (measured at 390 nm). After elution the column had a greyish colour, and it was cleaned with 1 M acetic acid, 2 M NaCl and 1 M NaOH. The brown coloured protein was further purified by gel filtration (two runs as described in section 10.1.2) on the HiLoad 16/60 Superdex 75 prep grade column. Upon elution one brown band was observed moving down the column and collected in one fraction (figure 10.3). The maximum absorption of the peak corresponding to the brown band was observed at 72.0 ml (measured at 390 nm). The brown coloured protein was finally purified by high resolution anion-exchange on the 16/10 Source 30Q column. Again, one brown band

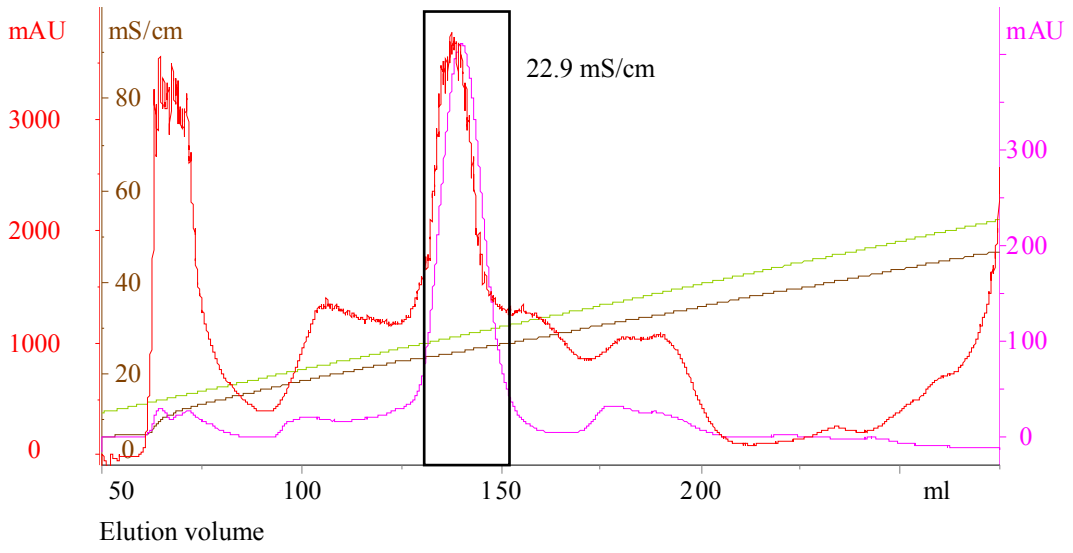


Figure 10.2 Chromatogram for purification of Ub-LCA ferredoxin obtained from HC1269 on the 16/10 Q Sepharose High Performance column. The collected fraction is indicated with a black box and the conductivity at maximum absorption (390 nm) is also indicated. The red and pink lines indicate the absorbance at 280 nm and 390 nm, respectively. The brown and the green lines indicate the conductivity and the NaCl gradient, respectively. The concentration of NaCl can be read from the conductivity axis employing the unit 10^{-2} M.

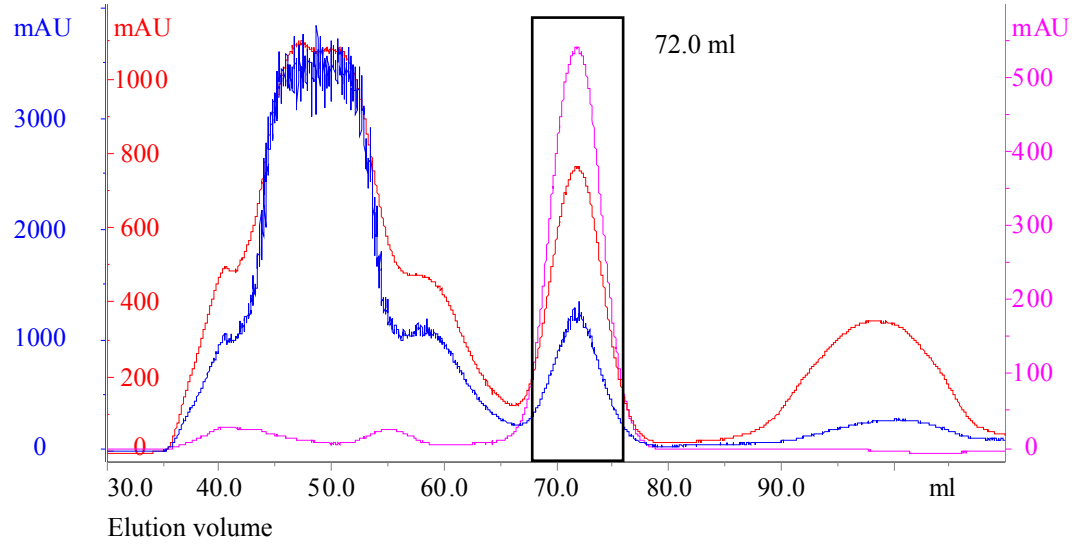


Figure 10.3 Chromatogram for purification of Ub-LCA ferredoxin obtained from HC1269 on the HiLoad 16/60 Superdex 75 prep grade column. The collected fraction is indicated with a black box and the elution volume at maximum absorption (390 nm) is also indicated. The blue, the red and the pink lines indicate the absorbance at 236 nm, 280 nm and 390 nm, respectively.

was observed upon elution and it was collected in one fraction (figure 10.4). The maximum absorption of the peak corresponding to the brown band was observed at a conductivity of 22.7 mS/cm (measured at 390 nm). A UV-vis spectrum recorded of purified Ub-LCA ferredoxin is shown in figure 10.5 A. An absorption maximum at approximately 390 nm was evident. The purity of Ub-LCA ferredoxin was analysed by SDS-PAGE (figure 10.5 B). The yield was 2.7 mg/L cell culture.

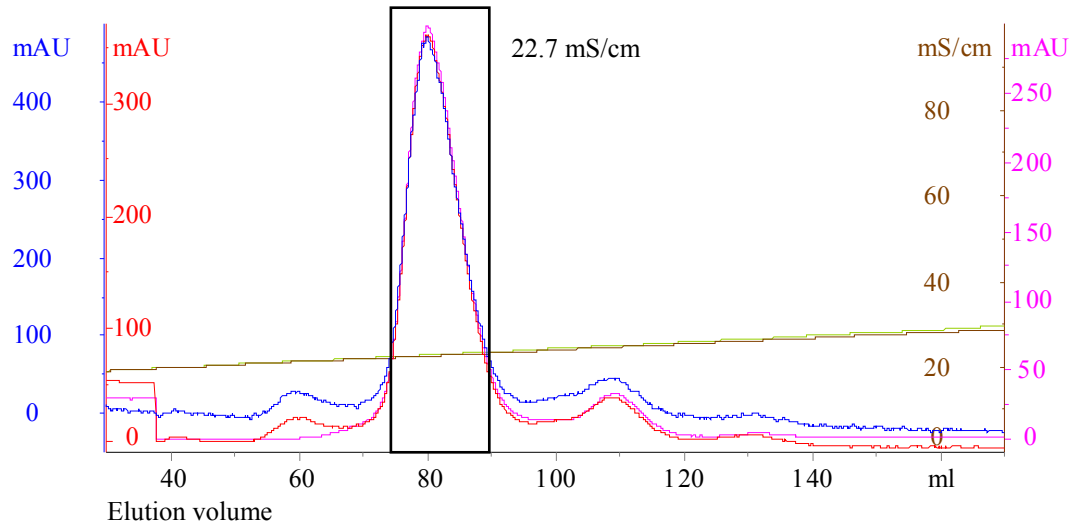


Figure 10.4 Chromatogram for purification of Ub-LCA ferredoxin obtained from HC1269 on the 16/10 Source 30Q column. The collected fraction is indicated with a black box and the conductivity at maximum absorption (390 nm) is also indicated. The blue, red and pink lines indicate the absorbance at 236 nm, 280 nm and 390 nm, respectively. The brown and the green lines indicate the conductivity and the NaCl gradient, respectively. The concentration of NaCl can be read from the conductivity axis employing the unit 10^{-2} M.

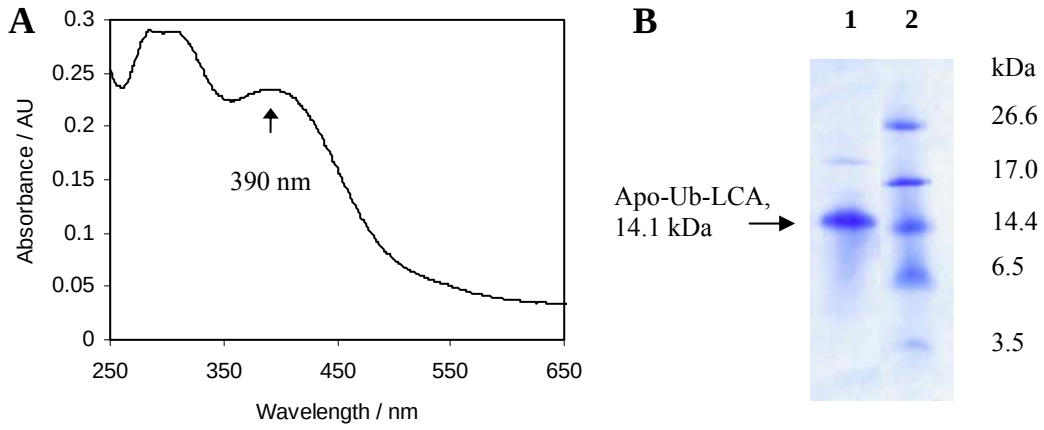


Figure 10.5 Identity of purified product from HC1269, Ub-LCA ferredoxin, analysed by UV-vis spectrophotometry and SDS-PAGE. A) UV-vis spectrum of purified Ub-LCA ferredoxin, a solution containing 0.4 mg/ml was diluted three times and the spectrum recorded. Absorption maximum at ca. 390 nm is indicated in the figure. B) SDS-PAGE results. Lane 1: 20 µl of a solution containing 0.4 mg/ml purified Ub-LCA ferredoxin. Lane 2: Marker, 1 µl.

10.2.1.2 Purification of LCA ferredoxin

When the protein extract (obtained from HC1419) was loaded onto the 16/10 Q Sepharose High Performance column a greenish brown band appeared close to the top of the column, which was also the case when purifying protein from HC1269. However, upon elution the brown band split into two and each band was collected in one fraction (figure 10.6). The first brown band eluted was assigned to Ub-LCA ferredoxin (see section 10.3.1 for a discussion) and the maximum absorption of the peak was observed at a conductivity of 22.1 mS/cm

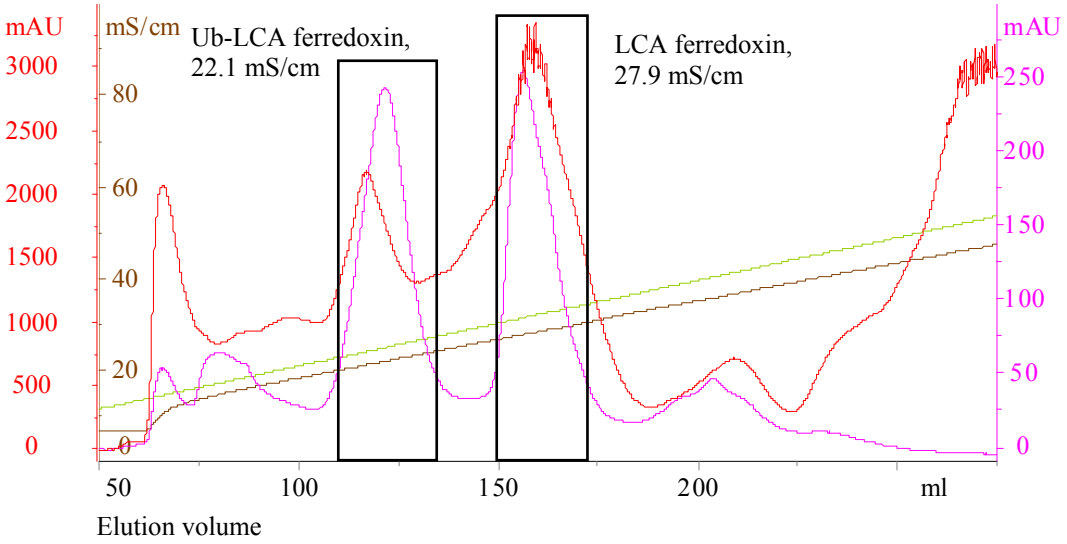


Figure 10.6 Chromatogram for purification of Ub-LCA and LCA ferredoxin obtained from HC1419 on the 16/10 Q Sepharose High Performance column. The collected fractions are indicated with black boxes and the conductivities at maximum absorption (390 nm) are also indicated. The red and pink lines indicate the absorbance at 280 nm and 390 nm, respectively. The brown and the green lines indicate the conductivity and the NaCl gradient, respectively. The concentration of NaCl can be read from the conductivity axis employing the unit 10^{-2} M.

(measured at 390 nm). The second brown band eluted was assigned to LCA ferredoxin (see section 10.3.1 for a discussion) and the maximum absorption of the peak was observed at a conductivity of 27.9 mS/cm (measured at 390 nm). Both were further purified by gel filtration on the HiLoad 16/60 Superdex 75 prep grade column. Upon elution one brown band was observed during each of the two purifications and it was collected in one fraction (figure 10.7). Ub-LCA and LCA ferredoxin eluted at 71.5 ml and 78.2 ml, respectively (measured at 390 nm). Both were purified by high resolution anion-exchange on the 16/10 Source 30Q column. Again, one brown band was observed upon elution in each of the purifications and it was collected in one fraction (figure 10.8). LCA and Ub-LCA ferredoxin eluted at 23.5 mS/cm and 29.2 mS/cm, respectively (measured at 390 nm). The purities of LCA and

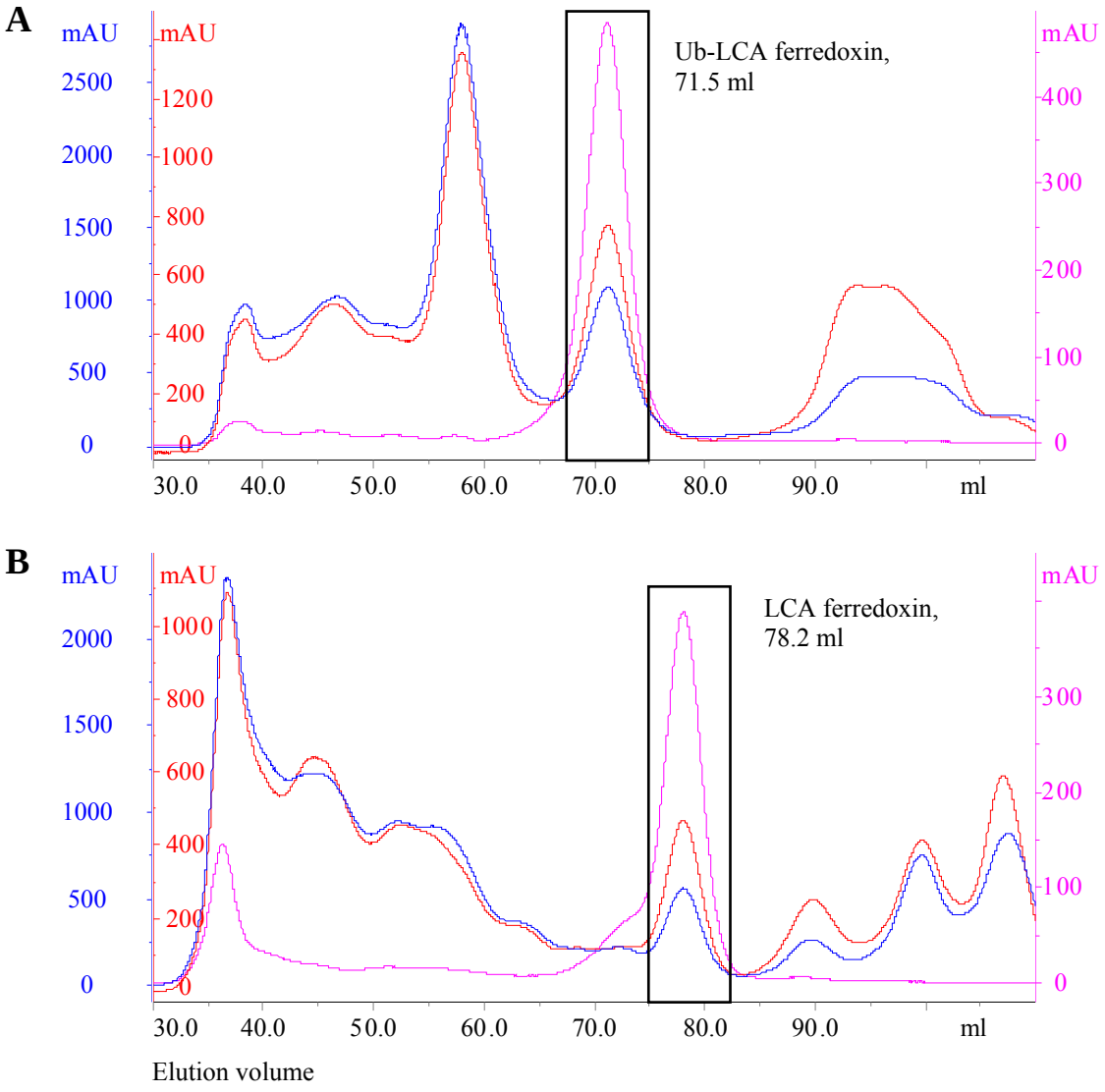


Figure 10.7 Chromatograms for purification of Ub-LCA ferredoxin (A) and LCA ferredoxin (B) obtained from HC1419 on the HiLoad 16/60 Superdex 75 prep grade column. In each chromatogram the collected fraction is indicated with a black box and the elution volume at maximum absorption (390 nm) is also indicated. The blue, the red and the pink lines indicate the absorbance at 236 nm, 280 nm and 390 nm, respectively.

Ub-LCA ferredoxin were analysed by SDS-PAGE (figure 10.9). UV-vis spectra of purified LCA and Ub-LCA ferredoxin have an absorption maximum at approximately 390 nm (figure 10.10). The yields of LCA and Ub-LCA ferredoxin were 0.3 mg/L cell culture and 1.6 mg/L cell culture, respectively.

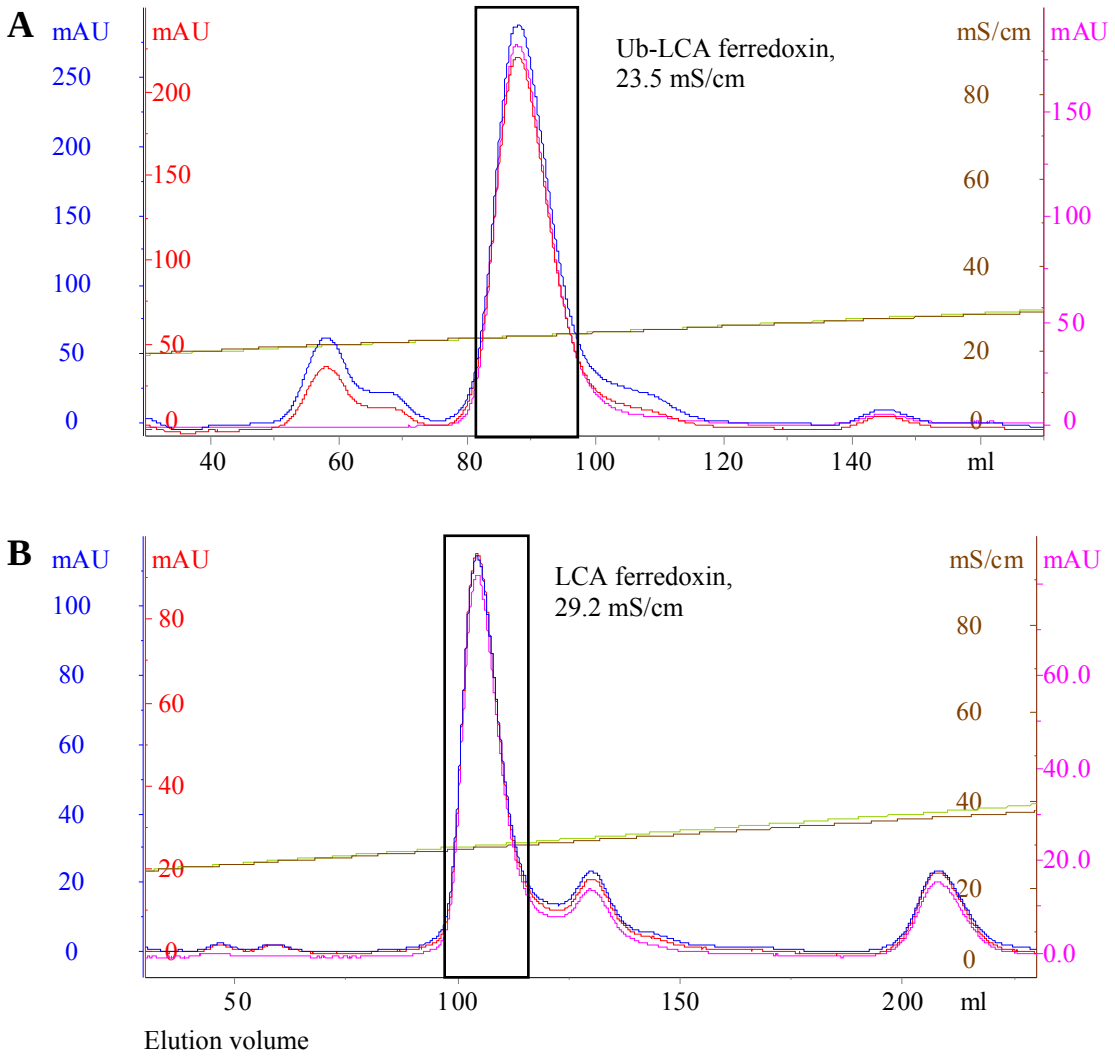


Figure 10.8 Chromatograms for purification of Ub-LCA ferredoxin (A) and LCA ferredoxin (B) obtained from HC1419 on the 16/10 Source 30Q column. In each chromatogram the collected fraction is indicated with a black box and the conductivity at maximum absorption (390 nm) is also indicated. The blue, red and pink lines indicate the absorbance at 236 nm, 280 nm and 390 nm, respectively. The brown and the green lines indicate the conductivity and the NaCl gradient, respectively. The concentration of NaCl can be read from the conductivity axis employing the unit 10^{-2} M.

10.2.2 Large scale purification of Ub-LCA ferredoxin

Observations made during the large scale purification of Ub-LCA obtained from the HC1269 strain were comparable to those observed from the small scale purification (section 10.2.1). Typical chromatograms obtained from the purification are shown in appendix A.3 along with purity analysis by SDS-PAGE. The yield of Ub-LCA ferredoxin was 1 to 2 mg/L cell culture.

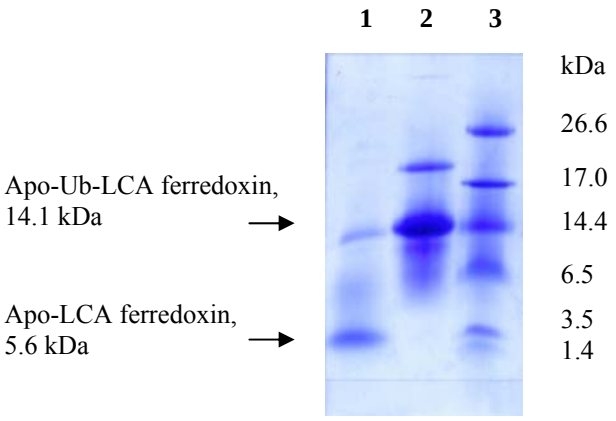


Figure 10.9 Purity of LCA and Ub-LCA ferredoxin obtained from HC1419 tested by SDS-PAGE. Lane 1: 20 µl of 1.8 mg/ml LCA ferredoxin. Lane 2: 5 µl of 6.4 mg/ml Ub-LCA ferredoxin. Lane 3: Marker, 1 µl.

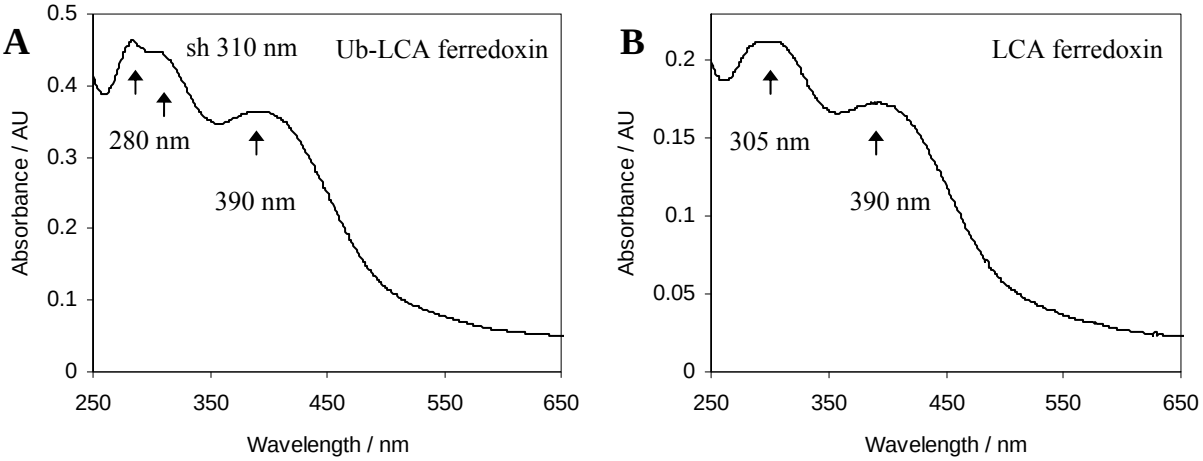


Figure 10.10 UV-vis spectra recorded of purified proteins obtained from HC1419. A) Purified Ub-LCA ferredoxin (0.2 mg/ml). B) Purified LCA ferredoxin (0.04 mg/ml). Maxima and shoulders (sh) are indicated in the figures.

10.2.3 Proteolytic cleavage of Ub-LCA ferredoxin

Compared to the test experiment (section 9.3.4) some changes were made in the first cleavage reaction. The Ub-LCA ferredoxin concentration was raised from 1.5 to 4.4 mg/ml and the DTT concentration was doubled from 4 to 8 mM. During the cleavage reaction no visible changes happened and the brown colour of the solution was unchanged. After the 22 hours of reaction, the buffer of the reaction solution was changed and YUH1 was separated from LCA ferredoxin and un-cleaved Ub-LCA ferredoxin by passing the solution through a HisTrap column (figure 10.11). The brown coloured proteins were seen to pass straight through the column without binding while uncoloured compounds were seen to elute upon washing with 0.5 M imidazole containing buffer. Ub-LCA and LCA ferredoxin were further purified by high-resolution anion-exchange (figure 10.12). The peak corresponding to Ub-LCA ferredoxin is dominating. Ub-LCA and LCA ferredoxin were collected separately. The UV-vis spectrum of Ub-LCA ferredoxin was comparable to the one shown in figure 10.10 A while

the spectrum of LCA ferredoxin was comparable to the one shown in figure 10.10 B. The yield of LCA ferredoxin was 1.8 mg which corresponds to 19 % (on a molar basis) and the un-cleaved Ub-LCA ferredoxin recovered was 14.0 mg (62 %).

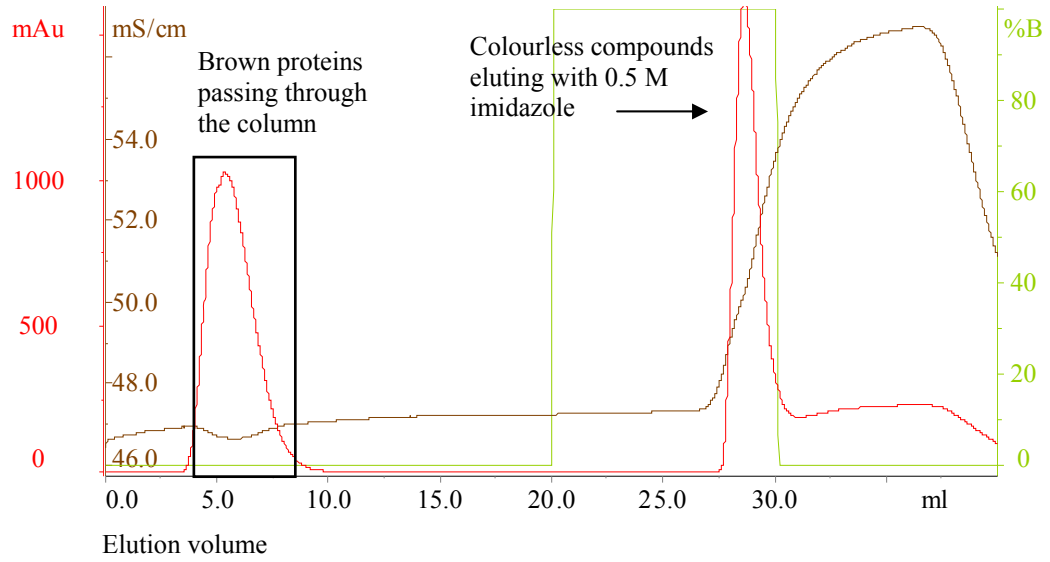


Figure 10.11 Chromatogram for separation of Ub-LCA and LCA ferredoxin from YUH1 on the HisTrap column. The collected fraction is indicated with a black box. The red line indicates the absorption at 280 nm and the brown line indicates the conductivity. The green line indicates the gradient where 0 % corresponds to buffer containing 20 mM imidazole and 100 % corresponds to buffer containing 0.5 M imidazole.

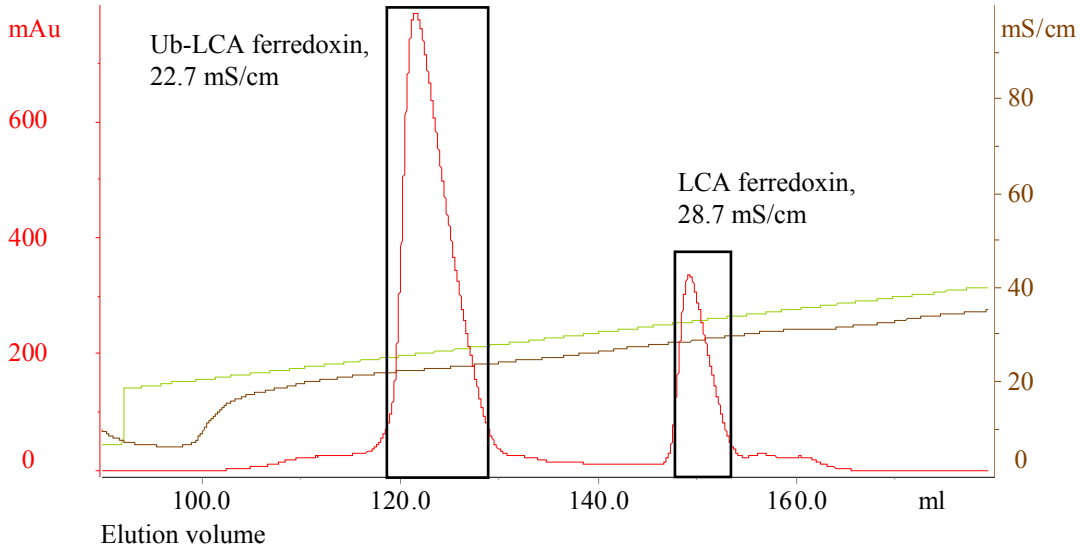


Figure 10.12 Chromatogram for separation of Ub-LCA and LCA ferredoxin obtained from the first cleavage experiment on the 6 ml Resource Q column. The collected fractions are indicated with black boxes and the conductivities at maximum absorption (280 nm) are also indicated. The red line indicates the absorbance at 280 nm. The brown and the green lines indicate the conductivity and the NaCl gradient, respectively. The concentration of NaCl can be read from the conductivity axis employing the unit 10^{-2} M.

Several changes were made in the second cleavage experiment. The DTT concentration was raised from 8 to 20 mM, the pH was lowered from 8.5 to 8.0 (also to include 0.15 M NaCl), the Ub-LCA ferredoxin concentration was raised from 4.4 mg/ml to 5.6 mg/ml and the molar ratio of Ub-LCA ferredoxin:YUH1 was raised from 1:1 to 1:2. Again, no visible changes happened during the cleavage reaction and separation of YUH1 from Ub-LCA and LCA ferredoxin was comparable to what was described above. Ub-LCA and LCA ferredoxin were separated by high-resolution anion-exchange (figure 10.13). Contrary to the previous separation (figure 10.12), the peaks corresponding to Ub-LCA and LCA ferredoxin were now comparable in size. The yield of LCA ferredoxin was 1.9 mg corresponding to 32 % (on a

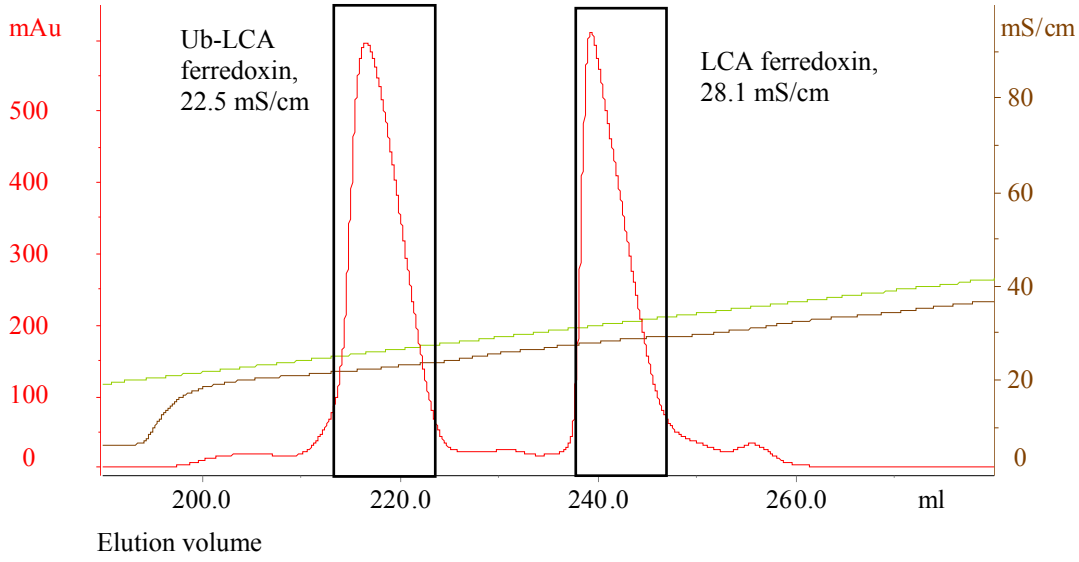


Figure 10.13 Chromatogram for separation of Ub-LCA and LCA ferredoxin obtained from the second cleavage experiment on the 6 ml Resource Q column. The collected fractions are indicated with black boxes and the conductivities at maximum absorption (280 nm) are also indicated. The red line indicates the absorbance at 280 nm. The brown and the green lines indicate the conductivity and the NaCl gradient, respectively. The concentration of NaCl can be read from the conductivity axis employing the unit 10^{-2} M.

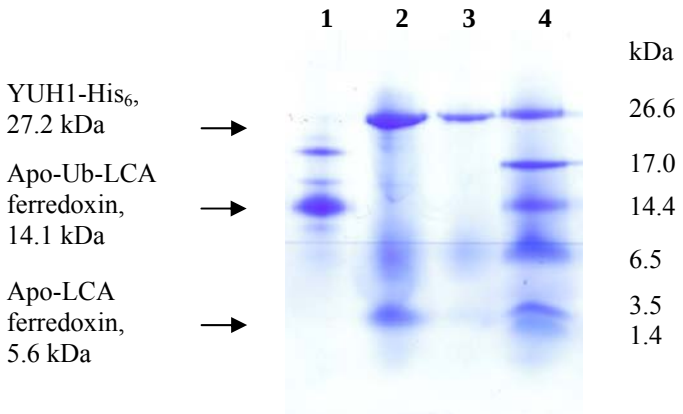


Figure 10.14 Purity of LCA ferredoxin obtained from cleavage reaction and recovered Ub-LCA ferredoxin tested by SDS-PAGE. Lane 1: 0.5 µl of 16.1 mg/ml Ub-LCA ferredoxin. Lane 2: 2 µl of 14.3 mg/ml LCA ferredoxin. Lane 3: 0.5 µl of 14.3 mg/ml LCA ferredoxin. Lane 4: Marker, 1 µl.

molar basis) and the un-cleaved Ub-LCA ferredoxin recovered was 5.5 mg (39 %). The total yield of LCA ferredoxin obtained from the first and second cleavage experiment was 0.1 and 0.2 mg/L cell culture, respectively. LCA ferredoxin obtained from the two cleavage reactions was pooled and concentrated to 14.3 mg/ml. The purity of LCA and Ub-LCA ferredoxin was tested by SDS-PAGE (figure 10.14). Recovered Ub-LCA ferredoxin gave several bands on the SDS-PAGE in addition to the band assigned to Ub-LCA ferredoxin, while LCA ferredoxin gave one band at ca. 26 kDa and a faint band assigned to LCA ferredoxin.

10.2.4 Solubility tests of LCA ferredoxin

When testing salts as precipitants at pH 4.5 heavy precipitations were observed within a day even at the low salt concentration. Phase separations were also observed at high salt concentrations. When using sodium acetate, ammonium dihydrogen phosphate and ammonium citrate, denaturation and colour loss were observed within two weeks.
When testing salts as precipitants at pH 9.0, no precipitation was observed when sodium acetate, magnesium sulfate and calcium chloride were used. With di-ammonium hydrogen phosphate and ammonium citrate, slight precipitation was observed at the low concentrations while heavy precipitation was evident at high concentrations. Ammonium sulfate resulted in heavy precipitation at high concentration while no precipitation was observed at low concentration. The drops in which heavy precipitation had been observed at high concentrations were then equilibrated over low concentration solutions. This resulted in dissolution of the precipitation obtained using ammonium sulfate and di-ammonium hydrogen phosphate while a marked reduction in the amount of precipitation was observed for ammonium citrate.
PEGs and MPD were tested as precipitants at pH 5.5. PEG 400 resulted in precipitation after one week at high concentration while slight precipitation was observed at low concentration after four weeks. PEG 4000 and 8000 resulted in slight precipitation at low concentrations and moderate precipitation at high concentration within one week. MPD caused colour loss and denaturation.

10.3 Discussion

10.3.1 Small scale purifications of Ub-LCA and LCA ferredoxin

It was first attempted to purify Ub-LCA ferredoxin since it is easily detectable by SDS-PAGE and this could prove advantageous if holo-protein was not formed in the *E. coli* cells. LCA ferredoxin is not expected to be easily detectible by SDS-PAGE (stained with Coomassie blue G-250) since it only contains two basic and no aromatic residues (figure 3.3) (Compton and Jones, 1985).
Protein extract obtained from the two strains did not show any distinct bands for Ub-LCA ferredoxin or Ub when analysing the extract by SDS-PAGE. This shows that the Ub-LCA ferredoxin has not been highly over-expressed in any of the strains and the cultivation protocols should be optimized.

10.3.1.1 Purification of Ub-LCA ferredoxin

When performing purification of Ub-LCA ferredoxin from HC1269, brown coloured protein was evident from the first purification step (section 10.2.1.1), which indicates the presence of an FeS protein. The brown coloured protein was purified and the identity verified by SDS-PAGE analysis (figure 10.5 B) and UV-vis spectrophotometric analysis (figure 10.5 A). The dominating band on the SDS-PAGE was at ca. 14.4 kDa and apo/holo-Ub-LCA ferredoxin has a molecular weight of 14.8/14.1 kDa. A minor impurity was observed at ca. 20 kDa. This impurity may be removed by further optimization of the purification procedure. From the UV-vis spectrum an absorption maximum at ca. 390 nm was observed (figure 10.5 A). This spectrum is very similar to the spectra obtained from holo-Ub-Ori-ISP (chapters 6 and 7), which were assigned to a $[4Fe\text{-}4S]^{2+}$ protein. Therefore it can be said with great certainty that Ub-LCA ferredoxin contains at least one $[4Fe\text{-}4S]^{2+}$ cluster. This proves that holo-Ub-LCA ferredoxin has been produced in *E. coli* cells. The yield of protein was however very low (2.7 mg/L cell culture) and the cultivation procedure should be optimized.

10.3.1.2 Purification of LCA ferredoxin

When purifying protein from HC1419 two peaks corresponding to the collected brown fractions were observed from the chromatogram for the first purification (figure 10.6). Comparing this chromatogram to the one obtained when purifying Ub-LCA ferredoxin from HC1269 (figure 10.2) it is evident that the peak at 22.1 mS/cm (figure 10.6) is comparable to the peak assigned to Ub-LCA ferredoxin eluting at 22.9 mS/cm (figure 10.2). The first peak was therefore assigned to Ub-LCA ferredoxin. The second peak at 27.9 mS/cm (figure 10.6) was assigned to LCA ferredoxin, since LCA ferredoxin is expected to elute at a higher conductivity due to the change in surface charge distribution. A shift in elution to higher conductivities was also observed when ubiquitin was cleaved off Ub-Ori-ISP (producing Ori-ISP, chapter 7). Both of the brown coloured proteins were purified and their identity verified by SDS-PAGE analysis (figure 10.9) and UV-vis spectrophotometric analysis (figure 10.10). Analysing Ub-LCA ferredoxin by SDS-PAGE gave a result (figure 10.9 lane 2) which was highly similar to the result obtained when purifying Ub-LCA ferredoxin from HC1269 (figure 10.5 B). Further, the UV-vis spectrum (figure 10.10 B) was almost identical to the one obtained from HC1269 Ub-LCA ferredoxin (figure 10.5 A). On SDS-PAGE (figure 10.9 lane 1) LCA ferredoxin gave a dominating band at ca. 3 kDa. Compared to the amount loaded, this band is very faint, which was also expected. This provides a hint that it in fact is LCA ferredoxin (molecular weight of apo/holo-protein is 5.6/6.3 kDa) even though the band is found at quite a low molecular weight. It is however my experience that low molecular weight proteins and peptides often do possess odd properties when analysed by SDS-PAGE. To obtain an indisputable verification, LCA ferredoxin should be analysed by mass spectrometry. An impurity at ca. 14.4 kDa was observed in the gel which may be removed by further optimization of the purification procedure. The UV-vis spectrum supports the identity of LCA ferredoxin (figure 10.10 B) since no absorption maximum at 280 nm associated with aromatic residues are present. Distinct absorption maxima are present at approximately 305 nm and 390 nm, which strongly indicate the presence of FeS clusters. The UV-vis spectrum is almost identical to the spectra obtained for holo-Ub-LCA ferredoxin (figure 10.10 A) except for the missing absorption at 280 nm. Purified LCA ferredoxin very likely contains one or two

$[4Fe\text{-}4S]^{2+}$ clusters. The yield of protein was however very low; 0.3 mg LCA ferredoxin and 1.6 mg Ub-LCA ferredoxin per litre of cell culture. This corresponds to that approximately 30 % of holo-protein is found as holo-LCA ferredoxin (on a molar basis). Therefore the cultivation procedure should be optimized.
In chapter 9 it was found that YUH1 was inefficient in cleaving Ub-LCA ferredoxin *in vitro* under the investigated conditions and based on the results presented in this chapter UBP1 also appears inefficient for *in vivo* cleavage of Ub-LCA ferredoxin.

10.3.1.3 The strategy for obtaining LCA ferredoxin for further characterization

From the HC1269 strain the yield of Ub-LCA ferredoxin was 2.7 mg/L cell culture. If all protein could be cleaved and retained this would result in a yield of LCA ferredoxin of 1.1 mg/L cell culture. Initial cleavage experiments were performed (see chapter 9), and here it was observed that cleavage was about 20 % under the conditions used, which would result in a yield of 0.2 mg/L cell culture. This is lower than the LCA ferredoxin yield from the HC1419 strain (0.3 mg/L cell culture). Since complete cleavage of Ub-Ori-ISP was obtained using much less YUH1 (see chapter 9) it is believed that the conditions for Ub-LCA ferredoxin cleavage can be optimized further. As stated in the beginning of the chapter it is not necessary to cleave off Ub for preliminary characterization. Therefore it was decided to proceed with purification of Ub-LCA ferredoxin from the HC1269 strain and then perform cleavage of the purified Ub-fusion using YUH1.

10.3.2 Large scale purification of Ub-LCA ferredoxin

During the purification procedure it was observed from UV-vis spectrophotometric measurements at 390 nm (results not shown) that about 50 % of the starting material was lost for each of the purification steps; gel filtration and high resolution anion-exchange. Normally, recovery is about 90 to 95 %, which accounts for some of the loss. As seen from the chromatograms (appendix A.3) the compound that is responsible for the absolute majority of absorption at 390 nm is Ub-LCA ferredoxin. Separation of coloured compounds cannot account for all of the remaining loss. Since the inert gas flushing was used for obtaining anaerobic conditions, complete removal of oxygen is not accomplished (see section 6.3.2.1). Breakdown due to oxidation may account for the remaining loss of protein since FeS proteins generally are oxygen sensitive (see chapter 2). The stability of Ub-LCA/LCA ferredoxin towards oxygen has not been examined. Therefore, it may be an advantage to perform purification inside an anaerobic chamber. The yield of the large scale purifications was found from 1 to 2 mg/L cell culture. This is lower than for the small scale purifications and may be attributed to the cultivation procedure and breakdown during purification. The same problems are expected when working with expression and purification of protein from the HC1419 strain.
The purity of the obtained product was analysed by SDS-PAGE (appendix A.3) and as seen previously (section 10.3.1) an impurity around 20 kDa is present.

10.3.3 Proteolytic cleavage of Ub-LCA ferredoxin

10.3.3.1 Separation of reaction products from proteolytic cleavage

Prior to proteolytic cleavage it was tested how the products of the reaction could be separated on analytical columns (these results are not shown). It was found that at pH 8.0 Ub-LCA and LCA ferredoxin were separable by anion-exchange (1 ml Resource Q column) (as previously shown, see section 10.3.1) and ubiquitin did not bind to the column. LCA ferredoxin and YUH1 eluted at the same conductivity rendering them inseparable by anion-exchange at this pH. Upon analytical gel filtration (13/31 Superdex 75 HR) it was found that YUH1 and LCA ferredoxin were separable but Ub-LCA ferredoxin and Ub could not be completely separated from LCA ferredoxin. YUH1 is equipped with a His_6-tag (YUH1-His_6) and it was found that the enzyme was able to completely bind to a 5 ml HisTrap HP column (10.6 mg, 20 mM Tris/HCl, 20 mM imidazole, 0.5 M NaCl pH 7.5).
Based on the observations the following purification strategy was adopted: YUH1 should be removed from other products and reactants by purification on a HisTrap HP column and then LCA ferredoxin should be separated from Ub and Ub-LCA ferredoxin by anion-exchange.

10.3.3.2 Cleavage followed by separation

In the first cleavage reaction some changes to the preliminary cleavage experiments (section 9.3.4) were made (section 10.2.3). The yields of the two procedures were similar. Therefore doubling the DTT concentration and tripling the Ub-LCA ferredoxin concentration did not seem to have any effect on the yields.
In the second experiment several changes were made (see section 10.2.3). The yield of LCA ferredoxin increased from 19 % to 32 %. The dramatic change was also observed in the anion-exchange chromatograms from each experiment (compare figures 10.12 and 10.13). Since several changes were made compared to the first experiment, it is difficult to single out one parameter responsible for the higher yield and it may be a combination of parameters. Since DTT and Ub-LCA ferredoxin concentrations did not seem to play any significant part as judged from the first experiment, the increased Ub-LCA ferredoxin:YUH1 molar ratio (from 1:1 to 1:2) and the decrease in pH (from 8.5 to 8.0) may be responsible. It is necessary to further investigate the effects of pH and molar excess of YUH1. To increase the yield even further, it may be advantageous to increase the temperature of the reaction mixture to 37 ºC (see chapter 9).
The purity of recovered Ub-LCA ferredoxin from the cleavage reaction was tested by SDS-PAGE (figure 10.14 lane 1). As observed previously (section 10.3.1) the impurity at 20 kDa was still present and even more prominent. Impurities at 15 and 10 kDa had now become apparent. These impurities had probably been present prior to cleavage, but now they were visible since they had remained (accumulated) during purification while some of the Ub-LCA ferredoxin had been cleaved to LCA ferredoxin.
The purity of LCA ferredoxin was also tested by SDS-PAGE (figure 10.14 lanes 2 and 3). Besides the faint band assigned to LCA ferredoxin, a dominating band at ca. 26 kD was present. This band was assigned to YUH1 (see chapter 9). Comparing the intensity of the band to gels containing known amounts of YUH1 (results not shown) a rough estimate of the

degree of contamination was ca. 10 to 15 % of the total mass. The presence of YUH1 must be a result of incomplete binding to the HisTrap HP column or that the column was overloaded. Purified YUH1 loaded in a comparable amount was however found to bind completely to the HisTrap HP column (section 10.3.3.1). The other compounds found in solution (Ub, LCA ferredoxin and uncleaved Ub-LCA ferredoxin) might induce incomplete binding. Since the cleavage reaction was incomplete it is imaginable that YUH1 could be in contact with Ub-LCA ferredoxin, which may somehow hinder contact between the His_6-tag and the column material. YUH1 and LCA ferredoxin can be separated by gel filtration (section 10.3.3.1). This would however require a third purification step and it may be better changing the strategy for separation of LCA ferredoxin from YUH1, Ub and Ub-LCA ferredoxin by skipping the HisTrap HP column and performing anion-exchange followed by gel filtration instead.

10.3.4 Re-evaluation of the procedure for obtaining LCA ferredoxin

The yield of LCA ferredoxin from the second cleavage experiment was low (ca. 30 %) and at least two purification steps are believed to be required (section 10.3.3.2) for optimal separation of products. The total yield of LCA ferredoxin from the procedure of first purifying Ub-LCA ferredoxin and then performing cleavage reactions and purifications was about 0.2 mg/L cell culture. This is comparable to what was obtained when purifying LCA ferredoxin directly from the cell extract (0.3 mg/L, see section 10.2.1.2). At the same time a huge amount of YUH1 was used (4 mg per 1 mg Ub-LCA ferredoxin) which makes this procedure time-consuming and expensive.

At present it seems more advantageous to return to the strategy of obtaining LCA ferredoxin directly from the *E. coli* cells (HC1419) and to optimize the cultivation procedure to yield more LCA ferredoxin. Alternatively, a new *E. coli* strain should be constructed which only expresses LCA ferredoxin.

10.3.5 Solubility tests of LCA ferredoxin

LCA ferredoxin was contaminated with YUH1 (see section 10.3.3.2) but since a very small amount was available it was decided not to further purify but instead test the concentration range for crystallization to be used for future experiments.

The Solubility Tool Kit (Molecular Dimensions) was applied to test solubility properties of LCA ferredoxin. It applies the Hofmeister series for testing different salts (Collins, 2004; Riès-Kautt and Ducruix, 1997) and tests solubility at pH (4.5) < pI, pH (9.0) > pI and pH ≈ pI. Apo-LCA ferredoxin has a theoretical pI of 4.0 (calculated using the ExPASy Proteomics Server, Bjellqvist *et al.*, 1993 and 1994) and therefore at pH 4.5 the holo-protein is very close to the pI. At pH ≈ pI the net charge of a protein is close to zero and the solubility is very low. This is also reflected by the results since at pH 4.5 massive precipitations occurred when using various salts as precipitants (section 10.2.4). The protein was also found to denature using certain precipitants, which may be due to a combination of the precipitant in question and the pH since the stability of LCA ferredoxin at various pH values has not been determined.

At pH 9.0 ammonium sulfate, di-ammonium hydrogen phosphate and ammonium citrate yielded brown precipitation while sodium acetate, magnesium sulfate and calcium chloride did not yield any precipitation. This is in agreement with the Hofmeister series since the first

three precipitants are more powerful. The precipitation was either completely or to a high extent soluble which shows that the precipitation was crystalline and provides hopes for future crystallization.
When using salts as precipitants the solubility is highly dependent on pH and the particular precipitant. It might be necessary to test the solubility and precipitation properties over a broad concentration interval e.g. from 5 to 20 mg/ml.
When pH ≈ pI organic solvents such as PEGs and MPD are applied in The Solubility tool Kit. The lowest pH available is pH 5.5 which was used in the experiments. MPD resulted in denaturation of the protein. This may be a combined effect of pH and precipitant. Applying PEG 4000 and 8000 as precipitants, precipitation was observed at the low concentration within a week which might suggest that the protein concentration is too high. PEG 400 did not result in precipitation at the lowest concentration until four weeks had passed indicating the concentration might be close to optimal under the given conditions.
Future crystallization experiments will probably include commercially available screens and it may be useful to test two or three different LCA ferredoxin concentrations.

10.4 Conclusion

Two approaches for obtaining holo-LCA ferredoxin have investigated: either directly from protein expressing cells (co-expression of Ub-LCA ferredoxin and UBP1) or from proteolytic cleavage of already purified Ub-LCA ferredoxin (from Ub-LCA expressing cells).
Procedures for expression of holo-Ub-LCA and holo-LCA ferredoxin in *E. coli* cells have been prepared. The yield of each protein was very low and the procedures should be further optimized.
Procedures for small scale purification of Ub-LCA and LCA ferredoxin along with a large scale procedure for purification of Ub-LCA ferredoxin have been developed.
Proteolytic cleavage of Ub-LCA ferredoxin using YUH1 to obtain LCA ferredoxin has been partly optimized. However further optimization is required to improve the yield and to minimize YUH1 consumption. Separation of the reaction products has been examined and a purification strategy put forward.
The yields of LCA ferredoxin obtained from the two approaches were comparable. However, obtaining LCA ferredoxin directly from protein expressing cells proved much less time consuming.

Chapter eleven

11 Crystallization trials for *Pyrococcus furiosus* [4Fe-4S] ferredoxin

This chapter describes trials to crystallize the $[4Fe\text{-}4S]^{2+}$ ferredoxin from the hyperthermophilic archaeon *P. furiosus* (section 2.5). Crystal structures of the native [3Fe-4S] ferredoxin along with the mutant D14C [4Fe-4S] ferredoxin have already been solved (Nielsen *et al.*, 2004; Johannesen *et al.*, in preparation). Earlier attempts to crystallize the native [4Fe-4S] ferredoxin in our research group (unpublished) have been unsuccessful. It was therefore decided to start from scratch and screen for crystallization conditions for the native [4Fe-4S] ferredoxin. Reasons as to why crystallization of the [4Fe-4S] ferredoxin appears to be difficult are also investigated.

Expression and purification of *P. furiosus* ferredoxin were performed under aerobic conditions by Laboratory Assistant, Stephanie Boy and Managing Laboratory Technician, Lise-Lotte Jespersen in accordance to existing procedures described in in-house protocols. In all the presented crystallization trials the protein used was in 20 mM Tris/HCl pH 8.0 and stored at -20ºC or in liquid nitrogen. The concentration of holo-protein was calculated using a molar absorptivity of 17 $mM^{-1} \cdot cm^{-1}$ at 390 nm (Conover *et al.*, 1990).

11.1 Experimental

Aerobic screening was performed at the University of Copenhagen, Department of Chemistry (section 11.1.1), while anaerobic screening and initial optimization was performed at IBS (Grenoble, France) (section 11.1.2.1). Further optimization was performed in our own laboratories (11.1.2.2).

11.1.1 Aerobic screening

Protein purified under aerobic conditions was re-purified by HPLC (ÄKTA^{TM} purifier 100) using high resolution anion-exchange under anaerobic conditions using Ar-flushing. This was performed by loading the protein onto a 20 mM Tris/HCl pH 8.0 equilibrated 16/10 Source 30Q column. The protein was eluted using a linear gradient from 0.15 M to 0.3 M NaCl in 10 CV. The buffer was changed to 20 mM Tris/HCl pH 8.0 via ultrafiltration and up-concentrated using a stirred Amicon cell equipped with a YM3 membrane.

Initial crystallization screening was carried out under aerobic conditions using an Oryx 8 robot system from Douglas Instruments at room temperature. Prior to set up, the protein solution was spun for 10 min at 10000 g. Three commercially available screens were employed, each containing 96 reagents: Index Screen (Hampton Research), JCSG+ Suite and PACT Suite (both from Quiagen). The sitting drop vapour diffusion technique was applied and the reservoir and total drop volumes were 100 µl and 0.3 µl, respectively. For each tested

condition, three drops were set up containing 25 %, 50 % and 75 % of protein solution, respectively. All three screens were set up using a protein stock of 14.5 mg/ml and all were stored at room temperature. The JCSG+ suite was also set up using a protein stock of 45.0 mg/ml. Conditions for set up are presented in table 11.1.

Table 11.1 Overview of aerobic screening.

Screen	Concentration of protein stock (mg/ml)	Storage temperature	Volume ratios of protein solution to reservoir solution		
Index Screen	14.5	Room temperature	1:3	1:1	3:1
PACT Suite	14.5	Room temperature	1:3	1:1	3:1
JCSG+ Suite	14.5	Room temperature	1:3	1:1	3:1
JCSG+ Suite	45.0	Room temperature	1:3	1:1	3:1

11.1.2 Crystallization trials performed under anaerobic conditions

Protein purified under aerobic conditions was re-purified as described in section 11.1.1 in an anaerobic chamber (see section 4.2). Ultrafiltration was performed using either a stirred Amicon cell equipped with a YM3 membrane or a Vivacell 70 equipped with a 5000 MWCO membrane. HPLC purifications were carried out on an ÄKTAFPLC™ system in France and on an ÄKTAprime™ plus system in Denmark. Prior to set up of crystallization experiments, the protein solution was filtered (0.22 µm, Ultrafree-MC, Millipore).

11.1.2.1 Anaerobic screening and initial optimization

Anaerobic screening and initial optimizations were performed (anaerobic chamber described in section 4.2.1) using a second batch of protein. Three commercially available screens were employed, each containing 96 reagents; Crystal Screen HT (Hampton Research), PEGs Suite and pH Clear Suite (both from Qiagen). The sitting drop vapour diffusion technique was applied and the reservoir volumes were 100 µl. The crystallization drops were set up using 1 µl of protein solution and 1 µl of reservoir solution. A protein stock of 10.0 mg/ml was used and set up and storage was at 20 ± 1ºC. Crystal Screen I (Hampton Research) along with a home made version of the Solubility Tool Kit (originally manufactured by Molecular Dimension) at high pH (see appendix A.4 section A.4.1) was also set up using the hanging drop vapour diffusion technique. The reservoir volumes were 500 µl and each crystallization drop consisted of 1 µl of protein solution and 1 µl of reservoir solution. A protein stock of 15.0 mg/ml was used and set up, and storage was at 20 ± 1ºC.
During screening, crystals appeared and several experiments to optimize the crystallization conditions were performed. The conditions tested in these experiments are outlined in appendix A.4 section A.4.2.

11.1.2.2 Further optimization

Further optimizations were performed (anaerobic chamber described in section 4.2.2) using a third batch of protein. The conditions tested in these experiments are outlined in appendix A.4 section A.4.3. For all the trials a protein stock of 23.5 mg/ml was used and set up and storage was at room temperature.

11.1.3 Monomer-dimer interconversion

It has been reported that *P. furiosus* ferredoxin can be found as a monomer or dimer depending on ionic strength (Hasan *et al.*, 2002). This was investigated by analytical gel filtration under aerobic conditions. These experiments were performed by Mette Jensen and Tina Vognsen (project students) under my supervision.
The purified [4Fe-4S] ferredoxin (8.5 mg/ml, 20 mM Tris/HCl pH 8.0) was diluted 17 times with NaCl-containing buffer. Six different NaCl concentrations were tested. The dilution buffers were:

1. 20 mM Tris/HCl, 0.15 M NaCl pH 8.0
2. 20 mM Tris/HCl, 1.0 M NaCl pH 8.0
3. 20 mM Tris/HCl, 2.0 M NaCl pH 8.0
4. 20 mM Tris/HCl, 3.0 M NaCl pH 8.0
5. 20 mM Na acetate 0.15 M NaCl pH 4.6
6. 20 mM Na acetate 3.0 M NaCl pH 4.6

The six solutions containing the diluted protein were incubated for 24 hours at 4ºC. Each solution was then analysed by analytical gel filtration at room temperature by HPLC (ÄKTA^{TM} purifier 100). 250 µl aliquots of sample (linear flow rate: 0.3 cm/min) were loaded onto a 13/31 Superdex 75 HR column equilibrated with the corresponding dilution buffer and protein was eluted with the same buffer.

11.2 Results and discussion

11.2.1 Aerobic screening

It is known that upon oxidation the $[4Fe\text{-}4S]^{2+}$ ferredoxin converts to the $[3Fe\text{-}4S]^{1+}$ ferredoxin (section 2.5). This is a slow process that takes place as a consequence of exposure to air. Upon re-purification of aerobically purified $[4Fe\text{-}4S]^{2+}$ ferredoxin that had been stored at -20ºC (section 11.1.2) it was found that approximately 10 % of the $[4Fe\text{-}4S]^{2+}$ ferredoxin had been oxidized to $[3Fe\text{-}4S]^{1+}$ ferredoxin (results not shown). It was also found that when the $[4Fe\text{-}4S]^{2+}$ ferredoxin was exposed to air and the pH had been lowered to 4.6, the dominating species was the $[3Fe\text{-}4S]^{1+}$ ferredoxin after one week of exposure (results not shown). This was evident from UV-vis spectrophotometric analysis since the absorption maximum had changed from 390 nm to 408 nm (figure 2.2). This is not surprising since iron-sulfur clusters are acid labile (section 2.1).
Even though slow oxidation was evident it was decided to perform rapid screening using an experimental set up performed by a robot. This was not possible to perform under anaerobic conditions.
No crystals appeared in any of the tested conditions. However, brown crystalline precipitation appeared in many of the tested conditions. Denaturation was observed in some cases when the pH was below 5 and especially when phosphate was present. A general trend was that the lower the pH, the heavier the precipitation. *P. furiosus* apo-ferredoxin has a theoretical pI of 3.9 (calculated using the ExPASy Proteomics Server, Bjellqvist *et al.*, 1993 and 1994) and

therefore at low pH the net charge of the protein is close to zero hence the solubility is very low, which is reflected by the results.

Index Screen
It was found that ammonium sulphate and phosphate salts resulted in precipitation or phase separation. Sodium chloride gave precipitation when pH was below 4.5 and magnesium formate did not yield any precipitation. PEG 3350 without salts only resulted in slight precipitation below pH 5.5. PEG 3350 in combination with salts resulted in precipitation at almost all conditions tested. MPD resulted in phase separation.

PACT Suite
Testing pH values in the interval 4 to 9 using four different buffer systems at each tested pH value and at constant PEG 1500 concentration, gave precipitation when the SPG buffer (contains succinic acid, Na dihydrogen phosphate and glycine) and the PCB buffer (contains Na propionate, Na cacodylate and Bis-Tris propane) were used. When testing the effect of cations (chloride salts) at pH 5 to 8 at constant PEG 6000 concentration, it was found that ammonium, zinc, magnesium and calcium caused precipitation at all pH values while lithium and sodium only caused precipitation at pH 5. When testing the effects of various anions (sodium salts) at a constant PEG 3350 concentration and at pH values from 6.5 to 8.5 as well as without any buffer, it was found that citrate and malonate did not yield any precipitation while tartrate resulted in loss of colour indicating cluster degradation. Phosphate, sulphate, acetate, formate, nitrate, thiocyanate, iodide, bromide and fluoride resulted in precipitation.

JCSG+ Suite
The observations were comparable to those made from the Index Screen and The PACT Suite. The JCSG+ Suite screens conditions for which other proteins and enzymes have been successfully crystallized. It is difficult to draw any conclusions since no hits were found.

11.2.2 Anaerobic screening

Holo-protein purified under strict anaerobic conditions was re-purified and it was found that less than 0.5 % had converted to the $[3Fe\text{-}4S]^{1+}$ ferredoxin at pH 8.0 (results not shown). Further, no conversion was evident after exposing the protein to pH 4.6 at room temperature for a period of four weeks as determined by UV-vis spectrophotometric analysis (results not shown). However, this only shows that the $[4Fe\text{-}4S]^{2+}$ ferredoxin is the dominating species.
Testing the Crystal Screen HT, the PEGs Suite and the pH Clear Suite using a protein stock of 10.0 mg/ml gave only slight precipitation in few of the drops, and no crystals were observed. The results obtained were comparable to those observed during aerobic screening (section 11.2.1). From the PEGs Suite, precipitation was only observed when PEG 3350 in combination with various salts was tested; calcium chloride, sodium thiocyanate, lithium nitrate, zinc/calcium/ammonium acetate, magnesium/ammonium sulfate, potassium/ ammonium phosphate and tri-lithium/di-ammonium citrate. From the pH Clear Suite, precipitation was only observed when ammonium sulphate was used as precipitant.
Since only slight precipitation had appeared, it was decided to increase the concentration of the protein stock to 15.0 mg/ml and to set up Crystal Screen I and a home made version of the

Solubility Tool Kit (see appendix A.4 section A.4.1). Using the solubility tool kit, precipitation was observed when ammonium sulphate and ammonium citrate were used. However, no crystals were observed. From Crystal Screen I, precipitation was observed in approximately 75 % of the drops. After 10 days, needles appeared to grow from the precipitation of condition no. 47 (2.0 M ammonium sulphate, 0.1 M Na acetate pH 4.6). These needles grew to an approximate size of ca. 150 × 20 × 20 µm and they had a brown colour (see figure 11.1). The needles were present for almost four weeks and in this period of time the drop containing the needles was used for seeding purposes. However, within a period of two days the needles suddenly dissolved before the diffraction properties could be tested. When the needles dissolved, the drop had not been disturbed for two weeks. However, there had been problems with the thermostat. During the first three weeks of storage the temperature had been unstable ranging from 20 to 25ºC. at the time of dissolution the thermostat had been repaired and the temperature was at a constant 20 ± 1ºC. It is likely that the crystals were sensitive towards temperature.

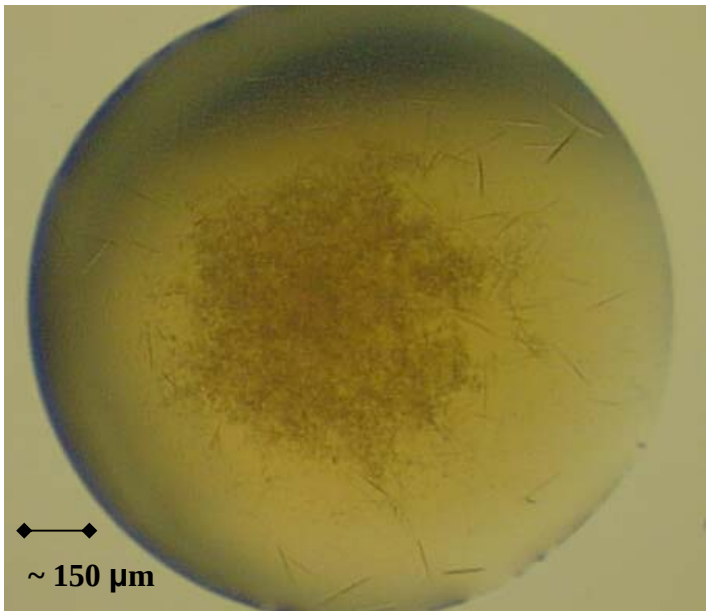


Figure 11.1 Needles observed during screening. The needles were observed when Crystal Screen I condition no. 47 (2.0 M ammonium sulphate, 0.1 M Na acetate pH 4.6) was applied. The drop was set up using 1 µl of reservoir solution and 1 µl of protein stock solution (15.0 mg/ml). The picture was taken a few days after the needles had appeared, when they were not fully grown.

11.2.3 Initial optimization

Optimization of the crystallization conditions (section 11.2.2) was attempted and the trials and observations of each condition are outlined in appendix A.4 section A.4.2.

In the first trial the pH (Na acetate pH 4.0 to 6.0) was tested against the concentration of the precipitant (1.8 to 2.4 M ammonium sulphate), but no crystals were observed.

The second trial was comparable to the first trial, but 2 % PEG 400 was added to the reservoir solution. In a few cases, seeding was also attempted. No crystals were obtained. Condition no. 47 (Crystal Screen I) was tested again with and without addition of 2 % PEG 400, and with

and without seeding, using two protein concentrations (8 conditions). This resulted in small needles (figure 11.2) when no PEG was added irrespective of seeding and protein stock.
In the third trial, the concentration of the buffer (50 mM to 150 mM Na acetate) was tested against pH (4.6, 5.0 and 6.0) at two ammonium sulphate concentrations (1.8 and 2.0 M) using two protein concentrations (15.0 and 20.0 mg/ml). In a few drops small needles were observed. This was generally observed using 100 mM Na acetate at pH 4.6 and a protein stock of 20.0 mg/ml.
Trial 4 exploited the effect of using Additive Screen (Hampton Research, conditions 1 to 24) in combination with a precipitant solution consisting of 2.0 M ammonium sulphate and 0.1 M Na acetate pH 4.6. In eight of the tested conditions small needles were observed (appendix A.4 section A.4.2, trial 4).
The large needles observed during screening (section 11.2.2) could not be reproduced. However, small needles were observed several times during the attempts to optimize the crystallization conditions.

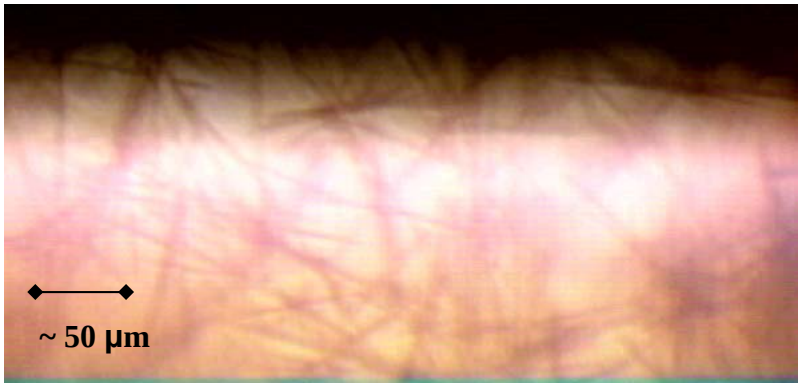


Figure 11.2 Example of the brown needles observed during initial optimization of the crystallization conditions.

11.2.4 Further optimization

Further optimization was attempted using a new batch of protein at a higher concentration (23.5 mg/ml). The trials and observations of each condition are outlined in appendix A.4 section A.4.3.
Since a new protein batch was used, the buffer pH (Na acetate pH 4.0 to 6.0) was tested against the concentration of the precipitant (1.6 to 2.4 M ammonium sulphate) again (trial 5). Needles appeared independent on the protein to precipitant volume ratio when the precipitant concentration was 1.6 M and the pH was 4.0. However, the largest needles appeared at the lowest protein to precipitant volume ratio (figure 11.3 A). It was decided not to lower the pH further since the FeS cluster is acid labile, but to test the precipitant concentration in the range from 1.2 M to 2.0 M at pH 4.0 and 4.6 and also test an even lower protein to precipitant volume ratio (trial 6). Needles only appeared at pH 4.0 in the precipitant concentration range; 1.4 to 1.6 M.
Since needles had been observed when The Additive Screen had been applied (section 11.2.3), it was decided to test seven different salts at pH 4.6 using two different precipitant concentrations; 1.8 and 2.0 M (trial 7). Li_2SO_4 resulted in needles at both precipitant

concentrations (figure 11.3 B). At 1.8 M needles were only observed at the low protein to precipitant volume ratio. $(NH_4)H_2PO_4$ and $ZnCl_2$ resulted in needles when the precipitant concentration was 2.0 M and the protein to precipitant volume ratio was low. It was decided to test the effect of Li_2SO_4 and $(NH_4)H_2PO_4$ at pH 4.0 and 4.6 in the precipitant concentration interval 1.2 M to 2.2 M (trial 8). At pH 4.0 Li_2SO_4 resulted in needles at a precipitation concentrations of 1.2 M and 1.4 M (figure 11.3 C). At pH 4.6 needles were observed at 1.8 M. At pH 4.0 $(NH_4)H_2PO_4$ resulted in needles at a precipitation concentration from 1.2 M to 1.6 M. At pH 4.6 needles were observed at 1.8 M.

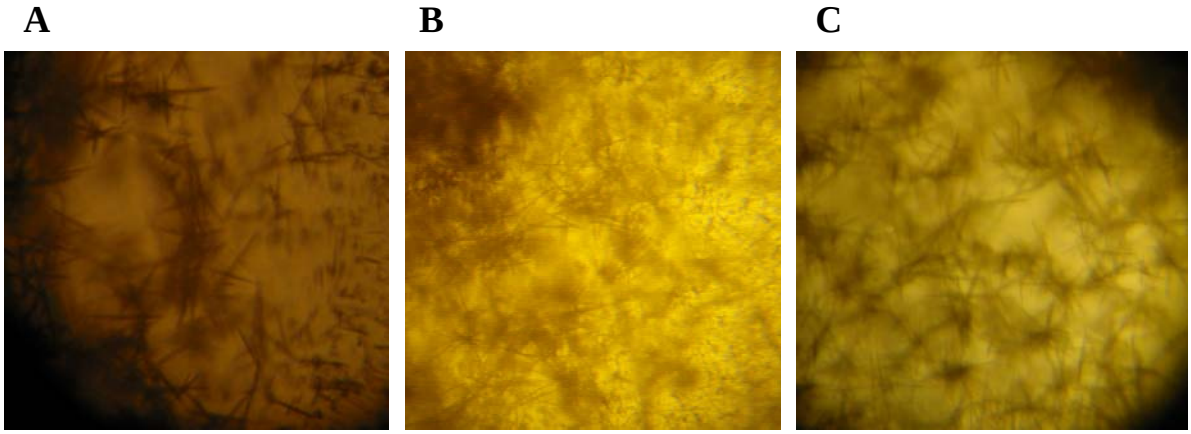


Figure 11.3 Examples of the brown needles observed during optimization of the crystallization conditions. In each of the panels the largest needles are approximately 80 µm long. A) The reservoir solution contained 1.6 M $(NH_4)_2SO_4$ and 0.1 M Na acetate/HCl pH 4.0. The drop was set up using 1 µl protein solution (23.5 mg/ml) and 2 µl reservoir solution. B) The reservoir solution contained 2.0 M $(NH_4)_2SO_4$, 0.1 M Na acetate/HCl pH 4.6 and 200 mM Li_2SO_4. The drop was set up using 1 µl protein solution (23.5 mg/ml) and 2 µl reservoir solution. C) The reservoir solution contained 1.2 M $(NH_4)_2SO_4$, 0.1 M Na acetate/HCl pH 4.0 and 200 mM Li_2SO_4. The drop was set up using 1.5 µl of protein (23.5 mg/ml) and reservoir solutions.

Due to the small size of the needles obtained during optimization, the diffraction properties were not tested. The needles were only found in drops containing moderate to heavy precipitation and in all cases they seemed to grow from the precipitation. The large single needles observed during screening (section 11.2.2) were not reproduced.
Precipitant concentration at which needles appear have been found but conditions should be further optimized to gain larger, single needles or three dimensional crystals suitable for diffraction experiments. Since the pH of the precipitant was low (pH 4.0 or 4.6) this might have caused partial degradation of the [4Fe-4S] cluster to a [3Fe-4S] cluster. It was attempted to dissolve the precipitation and needles in a few drops and then record the UV-vis spectrum to gain insight into cluster composition, but the obtained spectrum could not be convincingly interpreted (results not shown).

11.2.5 Monomer-dimer interconversion

It has been reported that a monomer/dimer equilibrium exists as a function of ionic strength for *P. furiosus* ferredoxin (Hasan *et al.*, 2002). At a low ionic strength (below 0.35 M NaCl, pH 7.8) the protein was predominantly found as a dimer, while at high ionic strengths (above

0.35 M NaCl, pH 7.8) the protein was present as a monomer. These observations were based on analytical gel filtrations since the elution volume of the protein was found to increase with increasing salt concentration compared to those of a reference protein (cytrochrome *c*) (Hasan *et al.*, 2002).

If a monomer/dimer equilibrium exists, this may greatly affect the crystallization properties since salts in various concentrations are often used as precipitants. Further, during equilibration of the crystallization drop, the compounds are up-concentrated which raises the ionic strength of the drop. If the protein in the crystallization drop is not homogeneous, then crystallization may prove impossible.

It was decided to verify the suggested monomer/dimer equilibrium by performing analytical gel filtration of *P. furiosus* [4Fe-4S] ferredoxin at different NaCl concentrations and pH values (section 11.1.3).

Figure 11.4 shows the obtained chromatograms when pH 8.0 was employed. It was observed that the elution volume increased with increasing NaCl concentrations signifying a decrease in size. Figure 11.5 presents a plot of the elution volume as a function of the NaCl concentration at the two tested pH values (4.6 and 8.0). At a given salt concentration the elution volume was highest at pH 4.6, again signifying a decrease of size.

Empirical considerations suggest that if a monomer/dimer equilibrium was present, then one would expect to see elution about two distinct elution volumes. One would also expect to see tailing from the elution point of one species toward the elution point of the other species or perhaps even two peaks if a salt concentration resulting in close to equal amounts of monomer and dimer was applied. However, this was not observed.

When performing gel filtration, the separation is affected by size and shape of the proteins. The radius of the molecule is not just determined by the amino acids constituting the protein, but also on the solvation shell/ionic radius of the protein. Therefore it is the hydrodynamic radius/volume, which in part determines the elution properties during gel filtration (Barth *et al.*, 1998).

The hydrodynamic radius of a given protein is highly dependent on the physicochemical properties of the solvent such as pH and ionic strength. It depends on the effective net charge of the protein (Piaggio *et al.*, 2002). It was found that the lower the charge of a particular protein, the lower was the resulting hydrodynamic radius. Further, when the ionic strength of a solution was increased (by adding a salt), then the resulting hydrodynamic radius decreased (Piaggio *et al.*, 2002). The obtained results from the *P. furiosus* ferredoxin (figures 11.4 and 11.5) fit the above stated observations. The ferredoxin is highly acidic and at pH 8.0 it has a high negative charge. At pH 4.6 the charge is low since the pH is close to pI of the protein (section 11.2.1).

Based on the results, it is unlikely that a monomer/dimer equilibrium exists. It is more likely that the changes observed in the elution volume when varying pH and salt concentration are due to changes of the hydrodynamic radius of the protein. The observations suggest that the ferredoxin is found merely as a dimer or merely as a monomer, but based on the elution profile it is difficult to establish whether the ferredoxin is in fact a dimer or a monomer.

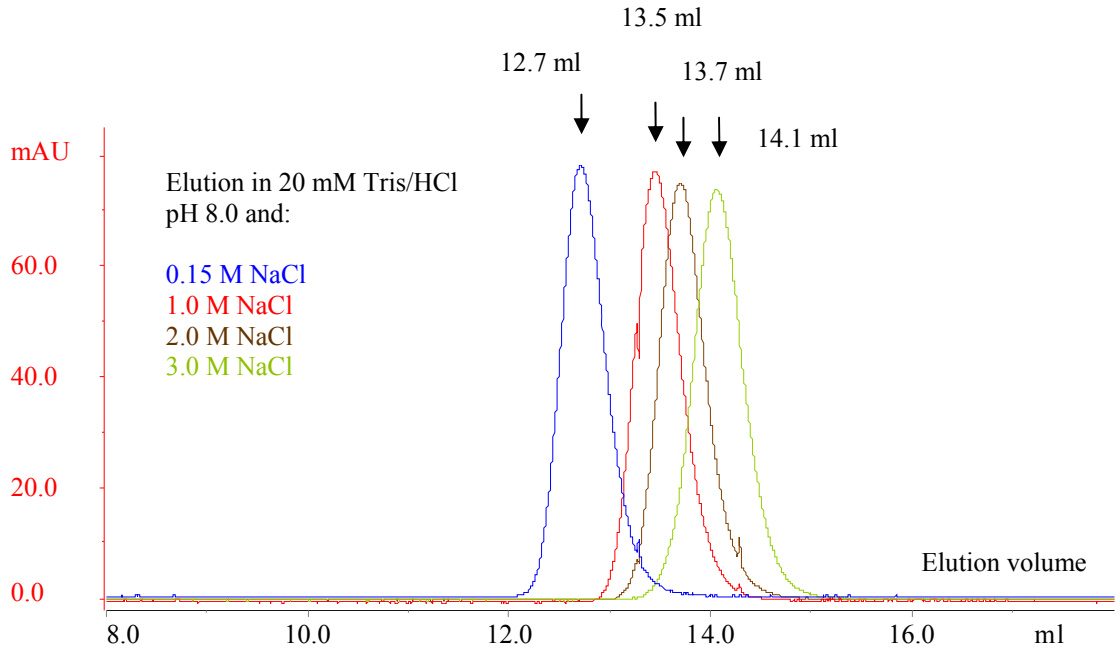


Figure 11.4 Chromatograms for analytical gel filtrations obtained when analysing elution of *P. furiosus* [4Fe-4S] ferredoxin at pH 8.0 (20 mM Tris/HCl) at various concentrations of NaCl. Absorption was measured at 390 nm. The blue, red, brown and green lines indicate overlaid chromatograms when 0.15 M, 1.0 M, 2.0 M and 3.0 M NaCl was applied, respectively.

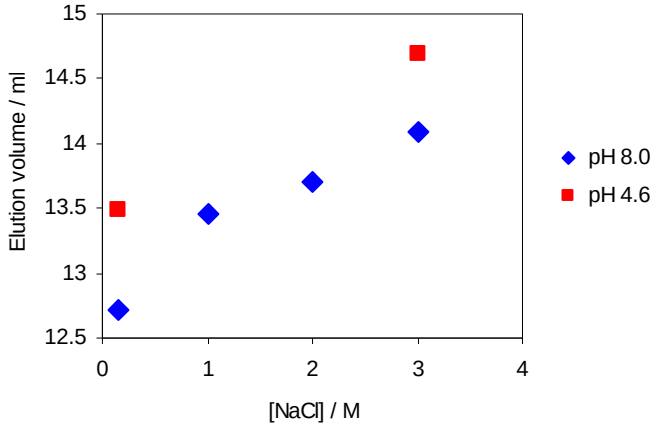


Figure 11.5 Elution volumes of *P. furiosus* [4Fe-4S] ferredoxin measured during analytical gel filtrations when different concentrations of NaCl were added to the elution buffer. Elution volumes were measured at two pH values: 4.6 (red) and 8.0 (blue). The elution buffers were 20 mM Tris/HCl pH 8.0 and 20 mM Na acetate/HCl pH 4.6.

11.2.6 Comments on crystallization

Based on the results it is clear that future crystallization trials should be carried out under strict anaerobic conditions due to oxidation of the FeS cluster. It appears to be quite a difficult task to obtain suitable crystals of the $[4Fe\text{-}4S]^{2+}$ ferredoxin even though anaerobic conditions were applied. One reason may be cluster degradation from a [4Fe-4S] to a [3Fe-4S] cluster

due to trace amounts of oxygen or low pH. Another reason however might be a monomer-dimer interconversion as discussed in section 11.2.5. Even though crystals (needles) have been obtained crystallization conditions are difficult to optimize and it appears that temperature plays an important role in the process of growing crystals.

11.3 Conclusion

It has been attempted to crystallize the $[4Fe\text{-}4S]^{2+}$ ferredoxin from *P. furiosus*. Crystals in the form of needles have been obtained and further optimization should be performed to gain crystals suitable for diffraction experiments. Reasons as to why the ferredoxin is difficult to crystallize have been discussed.

Chapter twelve

12 Concluding remarks and outlook

The first part of this thesis concerns the origin of life and iron-sulfur proteins. Three proteins have been investigated. Two of the proteins (Ori-ISP and LCA ferredoxin) originate from phylogenetic analysis of ferredoxins (Davis, 2002) and a theory concerning co-evolution of amino acids and the genetic code (Davis, 1999). The third protein is a [4Fe-4S] ferredoxin that originates from the hyperthermophilic archaeon *Pyrococcus furiosus*.

12.1 Reflections concerning the origin of life and Davis' theory

How the lifeless chemosphere of early Earth transformed into a living biosphere supporting thousands of different organisms has puzzled the human race for many years. Different theories on the origin of life have been put forward (chapter 3). The events preceding certain evolutionary steps might have been very complex and difficult to infer. A unanimous explanation of early evolution and evolutionary processes may never be unveiled.

The reconstruction and characterization of LCA ferredoxin and Ori-ISP is based on a theory by B. K. Davis (1999 and 2002), who is a supporter of an autotrophic origin of life situated at pyrite surfaces of hydrothermal vents. Thorough analysis of Davis' evolution theory and his phylogenetic analysis of ferredoxins is not within the scope of this thesis. However, some issues should be noted.

Among other important issues which Davis' theory (section 3.2) does not address are the origin of nucleotides, how they first polymerized and the handedness of amino acids. Neither does it take into account that some amino acids could have been added to the code and eliminated at a later stage in evolution. His theory provides an order of entrance of particular amino acids into the genetic code and hereby when they entered proteins. This is used when analysing the LCA ferredoxin to obtain the sequence of Ori-ISP.

Ancestral sequences for bacteria and archaea ferredoxins were constructed based on the Neighbour-Joining method. This is a distance-based method that provides no probabilities of the various residues in the sequences (Felsenstein, 2004). Davis recognises the internal repeat, which is a general feature of ferredoxins (section 2.3.3). To obtain the LCA ferredoxin, Davis takes this repeat into account, and hereby he invokes the element of personal choice in the resulting sequence since he cannot use probabilities as an argument. To go even further back, Davis only takes the residues of the repeat into account and a stage 10 peptide emerges. He finds that this peptide has a more ancient residue profile (only three residues have entered after stage 5) than extant ferredoxins, and then he argues that this peptide was in fact a stage 5 protein. Davis then forces mutations in these three positions (section 3.3.3) using the argument that the resultant residues were in fact present in the descendant ferredoxins. Again Davis has forced the element of personal choice on the sequence.

Based on the fact that an internal repeat is present in ferredoxins, it is highly likely that a short peptide in fact was the ancestor of ferredoxins, but I find it less likely that this sequence (or sequences) is in fact the sequence deduced by Davis. The reconstruction and characterization of Ori-ISP will however not prove Davis' theory, although it will make it more probable and thereby support the chemoautotrophic model for the origin of life.

12.2 Model system

Nonetheless, holo-Ori-ISP alone and in combination with holo-LCA ferredoxin provide a model system for learning more about the properties and the factors determining these properties of ferredoxins – extant as well as ancient. One interesting finding is the stability of the holo-proteins. LCA ferredoxin appears to be much more stable than dimeric Ori-ISP even though these proteins are very similar. They differ by the 6-residue insert, the 3-residue C-terminal tail and three mutations in each of the 23-residue repeats (figure 3.3). It would be interesting to produce the stage 10 protein and compare the properties.

Ori-ISP is believed to have been formed 3.5 billion years ago where the conditions on Earth were quite different from today, one of which was the absence of O_2. Hence, protection against O_2 was not necessary under the conditions which prevailed at that time. The negatively charged "attachment site" in the N-terminal region of Ori-ISP suggests that it might be able to bind to positively charged surfaces. Ori-ISP could have functioned on surfaces forming two-dimensional arrays providing directional function. It might have functioned in some of the early metabolic cycles as an electron carrier or as a catalytic site. Hence, Ori-ISP has plausibly played an important role in the very early stages of the development of life. An understanding of this protein and its chemical and biophysical properties might help define criteria for surface metabolism and the origin of biological life.

12.3 Ub-Ori-ISP and Ori-ISP

Cluster reconstitution into Ub-Ori-ISP and HPLC purification of the reaction products followed by characterization by analytical gel filtration and UV-vis spectrophotometry were successfully accomplished in anaerobic chambers (chapter 7).

When $FeCl_3$ was used as iron source, a monomeric $[4Fe\text{-}4S]^{2+}$ protein was eluted at low conductivities while a dimeric $[3Fe\text{-}4S]^{1+}$ protein or a protein containing two different clusters or multiple cluster types was eluted at higher conductivities. When Fe(II) was used as the iron source, only monomeric species were obtained. $[4Fe\text{-}4S]^{2+}$ protein was found to elute at low conductivity, while the reduced form, containing a $[4Fe\text{-}4S]^{1+}$ cluster, was found to elute at a higher conductivity.

A procedure for obtaining proteolytic cleavage of the Ub-Ori-ISP fusion protein followed by cluster reconstitution into Ori-ISP was developed. When using $FeCl_3$ as an iron source the obtained products were comparable to those obtained for Ub-Ori-ISP. The method has however not been optimized.

The cluster reconstitution reaction was optimized outside the anaerobic chamber using the Ar/N_2 flushing for obtaining anaerobic conditions. Much higher stability of the products was found when cluster reconstitution was performed inside the anaerobic chambers. The Ori-ISP system appears quite flexible concerning formation of monomers, dimers and the type of cluster obtained. Therefore the cluster reconstitution reaction should be further optimized to

control the outcome of reaction. Reaction times, molar excess of DTT, iron and sulfide are parameters that can be varied. In this connection the effects of iron source should also be investigated further.
It was found that even though purification and characterization was performed inside the anaerobic chambers degradation of the FeS cluster did seem to happen. Investigations into further stabilization of holo-protein should be performed. This might include addition of reducing agents and/or addition of glycerol to purification buffers.
Using a column with a higher resolution power during purification might be advantageous in exploring different species of the poorly resolved peaks that were found between main peaks.
To gain knowledge on the properties of the FeS cluster(s) (resulting spin, localization and delocalization patterns), EPR and Mössbauer spectroscopic studies should be performed on the obtained species. The redox properties of the holo-protein are also important in understanding what functions this protein could have undertaken. These properties can be investigated by cyclic voltammetry. The surface binding properties and organization of Ori-ISP can be studied using *in situ* scanning tunnelling microscopic techniques. Structural information is also important in order to understand the properties of the protein. However these may be difficult to obtain. Nuclear Magnetic Resonance analysis may provide some insight into solution structures. Crystallization of the holo-Ori-ISP may prove difficult due to the high flexibility of the protein and cluster degradation.

12.4 Ub-LCA and LCA ferredoxin

Two approaches for obtaining holo-LCA ferredoxin were investigated (chapter 10). In the first approach holo-Ub-LCA ferredoxin was expressed by *E. coli* cells, then the protein was purified and finally the Ub-fusion was cleaved off by YUH1 followed by further purification. The second approach was that holo-LCA ferredoxin was purified directly from *E. coli* cells co-expressing holo-Ub-LCA ferredoxin and the Ub-protease, UBP1.
The yields of purified holo-LCA ferredoxin obtained from either of the two methods were comparable, but very low. This was partly due to a low level of holo-protein expression in the *E. coli* cells and also due to low activity of YUH1 and UBP1 under the investigated conditions. Further, the first approach was much more time consuming than the second. Protein expression from the *E. coli* cells should be optimized to provide a higher yield of holo-protein. If the second approach for obtaining LCA ferredoxin is to be pursued further, then it should be investigated if the activity of UBP1 could be enhanced during the induction period (*in vivo*). If the first approach is to be pursued further, then *in vitro* cleavage of Ub-LCA ferredoxin using YUH1 should be optimized. It should be considered if it would be advantageous to express and purify UBP1 for *in vitro* cleavage.
Based on the obtained results the second approach is presently preferred. However, an alternative approach is to construct a third *E. coli* strain only expressing LCA ferredoxin.
Even though stability tests have not been performed, holo-LCA ferredoxin does appear to be sensitive towards O_2. However, holo-LCA ferredoxin is far more stable than holo-Ori-ISP.
Based on UV-vis spectrophotometric analysis it was found that both of the holo-proteins, Ub-LCA and LCA ferredoxin, contained at least one $[4Fe\text{-}4S]^{2+}$ cluster. To gain further knowledge on the LCA ferredoxin, it should be analysed employing the techniques suggested for Ori-ISP (section 12.3). The three-dimensional structure should be determined, which

could be obtained by crystallizing the protein followed by X-ray diffraction analysis. Finally the properties of holo-LCA ferredoxin should be compared to contemporary ferredoxins and those of holo-Ori-ISP.

12.5 *Pyrococcus furiosus* [4Fe-4S] ferredoxin

The $[4Fe\text{-}4S]^{2+}$ ferredoxin from *P. furiosus* was crystallized (chapter 11). However, the obtained needles need to be optimized before they are suitable for diffraction experiments. It was found that addition of salts appeared to advance crystal formation. Therefore, the effects of various salts should be further investigated. It may also prove advantageous to perform additive screening. It was argued (chapter 11) that the difficulties in crystallizing the ferredoxin may arise because of oxidation of the [4Fe-4S] cluster to a [3Fe-4S] cluster. It would be interesting to test the effects of adding a reducing agent and Fe(II) to the protein prior to performing crystallization trials. Temperature did also appear to be an important parameter for crystal growth, which needs to be further investigated. Finally, changing the buffer system and making some changes in the purification procedure may also change the crystallization properties of the enzyme.

Appendix one

A.1 Calibration curve for 13/31 Superdex 75 HR column

A calibration curve was determined for the 13/31 Superdex 75 HR column as described in the handbook, Gel Filtration, Principles and Methods. For calibration, the Low Molecular Weight Gel Filtration Calibration Kit (GE Healthcare, Amersham Biosciences) was applied.
The curve was prepared by plotting K_{av} for the standard proteins as a function of log M_w. K_{av} is defined as:

$$K_{av} = \frac{V_e - V_o}{V_t - V_o} \qquad \text{(A.1.1)}$$

V_o = column void volume = elution volume for Blue Dextran 2000
V_t = total bed volume
V_e = elution volume of the protein

V_t was determined from the dimensions of the column;
$V_t = 31 \text{ cm} \cdot \pi \cdot (0.65 \text{ cm})^2 = \underline{41.15 \text{ ml}}$

V_o and V_e values were determined as described in the following. Three samples containing the known standards (all dissolved in 20 mM Tris/HCl, 0.15 M NaCl pH 8.0) were prepared for analysis on the gel filtration column:

1) 1.0 mg/ml Blue Dextran 2000
2) 5.2 mg/ml Ovalbumin, 5.0 mg/ml Ribonuclease A
3) 4.8 mg/ml Chymotrypsinogen A, 5.7 mg/ml Albumin

250 µl of each of the three solutions were loaded onto the 20 mM Tris/HCl, 0.15 M NaCl pH 8.0 equilibrated column using a linear flow rate of 0.3 cm/min (chromatograms not shown). The volume of elution was determined for each of the five compounds and K_{av} values were calculated (presented in table A.1.1). K_{av} was plotted as a function of log M_w as presented in figure A.1.1 and a trend line was determined and presented in the figure. The R^2 value is not that good (the closer to one, the better the fit) and the calibration curve should be used with caution.
An approximate molecular weight can be determined for eluting species by first determining the volume of elution, then calculating K_{av} and finally use the equation for the trend line to calculate the molecular weight.

Table A.1.1 Elution points and K_{av} values for standards.

	V_0 / ml	V_e / ml	M_w / kDa	K_{av}
Blue Dextran 2000	7.86			
Ribonuclease A		13.63	13.7	0.17
Chymotrypsinogen A		12.56	25.0	0.14
Ovalbumin		10.52	43.0	0.08
Albumin		9.64	67.0	0.05

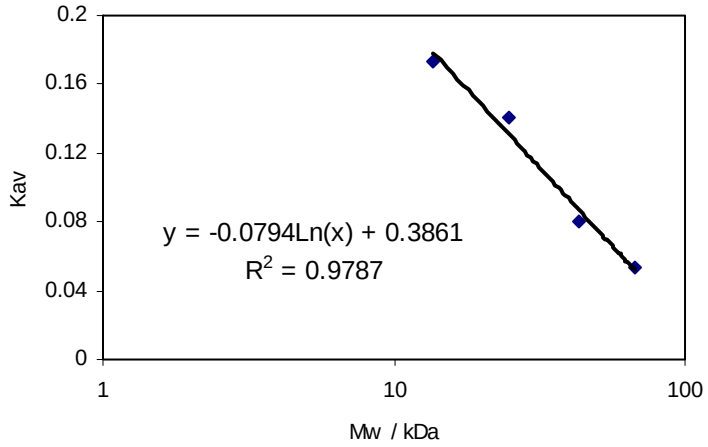


Figure A.1.1 Calibration curve for 13/31 Superdex 75 HR column. The trend line and the R-squared value is also shown.

Appendix two

A.2 Set up for crystallization of holo-Ub-Ori-ISP

Two experiments were set up for optimizing the most promising conditions obtained from the commercially available screens (chapter 8). Experiment 1 and 2 investigate the effects of ammonium sulphate and MPD, respectively. A third experiment was set up and it is a home made version of The Solubility Tool Kit (Molecular Dimensions) which applies the Hofmeiser series for testing the effects of different salts on solubility (Collins, 2004; Riès-Kautt and Ducruix, 1997). Here pH 8.0 was employed since this is the pH in which the protein was purified and known to be somewhat stable.

In each experiment the hanging drop vapour diffusion method was applied and the drop was equilibrated over 500 ml of reservoir solution. All drops were set up using 1 µl of protein solution and 1 µl of reservoir solution. The protein concentration was 100 mg/ml and the temperature during set up and storage was 20 ± 1ºC.

Experiment 1:

Number	Precipitant	Salt / additive	Buffer
1	2.0 M $(NH_4)_2SO_4$		100 mM MES pH 6.5
2	2.0 M $(NH_4)_2SO_4$		100 mM HEPES pH 7.0
3	2.0 M $(NH_4)_2SO_4$		100 mM HEPES pH 7.5
4	2.0 M $(NH_4)_2SO_4$		100 mM Tris/HCl pH 8.0
5	2.0 M $(NH_4)_2SO_4$		100 mM Tris/HCl pH 8.5
6	2.2 M $(NH_4)_2SO_4$		100 mM MES pH 6.5
7	2.2 M $(NH_4)_2SO_4$		100 mM HEPES pH 7.0
8	2.2 M $(NH_4)_2SO_4$		100 mM HEPES pH 7.5
9	2.2 M $(NH_4)_2SO_4$		100 mM Tris/HCl pH 8.0
10	2.2 M $(NH_4)_2SO_4$		100 mM Tris/HCl pH 8.5
11	2.4 M $(NH_4)_2SO_4$		100 mM MES pH 6.5
12	2.4 M $(NH_4)_2SO_4$		100 mM HEPES pH 7.0
13	2.4 M $(NH_4)_2SO_4$		100 mM HEPES pH 7.5
14	2.4 M $(NH_4)_2SO_4$		100 mM Tris/HCl pH 8.0
15	2.4 M $(NH_4)_2SO_4$		100 mM Tris/HCl pH 8.5
16	2.6 M $(NH_4)_2SO_4$		100 mM MES pH 6.5
17	2.6 M $(NH_4)_2SO_4$		100 mM HEPES pH 7.0
18	2.6 M $(NH_4)_2SO_4$		100 mM HEPES pH 7.5
19	2.6 M $(NH_4)_2SO_4$		100 mM Tris/HCl pH 8.0
20	2.6 M $(NH_4)_2SO_4$		100 mM Tris/HCl pH 8.5
21	2.8 M $(NH_4)_2SO_4$		100 mM MES pH 6.5
22	2.8 M $(NH_4)_2SO_4$		100 mM HEPES pH 7.0
23	2.8 M $(NH_4)_2SO_4$		100 mM HEPES pH 7.5
24	2.8 M $(NH_4)_2SO_4$		100 mM Tris/HCl pH 8.0
25	2.8 M $(NH_4)_2SO_4$		100 mM Tris/HCl pH 8.5

26	3.0 M $(NH_4)_2SO_4$		100 mM MES pH 6.5
27	3.0 M $(NH_4)_2SO_4$		100 mM HEPES pH 7.0
28	3.0 M $(NH_4)_2SO_4$		100 mM HEPES pH 7.5
29	3.0 M $(NH_4)_2SO_4$		100 mM Tris/HCl pH 8.0
30	3.0 M $(NH_4)_2SO_4$		100 mM Tris/HCl pH 8.5
31	2.0 M $(NH_4)_2SO_4$	2 % PEG 400	100 mM MES pH 6.5
32	2.0 M $(NH_4)_2SO_4$	2 % PEG 400	100 mM HEPES pH 7.5
33	2.0 M $(NH_4)_2SO_4$	2 % PEG 400	100 mM Tris/HCl pH 8.5
34	2.5 M $(NH_4)_2SO_4$	200 mM Li_2SO_4	100 mM HEPES pH 7.5
35	2.5 M $(NH_4)_2SO_4$	200 mM $MgSO_4$	100 mM HEPES pH 7.5
36	2.5 M $(NH_4)_2SO_4$	200 mM $MgCl_2$	100 mM HEPES pH 7.5

Experiment 2:

Number	Precipitant	Salt / additive	Buffer
1	40 % MPD		100 mM HEPES pH 7.0
2	40 % MPD		100 mM HEPES pH 7.5
3	40 % MPD		100 mM Tris/HCl pH 8.0
4	40 % MPD		100 mM Tris/HCl pH 8.5
5	45 % MPD		100 mM HEPES pH 7.0
6	45 % MPD		100 mM HEPES pH 7.5
7	45 % MPD		100 mM Tris/HCl pH 8.0
8	45 % MPD		100 mM Tris/HCl pH 8.5
9	50 % MPD		100 mM HEPES pH 7.0
10	50 % MPD		100 mM HEPES pH 7.5
11	50 % MPD		100 mM Tris/HCl pH 8.0
12	50 % MPD		100 mM Tris/HCl pH 8.5
13	55 % MPD		100 mM HEPES pH 7.0
14	55 % MPD		100 mM HEPES pH 7.5
15	55 % MPD		100 mM Tris/HCl pH 8.0
16	55 % MPD		100 mM Tris/HCl pH 8.5
17	60 % MPD		100 mM HEPES pH 7.0
18	60 % MPD		100 mM HEPES pH 7.5
19	60 % MPD		100 mM Tris/HCl pH 8.0
20	60 % MPD		100 mM Tris/HCl pH 8.5
21	45 % MPD	200 mM $MgCl_2$	100 mM Tris/HCl pH 8.0
22	50 % MPD	200 mM $MgCl_2$	100 mM Tris/HCl pH 8.0
23	50 % MPD	200 mM Li_2SO_4	100 mM Tris/HCl pH 8.0
24	50 % MPD	200 mM $MgSO_4$	100 mM Tris/HCl pH 8.0
25	50 % MPD	200 mM $CaCl_2$	100 mM Tris/HCl pH 8.0
26	55 % MPD	200 mM $MgCl_2$	100 mM Tris/HCl pH 8.0

Experiment 3:

Number	Precipitant	Salt / additive	Buffer
1	2.3 M $(NH_4)_2SO_4$		100 mM Tris/HCl pH 8.0
2	2.6 M $(NH_4)_2SO_4$		100 mM Tris/HCl pH 8.0
3	2.9 M $(NH_4)_2SO_4$		100 mM Tris/HCl pH 8.0
4	3.2 M $(NH_4)_2SO_4$		100 mM Tris/HCl pH 8.0
5	1.4 M (NH_4)-citrate		Adjusted to pH 8.0
6	1.7 M (NH_4)-citrate		Adjusted to pH 8.0

7	2.0 M (NH_4)-citrate		Adjusted to pH 8.0
8	2.3 M (NH_4)-citrate		Adjusted to pH 8.0
9	2.0 M Na-acetate		Adjusted to pH 8.0
10	2.5 M Na-acetate		Adjusted to pH 8.0
11	3.0 M Na-acetate		Adjusted to pH 8.0
12	3.5 M Na-acetate		Adjusted to pH 8.0
13	1.2 M Mg SO_4		100 mM Tris/HCl pH 8.0
14	1.4 M Mg SO_4		100 mM Tris/HCl pH 8.0
15	1.6 M Mg SO_4		100 mM Tris/HCl pH 8.0
16	1.8 M Mg SO_4		100 mM Tris/HCl pH 8.0
17	1.2 M $CaCl_2$		100 mM Tris/HCl pH 8.0
18	1.4 M $CaCl_2$		100 mM Tris/HCl pH 8.0
19	1.6 M $CaCl_2$		100 mM Tris/HCl pH 8.0
20	1.8 M $CaCl_2$		100 mM Tris/HCl pH 8.0

Appendix three

A.3 Chromatograms for large scale purification of Ub-LCA ferredoxin

This appendix presents typical chromatograms obtained for large scale purification of Ub-LCA ferredoxin as described in section 10.1.3. The chromatograms are from purifications on the 50/7 Q Sepharose Fast Flow column (figure A.3.1), the HiLoad 16/60 Superdex 75 prep

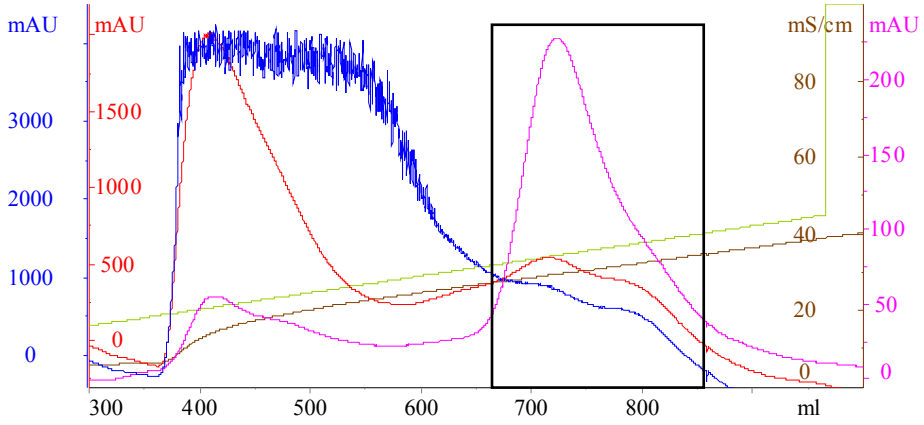


Figure A.3.1 Chromatogram for purification of Ub-LCA ferredoxin on the 50/7 Q Sepharose Fast Flow column. The collected fraction is indicated with a black box. The blue, red and pink lines indicate the absorbance at 236 nm, 280 nm and 390 nm, respectively. The brown and the green lines indicate the conductivity and the NaCl gradient, respectively. The concentration of NaCl can be read from the conductivity axis employing the unit 10^{-2} M.

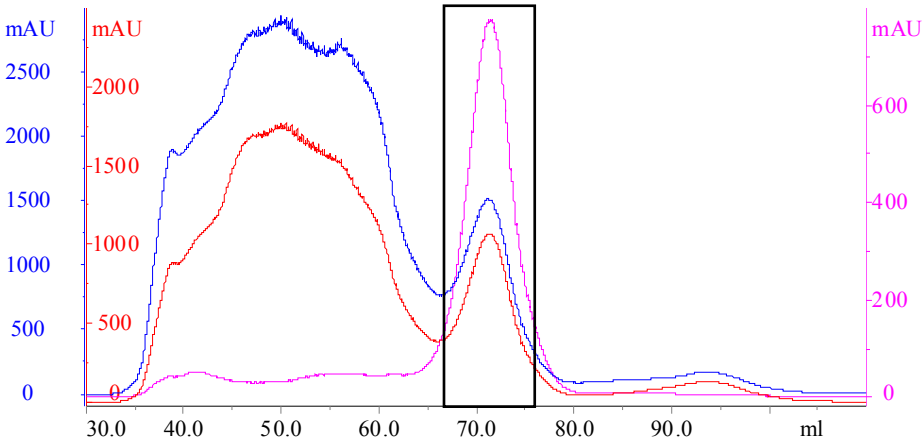


Figure A.3.2 Chromatogram for purification of Ub-LCA ferredoxin on the HiLoad 16/60 Superdex 75 prep grade column. The collected fraction is indicated with a black box. The blue, red and the pink lines indicate the absorbance at 236 nm, 280 nm and 390 nm, respectively.

grade column (figure A.3.2) and the 16/10 Source 30Q column (figure A.3.3). The purity of the obtained product, Ub-LCA ferredoxin, was tested by SDS-PAGE and the results are presented in figure A.3.4. The obtained results are discussed in chapter 10.

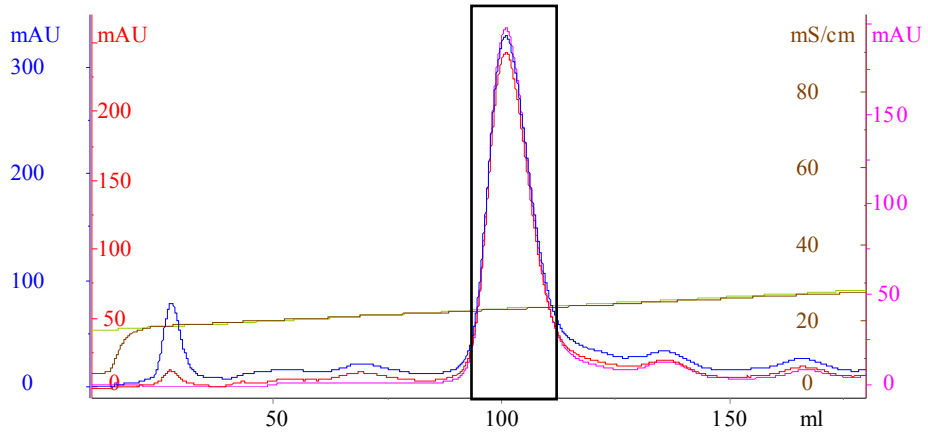


Figure A.3.3 Chromatogram for purification of Ub-LCA ferredoxin on the 16/10 Source 30Q column. The collected fraction is indicated with a black box. The blue, red and pink lines indicate the absorbance at 236 nm, 280 nm and 390 nm, respectively. The brown and the green lines indicate the conductivity and the NaCl gradient, respectively. The concentration of NaCl can be read from the conductivity axis employing the unit 10^{-2} M.

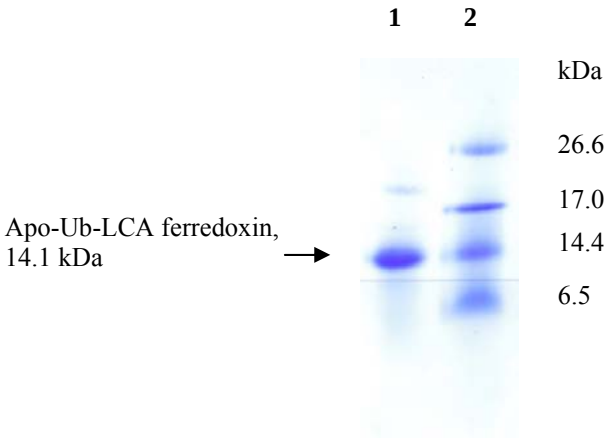


Figure A.3.4 Purity of Ub-LCA ferredoxin tested by SDS-PAGE. Lane 1: 5 µl of 2.3 mg/ml LCA. Lane 2: Marker, 1 µl.

Appendix four

A.4 Crystallization of *P. furiosus* [4Fe-4S] ferredoxin

This appendix presents conditions used for screening and optimizing crystallization of *P. furiosus* [4Fe-4S] ferredoxin (elaborated from chapter 11). In each of the outlined trials, the hanging drop vapour diffusion method was applied and each drop was equilibrated over 500 ml of reservoir solution. Observations made for each experiment are also outlined using the following abbreviations:

P1	Slight precipitate
P2	Moderate precipitate
P3	Heavy precipitate
F	Phase separation
D	Denaturation
S	Skin covering the drop
N	Needles
K	Quasi-crystalline material

A.4.1 Screening

A home made version of The Solubility Tool Kit (Molecular Dimensions) which applies the Hofmeiser series for testing the effects of different salts on solubility (Collins, 2004; Riès-Kautt and Ducruix, 1997) was set up. The set up was performed in an anaerobic chamber (section 4.2.1). The temperature during set up and storage was 20 ± 1ºC. 1 µl of protein stock solution (15.0 mg/ml) was mixed with 1 µl of reservoir solution.

No.	Precipitant	Buffer	Observations
1	2.3 M $(NH_4)_2SO_4$	100 mM Tris/HCl pH 8.0	
2	2.6 M $(NH_4)_2SO_4$	100 mM Tris/HCl pH 8.0	P1
3	2.9 M $(NH_4)_2SO_4$	100 mM Tris/HCl pH 8.0	P1
4	3.2 M $(NH_4)_2SO_4$	100 mM Tris/HCl pH 8.0	P2
5	1.4 M (NH_4)-citrate	Adjusted to pH 8.0	P1
6	1.7 M (NH_4)-citrate	Adjusted to pH 8.0	P2
7	2.0 M (NH_4)-citrate	Adjusted to pH 8.0	P2
8	2.3 M (NH_4)-citrate	Adjusted to pH 8.0	P2
9	2.0 M Na-acetate	Adjusted to pH 8.0	
10	2.5 M Na-acetate	Adjusted to pH 8.0	
11	3.0 M Na-acetate	Adjusted to pH 8.0	
12	3.5 M Na-acetate	Adjusted to pH 8.0	
13	1.2 M Mg SO_4	100 mM Tris/HCl pH 8.0	
14	1.4 M Mg SO_4	100 mM Tris/HCl pH 8.0	

15	1.6 M Mg SO_4	100 mM Tris/HCl pH 8.0	
16	1.8 M Mg SO_4	100 mM Tris/HCl pH 8.0	
17	1.2 M $CaCl_2$	100 mM Tris/HCl pH 8.0	
18	1.4 M $CaCl_2$	100 mM Tris/HCl pH 8.0	
19	1.6 M $CaCl_2$	100 mM Tris/HCl pH 8.0	
20	1.8 M $CaCl_2$	100 mM Tris/HCl pH 8.0	

A.4.2 Initial optimization

These trials were performed in an anaerobic chamber (section 4.2.1). The temperature during set up and storage was 20 ± 1ºC.

Trial 1:

1 µl of protein stock solution (15.0 mg/ml) was mixed with 1 µl of reservoir solution.

No.	Precipitant	Buffer	Observations
1	1.8 M $(NH_4)_2SO_4$	0.1 M Na acetate/HCl pH 4.0	P3, S, D
2	2.0 M $(NH_4)_2SO_4$	0.1 M Na acetate/HCl pH 4.0	P3, S, D
3	2.2 M $(NH_4)_2SO_4$	0.1 M Na acetate/HCl pH 4.0	P3, S, D
4	2.4 M $(NH_4)_2SO_4$	0.1 M Na acetate/HCl pH 4.0	P3, S, D
5	1.8 M $(NH_4)_2SO_4$	0.1 M Na acetate/HCl pH 4.6	P3, S
6	2.0 M $(NH_4)_2SO_4$	0.1 M Na acetate/HCl pH 4.6	P3, S
7	2.2 M $(NH_4)_2SO_4$	0.1 M Na acetate/HCl pH 4.6	P3, S
8	2.4 M $(NH_4)_2SO_4$	0.1 M Na acetate/HCl pH 4.6	P3, S
9	1.8 M $(NH_4)_2SO_4$	0.1 M Na acetate/HCl pH 5.0	P2
10	2.0 M $(NH_4)_2SO_4$	0.1 M Na acetate/HCl pH 5.0	P2
11	2.2 M $(NH_4)_2SO_4$	0.1 M Na acetate/HCl pH 5.0	P2, S
12	2.4 M $(NH_4)_2SO_4$	0.1 M Na acetate/HCl pH 5.0	P2, S
13	1.8 M $(NH_4)_2SO_4$	0.1 M Na acetate/HCl pH 6.0	P1
14	2.0 M $(NH_4)_2SO_4$	0.1 M Na acetate/HCl pH 6.0	P1
15	2.2 M $(NH_4)_2SO_4$	0.1 M Na acetate/HCl pH 6.0	P1
16	2.4 M $(NH_4)_2SO_4$	0.1 M Na acetate/HCl pH 6.0	P1

Trial 2:

1 µl of protein stock solution was mixed with 1 µl of reservoir solution. Two different protein stock concentrations (Conc. presented as mg/ml) were used.

No.	Conc.	Precipitant	Additive	Buffer	Observations
1	15	1.6 M $(NH_4)_2SO_4$	2 % PEG 400	0.1 M Na acetate/HCl pH 4.6	P2, S
2	15	1.8 M $(NH_4)_2SO_4$	2 % PEG 400	0.1 M Na acetate/HCl pH 4.6	P2
3	15	2.0 M $(NH_4)_2SO_4$	2 % PEG 400	0.1 M Na acetate/HCl pH 4.6	P3
4	15	2.2 M $(NH_4)_2SO_4$	2 % PEG 400	0.1 M Na acetate/HCl pH 4.6	P3, F
5	15	1.6 M $(NH_4)_2SO_4$	2 % PEG 400	0.1 M Na acetate/HCl pH 5.0	P1
6	15	1.8 M $(NH_4)_2SO_4$	2 % PEG 400	0.1 M Na acetate/HCl pH 5.0	P2
7	15	2.0 M $(NH_4)_2SO_4$	2 % PEG 400	0.1 M Na acetate/HCl pH 5.0	P2
8	15	2.2 M $(NH_4)_2SO_4$	2 % PEG 400	0.1 M Na acetate/HCl pH 5.0	P2
9	15	1.6 M $(NH_4)_2SO_4$	2 % PEG 400	0.1 M Na acetate/HCl pH 6.0	
10	15	1.8 M $(NH_4)_2SO_4$	2 % PEG 400	0.1 M Na acetate/HCl pH 6.0	P1
11	15	2.0 M $(NH_4)_2SO_4$	2 % PEG 400	0.1 M Na acetate/HCl pH 6.0	P1
12	15	2.2 M $(NH_4)_2SO_4$	2 % PEG 400	0.1 M Na acetate/HCl pH 6.0	P2
13	15	1.6 M $(NH_4)_2SO_4$	2 % PEG 400	0.1 M HEPES pH 7.5	

14	15	1.8 M $(NH_4)_2SO_4$	2 % PEG 400	0.1 M HEPES pH 7.5	
15	15	2.0 M $(NH_4)_2SO_4$	2 % PEG 400	0.1 M HEPES pH 7.5	P1
16	15	2.2 M $(NH_4)_2SO_4$	2 % PEG 400	0.1 M HEPES pH 7.5	P1
17*	15	1.8 M $(NH_4)_2SO_4$	2 % PEG 400	0.1 M Na acetate/HCl pH 4.6	P2
18*	15	1.8 M $(NH_4)_2SO_4$	2 % PEG 400	0.1 M Na acetate/HCl pH 5.0	P2
19*	15	1.8 M $(NH_4)_2SO_4$	2 % PEG 400	0.1 M Na acetate/HCl pH 6.0	P1
20*	15	1.8 M $(NH_4)_2SO_4$	2 % PEG 400	0.1 M HEPES pH 7.5	P1
21*	15 / 20	Crystal Screen I (Hampton Research) no. 47 + 2 % PEG 400			P3 / P3
22	15 / 20	Crystal Screen I (Hampton Research) no. 47 + 2 % PEG 400			P3 / P3
23*	15 / 20	Crystal Screen I (Hampton Research) no. 47			P3, N / P3, N
24	15 / 20	Crystal Screen I (Hampton Research) no. 47			P3, N / P3, N

* Seeding from drop containing needles; obtained using Crystal Screen I condition no. 47 (Hampton Research) during screening.

Trial 3:

2 µl of protein stock solution was mixed with 2 µl of reservoir solution. Two different protein stock concentrations 15.0 mg/ml and 20.0 mg/ml were tested for each condition.

No.	Precipitant	Buffer	Observations made at:	
			15 mg/ml	20 mg/ml
1	1.8 M $(NH_4)_2SO_4$	150 mM Na acetate/HCl pH 4.6	P3, S	P3, S
2	1.8 M $(NH_4)_2SO_4$	100 mM Na acetate/HCl pH 4.6	P3, S	P3, S, N
3	1.8 M $(NH_4)_2SO_4$	50 mM Na acetate/HCl pH 4.6	P3, S	P3, S, N
4	1.8 M $(NH_4)_2SO_4$	150 mM Na acetate/HCl pH 5.0	P2	P3, S
5	1.8 M $(NH_4)_2SO_4$	100 mM Na acetate/HCl pH 5.0	P2, N	P3, S
6	1.8 M $(NH_4)_2SO_4$	50 mM Na acetate/HCl pH 5.0	P2,	P3, S
7	1.8 M $(NH_4)_2SO_4$	150 mM Na acetate/HCl pH 6.0	P1	P2
8	1.8 M $(NH_4)_2SO_4$	100 mM Na acetate/HCl pH 6.0	P1	P2
9	1.8 M $(NH_4)_2SO_4$	50 mM Na acetate/HCl pH 6.0	P1	P2
10	2.0 M $(NH_4)_2SO_4$	150 mM Na acetate/HCl pH 4.6	P3, S	P3, S
11	2.0 M $(NH_4)_2SO_4$	100 mM Na acetate/HCl pH 4.6	P3, S	P3, S, N
12	2.0 M $(NH_4)_2SO_4$	50 mM Na acetate/HCl pH 4.6	P3, S	P3, S
13	2.0 M $(NH_4)_2SO_4$	150 mM Na acetate/HCl pH 5.0	P2	P3, S
14	2.0 M $(NH_4)_2SO_4$	100 mM Na acetate/HCl pH 5.0	P2	P3, S
15	2.0 M $(NH_4)_2SO_4$	50 mM Na acetate/HCl pH 5.0	P2	P3, S
16	2.0 M $(NH_4)_2SO_4$	150 mM Na acetate/HCl pH 6.0	P1	P2
17	2.0 M $(NH_4)_2SO_4$	100 mM Na acetate/HCl pH 6.0	P1	P2
18	2.0 M $(NH_4)_2SO_4$	50 mM Na acetate/HCl pH 6.0	P1	P2

Trial 4:

The first 24 conditions of Additive Screen (Hampton Research) were tested using the conditions at which needles had appeared during screening. 2 µl of protein stock solution (15.0 mg/ml) was mixed with 2.5 µl of reservoir solution and 0.5 µl of the additive.

No.	Reservoir solution		Additive (concentrations prior to mixing)	Observations
	Precipitant	Buffer		
1	2.0 M $(NH_4)_2SO_4$	0.1 M Na acetate/ HCl pH 4.6	0.1 M Barium chloride dihydrate	P3
2	Do	Do	0.1 M Cadmium chloride hydrate	P3, N
3	Do	Do	0.1 M Calcium chloride dihydrate	P3
4	Do	Do	0.1 M Cobalt(II) chloride hexahydrate	P3

5	Do	Do	0.1 M Copper(II) chloride dihydrate	P3, S
6	Do	Do	0.1 M Magnesium chloride hexahydrate	P3
7	Do	Do	0.1 M Manganese(II) chloride tetrahydrate	P3
8	Do	Do	0.1 M Strontium chloride hexahydrate	P3
9	Do	Do	0.1 M Yttrium(III) chloride hexahydrate	P3, N
10	Do	Do	0.1 M Zinc chloride	P3, N
11	Do	Do	0.1 M Iron(III) chloride hexahydrate	P3
12	Do	Do	0.1 M Nickel(II) chloride hexahydrate	P3
13	Do	Do	0.1 M Chromium(III) chloride hexahydrate	P3
14	Do	Do	0.1 M Praseodymium(III) acetate hydrate	P3
15	Do	Do	1.0 M Ammonium sulfate	P3, N
16	Do	Do	1.0 M Potassium chloride	P3, N
17	Do	Do	1.0 M Lithium chloride	P3, S, N
18	Do	Do	2.0 M Sodium chloride	P3
19	Do	Do	0.5 M Sodium fluoride	P3, N
20	Do	Do	1.0 M Sodium iodide	P3, F
21	Do	Do	2.0 M Sodium thiocyanate	P3
22	Do	Do	1.0 M Potassium sodium tartrate tetrahydrate	P3
23	Do	Do	1.0 M Sodium citrate tribasic dihydrate	P3, N
24	Do	Do	1.0 M Cecium chloride	P3

A.4.3 Further optimization

These trials were performed in an anaerobic chamber (section 4.2.2) at room temperature. A protein stock concentration of 23.5 mg/ml was applied. During set up, mixing of protein solution and reservoir solution was performed at three different volume ratios (protein:reservoir); 1:1 (total drop volume of 3 µl), 1:2 (total drop volume of 3 µl) and 1:3 (total drop volume of 4 µl). Not all three ratios were applied in each trial.

Trial 5:

No.	Precipitant	Buffer	Observations made at protein to reservoir volume:	
			1:2	1:1
1	1.6 M $(NH_4)_2SO_4$	0.1 M Na acetate/HCl pH 4.0	P2, N	P3, K
2	1.8 M $(NH_4)_2SO_4$	0.1 M Na acetate/HCl pH 4.0	P2	P3, F, K
3	2.0 M $(NH_4)_2SO_4$	0.1 M Na acetate/HCl pH 4.0	P3, S	P3, F, K, S
4	2.2 M $(NH_4)_2SO_4$	0.1 M Na acetate/HCl pH 4.0	P3, S, K	P3, F, K, S, D
5	1.6 M $(NH_4)_2SO_4$	0.1 M Na acetate/HCl pH 4.6	P1	P
6	1.8 M $(NH_4)_2SO_4$	0.1 M Na acetate/HCl pH 4.6	P1	P2
7	2.0 M $(NH_4)_2SO_4$	0.1 M Na acetate/HCl pH 4.6	P2, K, S, F	P3, K, F, S
8	2.2 M $(NH_4)_2SO_4$	0.1 M Na acetate/HCl pH 4.6	P3, F	P3, K, F, S
9	1.6 M $(NH_4)_2SO_4$	0.1 M Na acetate/HCl pH 5.0		
10	1.8 M $(NH_4)_2SO_4$	0.1 M Na acetate/HCl pH 5.0	P1	P1
11	2.0 M $(NH_4)_2SO_4$	0.1 M Na acetate/HCl pH 5.0	P2	P2
12	2.2 M $(NH_4)_2SO_4$	0.1 M Na acetate/HCl pH 5.0	P3	P3
13	1.6 M $(NH_4)_2SO_4$	0.1 M Na acetate/HCl pH 6.0		
14	1.8 M $(NH_4)_2SO_4$	0.1 M Na acetate/HCl pH 6.0		
15	2.0 M $(NH_4)_2SO_4$	0.1 M Na acetate/HCl pH 6.0	P1	P1
16	2.2 M $(NH_4)_2SO_4$	0.1 M Na acetate/HCl pH 6.0	P1	P1

Trial 6:

No.	Precipitant	Buffer	Observations made at protein to reservoir volume:		
			1:3	1:2	1:1
1	1.2 M $(NH_4)_2SO_4$	0.1 M Na acetate/HCl pH 4.0		P1	P2, S
2	1.4 M $(NH_4)_2SO_4$	0.1 M Na acetate/HCl pH 4.0	P1, N	P2, N	P3, N F, K
3	1.5 M $(NH_4)_2SO_4$	0.1 M Na acetate/HCl pH 4.0	P1, N	P2, N	P3, N, F, K
4	1.6 M $(NH_4)_2SO_4$	0.1 M Na acetate/HCl pH 4.0	P2, N	P2, N, K	P3, F, K
5	1.7 M $(NH_4)_2SO_4$	0.1 M Na acetate/HCl pH 4.0	P2	P2, K	P3, F, K
6	1.8 M $(NH_4)_2SO_4$	0.1 M Na acetate/HCl pH 4.0	P2	P2, K	P3, F, K
7	1.9 M $(NH_4)_2SO_4$	0.1 M Na acetate/HCl pH 4.0	P2	P2, K, F	P3, F, K
8	2.0 M $(NH_4)_2SO_4$	0.1 M Na acetate/HCl pH 4.0	P3, K	P3, K, F	P3, F, K
9	1.2 M $(NH_4)_2SO_4$	0.1 M Na acetate/HCl pH 4.6			P1
10	1.4 M $(NH_4)_2SO_4$	0.1 M Na acetate/HCl pH 4.6		P1	P1
11	1.5 M $(NH_4)_2SO_4$	0.1 M Na acetate/HCl pH 4.6	P1	P1	P1
12	1.6 M $(NH_4)_2SO_4$	0.1 M Na acetate/HCl pH 4.6	P1	P1	P1
13	1.7 M $(NH_4)_2SO_4$	0.1 M Na acetate/HCl pH 4.6	P1	P1	P2
14	1.8 M $(NH_4)_2SO_4$	0.1 M Na acetate/HCl pH 4.6	P2	P2	P3, F, K
15	1.9 M $(NH_4)_2SO_4$	0.1 M Na acetate/HCl pH 4.6	P2	P2	P3, F, K
16	2.0 M $(NH_4)_2SO_4$	0.1 M Na acetate/HCl pH 4.6	P2	P3, F, K	P3

Trial 7:

No.	Precipitant	Buffer	Additive	Observations made at protein to reservoir volume:	
				1:2	1:1
1	2.0 M $(NH_4)_2SO_4$	0.1 M Na acetate/HCl pH 4.6	0.2 M Li_2SO_4	P2, F, K, N	P3, F, K, N
2	2.0 M $(NH_4)_2SO_4$	Do	10 mM $ZnCl_2$	P2, N, K	P3, F
3	2.0 M $(NH_4)_2SO_4$	Do	10 mM Zn acetate	P2	P3
4	2.0 M $(NH_4)_2SO_4$	Do	10 mM Mg acetate	P2	P3
5	2.0 M $(NH_4)_2SO_4$	Do	0.2 M $(NH_4)H_2PO_4$	P3, F, K, N	P3, F, D
6	2.0 M $(NH_4)_2SO_4$	Do	10 mM $MgCl_2$	P2	P3
7	2.0 M $(NH_4)_2SO_4$	Do	0.2 M NH_4 acetate	P2	P3
8	1.8 M $(NH_4)_2SO_4$	Do	0.2 M Li_2SO_4	P2, N	P3, F, K
9	1.8 M $(NH_4)_2SO_4$	Do	10 mM $ZnCl_2$	P1	P2
10	1.8 M $(NH_4)_2SO_4$	Do	10 mM Zn acetate	P1	P2
11	1.8 M $(NH_4)_2SO_4$	Do	10 mM Mg acetate	P1	P2
12	1.8 M $(NH_4)_2SO_4$	Do	0.2 M $(NH_4)H_2PO_4$	P2, F, K, D	P2, F, D
13	1.8 M $(NH_4)_2SO_4$	Do	10 mM $MgCl_2$	P1	P2
14	1.8 M $(NH_4)_2SO_4$	Do	0.2 M NH_4 acetate	P1	P2

Trial 8:

No.	Precipitant	Buffer: 100 mM Na acetate	Additive	Observations made at protein to reservoir volume:		
				1:3	1:2	1:1
1	1.2 M $(NH_4)_2SO_4$	pH 4.0	0.2 M Li_2SO_4	P2, N	P3, N	P3, N, F, K
2	1.4 M $(NH_4)_2SO_4$	pH 4.0	0.2 M Li_2SO_4	P2, N	P3, N	P3, F, K
3	1.6 M $(NH_4)_2SO_4$	pH 4.0	0.2 M Li_2SO_4	P2	P3, D	P3, F, K, D
4	1.8 M $(NH_4)_2SO_4$	pH 4.0	0.2 M Li_2SO_4	P2	P3, F, K, D	P3, F, K, D
5	2.0 M $(NH_4)_2SO_4$	pH 4.0	0.2 M Li_2SO_4	P2	P3, F, K, D	P3, F, K, D
6	2.2 M $(NH_4)_2SO_4$	pH 4.0	0.2 M Li_2SO_4	P3, D	P3, D	P3, F, D

7	1.2 M $(NH_4)_2SO_4$	pH 4.6	0.2 M Li_2SO_4		P	P
8	1.4 M $(NH_4)_2SO_4$	pH 4.6	0.2 M Li_2SO_4	P1	P1	P1
9	1.6 M $(NH_4)_2SO_4$	pH 4.6	0.2 M Li_2SO_4	P1	P1	P2
10	1.8 M $(NH_4)_2SO_4$	pH 4.6	0.2 M Li_2SO_4	P1	P2, N, F, K	P3, N, F, K
11	2.0 M $(NH_4)_2SO_4$	pH 4.6	0.2 M Li_2SO_4	P2	P2, F, K	P3, F
12	2.2 M $(NH_4)_2SO_4$	pH 4.6	0.2 M Li_2SO_4	P3	P3, F, D	P3, F, D
13	1.2 M $(NH_4)_2SO_4$	pH 4.0	0.2 M $(NH_4)H_2PO_4$	P1	P2, N	P3, N, K
14	1.4 M $(NH_4)_2SO_4$	pH 4.0	0.2 M $(NH_4)H_2PO_4$	P2, N	P2, F, K	P3, F, K
15	1.6 M $(NH_4)_2SO_4$	pH 4.0	0.2 M $(NH_4)H_2PO_4$	P2, N	P2, F, K	P3, F, K
16	1.8 M $(NH_4)_2SO_4$	pH 4.0	0.2 M $(NH_4)H_2PO_4$	P2, D	P3, D	P3, F, K, D
17	2.0 M $(NH_4)_2SO_4$	pH 4.0	0.2 M $(NH_4)H_2PO_4$	P2, D	P3, F, K, D	P3, F
18	2.2 M $(NH_4)_2SO_4$	pH 4.0	0.2 M $(NH_4)H_2PO_4$	P3, D	P3, K, F, D	P3, F, D
19	1.2 M $(NH_4)_2SO_4$	pH 4.6	0.2 M $(NH_4)H_2PO_4$	P1	P1	P1
20	1.4 M $(NH_4)_2SO_4$	pH 4.6	0.2 M $(NH_4)H_2PO_4$	P1	P1	P1
21	1.6 M $(NH_4)_2SO_4$	pH 4.6	0.2 M $(NH_4)H_2PO_4$	P1	P1	P2
22	1.8 M $(NH_4)_2SO_4$	pH 4.6	0.2 M $(NH_4)H_2PO_4$	P2, N	P2, F, K	P3 F, K ,D
23	2.0 M $(NH_4)_2SO_4$	pH 4.6	0.2 M $(NH_4)H_2PO_4$	P2, F	P3, F	P3, F, D
24	2.2 M $(NH_4)_2SO_4$	pH 4.6	0.2 M $(NH_4)H_2PO_4$	P3, D	P3, F, D	P3, F, D

Bibliography

Altschul S.F., Gish W., Miller W., Myers E.W. and Lipman D.J., Basic local alignment search tool, *J. Mol. Biol.* 215 (1990) 403-410.

Aono S., Bryant F.O. and Adams M. W. W., A novel and remarkably thermostable ferredoxin from the hyperthermophilic achebacterium *Pyrococcus furiosus*, *J. Bacteriol.* 171 (1989) 3433-3439.

Barth H. G., Boyes B. E and Jackson C. Size exclusion chromatography and related separation techniques, *Anal. Chem.* 70 (1998) 251R-278R.

Barthelmes J., Ebeling C., Chang A., Schomburg I. and Schomburg D., BRENDA, AMENDA and FRENDA: the enzyme information system 2007, *Nucleic Acids Res.* (Database issue) 35 (2007) D511-D5114.

Beinert H., Holm R. H. and Münck E., Iron-sulfur clusters: nature's modular multipurpose structures, *Science* 277 (1997) 653-659.

Beinert H., Iron-sulfur proteins: ancient structures, still full of surprises, *J. Biol. Inorg. Chem.*, 5 (2000) 2-15.

Bentrop D., Capozzi F., and Luchinat C., Iron-Sulfur Proteins, Handbook on metalloproteins, p 357-460, edited by Bertini I., Siegel A., and Siegel H., Dekker, NY (2001).

Berry S., The chemical basis of membrane bioenergetics, *J. Mol. Evol.* 54 (2002) 595-613.

Bertini I., Donaire A., Feinberg B. A., Luchinat C., Piccioli M. and Yuan H., Solution structure of the oxidized 2[4Fe-4S] ferredoxin from *Clostridium pasteurianum*, *Eur. J. Biochem.* 232 (1995) 192-205.

Bertini I., Gray H. B., Stiefel E. I. and Valentive J. S., Biological inorganic chemistry: structure and reactivity (1st edition), University Science Books, USA, (2007) (ISBN: 1891389432).

Bian S. and Cowan J. A., Protein-bound iron-sulfur centers. Form, function and assembly, *Coord. Chem. Rev.*, 190-192 (1999) 1049-1066.

Bjellqvist B.,Hughes G. J., Pasquali C., Paquet N., Ravier F., Sanchez J.-C., Frutiger S. and Hochstrasser D.F. The focusing positions of polypeptides in immobilized pH gradients can be predicted from their amino acid sequences. *Electrophoresis* 14 (1993) 1023-1031.

Bjellqvist B., Basse B., Olsen E., and Celis J. E., Reference points for comparisons of two-dimensional maps of proteins from different human cell types defined in a pH scale where isoelectric points correlate with polypeptide compositions. *Electrophoresis* 15 (1994) 529–539.

Bolton D., Evans P. A., Stott K. and Broadhurst R. W., Structure and properties of a dimereic N-terminal fragment of human ubiquitin. *J. Mol. Biol.* 314 (2001) 773-787.

Braunschweig Enzyme Database: http://www.brenda-enzymes.info/ (accessed March 17th 2008).

Brereton P. S., Verhagen M. F. J., Zhou Z. H., and Adams M. W. W., Effect of iron-sulfur cluster environment in modulating the thermodynamic properties and biological function of ferredoxin from *Pyrococcus furiosus*, *Biochemistry* 37 (1998) 7351-7362.

Brereton P. S., Duderstadt R. E., Staples C. R., Johnson M. K., and Adams M. W. W., Effect of serine ligation at each of the iron sites of the $[Fe_4S_4]$ cluster of *Pyrococcus furiosus* ferredoxin on redox, spectroscopic, and biological properties, *Biochemistry* 38 (1999) 10594-10605.

Butt T. R., Jonnalagadda S., Monia B. P., Sternberg E. J., Marsh J. A., Stadel J. M, Ecker D. J. and Crooke S. T., Ubiquitin fusion augments the yield of cloned gene products in *Escherichia coli*, *Proc. Natl. Acad. Sci. USA* 86 (1989) 2540-2544.

Calzolai L., Gorst C. M., Zhao Z.H., Teng Q., Adams M. W. W. and La Mar G. N., ^{1}H NMR investigation of electronic and molecular structure of the four-iron cluster ferredoxin from the hyperthermophilic *Pyrococcus furiosus*. Identification of Asp 14 as cluster ligand in each of the four redox states, *Biochemistry* 34 (1995) 11373-11384.

Campbell M. K. and Farrell S. O., Biochemistry (4th edition), Thomson, USA (2003) (ISBN: 0030348498).

Case A. and Stein R. L., Mechanistic studies of ubiquitin C-terminal hydrolase L1, *Biochemistry* 45 (2006) 2443-2452,

Caspersen M. B., Bennett K. and Christensen H. E. M., Expression and characterization of recombinant *Rhodocyclus tenuis* high potential iron-sulfur protein, *Protein Expres. Purif.* 19 (2000) 259-264.

Chatelet C. Meyer J., The [2Fe-2S] protein I (Shetna protein I) from *Azotobacter vinelandii* is homologous to the [2Fe-2S] ferredoxin from *Clostridium pasteurianum*, *J. Biol. Inorg. Chem.* 4 (1999) 311-317.

Christophersen, S., Master of Science thesis, Reconstruction and preliminary characteization of the evolutionary origin of iron-sulfur proteins: the oldest known protein and its relation to the origin of life. Department of Chemistry, Technical university of Denmark (2004).

Coghlan V. M. and Vickery L. E., Site-specific mutations in human ferredoxin that affect binding to ferredoxin reductase and cytochrome $P450_{scc}$, *J. Biol. Chem.* 266 (1991) 18606-18612.

Collins K. D., Ions from the Hofmeister series and osmolytes: effect on proteins in solution and in the crystallization process, *Methods* 34 (2004) 300-311.

Compton S. J. and Jones C. G., Mechanism of dye response and interference in the Bradford protein assay, *Anal. Biochem.* 151 (1985) 369-374.

Conover R. C., Kowalt A. T., Fu W., Park J.-B., Aono S., Adams M. W. W., and Johnson M. K., Spectroscopic characterization of the novel iron-sulfur cluster in *Pyrococcus furiosus* ferredoxin, *J. Biol. Chem.* 265 (1990) 8533-8541.

Conover R. C., Park J. B., Adams M. W. W. and Johnson M. K. Formation and properties of a $NiFe_3S_4$ cluster in *Pyrococcus furiosus* ferredoxin. *J. Am. Chem. Soc.* 112 (1990b) 4562-4564.

Cooper S. J., Garner C. D., Hagen W. R., Lindley P. F. and Bailey S., Hybrid-cluster protein (HCP) from *Desulfovibrio vulgaris* (Hildeenborough) at 1.6 Å, *Biochemistry* 39 (2000) 15044-15054.

Cotton F. A., Wilkinson G., and Gaus P. L., Basic Inorganic Chemistry (3rd edition), Wiley, NY (1995) (ISBN: 0471505323).

Crack J., Green J. and Thomson A. J., Mechanism of oxygen sensing by the bacterial transcription factor fumarate-nitrate reduction (FNR), *J. Biol. Chem.* 279 (2004) 9278-9286.

Creighton T. E., Protein structure, a practical approach, IRL Press, Oxford (1989).

Dauter Z., Wilson K. S., Sieker L. C., Meyer J. and Moulis J.-M., Atomic resolution (0.94 Å) structure of *Clostridium acidurici* ferredoxin. Detailed geometry of [4Fe-4S] clusters in a protein, *Biochemistry* 36 (1997) 16065-16073.

Davis B. K., Evolution of the genetic code, *Prog. Biophys. Mol. Biol*. 72 (1999) 157-243.

Davis B. K., Molecular evolution before the origin of species, *Prog. Biophys. Mol. Biol*. 79 (2002) 77-133.

DeLano W.L., The PyMOL Molecular Graphics System (2002) DeLano Scientific, Palo Alto, CA, USA. http://www.pymol.org (accessed December 9th 2007).

Duerksen-Huges P. J., Williamson M. M. and Wilkinson K. D., Affinity chromatography using protein immobilized via arginine residues: purification of ubiquitin carboxyl-terminal hydrolases, *Biochemistry* 28 (1989) 8530-8536.

Duin E. C., Lafferty M. E., Crouse B. R., Allen R. M., Sanyal I., Flint D. H. and Johnson M. K., [2Fe-2S] to [4Fe-4S] cluster conversion in *Escherichia coli* biotin synthase, *Biochemistry* 36 (1997) 11811-11820.

Edwards M. R., From a soup or a seed? Pyritic metabolic complexes in the origin of life, *Trends Ecol. Evol.* 13 (1998) 178-181.

Einsle O., Tezcan F. A., Andrade S. L. A., Schmid B., Yoshida M., Howard J. B. and Rees D. C., Nitrogenase MoFe-protein at 1.16 Å resolution: a central ligand in the FeMo-cofactor, *Science* 297 (2002) 1696-1700.

ExPASy Proteomics Server, http://www.expasy.org/ (accessed March 17th, 2008).

Fang S. and Weissman A. M., A field guide to ubiquitylation, *Cell. Mol. Life. Sci.* 61 (2004) 1546-1561.

Fawcett S. E. J., Davis D., Breton J. L., Thomson A. J., and Armstrong F. A., Voltammetric studies of the reactions of iron-sulphur clusters ([3Fe-4S] or [M3Fe-4S]) formed in *Pyrococcus furiosus* ferredoxin, *Biochem. J.* 335 (1998) 357-368.

Felsenstein J., Inferring Phylogenies, Sinauer Associates Inc, Sunderland, Massachusetts, USA (2004) (ISBN: 0878931775).

Fiala G. and Stetter K. O., *Pyrococcus furiosus* sp. nov. represents a novel genus of marine heterotrophic archaebacteria growing optimally at 100ºC. *Arch. Microbiol.* 145 (1986) 56-61.

Finnegan M. G., Conover R. C., Park J. B, Zhou Z. H. Adams, M. W. W. and Johnson M. K., Electronic, magnetic, redox, and ligand-binding properties of [MFe_3S_4] clusters (M = Zn, Co, Mn) in *Pyrococcus furiosus* ferredoxin, *Inorg. Chem.* 34 (1995) 5358-5369.

Flint D. H. and Allen R. M., Iron-sulfur proteins with non-redox functions, *Chem. Rev.* 96 (1996) 2315-2334.

Frazzon J. and Dean D. R., Formation of iron-sulfur clusters in bacteria: an emerging field in bioinorganic chemistry, *Curr. Opin. Chem. Biol.* 7 (2003) 166-173.

Fu W., Telser J., Hoffman B. M., Smith E. T., Adams M. W. W., Finnegan M. G., Conover R. C. and Johnson M. K., 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.* 116 (1994) 5722-5729.

Fujii T., Hata Y., Wakagi T., Tanaka N. and Oshima T., Novel zinc-binding centre in thermoacidophilic archaeal ferredoxins, *Nat. Struct. Biol.* 3 (1996) 834-837.

Fukuyama K., Ferredoxins containing one [4Fe-4S] center, Handbook of metalloproteins, p 543-552, edited by Messerschmidt A., Huber R., Poulos T., and Wieghardt K., Wiley, NY (2001) (ISBN: 0471627437).

Fukuyama K., Okada T., Kakuta Y. and Takahash 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.* 315 (2002) 1155-1166.

Furnes H., Banerjee N. R., Muehlenbachs K., Staudigel H., and de Witt M., Early life recorded in archean pillow lavas. *Science* 304 (2004) 578-581.

Gasteiger E., Hoogland C., Gattiker A., Duvaud S., Wilkins M.R., Appel R.D. and Bairoch A., Protein identification and analysis tools on the ExPASy Server, The Proteomics Protocols Handbook, p 571-607, edited by Walker J. M., Humana Press, USA (2005) (ISBN: 1588293432).

Gel Filtration, Principles and Methods, Amersham Biosciences 2002 (18-1022-18, edition AI).

Giastas P., Pinotsis N., Efthymiou G., Wilmanns M., Kyritsis P., Moulis J.-M. and Mavridis, I. M., 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.* 11 (2006) 445-458.

Gibney B. R., Mulholland S. E., Rabanal F. and Dutton P. L., Ferredoxin and ferredoxin-heme maquettes. *Proc. Natl. Acad. Sci. USA* 93 (1996) 15041-15046.

Gohara D. W., Ha C. S., Ghosh S. K. B., Arnold J. J., Wisniewski T. J., and Cameron C. E., Production of "authentic" poliovirus RNA-dependent RNA polymerase ($3D^{pol}$) by ubiquitin-protease-mediated cleavage in *Escherichia coli*, *Protein Expres. Purif.*, 17 (1999) 128-138.

Gorst C. M., Zhou Z. H., Ma K., Teng Q., Howard J. B., Adams M. W. W. and La Mar G. N., Participation of the disulfide bridge in the redox cycle of the ferredoxin from the hyperthermophilic *Pyrococcus furiosus*: ^{1}H nuclear magnetic resonance time resolution of the four redox states at ambient temperature, *Biochemistry* 34 (1995) 8788-8795.

Haldane J. B. S., The origin of life, *Rationalist Annu.* 148 (1929) 3-10.

Hasan M. N, Hagedoorn P. L. and Hagen W. R., *Pyrococcus furiosus* ferredoxin is a functional dimer, *FEBS Lett.* 531 (2002) 335-338.

Hong J. S. and Rabinowitz J. C., The all-or-none mode of the reconstitution and the reactions of α,α'-bipyridyl and mercurials with clostridial ferredoxin, *J. Biol. Chem.* 245 (1970) 6574-6581.

Huber C. and Wächtershäuser G., Activated acetic acid by carbon fixation on (Fe,Ni)S under primordial conditions, *Science* 276 (1997) 245-247.

Huber C. and Wächtershäuser G., Peptides by activation of amino acids with CO on (Ni,Fe)S surfaces: implication for the origin of life, *Science* 281 (1998) 670-672.

Huber C., Eisenreich W., and Hecht S. and Wächtershäuser G., A possible primordial peptide cycle, *Science* 301 (2003) 938-940.

Huber C. and Wächtershäuser G., α-hydroxy and α-amino acids under possible hadean, volcanic origin-of-life conditions, *Science* 314 (2006) 630-632.

Imsande J., Iron-sulfur clusters: Formation, perturbation and physiological functions, *Plant Physiol. Biochem.* 37 (1999) 87-97.

Jackson S. E., Ubiquitin: a small folding paradigm, *Org. Biomol. Chem.* 4 (2006) 1845-1853.

Johannessen M. N., Nielsen M. S., Ooi B. L., Christensen H. E. M. and Harris P., The crystal structure of the all cysteinyl coordinated D14C variant of [4Fe-4S] *Pyrococcus fusiosus* ferredoxin reveals two distinctly different molecules, *In preparation*.

Johnson M. K., Iron-sulfur proteins: new roles for old clusters, *Curr. Opin. Chem. Biol.*, 2 (1998) 173-181.

Jørgensen A. M., Midterm Project, Protein specific dynamics in cytochrom c_4 studied by Mössbauer spectroscopy, Department of Chemistry, Technical University of Denmark (2003).

Kalapos M. P., The energetics of the reductive citrate acid cycle in the pyrite-pulled surface metabolism in thee early stage of evolution, *J. Theor. Biol.* 248 (2007) 251-258.

Kasting J. F., Earth's early atmosphere, *Science* 259 (1993) 920-926.

Kengen S. W. M., Stams. A. J. M. and de Vos W. M. Sugar metabolism of hyberthermophiles, *FEMS Microbiol. Rev.* 18 (1996) 119-137.

Khoroshilova N., Popescu C., Münck E., Beinert H. and Kiley P. J., Iron-sulfur cluster disassembly in the FNR protein of *Escherichia coli* by O_2: [4Fe-4S] to [2Fe-2S] conversion with loss of biological activity, *Proc. Natl. Acad. Sci. USA* 94 (1997) 6087-6092.

Kissinger C. R., Sieker L. C. Adman E.T and Jensen L. H., Refined crystal structure of ferredoxin II from *Desulfovibrio gigas* at 1.7 Å, *J. Mol. Biol.* 219 (1991) 693-715.

Kohno T., Kusunoki H., Sato K. and Wakamatsu K., A new general method for the biosynthesis of stable isotope-enriched peptides using a decahistidine-tagged ubiquitin fusion system: an application to the production of mastoparan-X uniformly enriched with ^{15}N and ^{15}N/^{13}C, *J. Biomol. NMR* 12 (1998) 109-121.

Leal S. S., Teixeira M. and Gomes C. M., Studies on the degradation pathway of iron-sulfur centers during unfolding of a hyperstable ferredoxin: cluster dissociation, iron release and protein stability, *J. Biol. Inorg. Chem.* 9 (2004) 987-996.

Lill R. and Mühlenhoff U., Iron-sulfur-protein biogenesis in eukaryotes, *Trends Biochem. Sci.* 30 (2005) 133-141.

Liu A. and Gräslund A., Electron paramagnetic resonance evidence for a novel interconversion of $[3Fe\text{-}4S]^{+}$ and $[4Fe\text{-}4S]^{+}$ clusters with endogenous iron and sulfide in anaerobic ribonucleotide reductase actives *in vitro*, *J. Biol. Chem.* 275 (2000) 12367-12373.

Liu C.-C., Miller H. I., Kohr W. J. and Silber J. I., Purification of a ubiquitin protein peptidase from yeast with efficient in vitro assays, *J. Biol. Chem.* 264 (1989) 20331-20338.

Lode E. T., Murray C. L., Sweeney W. V. and Rabinowitz J. C., Synthesis and properties of *Clostridium acidiurici* [Leu^2]-ferredoxin: a function of the peptide chain and evidence against the direct role of the aromatic residues in electron transfer, *Proc. Natl. Acad. Sci. USA* 71 (1974) 1361-1365.

de Lucreazia D., Anella F. and Chiiarabelli C., Question 5: on the chemical reality of the RNA world, *Orig. Life Evol. Biosph.* 37 (2007) 379-385.

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* 4 (1996) 1291-301.

Malkin R. and Rabinowitz J. C., The reconstitution of clostridial ferredoxin, *Biochem. Biophys. Res. Commun.* 23 (1966) 822-827.

Martin W. and Russel M. J., On the origin of cells: a hypothesis for the evolutionary transition from abiotic geochemistry to chemoautotrophic prokaryotes, and from prokaryotes to nucleated cells, *Phil. Trans. R. Soc. B* 358 (2003) 59-85.

Martínez-Espinosa R. M., Marhuenda-Egea F. C., Donaire A. and Boneta M. J. NMR studies of a ferredoxin from *Haloferax mediterranei* and its physiological role in nitrate assimilatory pathway. *Biochim. et Biophys. Acta* 1623 (2003) 47-51.

Mayhew S. G., The redox potential of dithionite and SO^{-}_{2} from equilibrium reactions with flavodoxins, methyl viologen and hydrogen plus hydrogenase, *Eur. J. Biochem.* 85 (1978) 535-547.

Meyer J., Iron-sulfur protein folds, iron-sulfur chemistry, and evolution, *J. Biol. Inorg. Chem.* 13 (2008) 157-170.

Miller H. I., Henzel W. J., Ridgway J. B., Kuang W.-J., Chisholm V. and Lui C.-C., Cloning and expression of a yeast ubiquitin-protein cleaving activity in *Escherichia Coli*, *Nat. Biotechnol.* 7 (1989) 698-704.

Miller S. L., A production of amino acids under possible primitive earth conditions. *Science* 117 (1953) 528-529.

Mouesca J.-M. and Lamotte B., Iron-sulfur clusters and their electronic and magnetic properties. *Coord. Chem. Rev.* 178-180 (1998) 1573-1614.

Moulis J.-M., Davasse V., Golinelli M.-P., Meyer J., and Quinkal I., The coordination sphere of iron-sulfur clusters: lessons from site-directed mutagenesis experiments, *J. Biol. Inorg. Chem.* 1 (1996) 2-14.

Moura J. J. G., Macedo A. L., Goodfellow B. J., and Moura I., Ferredoxins containing one [3Fe-4S] cluster. *Desulfovibrio gigas* ferredoxin II – solution structure, Handbook of metalloproteins, p 553-559, edited by Messerschmidt A., Huber R., Poulos T., and Wieghardt K., Wiley, NY (2001) (ISBN: 0471627437).

Mulholland S. E., Gibney B. R., Rabanal F. and Dutton P. L., Characterization of the fundamental protein ligand requirements of $[4Fe\text{-}4S]^{2+/+}$ clusters with sixteen amino acid maquettes, *J. Am. Chem. Soc.* 120 (1998) 10296-10302.

Mulholland S. E., Gibney B. R., Rabanal F. and Dutton P. L., Determination of nonligand amino acids critical to $[4Fe\text{-}4S]^{2+/+}$ assembly in ferredoxin maquettes. *Biochemistry* 38 (1999) 10442-10448.

National Center for Biotechnology Information, http://www.ncbi.nlm.nih.gov/ (accessed, March 17th 2008).

Nielsen M. S., Harris P., Ooi B. L., Christensen H. E. M., The 1.5 Å resolution cystal structure of $[Fe_3S_4]$-ferredoxin from the hyperthermophilic archaeon *Pyrucoccus furiosus*, *Biochemistry* 43 (2004) 5188-5194.

Nisbet E. G. and Sleep N. H., The habitat and nature of early life. *Nature* 409 (2001) 1083-1091.

Noodleman L., Peng C. Y., Case D. A. and Mouesca J.-M., Orbital interactions, electron delocallization and spin coupling in iron-sulfur clusters. *Coord. Chem. Rev.* 144 (1995) 199-244.

Nørgaard H., Master of Science thesis, Initial spectroscopic characterization of the evolutionary oldest known protein by electron paramagnetic resonance spectroscopy and Mössbauer spectroscopy, Department of Chemistry, Technical University of Denmark (2006).

Ollagnier S., Meier C., Mulliez E., Gaillard J., Schuenemann V., Trautwein A., Mattioli T., Lutz M. and Fontecave M., Assembly of 2Fe-2S and 4Fe-4S clusters in anaerobic ribonucleotide reductase from *Escherichia coli*, *J. Am. Chem. Soc.* 121 (1999) 6344-6350.

Oparin A. I., The origin of life. MacMillian, New York, USA (1938).

Orgel L. E., The origin of life – a review of facts and speculations, *Trends Biochem. Sci.* 23 (1998) 491-495.

Orme-Johnson W. H., Metalloenzymes Conf., Oxford, Tryptic cleavage of *Clostridium acidi-urici* apoferredoxin, and reconstruction of the separated fragments. *Biochem. Soc. Trans.* 1 (1973) 30-31.

Pace C. N., Vajdos F., Fee L., Grimsley G., and Gray T., How to measure and predict the molar absorption coefficient of a protein, *Protein Sci.* 4 (1995) 2411-2423.

Pan G., Menon A. L. and Adams M. W. W., Characterization of a [2Fe-2S] protein encoded in the iron-hydrogenase operon of *Thermotoga maritime*, *J. Biol. Inorg. Chem.* 8 (2003) 469-47.

Piaggio M. V., Peirotto M. B. and Deiber J. A., Effect of background electrolyte on the estimation of protein hydrodynamic radius and net charge through capillary zone electrophoresis, *Electrophoresis* 26 (2005) 3232-3246.

Pierrel F., Hernandez H. L., Johnson M. K., Fontecave M. and Atta M., MiaB protein from *Thermotoga maritima*, characterization of an extremely thermophilic tRNA-methylthiotransferase, *J. Biol. Chem.* 278 (2003) 29515-29524.

Pilon A., Yost P., Chase T. E., Lohnas G., Burkett T., Roberts S., and Bentley W. E., Ubiquitin fusion technology: Bioprocessing of peptides, *Biotechnol. Prog.* 13 (1997) 374-379.

Rao P. V. and Holm R. H., Synthetic analogues of the active site of iron-sulfur proteins, *Chem. Rev.* 104 (2004) 527-559.

Rappe A. K., Casewit C. J., Colwell K. S., Goddard III W. A. and Skiff W. M., UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations, *J. Am. Chem. Soc.* 114 (1992) 10024-10035.

Riès-Kautt M. and Ducruix A, Inferences from physico-chemical studies of crystallogenesis and the precrystalline stage, *Methods. Enzymol.* 276 (1997) 23-59.

Rode B. M., Peptides and the origin of life, *Peptides* 20 (1999) 773-786.

Rypniewski W. R., Breiter D. R., Benning M. M., Wesenberg G., Oh B. H., Markley J. L., Rayment I. and Holden H. M., Crystallization and structure determination to 2.5 Å resolution of the oxidized [2Fe-2S] ferredoxin isolated from *Anabaena* 7120, *Biochemistry* 30 (1991) 4126-4131.

Sakamoto T., Tanaka T., Ito Y., Rajesh S., Iwamoto-Sugai M., Kodera Y., Tsuchida N., Shibata T. and Khono T., An NMR analysis of ubiquitin recognition by yeast ubiquitin hydrolase: evidence for novel substrate recognition by a cysteine protease, *Biochemistry* 38 (1999) 11634-11642.

Schünemann V. and Winkler H., Structure and dynamics of biomolecules studies by Mössbauer spectroscopy, *Rep. Prog. Phys.* 63 (2000) 263-353.

Scott T. A., Berlinguette C. P., Holm R. H. and Hong-Cai Zhou, Initial synthesis and structure of an all-ferrus analogue of the fully reduced $[Fe_4S_4]^0$ cluster of the nitrogenise iron protein, *Proc. Natl. Acad. Sci. USA* 201 (2005) 9741-9744.

Sham S., Calzolai L., Wang P.-L., Bren K., Haarklau H.,, Brereton P. S., Adams M. W. W. and La Mar G. N., A solution NMR molecular model for the aspartate-ligated, cubane cluster containing ferredoxin from the hyperthermophilic acheaon *Pyrococcus furiosus*. *Biochemistry* 41 (2002) 12498-12508.

Sieker L. C. and Adman E. T., The 2[4Fe-4S] ferredoxins, Handbook of metalloproteins, p 574-592, edited by Messerschmidt A., Huber R., Poulos T., and Wieghardt K., Wiley, NY (2001) (ISBN: 0471627437).

Silva P. J., van den Ban E. C. D., Wassink H., Haaker H., de Castro B., Robb F. T. and Hagen W. R., Enzymes of hydrogen metabolism in *Pyrococcus furiosus*. *Eur. J. Biochem.* 167 (2000) 6541-6555.

Skoog D. A., West D. M., and Holler F. J., Fundamentals of Analytical Chemistry (7th edition), Saunders College Publishing, USA, (1996) (ISBN: 0030059380).

Sleep N. H., Zahnle K. J. Kasting J. F. and Morowitz H. J., Annihilation of ecosystems by large asteroid impacts on the early Earth, *Nature* 342 (1989) 139-142.

Sow T.-C., Pedersen M. V., Christensen H. E. M. and Ooi B.-L., Total synthesis of a miniferredoxin, *Biochem. Biophys. Res. Commun.* 223 (1996) 360-364.

Sow T.-C., Ph.D. thesis, Design, synthesis and characterisation of iron-sulfur proteins and a novel molybdenum analogue, National University of Singapore (1997).

Staples C. R., Dawan I. K., Finnegan M. G., Dwinell D. A., Zhou Z. H., Huang H., Verhagen M. F. J. M., Adams M. W. W. and Johnson M. K., Electronic, magnetic, redox properties of [MFe_3S_4] clusters (M = Cd, Cu, Cr) in *Pyrococcus furiosus* ferredoxin, *Inorg. Chem.* 36 (1997) 5740-574.

Sticht H. and Rösch P., The structure of iron-sulfur proteins, *Prog. Biophys. Mol. Biol*. 70 (1998) 95-136.

Stout C. D., Ferredoxins containing two different Fe/S centers of the forms [4Fe-4S] and [3Fe-4S], Handbook of metalloproteins, p 560-573, edited by Messerschmidt A., Huber R., Poulos T., and Wieghardt K., Wiley, NY (2001) (ISBN: 0471627437).

Studier F. W., Rosenberg A. H., Dunn J. J., and Dubendorff J. W., Use of T7 RNA polymerase to direct expression of cloned genes, *Methods Enzymol.* 185 (1990) 60-89.

Sweet C. R., Expression of recombinant proteins from *lac* promoters, *Methods in molecular biology* 235 (2003) 277-288.

Thompson M., 2004, Arguslab 4.0.1, Planaria Software LLC.

Tilley G. J., Camba R., Burgess B. K. and Armstrong F. A., Influence of electrochemical properties in determining the sensitivity of [4Fe-4S] clusters in proteins to oxidative damage, *Biochem. J.* 360 (2001) 717-726.

Tobias J. W. and Varshavsky A., Cloning and functional analysis of the ubiquitin-specific protease gene UBP1 of *Saccharomyces ceravisiae*, *J. Biol. Chem.* 266 (1991) 12021-12028.

Tributsch H., Fiechter S., Jokisch D., Rojas-Chapana J. and Ellmer K., Photoelectrochemical power, chemical energy and catalytic activity for organic evolution on natural pyrite interfaces, *Origins Life Evol. Biosphere* 32 (2003) 129-162.

Varshavsky A., Ubiquitin fusion techniques and its descendants, *Methods Enzymol.* 327 (2000) 578-593.

Vernède X. and Fontecilla-Camps J. C., A method to stabilize reduced and/or gas-treated protein crystals by flash-cooling under a controlled atmosphere. *J. Appl. Cryst.* 32 (1999) 505-509.

Vijay-Kumar S., Bugg C. E., Wilkinson K. D., Vierstra R. D., Hatfield P. M. and Cook W. J., Comparison of the three-dimensional structures of human, yeast and oat ubiquitin, *J. Biol. Chem.* 262 (1987) 6396-6399.

Voet D. and Voet J. G., Biochemistry (3rd edition), Wiley, USA (2003) (ISBN: 0471392235).

Wächtershäuser G., Evolution of the first metabolic cycles, *Proc. Natl. Acad. Sci. USA* 87 (1990) 200-204.

Wächtershäuser G., Groundworks for an evolutionary biochemistry: the iron-sulfur world, *Prog. Biophys. Mol. Biol*. 58 (1992) 85-201.

Wächtershäuser G., From volcanic origins of chemoautotrophic life to bacteria, archaea and eukarya. *Phil. Trans. R. Soc. B* 361 (2006) 1787-1808.

Wilkinson K. D., Cox M. J., O'Connor L. B and Shapira R., Structure and activities of a variant ubiquitin sequence from bakers yeast, *Biochemistry* 25 (1986) 4999-5004.

Wilkinson K. D., Ubiquitination and deubiquitination: target of proteins for degradation by the proteasome, *Seminars in Cell and Developmental Biology*, 11 (2000) 141-148.

Zanetti G., Binda C., and Aliverti A., The [2Fe-2S] ferredoxins, Handbook of metalloproteins, p 532-542, edited by Messerschmidt A., Huber R., Poulos T., and Wieghardt K., Wiley, NY (2001) (ISBN: 0471627437).

van Zuilen M. A., Lepland A., and Arrhenius G., Reassessing the evidence for the earliest traces of life, *Nature*, 418 (2002) 627-630.

Part II

Mycobacterium tuberculosis dCTP deaminase:dUTPase: Structural implications for inhibition and mechanism of hydrolysis

Chapter one

1 Introduction

The second part of this dissertation concerns crystallization of the bifunctional enzyme dCTP deaminase:dUTPase from *Mycobacterium tuberculosis* (*M. tuberculosis*). This enzyme is involved in the reaction pathway for biosynthesis of the nucleotide dTTP and catalyses the conversion of dCTP to dUMP (Helt *et al.*, 2008). dCTP deaminase:dUTPase belongs to a family of enzymes also including two other members; the monofunctional dCTPdeaminase and dUTPase (Johansson *et. al*, 2003 and 2005).

Studying this enzyme family will provide an understanding of the evolutionary relations between the dUTPases and the dCTP deaminases, which will contribute to explain evolutionary relations between different organisms and the flow of genetic information. Also, studies of enzymes involved in pathways for biosynthesis of dTTP might provide clues concerning the evolution of DNA, e.g. clues as to how DNA evolved to contain thymidylate instead of deoxyuridylate. Finally, tuberculosis continues to be a worldwide problem and characterizing dCTP deaminase:dUTPase from *M. tuberculosis* might lead to a new potential drug target.

I learned about ongoing research on *M. tuberculosis* dCTP deaminase:dUTPase from my supervisor, Associate Professor, Pernille Harris, who had been contacted by Associate Professor, Martin Willemoës (Department of Biology, University of Copenhagen). I accepted the task of trying to crystallize this enzyme. The goal was to determine the structure of the enzyme as an apo-enzyme and also in complex with the inhibitor, dTTP, and a substrate analogue, α,β-imido dUTP.

1.1 Outline of part II

Chapter 2 provides an introduction to the enzyme family including dCTP deaminase, dCTP deaminase:dUTPase and dUTPase. Also included is a description of the biosynthesis of the deoxynucleotides used for incorporation into DNA and tuberculosis treatment aimed at the nucleotide biosynthesis. Chapter 3 describes crystallization trials, data collection and structure determination. The structures of the apo-enzyme and the enzyme in complex with dTTP are presented and compared. These are further discussed in relation to other members of the enzyme family and implications concerning the dTTP inhibition mode and hydrolysis mechanism are discussed. Finally chapter 4 gives an outlook discussing further experiments based on the results presented in this part of the dissertation.

Chapter two

2 dCTP deaminases, dUTPases and nucleotide synthesis

2.1 The deoxyribonucleotide synthesis

The building blocks from which DNA is synthesized are the deoxyribonucleotides, dCTP, dATP, dGTP and dTTP. Except for dTTP all of these are synthesized from the corresponding ribonucleotides by the action of ribonucleotide reductase at the diphosphate level (see figure 2.1). This enzyme reduces the 2'-hydroxyl group to a hydrogen atom. The ribonucleotide UDP is also reduced by ribonucleotide reductase but, the product dUDP is not used for incorporation into DNA. The reason for not including urasil in DNA is probably that cytosine in DNA can deaminate spontaneously yielding urasil. This deamination is mutagenic since urasil pairs with adenine instead of guanine. The cell contains a repair system in which it recognises uracil (it recognises the missing methyl group compared to thymine) and cuts it out of the sequence. If DNA contained urasil there would be no way of correcting the spontaneous deamination of cytosine (Berg *et al.*, 2002).

2.2 Synthesis of dTTP

DNA contains thymine and the corresponding deoxyribonucleotide used for incorporation into DNA is dTTP. In all organisms dUMP is the origin of dTTP. dUMP is methylated by thymidylate synthetase to yield dTMP, which is further phosphorylated in two steps by the action of dTMP kinase and nucleoside diphosphate kinase to yield dTTP (see figure 2.1) (Myllykallio *et al.*, 2002; Neuhard and Kelln, 1996).

There are two ways dUMP can be synthesized. A minor part of dUMP is obtained from reduction of UDP to dUDP by the action of ribonucleotide reductase, which is further phosphorylated to dUTP and then hydrolysed (reaction 2 figure 2.1) to dUMP by dUTPase (Bianchi *et al.*, 1987; Møllgaard and Neuhard, 1978; Neuhard and Thomassen, 1971). The major supply of dUMP originates from deamination of a cytosine deoxyribonucleotide (see fig 2.1). In Gram-positive bacteria and eukaryotes dUMP is obtained directly from dCMP by the action of dCMP deaminases (reaction 4 figure 2.1), which are zinc-containing metallo-enzymes (Almog *et al.*, 2004; Weiner *et al.*, 1993). In Gram-negative bacteria and some achaea dUMP is obtained from dCTP, which is deaminated and dephosphorylated. In *E. coli* and *Salmonella typhimurium* (*S. typhimurium*) this happens in two consecutive steps where dCTP is deaminated (reaction 1 figure 2.1) by dCTP deaminase to generate dUTP, which is further hydrolysed (reaction 2 figure 2.1) by dUTPase to generate dUMP (Johansson *et al.*, 2003; Neuhard and Nygaard, 1987). However, it was found a few years back that in the hyperthermophilic archeon *Methanocaldococcus jannaschii* (*M. jannaschii*), one enzyme, dCTP deaminase:dUTPase, carries out both functions in one active site (reaction 3 figure 2.1)

(Björnberg *et al.*, 2003; Li *et al.*, 2003). In both cases the deamination occurs without the aid of a metal ion (Johansson *et al.*, 2003 and 2005). Until very recently the *M. jannaschii* enzyme was the only reported enzyme to carry out both functions, but now, a second dCTPdeaminase:dUTPase from the bacteria *M. tuberculosis* has been discovered. This is interesting as it demonstrates that the bifunctional enzyme can be found outside the archaean kingdom (Helt *et al.*, 2008).

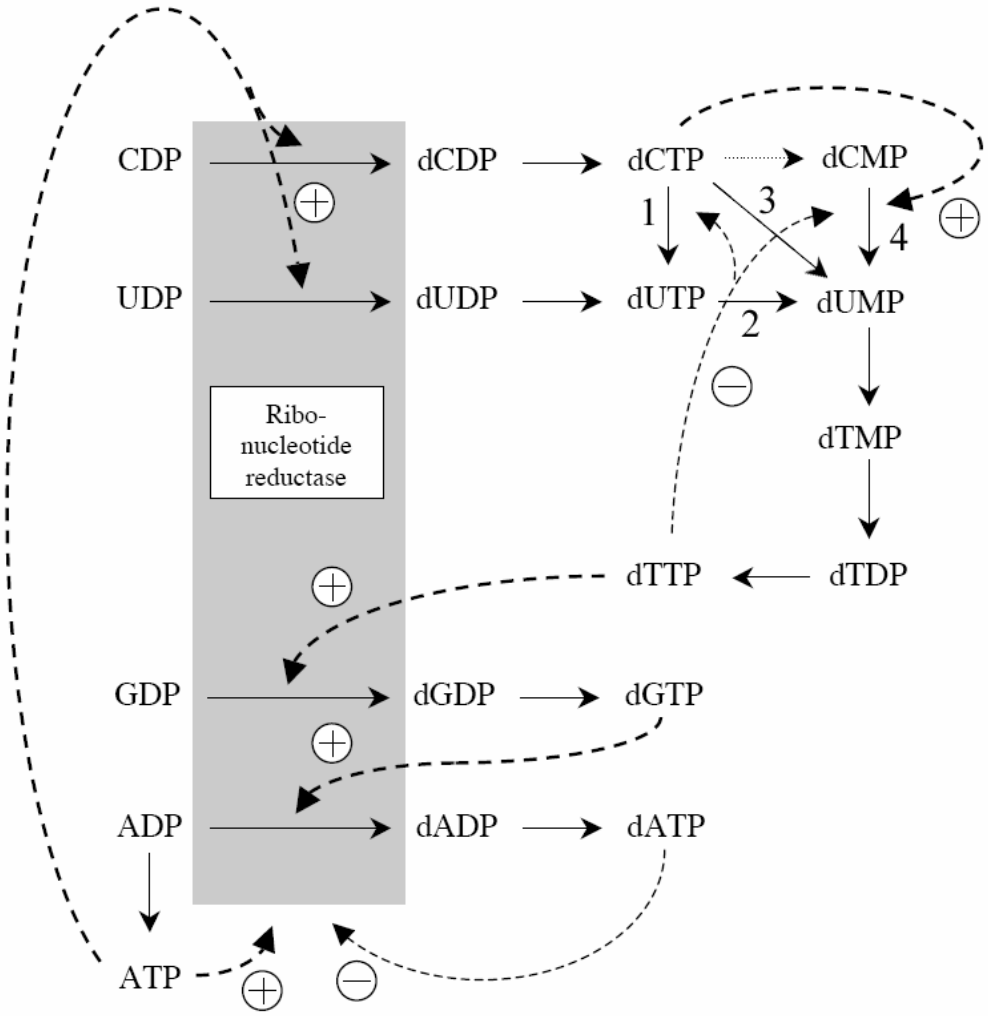


Figure 2.1 Overview of synthesis of deoxyribonucleotides used for incorporation into DNA. dGTP, dATP and dCTP are obtained directly at the diphosphate level from reduction of the corresponding oxynucleotides. dTTP is synthesised from dUMP via three enzymatic reactions. The synthesis of dUMP is dependent on the organism (see text). Solid arrows indicate enzymatic reactions while dotted arrows indicate how a particular nucleotide affects the reaction to which the arrow points. Activators are indicated with (+) while inhibitors are indicated with (-). Reaction 1 is carried out by dCTP deaminase, reaction 2 is carried out by dUTPase, reaction 3 is carried out by the bifunctional dCTP deaminase:dUTPase and reaction 4 is carried out by dCMP deaminase. The figure was created by Associate Professor Martin Willemoës (Department of Biology, University of Copenhagen) and presented with permission.

dTTP inhibits dCTP deminases and dCTP deaminase:dUTPases by binding in the active site (Johansson *et al.*, 2007,) while, dUTPases only show negligible affinity towards dTTP (Climie *et al.*, 1994; Nord *et al.*, 1997). dTTP also acts as an inhibitor of dCMP deaminases, where it binds in an allosteric site in competition with the activator, dCTP (Almog *et al.*, 2004) (see figure 2.1). Enzymes having dCTP deaminase activity are probably regulated by dTTP to avoid initiation of dCTP breakdown to dUMP in the presence of excess dTTP.

2.3 dCTP deaminases and dUTPases

dCTP deaminases, dUTPases and the bifunctional dCTPdeaminase:dUTPase belong to the same superfamily and they are all homo-trimers in their active form (Chan *et al.*, 2004; Johansson *et al.*, 2003 and 2005) (dUTPases can also be found as monomers or dimers, Gonzáles *et al.*, 2001). The dCTP deaminases and the bifunctional enzymes share the highest structural similarity. However, all enzymes share an almost identical core structure and three active sites are contained in the active trimer. Figure 2.2 presents representative crystal structures of the three types of enzymes.

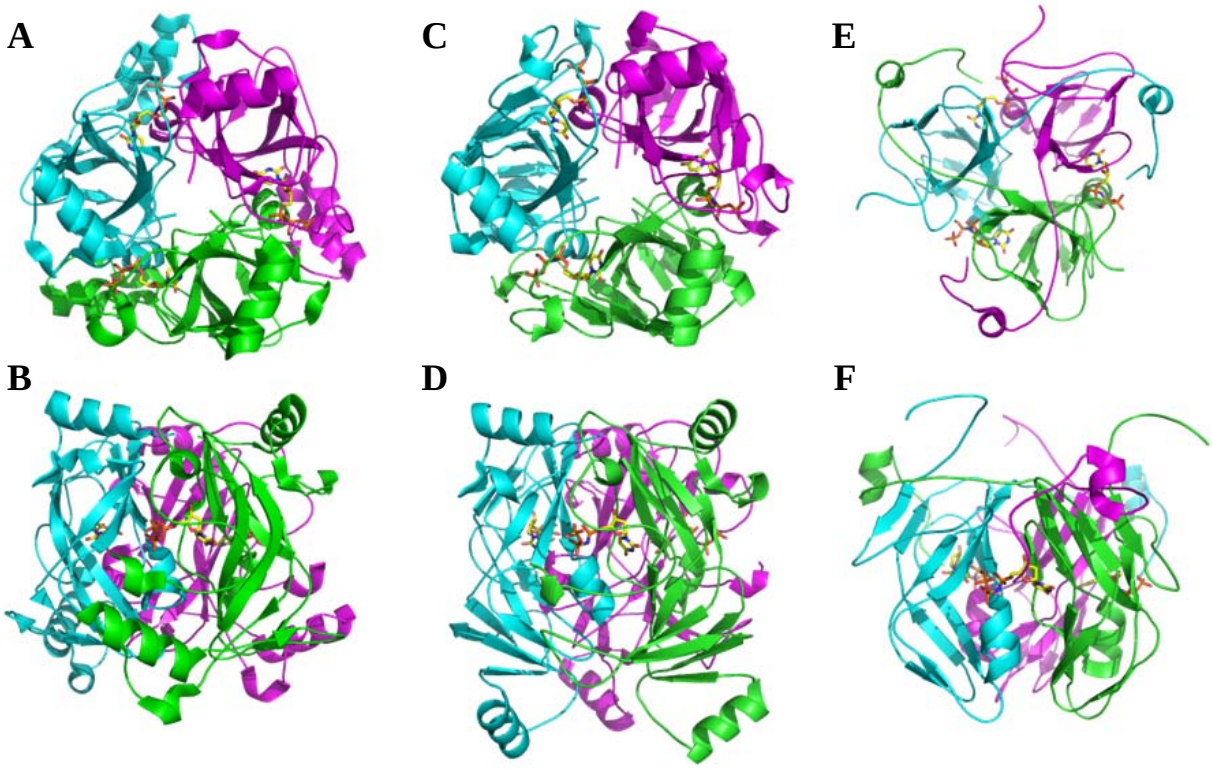


Figure 2.2 Ribbon views of dCTP deaminase and dUTPase crystal structures. A and B) Structure of *E. coli* dCTP deaminase-dUTP complex (PDB ID: 1XS1, Johansson *et al.*, 2005). C and D) Structure of *M. jannaschii* E145A dCTP deaminase:dUTPase-α,β-imido dUTP complex (PDB ID: 2HXD, Siggaard *et al.*, 2008). E and F) Structure of *M. tuberculosis* dUTPase-α,β-imido dUTP complex (PDB ID: 1SIX, Chan *et al.*, 2004). The structures are viewed along the threefold axis (A, C and E) and perpendicular to the threefold axis (B, D and F). The subunits in each of the three trimeric structures are shown in pink, cyan and green and in each structure the nucleotides are presented in ball and stick. The figure was prepared in PyMOL (DeLano W. L., 2002).

In the enzymes having dCTP deaminase activity the active site is composed of residues from two subunits and the C-terminal folds back upon the enzyme generating a lid covering the active site (Johansson *et al.*, 2005; Helt *et al.*, 2008; Siggaard *et al.*, 2008). In dUTPases the active site is composed of residues from all three monomers as a result of the C-terminal reaching across the trimer and interacting with the γ-phosphate of the nucleotide placed in a pocket between the two other subunits (figure 2.2) (Chan *et al.*, 2004).
Figure 2.3 shows a schematic view of the reactions these enzymes catalyse and points out the differences between the nucleotides. Whether the reactions in the bifunctional enzyme are concerted or happen in two consecutive steps have not yet been resolved. The dUTP analogue, α,β-imido dUTP, is an often used nucleotide in imitating the interactions between the substrate and enzyme in dUTPase. The dUTP analogue and the inhibitor, dTTP, of enzymes having dCTP deaminase activity are also shown in figure 2.3.

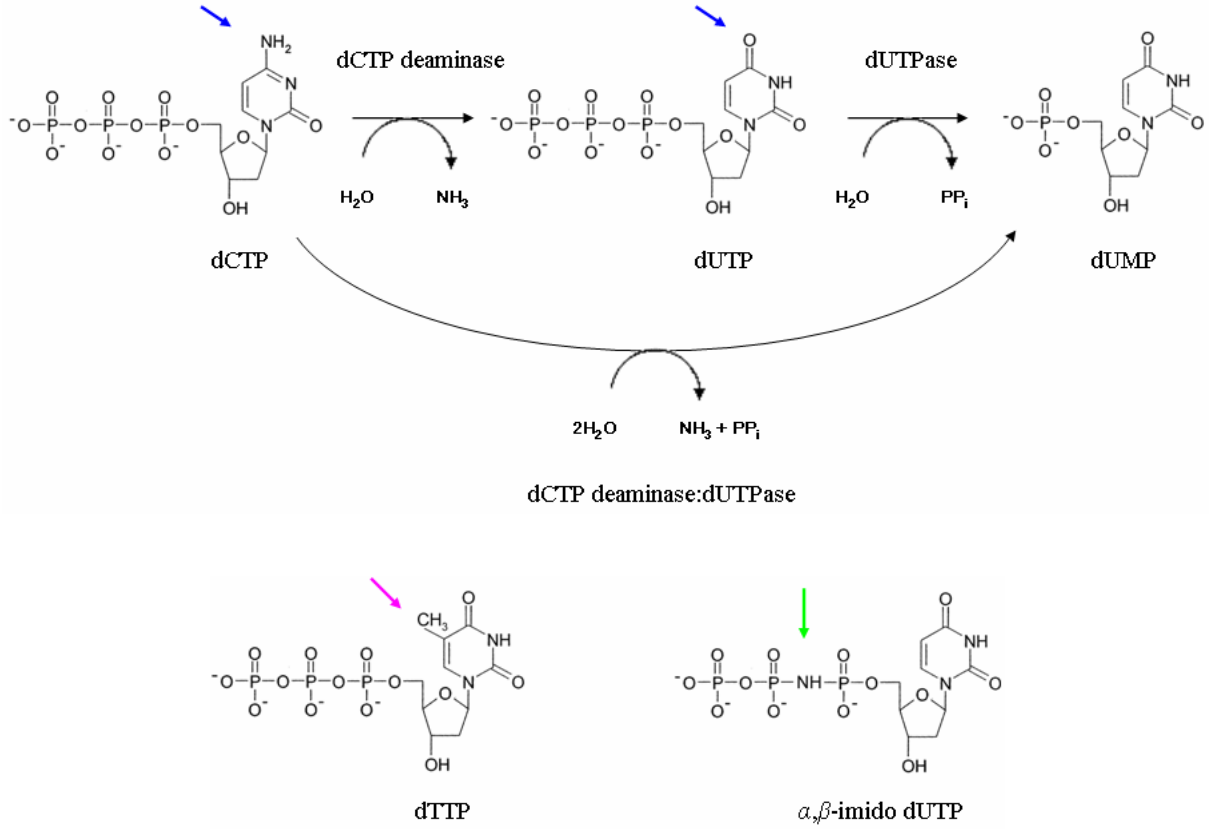


Figure 2.3 Schematic views of the reactions catalysed by the enzymes and differences in nucleotides. Upper part of the figure: dCTP deaminase catalyses the deamination of dCTP to yield dUTP and ammonia (the change indicated by the blue arrows). dUTP is then hydrolysed by dUTPase to yield dUMP and pyrophosphate. The bifunctional enzyme dCTP deaminase:dUTPase catalyses both of these reactions in one active site. Lower part of the figure: The inhibitor of dCTP deaminase and dCTPdeaminase:dUTPase, dTTP, only differs from dUTP by a single methyl group (indicated by pink arrow). In the dUTP analogue, α,β-imido dUTP, the α,β-bridging oxygen has been replaced with an imido group (green arrow) rendering the nucleotide non-hydrolysable by the mono-functional dUTPase.

The members of this enzyme family are non-metalloenzymes as opposed to dCMP deaminases (see section 2.2). The deamination of dCTP therefore happens by means of a different mechanism. However, for the reaction to take place dCTP deaminases and dUTPases require a magnesium ion (Neuhard, 1978; Shao *et al*., 1997), which has no catalytic effect but is coordinated to the phosphate moiety, rendering the true substrates $dCTP{\cdot}Mg^{2+}$ and $dUTP{\cdot}Mg^{2+}$, respectively, and the true inhibitor of the enzymes with dCTP deaminase function $dTTP{\cdot}Mg^{2+}$ (Johansson *et al.*, 2007).
Mutational analyses have been performed on enzymes from the family in order to establish which residues play key roles in the enzymatic reactions. These are outlined in section 3.2.6. Based on mutational analyses, obtained crystal structures and structure based sequence alignments, mechanisms for the deamination of dCTP (Johansson *et al.*, 2005) and hydrolysis of dUTP (Chan *et al.*, 2004) have been suggested and they are presented and discussed in section 3.2.6.2.
Identification of residues common to this group of enzymes and residues specific for the deaminase activity or the hydrolysis activity will be of great value in predicting the function of new enzymes belonging to this family just based on sequence searches. The more new enzymes that are determined to belong to this family the stronger the evidence will be for how this group of enzymes evolved and the connection between dCTPdeaminases and dUTPases.

2.4 Cooperativity in the active trimer

Kinetic studies on *S. typhimurium* and *E. coli* dCTP deaminase have revealed cooperativity of substrate binding and that dTTP inhibition happens by a non-trivial competition between substrate and inhibitor for binding to the same site (Beck *et al.*, 1975; Johanson *et al.*, 2007). Structural studies on *E. coli* dCTP deaminase have shown the enzyme capable of adopting two different conformations in the active site dependent on the bound nucleotide. The part of the active site that undergoes conformational changes is found close to the threefold axis of the homo-trimer rendering the two conformations mutually exclusive (see figure 2.4). The side chains of two residues from neighbouring subunits are likely to clash unless all three subunits are found in the same conformation (elaborated in section 3.2.7) (Johanson *et al.*, 2007). The same phenomenon has been observed in the *M. tuberculosis* dCTP deaminase:dUTPase (discussed in section 3.2.7) while dTTP inhibition in the *M. jannaschii* bifunctional enzyme is suggested to be of simple competition with dCTP for binding in the active site. dUTPases kinetic studies have not indicated any signs of cooperativity (Larsson *et al.*, 1996; Nord *et al.*, 1997).

2.5 The importance of dUTPases

dUTP is a toxic nucleotide in the cell. dUTP and dTTP differ only by a single methyl group (see figure 2.3) and therefore dUTP can readily be accepted by DNA-polymerase instead of dTTP and incorporated into DNA. The DNA repair system will correct the misincorporation. However, when the concentration of dUTP is high, inducing excessive incorporation of uracil into DNA, the repair system goes into a deleterious cycle of multiple excisions leading to multiple strand breakage causing the cell to die. This process is termed thymineless cell death. Consequently dUTPase is an essential enzyme not only required for synthesis of dTTP, but

also for keeping a vital low ratio of dUTP:dTTP, hence maintaining the integrity of the genomic DNA (Richards *et al.*, 1984; Gadsden *et al.*, 1993).
The trimeric dUTPases are not regulated by dTTP and they show extremely low affinity for dTTP (section 2.2). The reason for this selectivity against dTTP and nucleotides meant for incorporation into RNA is probably to avoid hydrolysis of non-lethal nucleotides at the expense of eliminating the highly toxic dUTP. As a consequence the nucleotide binding site has evolved to only recognise the uracil base (Chan *et al.*, 2004; Johanson *et al.* 2007).

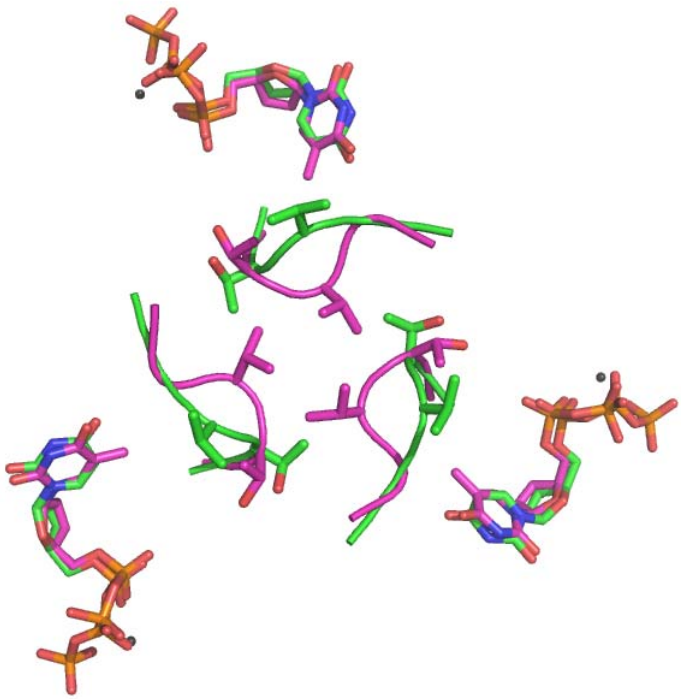


Figure 2.4 Active and inactive conformation of *E. coli* dCTP deaminase. Close-up of residues 120-124 from superposition of the *E. coli* dCTP daminase-dUTP complex structure (green, PDB ID:1XS1, Johansson *et al.*, 2005) and E138A-dTTP complex structure (pink, PDB ID: 2J4Q, Johansson *et al.*, 2007). The backbone is shown in ribbon view while the side chains of Val-122 and Thr-123 in all three subunits in the two structures are shown in sticks. Nucleotides are (shown in ball and stick) presented in standard atom colours. However, carbon is shown in colours corresponding to the colour of the matching enzyme. Mg^{2+} is shown in grey. The superposition is viewed along the three-fold axis. The E138A dTTP complex structure is found in the inactive conformation while the dUTP complex structure is found in the active conformation. It is evident that Val and Thr from neighbouring subunits will clash unless all subunits in the trimer are found in the same conformation (Johansson *et al.*, 2007). The figure was prepared in PyMOL (DeLano W. L., 2002) in a similar mode as presented by Johansson *et al.* (2007).

2.6 Tuberculosis treatment aimed at the nucleotide synthesis

Tuberculosis continues to be a burden in developing as well as developed countries, infecting and killing millions of people worldwide every year. Tuberculosis is also a major cause of death among people infected with HIV/AIDS. Multi-drug resistant strains are a growing problem and new drug targets are therefore necessary (World Health Organization). Microorganisms like *M. tuberculosis* need an abundant supply of nucleotides for the DNA

synthesis in the rapidly dividing cells. Therefore, drugs aimed at enzymes involved in nucleotide synthesis are a growing area of research.
A used approach in cancer treatment is to target enzymes of the thymidylate biosynthesis (Barabás *et al.*, 2004), and inhibition of dUTPases has also been considered as a potential target for inhibiting cancer growth and viral replication (McIntosh and Hanes, 1997; Grässer *et al*., 2001; Ladner, 2001).
Now, dUTPase has also been considered a drug target in the battle against tuberculosis. Human and *M. tuberculosis* dUTPases share 34 % sequence identity, which have to be taken into consideration when designing an inhibitor. Since the active sites of these two enzymes only show minor differences, a drug designed to inhibit *M. tuberculosis* dUTPase could easily inhibit the human form. The more different the considered enzymes are, the easier it is to design a species specific drug. In this connection knowing the structures of each enzyme is a key factor (Chan *et al.*, 2004).
The emergence of the bifunctional enzyme dCTP deaminase:dUTPase from *M. tuberculosis* yields a potential new drug target since this enzyme is not found in eukaryotes (Helt *et al.*, 2008). Obtaining the crystal structure of the enzyme and knowing how the enzyme is inhibited by dTTP is a necessary step in order to synthesize a drug targeting the bifunctional enzyme. The crystal structure of *M. tuberculosis* dCTP deaminase:dUTPase and the dTTP inhibition mode are presented and discussed in chapter 3 and have also been published (Helt *et al.*, 2008).

2.7 *M. tuberculosis* dCTP deaminase:dUTPase

The enzyme was found in a database search for probable bifunctional enzymes. The primary sequence of dCTP deaminase:dUTPase from *M. tuberculosis* is composed of 190 amino acid residues corresponding to a molecular weight of 20.9 kDa for each subunit in the trimeric enzyme (Helt *et al.*, 2008). The enzyme was expressed as a recombinant enzyme in *E. coli* and purified using standard purification techniques. Kinetic experiments has proved the enzyme to have both dCTPdeaminase and dUTPase activity with equal affinity towards dCTP and dUTP (Helt *et al.*, 2008). This contrasts findings for the bifunctional enzyme from *M. jannaschii*, which has a lower affinity for dUTP compared to dCTP (Björnberg *et al.*, 2003).

Chapter three

3 Crystal structure of dCTP deaminase dUTPase

3.1 Experimental

3.1.1 Crystallization

Enzyme used for crystallization trials, was generously supplied by Associate Professor Martin Willemoës (Department of Biology, University of Copenhagen). The enzyme was provided in two stock concentrations; 1.9 mg/ml and 4.3 mg/ml in 50 mM HEPES (4-(2-Hydroxy-ethyl)piperazine-1-ethanesulfonic acid) pH 6.8, 2 mM DTT, and kept at -20 ºC for storage.
Initial screening using commercially available screens was not performed. dCTP deaminase:dUTPase from *M. tuberculosis* bears a close resemblance to dCTP deaminase from *E. coli* which has previously been crystallized (Johansson *et al*., 2005). Consequently, initial screening was performed using the *E. coli* dCTP deaminase crystallization conditions (5-8 mg/ml enzyme, 5 mM dCTP or dUTP, 20 mM $MgCl_2$ in 50 mM HEPES pH 7.5 mixed with reservoir solution; 28-30 % (v/v) PEG 400, 0.2 M $MgCl_2$ in 0.1M HEPES pH 7.5) as a starting point. In all experiments the hanging drop vapour diffusion method was applied and the drop was equilibrated over 1 ml of reservoir solution. The concentration of enzyme was lowered compared to the *E. coli* conditions as it otherwise precipitated when concentrations above 4.3 mg/ml were attempted.
Crystallization trials attempting to crystallize the apo-enzyme, the enzyme in complex with inhibitor, dTTP, and the enzyme in complex with substrate analogue, α,β-imido dUTP, were all begun by testing PEG 400 (20 to 40 %) against $MgCl_2$ concentration (50 to 200 mM).

3.1.1.1 Crystallization of dCTP deaminase:dUTPase in the presence of dTTP

All the conditions for crystallization experiments are described in appendix A.1. Just prior to setting up each experiment, the enzyme solution was mixed with dTTP (100 mM stock solution) and $MgCl_2$ (2 M stock solution) to obtain suitable concentrations and kept at 0ºC during set up.
The crystal used for structure determination of dCTP deaminase:dUTPase in complex with its inhibitor, dTTP, was crystallized at 15ºC (appendix A.1, plate A3 number C6). 4 µl of enzyme solution containing 1.8 mg/ml enzyme, 20 mM $MgCl_2$, 5 mM dTTP, 50 mM HEPES pH 6.8 was mixed with 2 µl of reservoir solution, 45 % PEG 400, 200 mM $MgCl_2$, 100 mM HEPES pH 7.5, and allowed to equilibrate. Crystals appeared within one day and grew to a size of 0.11 × 0.05 × 0.05 mm in four days. After 10 days crystals were harvested and cryocooled for

storage and data collection, they were transferred directly from the crystallization drop to liquid nitrogen and kept herein for several weeks until data-collection.

3.1.1.2 Crystallization of apo-dCTP deaminase:dUTPase

All the conditions for crystallization experiments are described in appendix A.2. In the first experiment the enzyme was mixed with $MgCl_2$ (2 M stock solution) just prior to set up. In each experiment the enzyme was kept at 0ºC during set up.

The crystal used for structure determination of apo-dCTP deaminase:dUTPase was crystallized at 4ºC (appendix A.2, plate B5 number C6). 2 µl of enzyme stock solution containing 1.9 mg/ml enzyme, 50 mM HEPES pH 6.8 was mixed with 2 µl of reservoir solution, 20 % PEG 8000, 50 mM $MgCl_2$, 100 mM HEPES pH 7.5, and allowed to equilibrate. Micro-crystals and precipitation appeared after one day and after six weeks one crystal had grown to the size of 0.06 × 0.03 × 0.03 mm. The crystal was cryocooled for storage and data collection, it was transferred from the crystallization drop and soaked in a cryo-solution containing 75 % reservoir solution and 25 % PEG 400 for ca. ten seconds. The diffraction-properties were tested and the crystal was annealed (the cryo-stream was blocked) for five seconds. Diffraction properties were tested again and since the diffraction limit was about 4 Å a dataset was not collected. The crystal was stored in liquid nitrogen for several months and at a later synchrotron trip a complete dataset was collected (section 3.1.2).

3.1.1.3 Crystallization of dCTP deaminase:dUTPase in the presence of α,β-imido dUTP

All the conditions for crystallization experiments are described in appendix A.3. Just prior to setting up each experiment, the enzyme was mixed with α,β-imido dUTP (10 mM stock solution) and $MgCl_2$ (2 M stock solution) to obtain suitable concentrations and kept at 0ºC during set up.

The crystal used for structure determination of dCTP deaminase:dUTPase in complex with substrate-analogue, α,β-imido dUTP, was crystallized at 4ºC (appendix A.3, plate C2 number A5). 2.4 µl of enzyme solution containing 1.6 mg/ml, 17 mM $MgCl_2$, 1.7 mM α,β-imido dUTP, 50 mM HEPES pH 6.8 was mixed with 2 µl of reservoir solution, 30 % PEG 400, 200 mM $MgCl_2$, 100 mM HEPES pH 7.5, and allowed to equilibrate. Crystalline precipitation appeared after one day and after one month tiny crystals had appeared and grown to the size of 0.04 × 0.04 × 0.02 mm. After four months crystals were harvested and cryocooled for storage and data collection, they were transferred directly from the crystallization drop to liquid nitrogen and kept herein until data-collection.

It was also attempted to soak α,β-imido dUTP into crystals of apo-enzyme. 0.5 µl of a 10 mM α,β-imido dUTP solution was added to a crystallization drop containing crystals of apo-enzyme (appendix A.2, plate B4 number D2). After one hour another 0.5 µl was added and after a total two hours crystals were harvested and cryocooled. To another drop containing crystals of apo-enzyme (appendix A.2, plate B3 number D1) 1 µl of the 10 mM α,β-imido dUTP solution was added and crystals were harvested and cryocooled after two weeks. In both cases crystals were transferred directly from the crystallization drop to liquid nitrogen and kept herein until the diffraction properties could be tested.

3.1.2 Data collection and processing

Diffraction properties of crystals were tested at MaxLab (Lund, Sweden) beamline 9-11-3 and at ESRF (Grenoble, France) beamline ID14-3. Diffraction data were collected at the ESRF using a wavelength and temperature of 0.931 Å and 100 K, respectively. Data were collected from 0º to 180º using 1º increments. The diffracting properties of several crystals were tested before finding those suitable for data collection. Four trips to the ESRF were made and the data set for the apo-enzyme was collected by Managing Laboratory Technician Flemming Hansen (Department of Chemistry, University of Copenhagen) while the two others were collected by the author.

The diffraction data were indexed, integrated, scaled and merged using *XDS* and *XSCALE* (Kabsch, 1993). Details on data collection, processing and statistics are presented in table 3.1.

Table 3.1 Data collection, processing and statistics [a].

	dCTP deaminase:dUTPase-dTTP complex	Apo-dCTP deaminase:dUTPase	dCTP deaminase:dUTPase crystallized with α,β-imido dUTP
Synchrotron source	ESRF ID14-3	ESRF ID14-3	ESRF ID14-3
Detector	ADSC Quantum 4	ADSC Quantum 4	ADSC Quantum 4
Data collection temperature (K)	100	100	100
Wavelength (Å)	0.931	0.931	0.931
Resolution limits (Å)	19.6 - 2.0 (2.1 - 2.0)	19.9 - 2.5 (2.8 - 2.5)	20.0 - 3.0 (3.5-3.0)
No. of reflections	247,075 (33,274)	118,238 (33,612)	234,584 (85,651)
No. of unique reflections	22,235 (2,973)	32,512 (9,296)	42,308 (15,761)
Data redundancy	11.1 (11.2)	3.6 (3.6)	5.5 (5.4)
Completeness (%)	98.1 (97.8)	99.1 (99.0)	99.5 (99.9)
R_{merge} (%)[b]	15.9 (59.5)	19.3 (55.1)	31.1 (63.7)
Mean $I/\sigma(I)$	15.01 (4.46)	6.81 (2.24)	5.43 (2.56)
Space group	*R*32	$P2_1$	*R*3 (hexagonal settings)
Unit-cell parameters (Å,º)	a = b = c = 82.640 α = β = γ = 70.750	a = 62.800, b = 96.700, c = 78.00, α = γ = 90.0, β = 97.700	a = b = 97.700, c = 596.300 α = γ = 90.0, β = 120.0
Matthews coefficient, V_M ($Å^3$/Da)	1.95	1.87	2.19
Solvent content (%)	36.83	34.42	43.79

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

[b] $R_{merge} = \Sigma|I - \langle I \rangle|/\Sigma I$, where I is observed intensity and $\langle I \rangle$ is average intensity obtained from multiple observations of symmetry related reflections.

3.1.3 Structure determination and refinement

3.1.3.1 dCTP deaminase:dUTPase in complex with dTTP

The crystal structure of dCTP deaminase:dUTPase in complex with dTTP was solved by molecular replacement. Due to the high resemblance of *M. tuberculosis* dCTP deaminase:dUTPase to *M. jannaschii* dCTP deaminase:dUTPase and *E. coli* dCTP deaminase (see section 3.2.6) both were tested as search models for molecular replacement. Using the peptide chain of the *M. jannaschii* structure (PDB ID: 1OGH, Johansson *et al*., 2003) no solution could be obtained. However, using the peptide chain of the *E. coli* dCTP deaminase structure (PDB ID: 1XS1, Johansson *et al*., 2005) a solution was found. Molecular replacement was performed by *MOLREP* (Vagin and Teplyakov, 1997) from the *CCP*4 program suite (Collaborative Computational Project, Number 4, 1994) and a solution was obtained in space group *R*32 with two molecules (A and B) per asymmetric unit. Multiple rounds of model building in the σ_A-weighted $2\boldsymbol{F}_{obs}$-$\boldsymbol{F}_{calc}$ and $\boldsymbol{F}_{obs}$-$\boldsymbol{F}_{calc}$ electron density map and refinement were carried out using the programs *WinCoot* (Emsley and Cowtan, 2004) and *REFMAC*5 (Murshudov *et al.,* 1997, from the CCP4 program suite), respectively. Refinement progress was observed by the development in the *R*- and R_{free}-factor based on 5 % independent reflections. In both subunits residues 1-189 could be modelled and non-crystallographic symmetry was imposed on residues 1-112 and 118-180 during refinement. During refinement clear density for dTTP coordinated to Mg^{2+} appeared in both subunits. At the end of the refinement process a total of 169 water molecules were added to the model using the automated *ARP/wARP* (Lamzin *et al.*, 2001) procedure in *REFMAC*5 and inspected manually in *WinCoot*. Also, density corresponding to three PEG fragments appeared in the model. Finally, the model was validated using *WHAT IF* (Vriend, 1990) and *PROCHECK* (Laskowski *et al.*, 1993). Refinement statistics are presented in table 3.2.

3.1.3.2 Apo-dCTP deaminase:dUTPase

The crystal structure of apo-dCTP deaminase:dUTPase from *M. tuberculosis* was solved by molecular replacement using the peptide chain A of the enzyme in complex with dTTP (see section 3.1.3.1) as a search model. A solution was found in space group $P2_1$ with six molecules (A, B, C, D, E and F) per asymmetric unit. Model building and refinement was carried out as described (section 3.1.3.1). In the six subunits residues 1 to 161 (A), 159 (B, C, E), 160 (D), 158 (F) could be modelled and non-crystallographic symmetry was imposed on residues 1-55 and 75-158 during refinement. 120 water molecules and seven PEG fragments were added to the model during refinement and it was finally validated as described (section 3.1.3.1). Refinement statistics are presented in table 3.2.

3.1.3.3 dCTP deaminase:dUTPase crystallized in the presence of α,β-imido dUTP

The crystal structure of the enzyme crystallized in contact with α,β-imido dUTP was solved by molecular replacement using the peptide chain A of the enzyme in complex with dTTP (see section 3.1.3.1) as a search model. A solution was found in space group *R*3 (hexagonal setting) with 12 molecules (A, B, C, D, E, F, G, H, I, J, K and L) per asymmetric unit. Model

building and refinement was carried out as described (section 3.1.3.1). Non-crystallographic symmetry was imposed on residues 1-110, 119-161 and 178-183 during refinement. In all 12 subunits residues 1-161 could be modelled while residue 162 could be modelled in ten subunits and in subunit A, residue 163 could also be modelled. A part of the C-terminal could be modelled in all subunits, the shortest stretch (residue 177 to 184) found in subunit L and the longest stretch (residue 173 to 187) found in subunit C. In none of the 12 subunits density corresponding to the nucleotide, α,β-imido dUTP, was visible. However, in subunits B and G density corresponding to the triphosphoryl moiety of the nucleotide appeared in the model and in subunits A, D, E, H, J and K density corresponding to mono-phosphate appeared. Six water molecules were added to the model and it was validated using *PROCHECK* (Laskowski *et al.*, 1993). Refinement statistics are presented in table 3.2.

Table 3.2 Refinement statistics.

	dCTP deaminase:dUTPase-dTTP complex	Apo-dCTP deaminase:dUTPase	dCTP deaminase:dUTPase crystallized with α,β-imido dUTP
No. of reflections in working set	21,122	30,888	40,242
No. of reflections in test set	1,113	1,624	2,224
R (%)[a]	16.13	22.82	25.66
R_{free} (%)[b]	21.70	29.48	31.37
Molecules in asymmetric unit (asu)	2	6	12
No. of water molecules in asu	169	120	6
No of PEG molecules in asu	3	7	0
Rmsd bond length (Å)	0.010	0.011	-
Rmsd angles (º)	1.39	1.56	-
Rmsd dihedral angles (º)	1.81	2.45	-
Average B factor, main chain A/B/… (Å^2)	15.9/15.0	19.1/19.1/18.3/19.7/19.4/19.5	29.0 (average, all chains)
Average B factor, side chain A/B/… (Å^2)	18.3/17.4	19.6/19.7/19.0/20.5/20.0/20.1	31.6 (average, all chains)
Average B factor, protein A/B/… (Å^2)	17.1/16.2	19.5/19.4/18.7/20.1/19.7/19.8	30.2 (average, all chains)
Rmsd B factor, main chain A/B/… (Å^2)	1.37/1.43	0.50/0.52/0.47/0.53/0.51/0.51	0.82 (average, all chains)
Rmsd B factor, side chain A/B/… (Å^2)	1.81/1.80	0.55/0.53/0.55/0.54/0.54/0.53	0.96 (average, all chains)
Average B factor, water molecules (Å^2)	21.6	11.7	3.6
Avarage B factor, dTTP A/B (Å^2)	14.3/13.0	-	-
B factor, Mg A/B (Å^2)	9.8/7.7	-	-
Avarage B factor, PEG (Å^2)	33.4	25.0	-

[a] $R = \Sigma_{work}||\boldsymbol{F}_{obs}|— k|\boldsymbol{F}_{calc}||/\Sigma_{work}\ \boldsymbol{F}_{obs}$.

[b] $R_{free} = \Sigma_{test}||\boldsymbol{F}_{obs}|— k|\boldsymbol{F}_{calc}||/\Sigma_{test}\ \boldsymbol{F}_{obs}$, where $\boldsymbol{F}_{obs}$ and $\boldsymbol{F}_{calc}$ are observed and calculated structure factors, respectively, k is the scale factor, and the sums are the overall reflections in the working set and test set, respectively.

3.1.4 PDB code entries

The coordinates and structure factors have been deposited in the Protein Data Bank with accession codes 2QLP and 2QXX for the apo-enzyme and the enzyme-dTTP complex, respectively. The structure of the enzyme in complex with the substrate analogue has not been published.

3.2 Results and discussion

3.2.1 Crystallization

3.2.1.1 Crystallization of dCTP deaminase:dUTPase in the presence of dTTP

From the first experiment (appendix A.1, plate A1) rod shaped crystal appeared. It was clear that the higher the PEG concentration in the reservoir solution the larger the crystals. 20 % PEG 400 yielded clear drops. 25-30 % PEG 400 produced crystalline precipitation within one week. 35 % PEG produced crystalline precipitation from which needles grew within one week. 40 % PEG 400 produced micro crystals within two days and within one week rod shaped crystals and crystalline precipitation appeared when the $MgCl_2$ concentration was between 100 and 200 mM; the largest crystals appearing in 150 mM $MgCl_2$, having an approximate size of 0.10 × 0.02 × 0.02 mm (figure 3.1 A). One of these diffracted to 2.4 Å
In the second experiment (appendix A.1, plate A2) the PEG 400 concentration interval was raised to 35-55 % and also the $MgCl_2$ concentration interval was raised to 100-250 mM. The experiment was set up at room temperature as the first experiment, but precipitation was observed in all drops just after set up with no change over time. I was puzzled but led to believe that the crystallization conditions were dependent on temperature.

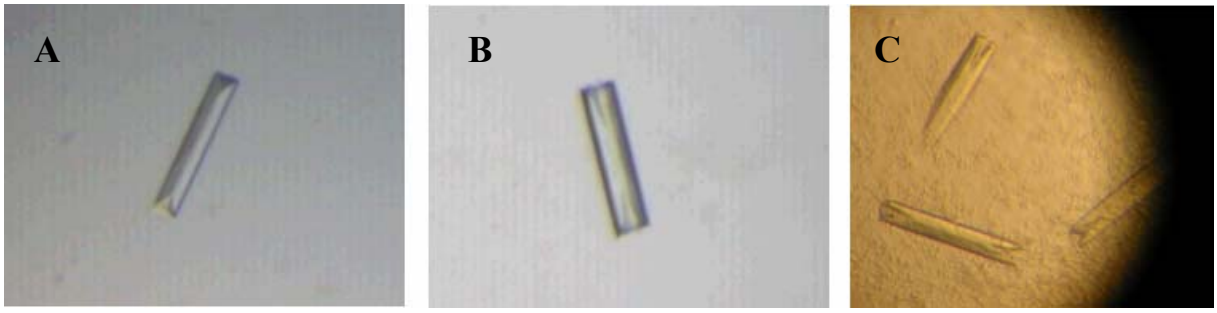


Figure 3.1 Crystals of dCTP deaminase dUTPase crystallized in the presence of dTTP. Conditions from which each of the crystals was grown are found in appendix A.1. A) Plate A1 number C5. The crystal is ca. 100 µm long. B and C) Plate A3 number C6. The largest crystals are ca. 110 µm long.

Therefore in experiment three (appendix A.1, plate A3), the same conditions were set up as in the first experiment. However, the plate was set up on ice and stored at 15ºC. Within one day micro crystals appeared in all drops. Within four days some of the crystals had grow in some of the drops. The crystallization pattern was not the same as observed in the first experiment, but this might be due to the change in set up and storage temperature. The most beautiful

crystals appeared in 35 % PEG 400, 150-250 mM $MgCl_2$ and 40 % PEG 400, 200 mM $MgCl_2$ (0.1 × 0.01 × 0.01 mm), while the largest in 45 % PEG 400, 200 mM $MgCl_2$, the volume of enzyme solution in the crystallization drop was twice that of the reservoir solution (0.11 × 0.05 × 0.05 mm) (figure 3.1 B and C). One of these was used to collect data to 2.0 Å.
In the fourth and fifth experiment (appendix A.1, plates A4 and A5) it was tried to reproduce the results and further optimize by lowering the temperature to 4ºC. However, only crystalline precipitation and smaller crystals were found.

3.2.1.2 Crystallization of apo-dCTP deaminase:dUTPase

From the first experiment (appendix A.2, plate B1) crystals appeared at a $MgCl_2$ concentration of 50 mM in the reservoir solution and at all examined PEG 400 concentrations. These crystals were the shape of long (100 µm) thick needles, the largest ones present at PEG concentrations of 25 % (figure 3.2 A). A PEG 400 concentration of 20 % also yielded some very small crystals the shape of "hexagonal tents". The diffracting properties of some of the needles were tested, but none diffracted.
A second experiment (appendix A.2, plate B2) was set up to optimize these conditions, but precipitation was observed in all drops just after set up with no change over time. This was believed to be due to high temperatures (see section 3.2.1.1).
Two experiments were set up at 4ºC and 15ºC, respectively (identical except for temperature, appendix A.2 plates B3 and B4). At 4ºC plate-like crystals appeared when the PEG concentration in the reservoir was 25 % and the $MgCl_2$ concentration was either 0 or 50 mM (figure 3.2 D to F shows crystals obtained at 50 mM $MgCl_2$). The largest crystals (0.1 × 0.07 × 0.02 mm) were found when the volume of enzyme solution in the crystallization drop was twice the volume of the reservoir solution. At 15ºC crystals were also found at a reservoir concentration of 25 % PEG 400, but only when the volume of enzyme solution in the crystallization drop was twice the volume of the reservoir solution. Plate-like crystals comparable to those obtained at 4ºC were also found at 15ºC, but "hexagonal tent"-shaped crystals also appeared (0.06 × 0.03 × 0.03 mm) (figure 3.2 B and C). The diffracting properties of some of the crystals were tested, but none diffracted.
It was decided to test if raising the PEG 400 concentration and also raising the concentration of the enzyme would yield better crystals (appendix A2 plate B5). However, this only produced precipitation and micro crystals. PEG 1000, PEG 4000 and PEG 8000 were also tested in combination with 50 mM $MgCl_2$. After one day precipitation and microcrystalline material was found in all drops. But after one month a few crystals (figure 3.2 G to I) appeared in one drop equilibrated over reservoir solution containing 20 % PEG 8000 and 50 mM $MgCl_2$. After six weeks the crystals had increased in size and the largest was 0.06 × 0.03 × 0.03 mm. When the diffraction properties were tested at first, high mosaicity and low resolution (about 4 Å) was observed. After annealing the crystal (section 3.1.1.2) low mosaicity was observed but the resolution remained low. Since this was the only diffracting crystal it was saved. After several months the diffraction properties were tested once more and surprisingly enough the resolution had improved markedly and a data set could be collected to 2.5 Å.

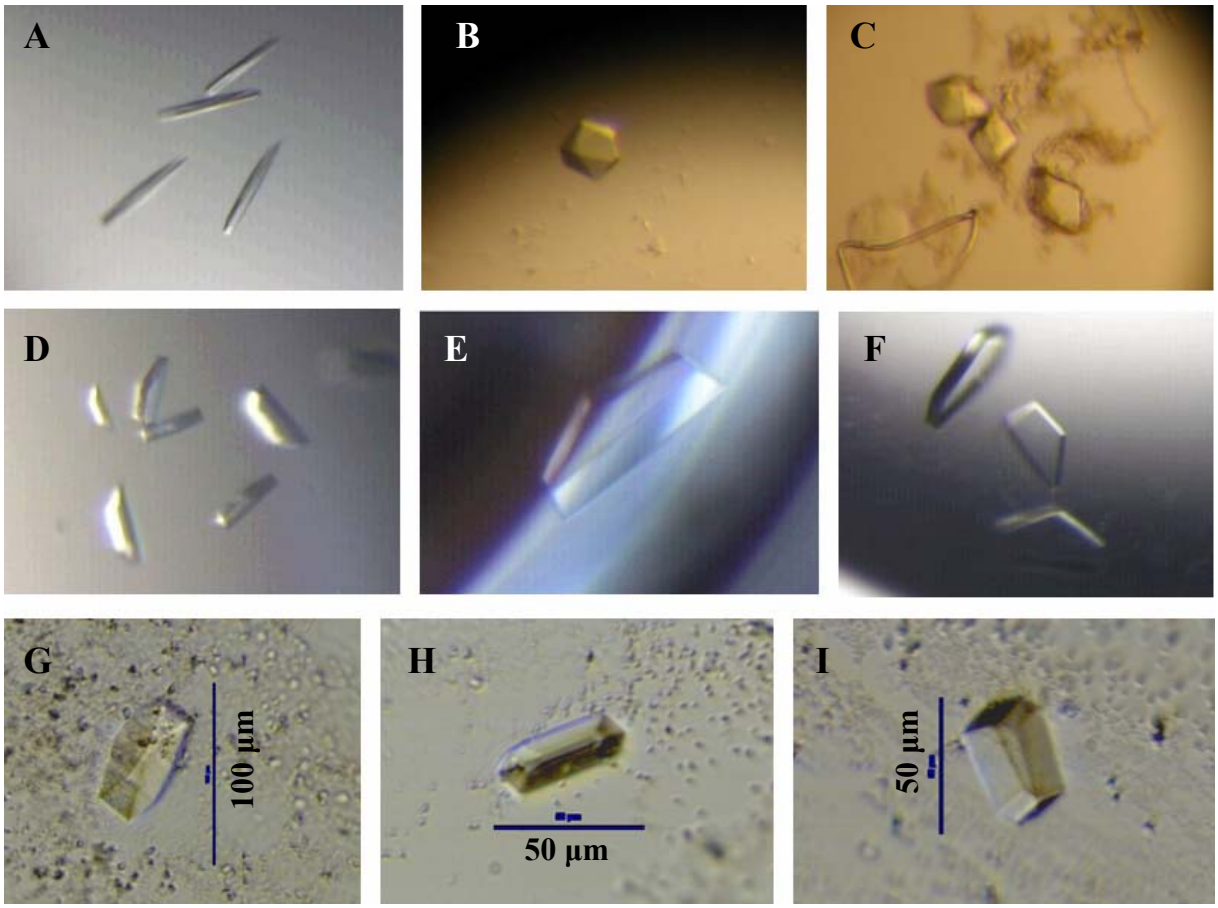


Figure 3.2 Crystals of apo-dCTP deaminase dUTPase. Conditions from which each of the crystals was grown are found in appendix A.2. A) Plate B1 number A2. The crystals are ca. 100 µm long. B) Plate B4 number D2. The diameter is ca. 30 µm. C) Plate B4 number A2. On the longest axis the crystals are ca. 60 µm. D) Plate B3 number D1. The largest crystals are ca. 80 µm long. E and F) Plate B3 number D2. The largest crystals are ca. 100 µm long. G, H and I) Plate B5 number C6. In each panel a bar indicates the size of the crystal. The crystal shown in G was used for data collection.

3.2.1.3 Crystallization of dCTP deaminase:dUTPase in the presence of α,β-imido dUTP

From the first experiment (appendix A.3, plate C1) only precipitation was observed, which was believed to be an effect of temperature (section 3.2.1.1).

In the second and third experiment (identical except for temperature, appendix A.3, plates C2 and C3) the temperature was lowered to 4ºC and 15ºC, respectively. At 15ºC only crystalline precipitation was observed while small crystals grew from the crystalline precipitation at 4ºC when the volume of enzyme solution in the crystallization drop was twice the volume of the reservoir solution. Crystals also appeared in the drop set up with 2.4 µl enzyme solution and 2 µl of reservoir solution containing 30 % PEG 400, 200 mM $MgCl_2$ (figure 3.3). The largest of the crystals grew to a size of 0.04 × 0.04 × 0.02 mm and this was used to collect data to 3 Å. Three more experiments were set up (appendix A.3, plates C4, C5 and C6) to optimize the conditions and test other PEGs as precipitants, but only micro crystals and precipitation

appeared. Soaking apo-crystals in α,β-imido dUTP did not yield diffracting crystals and attempts to anneal crystals resulted in breakdown of the crystal.

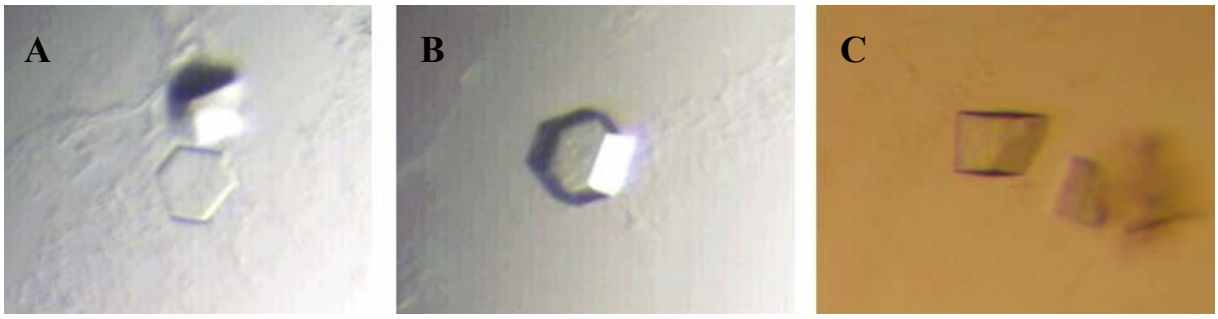


Figure 3.3 Crystals of dCTP deaminase dUTPase crystallized in the presence of α,β-imido dUTP. All crystals are from plate C2 number A5 (appendix A.3). The crystals are ca 40 μm on the longest axis. The crystal shown in B was used for data collection.

3.2.1.4 Comments on crystallization

Getting the enzyme to form crystals seems to be fairly easy, but it is difficult to get the crystals to grow to a size suitable for diffraction. Comparing diffracting properties of the obtained crystals to crystal size and the state of the crystallization drop, it is clear that the larger the crystal the better the diffraction. Also, diffracting crystals were found to either grow from crystalline precipitation or be present in drops with a considerable amount of precipitation.

The reason for poor diffraction and difficulty in obtaining large crystals especially observed in crystals of apo-enzyme and crystals crystallized in the presence α,β-imido dUTP might be due to disordered regions in the enzyme and difficulties in binding of α,β-imido dUTP (see sections 3.2.3 and 3.2.5).

3.2.2 The model of the dCTP deaminase:dUTPase-dTTP complex

The crystal structure was solved in space group *R*32 with two subunits, A and B, in the asymmetric unit. In both subunits residues 1-189 could be modelled leaving only one residue un-modelled. dTTP coordinated to one magnesium ion was found to be present in both A and B. The Ramachandran plot (figure 3.4) produced by *PROCHECK* shows that 89.4 % of the residues are found in the most favoured regions while 10.3 % of the residues are in the additional allowed regions. 0.3 % of the residues corresponding to one residue, Asn-185 A, are found in the generously allowed regions. The precise spatial position of this residue is questionable due to lack of electron density. The backbone of each chain fits nicely into the σ_A-weighted $2\boldsymbol{F}_{obs}$-$\boldsymbol{F}_{calc}$ electron density map except for residue Ile-187-A, which is barely visible. Also a few residues show weak side chain density: Asp-68-B, Arg-166-B and Asn-187-A. There is slight evidence from the $\boldsymbol{F}_{obs}$-$\boldsymbol{F}_{calc}$ electron density map that Ile-32 (not shown) and Thr-114 (figure 3.5) from both chains are present in more than one conformation (rotation around Cα-Cβ bond). However, imposing double conformation onto these residues does not improve the quality of the model and has not been included.

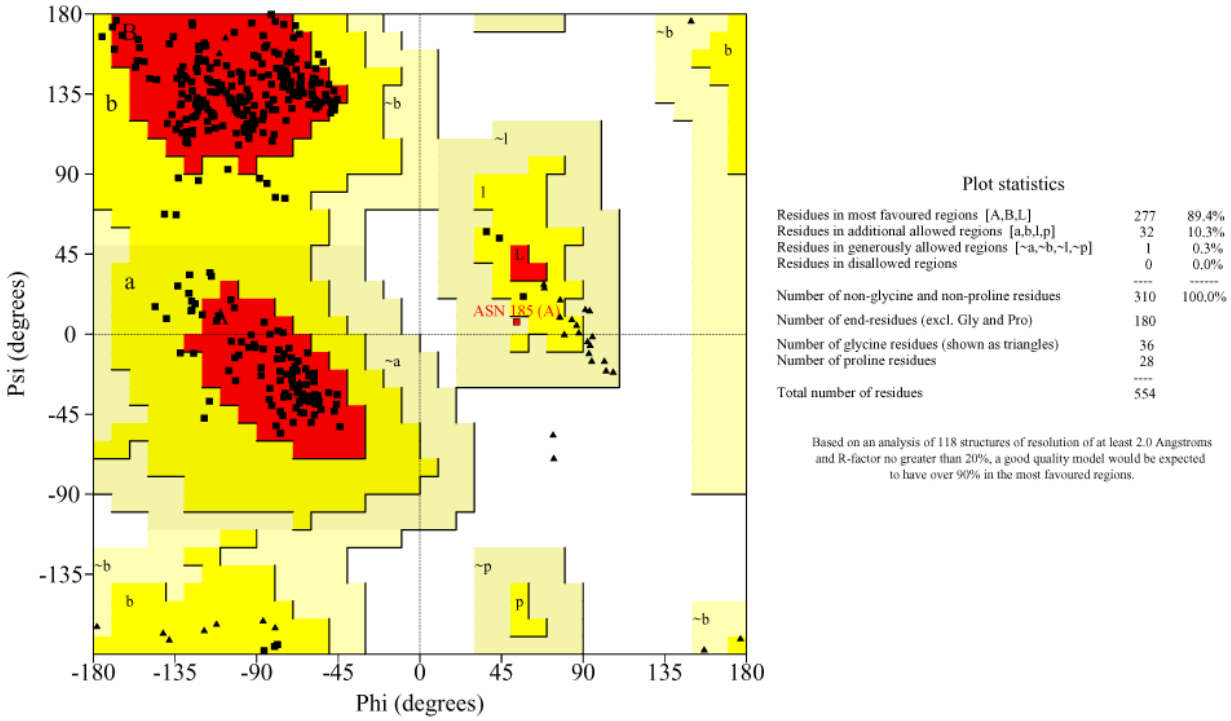


Figure 3.4 Ramachandran plot and plot statistics (produced by *PROCHECK,* Laskowski *et al.*, 1993) of the *M. tuberculosis* dCTP deaminase:dUTPase-dTTP complex structure.

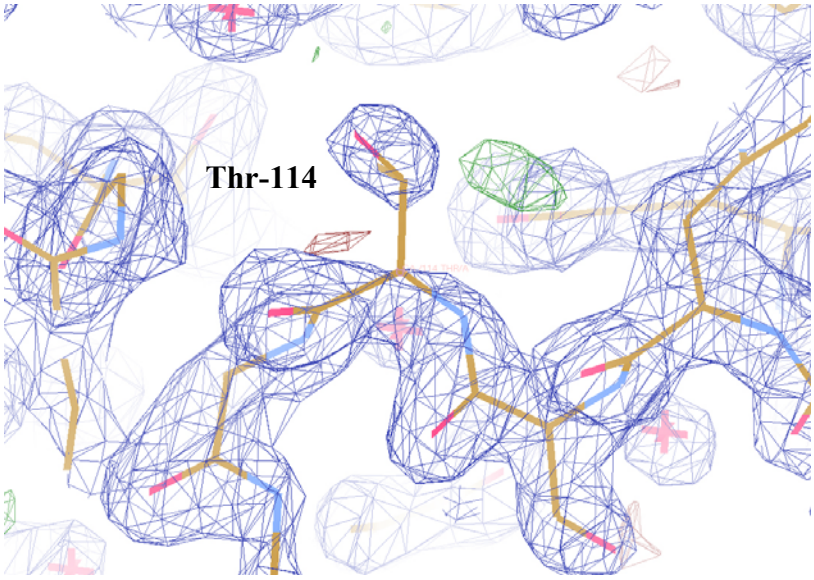


Figure 3.5 Electron density map presenting Thr-114 and the surrounding residues from subunit A of the *M. tuberculosis* dCTP deaminase:dUTPase-dTTP complex structure. The blue mesh represents the $2\boldsymbol{F}_{obs}$-$\boldsymbol{F}_{calc}$ map contoured at 1 σ. The green (positive) and red (negative) mesh represent the $\boldsymbol{F}_{obs}$-$\boldsymbol{F}_{calc}$ map contoured at 3 σ. There is slight evidence that Thr-114 is found in two conformations. The figure was prepared using *WinCoot* (Emsley and Cowtan, 2004).

The Cα atoms of each subunit can be superposed with a root mean square deviation (rmsd) of 0.22 Å while all residues of each subunit can be superposed with a rmsd value of 0.39 Å.

The threefold crystallographic symmetry generates two independent homotrimers of A and B subunits, respectively (see figure 3.6). The height of each trimer along the threefold axis is about 45 Å, while the diameter of the trimer perpendicular to the threefold axis is about 50 Å.

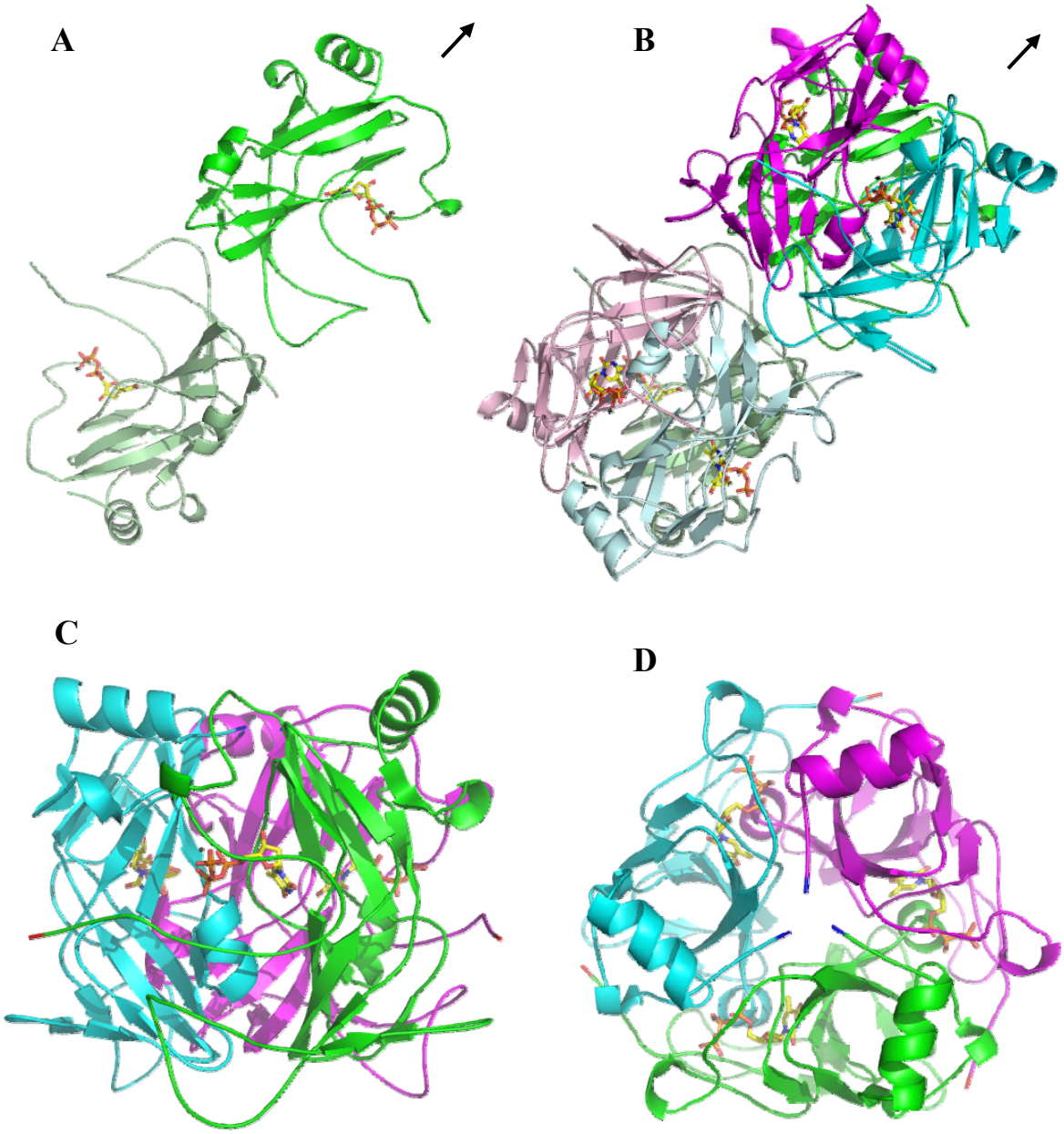


Figure 3.6 Ribbon views of the structure of *M. tuberculosis* dCTP deaminase:dUTPase-dTTP complex. The asymmetric unit contains two subunits (panel A), A and B, shown in green and pale green, respectively. The threefold symmetry-axis (black arrow) generates two homotrimers stacked on top of each other (panel B). In the trimer generated of three A subunits each subunit is shown in green, pink and cyan, respectively. In the trimer generated of three B subunits each subunit is shown in pale green, pale pink and pale cyan, respectively. Panels C and D show the A-trimer perpendicular to the threefold axis and viewed along the threefold axis, respectively. The N- and C-terminals are shown in blue and red, respectively. In each panel dTTP is shown in yellow in ball and stick and Mg^{2+} is shown in grey. The figure was prepared with PyMOL (DeLano W. L., 2002).

Each subunit of *M. tuberculosis* dCTP deaminase:dUTPase is composed of 13 β-strands (β1-β13), two α-helices (α1-α2) and one 3_{10}-helix (γ1) (see figure 3.7). Several turns and loops are present in the structure. The β-strands are arranged in four anti-parallel β-sheets (the position of a single string in the sheet is indicated by the order): S1 (β1, β7 and β10), S2 (β2, β13 and β8), S3 (β4, β3, β6, β11 and β9) and S4 (β5 and β12). The structure of each subunit has a core of β-sheets flanked by α-helices. β-strings 6 and 10 almost extends into strings 7 and 11 respectively. This generates a core structure of a distorted β-barrel, which is common to enzymes having dCTP deaminase activity (Johansson *et al.*, 2005).

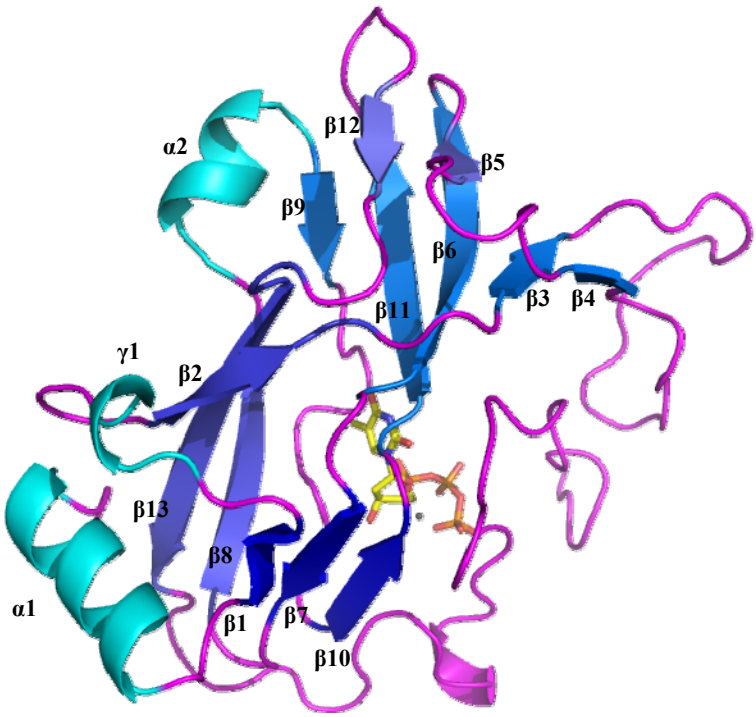


Figure 3.7 Secondary structure of *M. tuberculosis* dCTP deaminase:dUTPase. Ribbon view of subunit A of the *M. tuberculosis* dCTP deaminase:dUTPase-dTTP complex structure where helices are shown in cyan, while the four β-sheets are displayed in four different shades of blue. The inhibitor, dTTP, in complex with magnesium (grey) is shown in yellow in ball and stick. The figure was prepared in PyMOL (DeLano W. L., 2002).

Assembling the trimer also generates three active sites common to the enzyme family (see section 2.3). Each active site is found in a pocket between two monomers. The C-terminal folds back upon the active site like a lid rendering the nucleotide almost completely shielded from the solvent (see figure 3.8).
The nucleotide present in the active site shows an extensive hydrogen bonding network with the surroundings which is shown in figure 3.9. The $2\boldsymbol{F}_{obs}$-$\boldsymbol{F}_{calc}$ electron density map demonstrating the nucleotide is also presented in the figure. Since the active site is composed of residues from neighboring subunits, residues from one of the two are marked with asterisk (e.g. Ala*) in the following description. The numbering of water molecules are those from subunit A. The positions of the water molecules in subunit B are equivalent however, the numbers differ.

From the thymidine moiety O2, N3 and O4 form hydrogen bonds. O2 forms hydrogen bonds with Arg*-106, Gln-174 and Thr-127. N3 forms one hydrogen bond to Thr-127 while O4 forms hydrogen bonds to water molecules 51 and 66. Water 66 also forms a hydrogen bond to Glu-129 and Ser-113. Further, water molecule 51 is hydrogen bonded to Ala-115 and Glu-129. Also, Glu-129 is hydrogen bonded to water molecule 15, which is further hydrogen bonded to Ser*-102. Ser*-102 is also hydrogen bonded to Arg*-106 and it is placed directly above the pyrimidine plane, which is supported by Ile-126 on the other side.

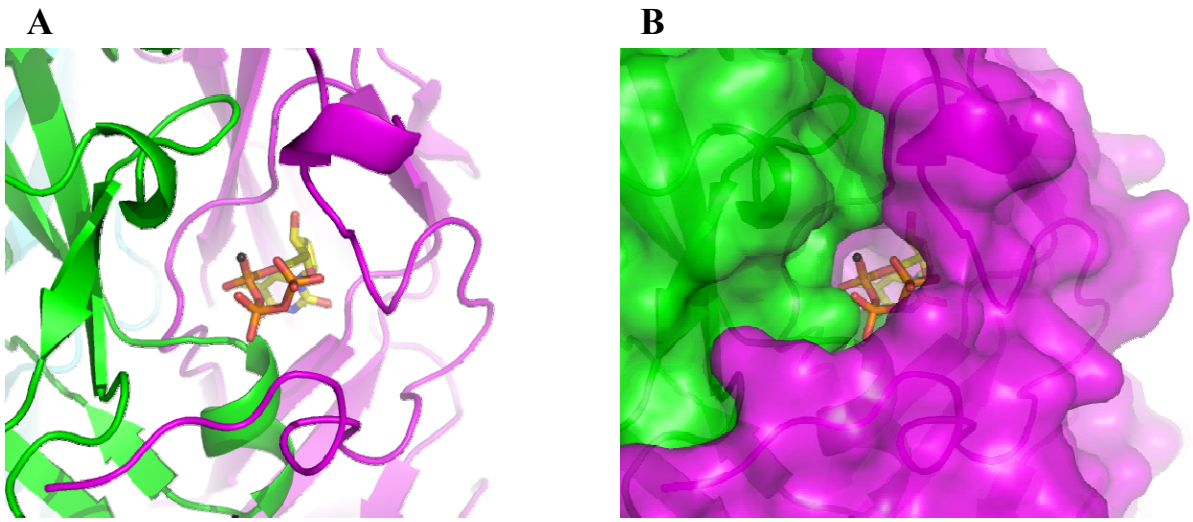


Figure 3.8 Active site architecture of *M. tuberculosis* dCTP deaminase:dUTPase-dTTP complex. dTTP is bound in the active site pocket between two of the three subunits in the active trimer. The two subunits are shown in ribbon view in pink and green, respectively. dTTP and Mg^{2+} are shown in ball and stick (yellow and grey). The C-terminal folds as a lid covering the active site pocket (A) and when the surfaces of both subunits are generated (B) it is evident that the nucleotide is almost completely shielded from solvent only exposed to the solvent by means of a narrow channel. The figure was prepared with PyMOL (DeLano W. L., 2002).

The deoxy-ribosyl moiety is only supported by one hydrogen bond which is O3 bound to Asp-119 and the deoxy-ribosyl moiety is placed in a hydrophobic cleft in a way so that C2 is pointing between Ile-118 and Phe-122. At the same time Asp-119 forms two hydrogen bonds to water molecule 72. Also, the deoxy-ribosyl moiety is supported by Tyr-162 and Tyr-171. The nucleotide is resting on a hydrophobic surface generated by residues Ile-118, Phe-122, Ile-126, Tyr-162 and Tyr-171.

The triphosphoryl moiety of the nucleotide acts as a tri-dentate ligand and is coordinated to a magnesium ion via O1, O2 and O3 from α-, β-, γ-phosphoryl, respectively. Mg^{2+} is also coordinated by water molecules 60, 65 and 166 generating an octahedral coordination sphere. O1 from the α-phosphoryl is further hydrogen bonded to water molecules 60, 65 and 72 while O2 is hydrogen bonded to Ser*-102 and Gln*-148. From the β-phosphoryl, O2 is further hydrogen bonded to water molecules 60 and 116 and Lys-101 while O1 is hydrogen bonded to Ser*-103. From the γ-phosphoryl, O3 is further hydrogen bonded to water molecules 53, 65 and 116 while O2 is bound to Tyr-162 and Lys-170. O1 is hydrogen bonded to Lys-170 and water molecule 163.

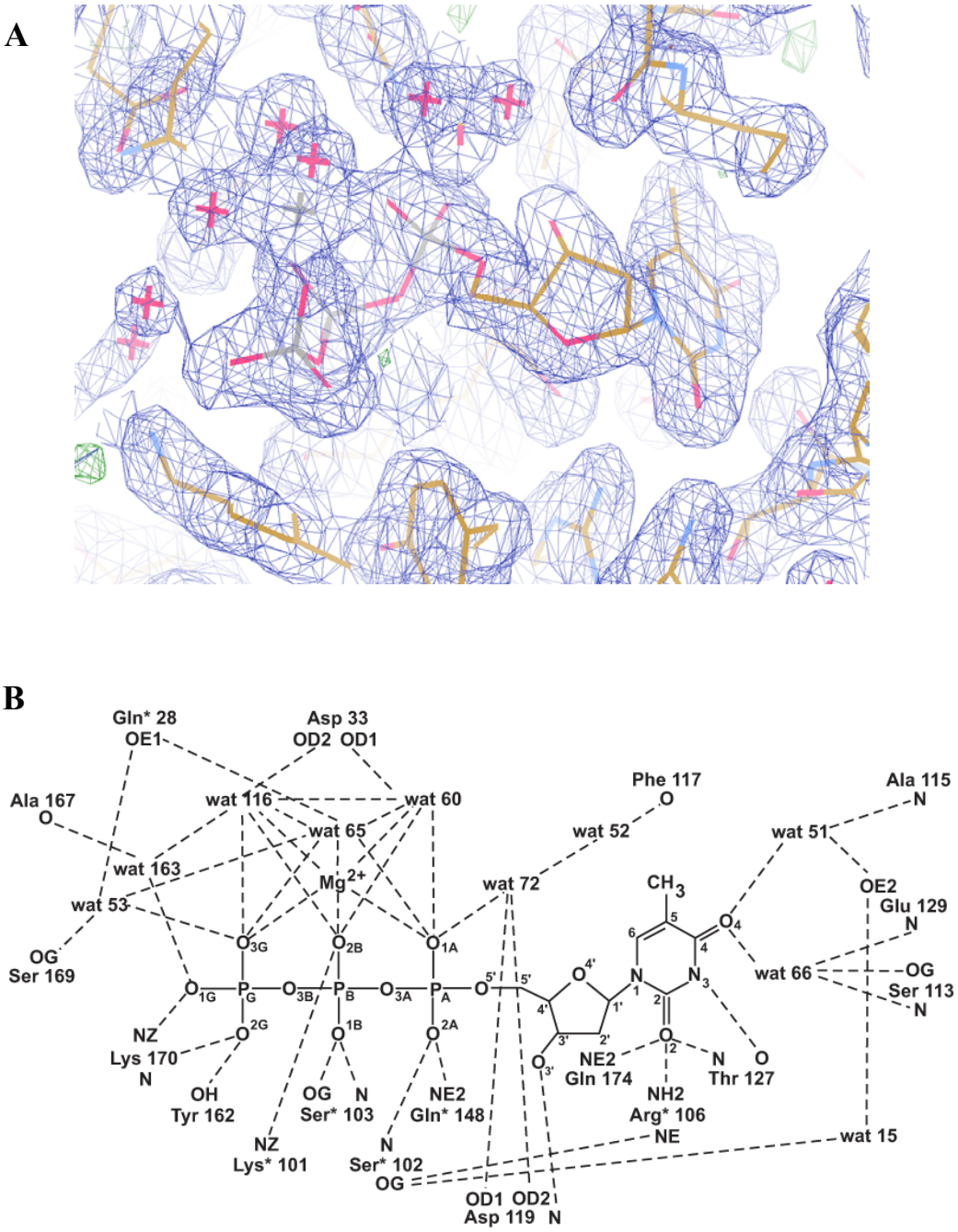


Fig 3.9 The inhibitor and the surrounding residues in the *M. tuberculosis* dCTP deaminase:dUTPase-dTTP complex structure. A) Electron density maps presenting the nucleotide in subunit A. The blue mesh represents the $2\boldsymbol{F}_{obs}$-$\boldsymbol{F}_{calc}$ map contoured at 1 σ and the green mesh represents the positive $\boldsymbol{F}_{obs}$-$\boldsymbol{F}_{calc}$ map contoured at 3 σ. B) Schematic view of the hydrogen bonding network surrounding dTTP. The active site is composed of residues from two subunits in the trimer. Residues without asterisk (e.g. Ala) are from the first subunit while residues marked with asterisk (e.g. Ala*) are from the second subunit. The dotted lines represent hydrogen bonds between atoms. Panel A was prepared in *WinCoot* (Emsley and Cowtan, 2004).

3.2.3 The model of the apo-dCTP deaminase:dUTPase

The structure of the apo-enzyme was solved in space group $P2_1$ with six molecules (A, B, C, D, E and F) in the asymmetric unit. The structure of the asymmetric unit is that of two homotrimers (as presented for the enzyme-dTTP complex structure in figure 3.6 B). However, as evident from the electron density map each subunit is missing approximately 30 residues in the C-terminal. The Ramachandran plot (figure 3.10) produced by *PROCHECK* shows that 89.7 % of the residues are found in the most favoured regions while 9.2 % of the residues are found in the additional allowed regions. 0.4 % of the residues corresponding to residues Glu-59-F and His-112-A, E, are found in the generously allowed regions. Finally, 0.8 % of the residues are found in the disallowed regions corresponding to Thr-114 (A to F). The backbone of each chain fits nicely into the σ_A-weighted $2\boldsymbol{F}_{obs}$-$\boldsymbol{F}_{calc}$ electron density map, but a few residues show weak side chain density: Leu-26-B, Arg-47-A,C, Gln-56-A,E, Asp-58-A,C, Glu-59-C,F, Asp-68-D and Ser-158-B. From the Ramachandran plot Glu-59-F is found in the generously allowed region, but the precise spatial position of this residue is questionable due to the lack of electron density. His-112-A,E are also in the generously allowed region and Thr 114 (A to F) are in the disallowed region. These residues are found in the active site and involved in rearrangements upon binding of the inhibitor (se section 3.2.4) and inspection of the $2\boldsymbol{F}_{obs}$-$\boldsymbol{F}_{calc}$ electron density map shows an excellent fit between spatial arrangement and electron density (not shown). The Cα atoms of each subunit can be superposed with rmsd values ranging from 0.15 Å to 0.25 Å, while all atoms in the subunits can be superposed with rmsd values ranging from 0.50 Å to 0.62 Å.

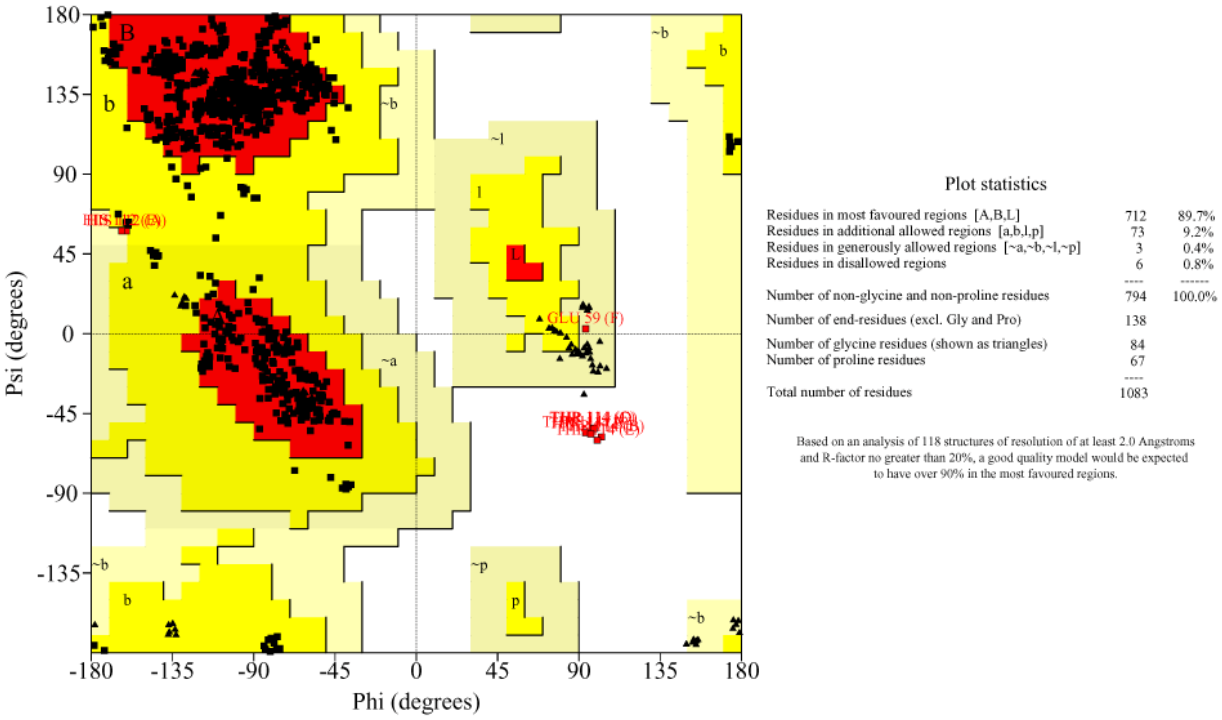


Figure 3.10 Ramachandran plot and plot statistics (produced by *PROCHECK,* Laskowski *et al.*, 1993) of the *M. tuberculosis* apo-dCTP deaminase:dUTPase structure.

3.2.4 Comparison of the structures of the apo-enzyme and the enzyme-dTTP complex

The overall fold of the apo-dCTP deaminase:dUTPase structure is the same as for the structure with the inhibitor bound (see section 3.2.2) and the structural elements are preserved (inspected in *WinCoot*, Emsley and Cowtan, 2004). Superposing monomers from each of the two structures it is apparent that six of the PEG fragments (all but number three in the structure file) from the apo-structure superpose with a part of the C-terminal (residues 177-179) from the dTTP complex structure (figure 3.11). This indicates that these PEG fragments could be wrongly modelled into the electron density and that the electron densities in these parts of the $2\boldsymbol{F}_{obs}$-$\boldsymbol{F}_{calc}$ electron density map may represent a short stretch of the C-terminal backbone. These PEG fragments have not been replaced with amino acids since no side chain density is visible in the $2\boldsymbol{F}_{obs}$-$\boldsymbol{F}_{calc}$ and $\boldsymbol{F}_{obs}$-$\boldsymbol{F}_{calc}$ electron density map.

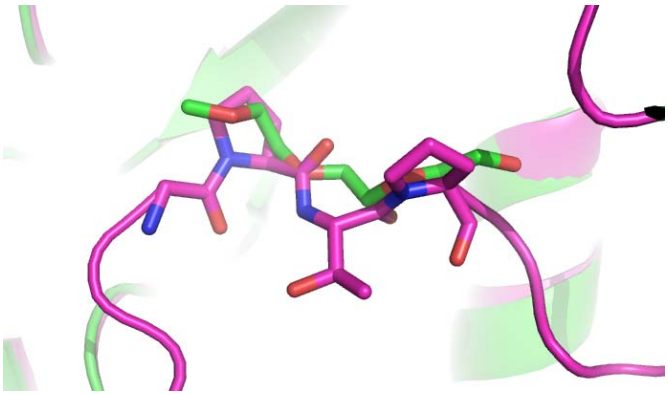


Figure 3.11 Position of PEG fragments in the apo-enzyme structure. Superposing the apo-enzyme structure (green) with the enzyme-dTTP complex structure (pink) shows that some of the PEG fragments from the apo-enzyme structure superpose with residues 177-179 from the C-terminal of the enzyme-dTTP complex (only one of six corresponding superpositions is shown). The figure was prepared in PyMOL (DeLano W. L., 2002).

In figure 3.12 subunit A from the structure of the apo-enzyme has been superposed with subunit A from the structure of the enzyme with the inhibitor bound. The Cα atoms of the subunits from the apo-structure and those of the enzyme-dTTP complex structure can be superposed with rmsd values ranging from 0.97 Å to 1.03 Å, while all atoms in the subunits can be superposed with rmsd values ranging from 1.35 Å to 1.46 Å. Comparing the two structures two major differences are apparent and these are described in the following sections (3.2.4.1 and 3.2.4.2).

3.2.4.1 Closing of the binding pocket

The first difference is that upon binding of the inhibitor, the C-terminal folds back on the enzyme generating a lid almost covering the binding pocket (see figure 3.12). Four residues from the C-terminal are involved in binding to the inhibitor either directly (Tyr-162 and Lys-170) or via water molecules (Ala-167 and Ser-169), see section 3.2.2. The structure of the apo-enzyme on the other hand has a disordered C-terminal. The inhibitor is actually a dTTP-Mg^{2+} complex and the magnesium ion appears to play an important role in positioning the tri-

phophoryl moiety correctly in the active site. It also plays an important role in the hydrogen-bonding network surrounding the nucleotide hereby also in closing the lid of the active site on top of the nucleotide.

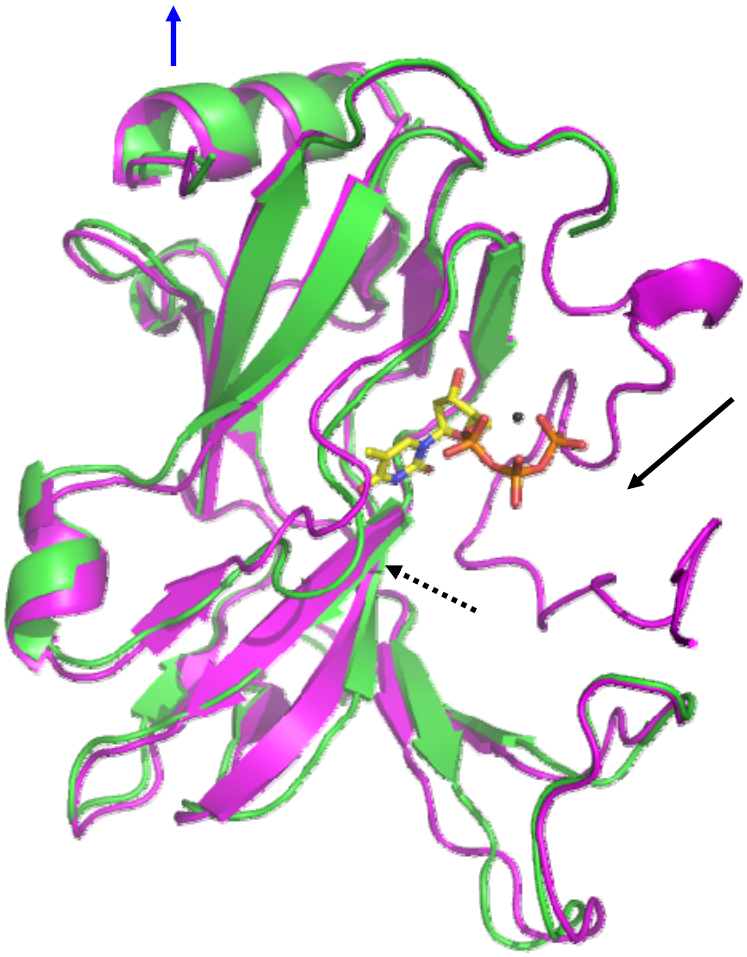


Figure 3.12 Comparison of the apo- and the dTTP-complex structures of *M. tuberculosis* dCTP deaminase:dUTPase. Ribbon view of superimposed subunits of the *M. tuberculosis* dCTP deaminase:dUTPase apo-form (green) and dTTP-complex (pink) with dTTP and Mg^{2+} shown in ball and stick (yellow and grey). The solid black arrow points to the C-terminal, which is only visible in the dTTP complex. The dotted arrow points to changes in the active site induced by binding of dTTP. The blue arrow represents the position of the threefold axis. The figure was prepared in PyMOL (DeLano W. L., 2002).

From a superposition (not shown) of the *M. tuberculosis* dCTP deaminase:dUTPase-dTTP complex structure with the *E. coli* dCTP deaminase structures when the enzyme is in contact with dCTP (PDB ID: 1XS4, E138A, Johansson *et al*., 2005) or dUTP (PDB ID: 1XS1, Johansson *et al*., 2005) an almost identical closing of the C-terminal lid is observed. This is also observed from a superposition with the *M. jannaschii* E145A dCTP deaminase:dUTPase-α,β-imido dUTP complex structure (PDB ID: 2HXD, Siggaard *et al.*, 2008.). These findings strongly indicate that the observed C-terminal fold is also present when the substrate dCTP binds to *M. tuberculosis* dCTP deaminase:dUTPase and probably also during the enzymatic reactions. An intriguing difference is that the *E. coli* E138A dCTP deaminase-dTTP complex structure (PDB ID: 2J4Q, Johansson *et al*., 2007) does not have an ordered C-terminal.

3.2.4.2 Conformational changes of residues 110-117

The second difference is in the binding pocket in the region where the thymidine moiety binds (figure 3.12). Figure 3.13 shows a close-up of the residues 109-118 from the superposition of the apo-enzyme and the enzyme in complex with the inhibitor. The situation is shown from two different angles. Upon binding dTTP, a major change is apparent in this stretch. Due to steric hindrance in accommodating the methyl group of the thymidine moiety, Gly-116 is

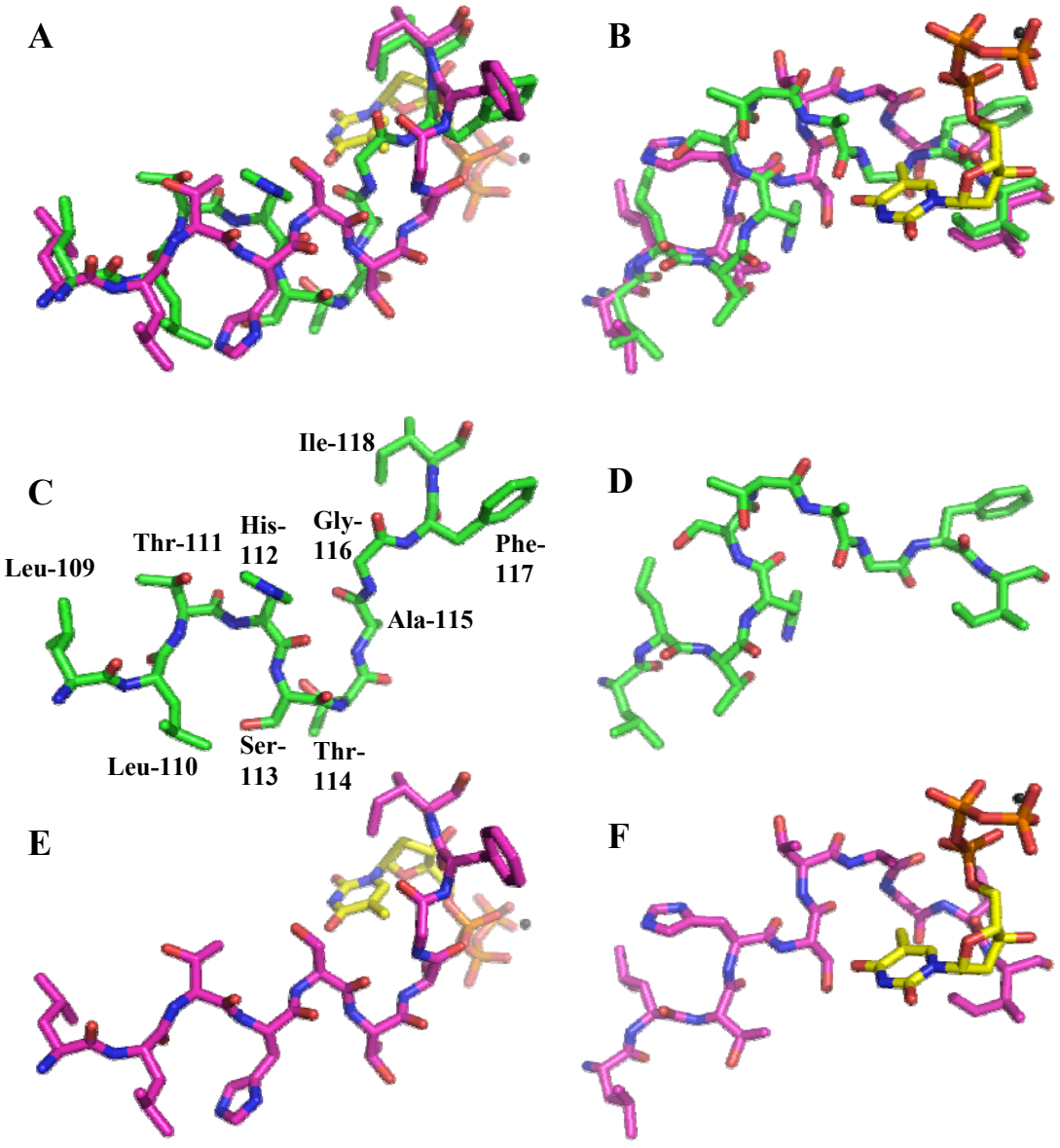


Figure 3.13 Structural changes induced by binding of the inhibitor, dTTP. Close-up of residues 109-118 from a superposition of the apo-structure (green) and dTTP-complex structure (pink), the residues are shown as sticks. The situation is shown from two different directions (A and B). The nucleotide and Mg^{2+} are shown in ball and stick (yellow and grey). To outline the changes, this residue stretch is also shown for each the two structures; C and D presents the apo-structure while E and F presents the dTTP-complex structure. Panels A, C, E are from one direction and B, D, F are from the second direction. In all panels the residues are (from left) Leu-109, Leu-110, Thr-111, His-112, Ser-113, Thr-114, Ala-115, Gly-116, Phe-117 and Ile-118 as shown in C. The figure was prepared with PyMOL (DeLano W. L., 2002).

pushed away. In the C-terminal direction this causes a shift in Phe-117 while Ile-118 has almost the same position. In the N-terminal direction binding of dTTP can almost be described as a wave-effect in the backbone. The side-chain position of Ala-115 is changed by 2.8 Å and it is now close to the α-phosphoryl of the nucleotide. Thr-114 is also shifted markedly and Ser-113 and His-112 almost switch positions. Thr-111 changes position slightly, while Leu-110 is pushed backwards due to the new position of His-112. Finally, almost no change is observed in Leu-109. The hydrogen bonding network in this stretch (109-118) changes as a consequence of the conformational changes. The structure of the apo-enzyme has a smaller number of hydrogen bonds, but this might be due to the lower resolution and few water molecules in the structure. In the trimer structure of the apo-enzyme no hydrogen bonds are found between separate monomers in this region. However, in the trimer structure of the enzyme-dTTP complex interactions between Thr-111 and Thr-114 in separate monomers causes a connection between all three monomers in this region.
The impact these structural changes have on the trimer and the regulation of enzyme activity is presented and discussed in section 3.2.7.

3.2.5 The model of dCTP deaminase:dUTPase crystallized in the presence of α,β-imido dUTP

The structure of the *M. tuberculosis* dCTP deaminase:dUTPase crystallized in the presence of α,β-imido dUTP was solved in spacegroup *R*3 with 12 subunits in the asymmetric unit. Due to the poor resolution (3 Å) and the high number of subunits in the asymmetric unit every single residue was not individually examined. During refinement emphasis was made to refine subunits A to D and residues in the vicinity of the active site in all subunits. Even though the model is poor, some conclusions can still be made.
The Ramachandran plot (figure 3.14) produced by *PROCHECK* shows that 85.2 % of the residues are found in the most favoured regions while 12.7 % of the residues are found in the additional allowed regions. 1.2 % of the residues are found in the generously allowed regions while 0.8 % of the residues are found in the disallowed regions. From the plot it is obvious that the residues 110-115 are those which are dominating the generously allowed and disallowed regions. This is interesting since these residues are found in the active site and they are observed to change conformation upon binding of nucleotide (see section 3.2.4.2).
In all subunits the backbone of each chain fits nicely into the σ_A-weighted $2\boldsymbol{F}_{obs}$-$\boldsymbol{F}_{calc}$ electron density map. In none of the subunits the entire C-terminal could be modelled as it could in the enzyme-dTTP complex structure (see section 3.2.2). However, 7 to 15 of the C-terminal residues could be modelled in the subunits. This contrasts the apo-structure (see sections 3.2.3. and 3.2.4), in which none of the C-terminal residues could be convincingly modelled. In two out of the 12 subunits the triphosphoryl moiety of α,β-imido dUTP could be modelled. Comparing subunits A to D with subunits A from the apo- and dTTP-complex structures differences are apparent in the region 112 to 116 (figure 3.15). This new structure does not resemble either of the apo- or dTTP-complex structures. In this structure the region 112-116 appears as if it is pushed into the active site and residues 114 and 115 are seen to overlap with the nucleotide of the dTTP-complex structure.
These findings indicate that the α,β-imido dUTP is partially bound in some of the subunits and the lid found to cover the active site (see section 3.2.2) is partially closed on top of the

active site. However, since it was very difficult to obtain crystals when α,β-imido dUTP was present in solution it might be that the nucleotide is not able to bind in the active site. Since dTTP binds very convincingly in the enzyme (see figure 3.9 A) only two factors can be responsible for this. The first is the methyl group of dTTP, not present in dUTP (figure 2.3). Since dUMP is the product, the uridine moiety might show lower binding capability in the active site than the thymidine moiety. However, kinetic results indicate that the enzyme has equal affinity for dCTP and dUTP (Helt *et al.*, 2008). The second is the replacement of

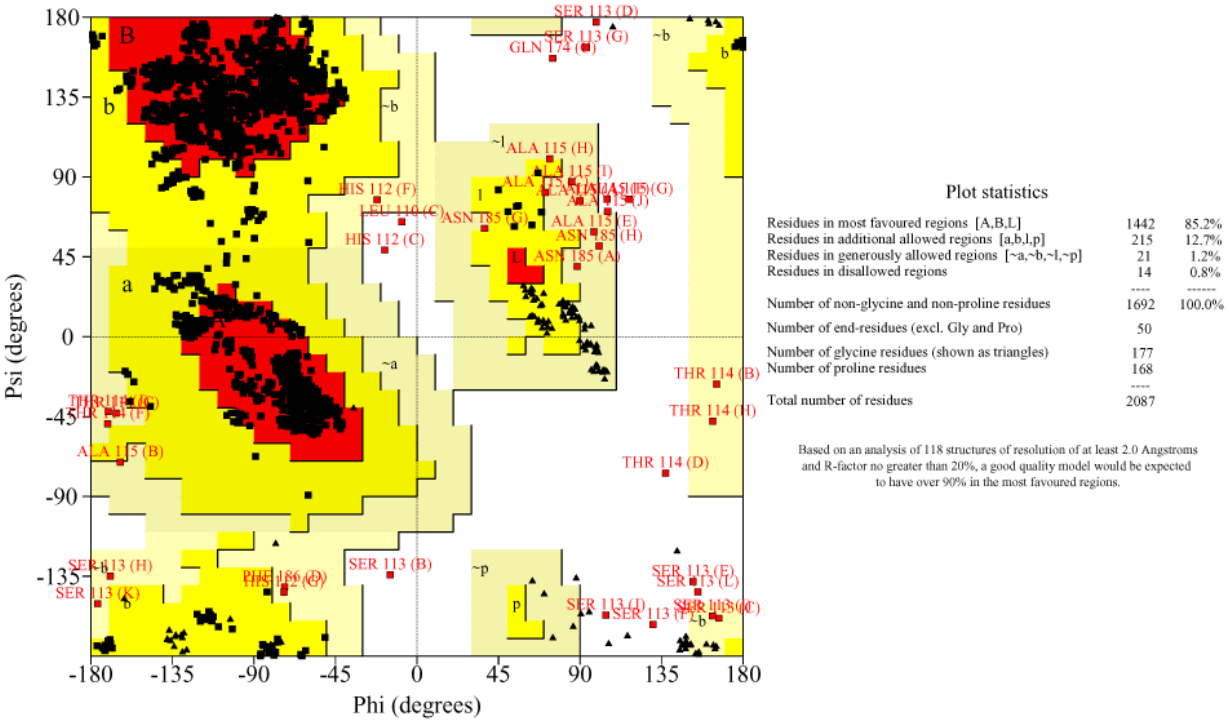


Figure 3.14 Ramachandran plot and plot statistics (produced by *PROCHECK,* Laskowski *et al.*, 1993) of the *M. tuberculosis* dCTP deaminase:dUTPase structure obtained from crystals grown in the presence of α,β-imido dUTP.

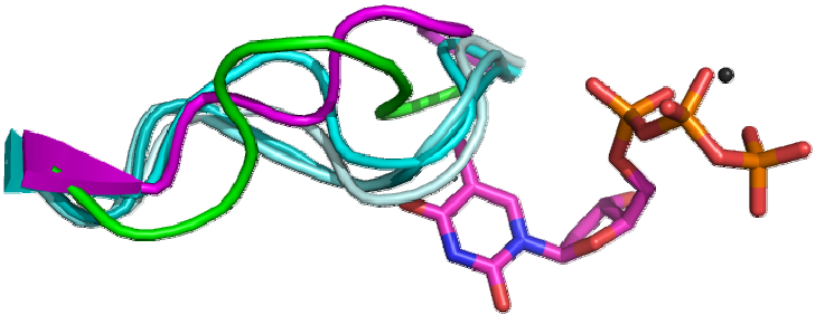


Figure 3.15 Differences in the binding pocket. Superposition of subunit A from the dTTP-complex structure (pink), subunit A from the apo structure (green) and subunits A to D from the structure obtained when the enzyme was crystallized in contact with α,β-imido dUTP (four shades of cyan). The backbone of residue 110 to 117 of each subunit is shown and dTTP from the dTTP-complex structure is also shown in ball and stick. The backbones of the subunits from the structure obtained from crystallizing the enzyme with α,β-imido dUTP are seen to overlap with dTTP. The figure was prepared with PyMOL (DeLano W. L., 2002).

the α,β-bridging oxygen with an imido group. In dUTPases α,β-imido dUTP is found to bind correctly in the active site and it is not hydrolysed (Chan *et al.*, 2004). In the substrate analogue O3 of α-phosphoryl is replaced by NH, which changes the direction of a possible hydrogen bond. This might prove unfavorable for nucleotide binding in the bifunctional enzymes since there are differences in the active site (see section 3.2.6.2). Another possibility is that α,β-imido dUTP is not resistant to hydrolysis in the bifunctional enzymes (nothing has yet been published on this matter). In this connection it is worth noting that the native enzyme in contact with α,β-imido dUTP has not been crystallized, neither from *E. coli* dCTP deaminase nor from *M. jannaschii* dCTP deaminase:dUTPase.

3.2.6 Structural comparisons with other members of the enzyme family

Structures of two dCTPdeaminase:dUTPases have now been solved, namely those of *M. jannaschii* (Johansson *et al.*, 2003; Siggaard *et. al.*, 2008) and *M. tuberculosis* (Helt *et al.*, 2008). Structures of only one dCTP deaminase have been solved, which is that of *E. coli* (Johansson *et al.*, 2005 and 2007). Several structures of dUTPases have been solved.

Mutational analyses have been performed on some of the enzymes from the family to establish which residues play key roles in the enzymatic reactions (Barabás *et al.*, 2004; Chan *et al.*, 2004; Johansson *et al.*, 2003, 2005 and 2007). From the *E. coli* dCTP deaminase the mutations E138A, E138Q, R115A (Johansson *et al.*, 2005), H121A and V122G (Johansson *et al.*, 2007) were found to inactivate the enzyme. From the *M. jannaschii* bifunctional enzyme it was found that the D135N mutant was inactive concerning both functions while the E145Q was an inactive deaminase but retained 25 % of the dUTPase activity (Johansson *et al.*, 2003). The glutamate found at position 138 in the *E. coli* enzyme corresponds to the glutamates found in position 145 in the *M. jannaschii* enzyme and in position 129 in the *M. tuberculosis* bifunctional enzyme (figure 3.16 A). The residues E138, R115 and H121 from the *E. coli* enzyme are conserved in corresponding positions in enzymes having dCTP deaminase activity (figure 3.16 A) and E138 and R115 are suggested to participate in the deamination reaction (see section 3.2.6.2 and Johansson *et al.*, 2005). From *E. coli* dUTPase the mutation D90N rendered the enzyme inactive (Barabás *et al.*, 2004). The *E. coli* residue D90 corresponds to *M. jannaschii* D135 and *M. tuberculosis* D119 (bifunctional enzyme). It is conserved in all members of the family (figure 3.16) and believed to play a key role in positioning the nucleotide correctly in the active site (Chan *et al.*, 2004).

Throughout the remaining part of chapter 3, amino acid numbers will refer to *M. tuberculosis* dCTP deaminase:dUTPase unless otherwise stated.

3.2.6.1 Overall structure

Figure 3.16 A shows a structure based sequence alignment of the *M. tuberculosis* dCTP deaminase:dUTPase-dTTP complex structure, the *M. jannaschii* E145A dCTP deaminase: dUTPase-α,β-imido dUTP complex structure (PDB ID: 2HXD, Siggaard *et. al.*, 2008, the mutation has been excluded in the alignment) and the *E. coli* dCTP deaminase-dUTP complex structure (PDB ID: 1XS1, Johansson *et al.*, 2005). Figure 3.17 shows a superposition of subunit A from each of the three structures. From the figures it is apparent that there is a very high structural similarity and the overall fold is almost conserved. There are minor

A

```
                                                                  *
1- Bifunc  M. tube   MLLSDRDLRAEISSGRLGIDPFDD-TLVQPSSIDVRLDCLFRVFNN-TRYTHIDPA---- 54
2- Bifunc  M. jann   MILSDKDIIDYVTSKRIIIKPFNK-DFVGPCSYDVTLGDEFIIY-DD---EVYDLS---- 51
3- dCTPdeam E. coli  MRLCDRDIEAWLDEGRLSINPRPPVERINGATVDVRLGNKFRTF-RGHTAAFIDLSGPKD 59

                                                                              *
1- Bifunc  M. tube   -KQ--QDELTS-LVQPVDGEPFVLHPG----------------------EFVLGSTLEL 87
2- Bifunc  M. jann   -KE--L-N-YK-RIKI--KNSILVCPLNYNLTEEKINYFKEKYNVDYVVEGGVLGTTNEY 103
3- dCTPdeam E. coli  EVSAALDRVMSDEIVLDEGEAFYLHPG----------------------ELALAVTLES 96

                                §
                                #                   §
                                *   #         #   **    *    #
1- Bifunc  M. tube   FTLPDNLAGRLEGKSSLGRLGLLTHSTAGFIDPGFSGHITLELSNVANLPITLWPGMKIG 147
2- Bifunc  M. jann   IELPNDISAQYQGRSSLGRVFLTSHQTAGWIDAGFKGKITLEIVAFDK-PVILYKNQRIG 162
3- dCTPdeam E. coli  VTLPADLVGWLDGRSSLARLGLMVHVTAHRIDPGWSGCIVLEFYNSGKLPLALRPGMLIG 156

                     §
1- Bifunc  M. tube   QLCMLRLTSPSEHPYG-SSRAGSKYQGQRGPTPSRSYQNFIRST 190
2- Bifunc  M. jann   QLIFSKLLSPADVGYSERK--TSKYAYQKSVMPSLIHLDNHKKD 204
3- dCTPdeam E. coli  ALSFEPLSGPAVRPYNRRE--DAKYRNQQGAVASRIDKD----- 193
```

B

```
                                                          *
1- Bifunc  M. tube   MLLSDRDLRAEISSGRLGIDP----------FDDTLVQPSSIDVRLDCLFRVFNNTRYT 49
2- dUTPase M. tube   ------------MSTTLAIVRLDPG----LPLPSRAHDGDAGVDLYSAE---------- 33
3- dUTPase E. coli   ------------MMKKIDVKILDPRVGKEFPLPTYATSGSAGLDLRACLN---------- 38
4- dUTPase Human     --------------MQLRFARLSEHA---T-APTRGSARAAGYDLYSAY---------- 31
5- dUTPase EIAV      ---------------MLAYQG--------TQIKEK-RDEDAGFDLCVPY---------- 25
6- dUTPase FIV       ----------------MIIEG--------DGILDK-RSEDAGYDLLAAK---------- 24

                                                                  §
                                                                  #
                                                  *               *   #
1- Bifunc  M. tube   HIDPAKQQDELTSLVQPVDGEPFVLHPGEFVLGSTLELFTLPD-NLAGRLEGKSSLGRL- 107
2- dUTPase M. tube   ---------------------DVELAPGRRALVRTGVAVAVPF-GMVGLVHPRSGLATRV 71
3- dUTPase E. coli   ---------------------DAVELAPGDTTLVPTGLAIHIADPSLAAMMLPRSGLGHKH 78
4- dUTPase Human     ---------------------DYTIPPMEKAVVKTDIQIALPS-GCYGRVAPRSGLAAKH 69
5- dUTPase EIAV      ---------------------DIMIPVSDTKIIPTDVKIQVPP-NSFGWVTGKSSMAKQ- 62
6- dUTPase FIV       ---------------------EIHLLPGEVKVIPTGVKLMLPK-GYWGLIIGKSSIGSK- 61

                           §
                     #  **    *    #                   §
1- Bifunc  M. tube   GLLTHSTAGFIDPGFSGHITLELSNVA-NLPITLWPGMKIGQLCMLRLTSPSEHPYGSSR 166
2- dUTPase M. tube   GLSIVNSPGTIDAGYRGEIKVALINLDPAAPIVVHRGDRIAQLLVQRVELVELVEVSSFD 131
3- dUTPase E. coli   GIVLGNLVGLIDSDYQGQLMISVWNRG-QDSFTIQPGERIAQMIFVPVVQAEFNLVEDFD 137
4- dUTPase Human     FIDVG--AGVIDEDYRGNVGVVLFNFG-KEKFEVKKGDRIAQLICERIFYPEIEEVQALD 126
5- dUTPase EIAV      GLLIN--GGIIDEGYTGEIQVICTNIG-KSNIKLIEGQKFAQLIILQHHSNSRQPWDENK 119
6- dUTPase FIV       GLDVL--GGVIDEGYRGEIGVIMINVS-RKSITLMERQKIAQLIILPCKHEVLEQGKVVM 118


1- Bifunc  M. tube   AGSKYQGQRGPTPSRSYQNFIRST 190
2- dUTPase M. tube   EAGLASTSRGDGGHGSSGGHASL- 154
3- dUTPase E. coli   A-----TDRGEGGFGHSGRQ---- 152
4- dUTPase Human     D-----TERGSGGFGSTGKN---- 141
5- dUTPase EIAV      I-----SQRGDKGFGSTGVF---- 134
6- dUTPase FIV       D-----SERGDNGYGSTGVFSSW- 136
```

*(**Figure 3.16**: figure legend found on the next page)*

Figure 3.16 Structure-based sequence alignments. A) *M. tuberculosis* dCTP deaminase:dUTPase-dTTP complex compared to *M. jannaschii* E145A dCTP deaminase:dUTPase-α,β-imido dUTP complex (PDB ID: 2HXD, Siggaard *et. al.*, 2008, the mutation was not included) and *E. coli* dCTPdeaminase-dUTP complex (PDB ID: 1XS1, Johansson *et al*., 2005). B) *M. tuberculosis* dCTP deaminase:dUTPase-dTTP complex compared to *M. tuberculosis* dUTPase-α,β-imido dUTP complex (PDB ID: 1SIX, Chan *et al.*, 2004), *E. coli* dUTPase-α,β-imido dUTP complex (PDB ID: 1RN8, Barabas *et al.*, 2004), dUTPase-dUDP complex from human (PDB ID: 1Q5H, Mol *et al.*, 1996), dUTPase from equine infectious anaemia virus (EIAV, PDB ID: 1DUN, Dauter *et al.*, 1999) and dUTPase-dUMP complex from feline immunodeficiency virus (FIV, PDB ID: 1F7N, Prasad *et al.*, 2000).
In both panels six residues conserved in all eight examined structures are marked above the alignment with *****. Residues believed to play an active part the deamination reaction are marked with **#**, while residues believed to play an active part the hydrolysis are marked with **§**. The structure-based sequence alignments were prepared with INDONESIA (Madsen *et al.*, 2002), manually inspected in *WinCoot* (Emsley and Cowtan, 2004) and corrected in BioEdit (Hall, Ibis Biosciences).

differences in the position, lengths and presence of β-strings within the three structures. However, the central core with the distorted β-barrel is preserved. α-helices found in the *M. tuberculosis* dCTP deaminase:dUTPase-dTTP complex structure are also present in the other structures. Both the *E. coli* dCTP deaminase and the *M. jannaschii* dCTP deaminase:dUTPase structures contain an α-helix which is not present in *M. tuberculosis* dCTP deaminase:dUTPase structure, but these helices are not found in corresponding positions. The extra helix found in *E. coli* dCTP deaminase is present in a region where both of the bifunctional enzymes show no secondary structural elements. In this region the E. coli structure shows 8 and 12 residue inserts compared to the *M. tuberculosis* and the *M. jannaschii* dCTP deaminase:dUTPase structures, respectively. *M. jannaschii* dCTP deaminase:dUTPase contain a 23-residue stretch not present in the other enzymes. The stretch containing the α-helix is present as a bulge in the "bottom" of the subunit. It is probably due to this stretch that the *M. jannaschii* dCTP deaminase:dUTPase structure could not perform as a search model in solving the structure of *M. tuberculosis* dCTP deaminase:dUTPase (see section 3.1.3.1). When superposing the structures of the bifunctional enzymes the bulge in the *M. jannaschii* structure clashes with neighbouring subunits from the *M. tuberculosis* structure (results not shown). From the structural sequence alignment it is clear that 46 residues are identical within the three structures. Comparing the structures two on two; *M. tuberculosis* vs. *E. coli*, *M. tuberculosis* vs. *M. jannaschii* and *E. coli* vs. *M. jannaschii*, show an additional 34, 17 and 14 identical residues, respectively. This indicates that the bifunctional enzyme from *M. tuberculosis* is more closely related to the dCTP deaminase from *E. coli* than to the bifunctional enzyme from *M. jannaschii*.
The *M. tuberculosis* dCTP deaminase:dUTPase also shows a high resemblance to dUTPases. Figure 3.18 shows a superposition of one subunit from each of the two structures; the *M. tuberculosis* dCTP deaminase:dUTPase-dTTP complex and the *M. tuberculosis* dUTPase-α,β-imido dUTP complex (PDB ID: 1SIX, Chan *et al.*, 2004). Figure 3.16 B shows a structure based sequence alignment between the bifunctional enzyme and five dUTPases. The enzymes share an almost identical core structure from residue Pro-71 to Thr-155. Only few residues from the N-terminal of *M. tuberculosis* dCTP deaminase:dUTPase, Pro-29 to Cys-38, superpose well with the N-terminal of the dUTPases. Also, *M. tuberculosis* dCTP deaminase:dUTPase contains a 30-residue stretch not found in the dUTPases. The structure of

the C-terminal differs markedly because the C-terminal of dUTPases reaches across the trimer so that the active site is composed of residues from all three monomers (Chan *et al.*, 2004).

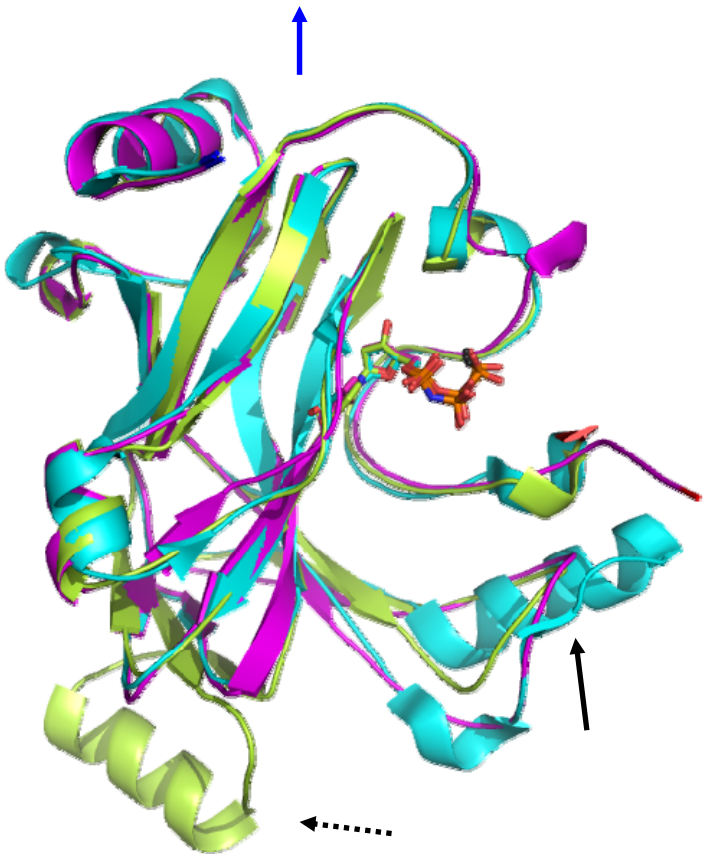


Figure 3.17 Comparison of enzymes having dCTP deaminase activity. Ribbon view of superimposed subunits of *M. tuberculosis* dCTP deaminase:dUTPase-dTTP complex (pink), *M. jannaschii* E145A dCTP deaminase:dUTPase-α,β-imido dUTP complex (lime, PDB ID: 2HXD, Siggaard *et. al.*, 2008) and *E. coli* dCTP deaminase-dUTP complex (cyan, PDB ID: 1XS1, Johansson *et al*., 2005). The enzymes share the same overall fold. The solid black arrow indicates an α-helix only found in the *E. coli* enzyme, while the dotted arrow indicates an α-helix only found in the *M. jannaschii* enzyme. The position of threefold symmetry-axis is indicated by the blue arrow. Nucleotides are (shown in ball and stick) presented in standard atom colours. However, carbon is shown in colours corresponding to the colour of the matching enzyme. Mg^{2+} is shown in grey. The N- and C-terminals are shown in blue and red, respectively. The figure was prepared using PyMOL (DeLano W. L., 2002).

3.2.6.2 Active site

As seen in figures 3.17 and 3.18 the position of the active site is conserved within the examined structures. There is a high degree of conservations in the residues constituting the active sites and their corresponding positions. It is furthermore seen that the position of the nucleotides is almost invariant in the enzyme-nucleotide complex structures.

Comparing all the examined structures (figure 3.16 A and B) it is clear that only six residues (Asp-33, Thr-84, Ser-102, Ile-118, Asp-119 and Gly-124) are strictly conserved not only in the sequence, but also in spatial arrangement (results not shown). This proves that these residues are crucial for correct binding of the nucleotide and active site architecture and they are indeed all found in vicinity of the active site. Ser-102 and Ile-118 are stacking with the

pyrimidine ring on either side while Thr-84 is involved in correct positioning of Ile-118. Asp-119 hydrogen bonds with the deoxyribosyl moiety while Asp-33 hydrogen bonds to water molecules coordinated to Mg^{2+}. Gly-124 is squeezed in between two hydrophobic residues one of which is involved in supporting the nucleotide, leaving no room for a side chain.

Figure 3.18 Comparison of the bifunctional enzyme and dUTPase. Ribbon view of superimposed subunits of *M. tuberculosis* dCTP deaminase:dUTPase-dTTP complex (pink) and *M. tuberculosis* dUTPase-α,β-imido dUTP complex (orange, PDB ID: 1SIX, Chan *et al.*, 2004). Nucleotides are (shown in ball and stick) presented in standard atom colours. However, carbon is shown in colours corresponding to the colour of the matching enzyme. N-and C- terminals are presented in blue and cyan, respectively. Mg^{2+} is shown in grey. The blue arrow represents the position of the threefold axis. The figure was prepared using PyMOL (DeLano W. L., 2002).

The hydrophobic surface on which the nucleotide rests (see section 3.2.2) is completely conserved in the bifunctional enzymes while in the *E. coli* dCTP deaminase structure four out of the five residues are conserved compared to the bifunctional enzymes. A conservative substitution (F to W) has taken place in the residue corresponding to position 122. Comparing the dUTPases to the bifunctional enzymes, the hydrophobic surface is conserved but, conservative substitutions are present in all positions except for Ile-118 (figure 3.16 B).
Residues from the *M. tuberculosis* dCTP deaminase:dUTPase-dTTP complex structure directly involved in hydrogen bonding to the thymidine moiety of the nucleotide (Arg-106, Thr-127 and Gln-174) are conserved in the bifunctional enzymes while Thr is substituted with Val in the *E. coli* dCTP deaminase. Several residues found very close to the pyrimidine

moiety are conserved in all the examined enzymes having dCTP deaminase activity. The same degree of conservation is not observed in the dUTPases (comparing figure 16 A and B), consistent with no dCTP deaminase activity.

Two mechanisms for the deamination of dCTP in *E. coli* dCTP deaminase have been suggested (Johansson *et al.*, 2005). In both mechanisms the same residues are considered to play a part in the reaction and the central point is activation of a water molecule to attack C4 in the pymidine ring. The side chains of residues corresponding to Ser-102, Arg-106 and Glu-129 and the backbone O of Ala-115 are believed to play an active part in the reaction. All four residues are conserved in the examined enzymes having dCTP deaminase activity. This conservation together with fact that Arg, Ala and Glu are not conserved in the monofunctional dUTPases (see figure 3.16 B) brings evidence to the suggested mechanisms and it is reasonable to suggest that the deamination reaction happens by means of the same mechanism in enzymes having dCTP deaminase function.

Residues from the *M. tuberculosis* dCTP deaminase:dUTPase-dTTP complex structure directly involved in hydrogen bonding to the phosphoryl part of the nucleotide are Lys-101, Ser-102, Ser-103, Gln-148, Tyr-162 and Lys-170. The serines, Tyr-162 and Lys-170 are conserved in the three enzymes having dCTP deaminase activity. In both *M. jannaschii* dCTP deaminase:dUTPase and *E. coli* dCTP deaminase Lys-101 is substituted with arginine. Gln-148 is conserved in *M. jannaschii* dCTP deaminase:dUTPase while it is substituted with alanine in *E. coli* dCTP deaminase (see figure 3.16 A). In the dUTPase structures Ser-102 and Gln-148 are strictly conserved. Lys-101 is found to be either lysine or arginine while Ser-103 is found to be either serine or glycine. In all dUTPases an arginine is found in the position of Tyr-162 while no residue is visible in the position corresponding to Lys-170.

From the active site comparisons it is clear that *M. tuberculosis* dCTP deaminase:dUTPase is more closely related to *M. jannaschii* dCTP deaminase:dUTPase than to *E. coli* dCTP deaminase even though the opposite was the case when comparing the overall sequence (see section 3.2.6.1)

A mechanism for the hydrolysis of dUTP in dUTPase has been suggested (Chan *et al.*, 2004) directly involving three residues corresponding to Ser-102, Asp-119 and Gln-148, which are strictly conserved in all examined dUTPases and the bifunctional enzymes from *M. tuberculosis* and *M. jannaschii*. The reaction involves nucleophilic attack on α-phosphoryl by an activated water molecule held in place by the side chains of aspartic acid and glutamine. During the reaction the leaving group, pyrophosphate, is supported by the side chain of serine. In *E. coli* dCTP deaminase glutamine is substituted with alanine and at the same time an arginine side chain (Arg-126, *E. coli* numbering) occupies the position of the nucleophilic water molecule (Johansson *et al*., 2005). In all the examined enzymes Gly-116 is conserved except in *E. coli* dCTP deaminase where a histidine (His-125 *E. coli* numbering) is present. A structural comparison (figure 3.19) between the *E. coli* dCTP deaminase-dUTP complex (PDB ID: 1XS1, Johansson *et al*., 2005) and the *M. tuberculosis* apo-dCTP deaminase:dUTPase shows that the histidine side chain would clash with the bulky side chain of residue 97 (Trp-106 *E. coli* numbering), which explains why this side chain is found in another conformation. The new position of the side chain corresponding to residue 97 would clash with the bulky side chain of residue 117 (Arg-126, *E. coli* numbering), which explains why this side chain in the *E. coli* enzyme is directed into the active site. Most likely, the

glycine at position 116 is a fingerprint residue for the group of bifunctional enzymes amongst dCTP deaminases.

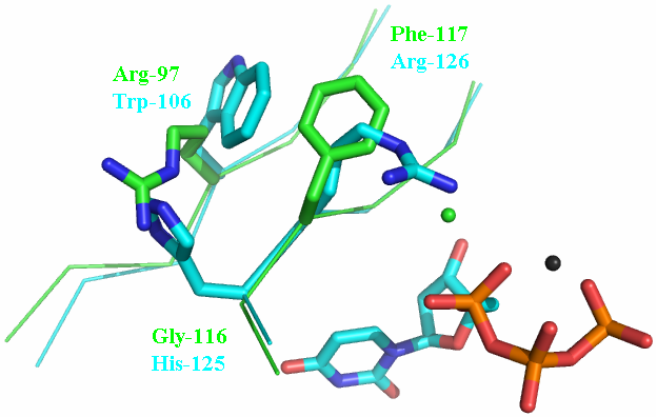


Figure 3.19 Mono-functionality of *E. coli* dCTP deaminase. A superposition of the *E. coli* dCTPdeaminase-dUTP complex (cyan, PDB ID: 1XS1, Johansson *et al.*, 2005) and the *M. tuberculosis* apo-dCTP deaminase:dUTPase (green) shows that the side chain of His-125 (*E. coli*) sets of a domino effect positioning the side chain of Arg-126 (*E. coli*) into the active site pushing away the water molecule which performs the nucleophilic attack (green, present in the apo-structure, see section 3.2.8). Nucleotide and side chains (shown in ball and stick) are presented in standard atom colours. However, carbon is shown in colours corresponding to the colour of the matching enzyme. Mg^{2+} is shown in grey. The figure was prepared in PyMOL (DeLano W. L., 2002).

3.2.7 dTTP regulation of enzyme activity

As discussed in section 3.2.4.2 major structural rearrangements are observed in the region 110-117 upon binding of the inhibitor in *M. tuberculosis* dCTP deaminase:dUTPase. These residues are found very close to the threefold axis and consequently, binding of one dTTP molecule in the trimer causes changes in the neighbouring binding pocket (see figure 3.20). Upon binding of a dTTP molecule in one of the three active sites, the position of His-112 is shifted, which causes the side chain to push the side chain of Leu-110 backwards. To avoid a clash, Thr-114 from the second binding pocket has to change position. This changes the conformation in the second subunit probably leaving the binding pocket more susceptible to binding of a second dTTP molecule. Unless all three subunits of the trimer are in the same conformation Leu-110 and Thr-114 of neighbouring subunits are expected to clash. Consequently, these two residues mediate a "switch" between the apo- and the inhibited conformation. These findings are in agreement with recently reported findings from the *E. coli* dCTP deaminase (Johansson *et al.*, 2007), where dTTP regulation of *E. coli* dCTP deaminase has been shown to involve the transition of the enzyme between two conformations (active and inactive) in a concerted mechanism as described in section 2.4. The enzyme is in an active conformation when complexed to dCTP or dUTP and in an inactive conformation when complexed to dTTP. The conformational change is mediated by amino acid residues Val-122 and Thr-123 (*E. coli* numbering), situated close to the threefold axis in the center of the trimer, and the interaction of this region with the additional 5-methyl group of dTTP compared to dCTP and dUTP (figure 2.4). Thr-114 in *M. tuberculosis* dCTP

deaminase:dUTPase corresponds to Thr-123 in *E. coli* dCTP deaminase and threonine is also found in the corresponding position in *M. jannaschii* dCTP deaminase:dUTPase (figure 3.16 A). The inactive conformation of *M. tuberculosis* dCTP deaminase:dUTPase is found when the enzyme is inhibited by dTTP, but is the apo-form of the enzyme then found in an active conformation? Comparing the position of the nucleotide and the backbone of residues 109-118 from the *M. tuberculosis* dCTP deaminase:dUTPase-dTTP complex structure and the apo-enzyme structure to the corresponding residues (118-127) from the *E. coli* dCTP deaminase-dUTP complex structure (PDB ID: 1XS1, Johansson *et al*., 2005) it is seen, that the position of the nucleotide is almost identical (see figure 3.21). The *E. coli* dCTP deaminase structure has a much higher resemblance to the apo-dCTP deaminase:dUTPase structure (rmsd values 1.14 Å from comparison of main chain atoms in the 10-residue stretch) than to the dCTP deaminase:dUTPase-dTTP complex structure (rmsd values 2.03 Å from comparison of main chain atoms in the 10-residue stretch). Therefore it is a reasonable assumption that the apo-conformation of dCTP deaminase:dUTPase is the active conformation, which is more susceptible to substrate than inhibitor.

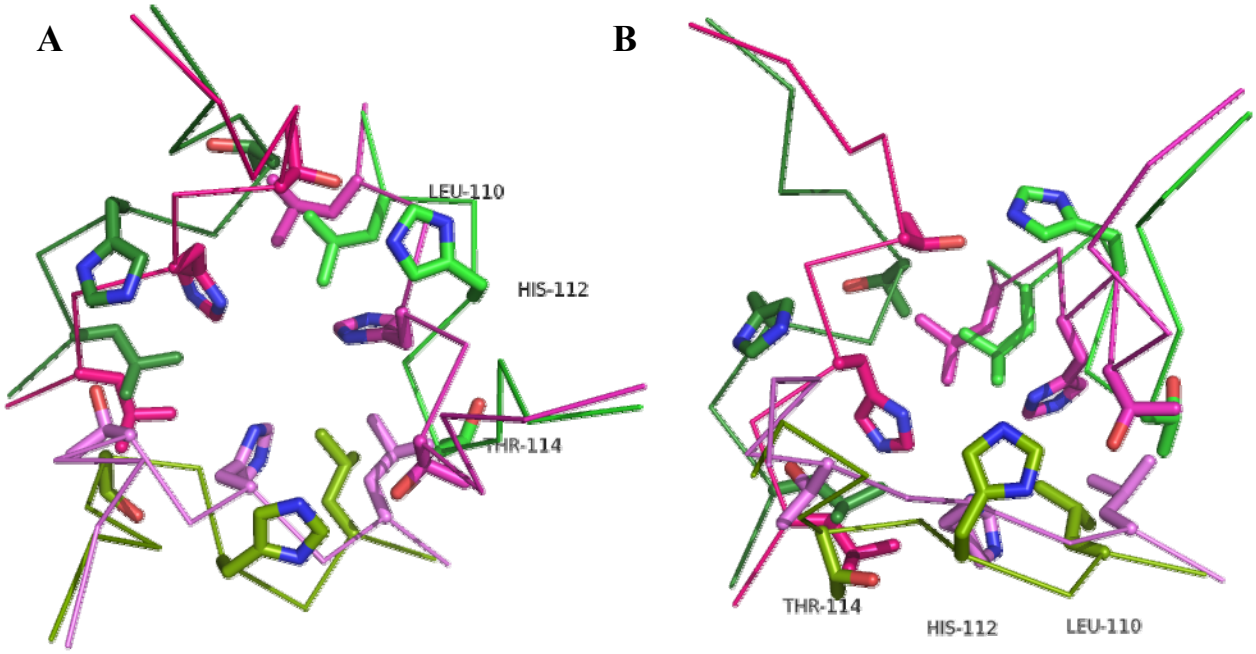


Figure 3.20 Changes in the trimer of *M. tuberculosis* dCTP deaminase:dUTPase as a consequence of dTTP Binding. Cα-trace of residues 109-118 in all three subunits from superposition of the apo-form of *M. tuberculosis* dCTP deaminase:dUTPase (three shades of green) and the dTTP-complex (three shades of pink). In all subunits the side chain of residues Leu-110, His-112 and Thr-114 are shown. A) The superposition viewed along the threefold axis. B) Same situation as in A, but sideways view. In both panels residues from one subunit are labelled. Panels were prepared with PyMOL (DeLano W. L., 2002).

The inactive conformation might be able to accommodate dCTP and dUTP. However, the reaction cannot take place due to non-optimal conditions (discussed in section 3.2.8). Since the same conformation (active or inactive) must be present in all three subunits of the trimer it is plausible to assume that dTTP could be found in one of the active sites while the two others might be occupied by dCTP. This might explain the observed kinetic data indicating a non-

competitive mechanism, where increasing the substrate concentration cannot fully reduce the dTTP inhibition (Helt *et al.*, 2008).

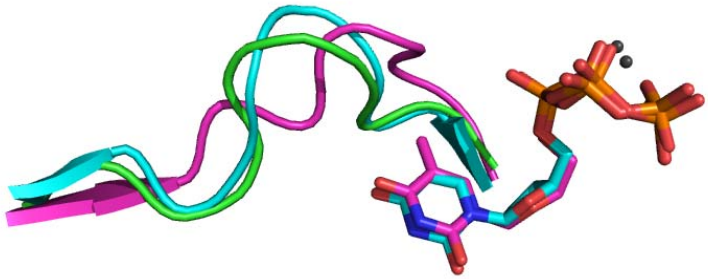


Figure 3.21 Active and inactive conformation of *M. tuberculosis* dCTP deaminase:dUTPase. Ribbon view of residues 109-118 from superposition of subunits from the *M. tuberculosis* dCTP deaminase:dUTPase apo-structure (green) and dTTP-complex structure (pink) with the *E. coli* dCTP deaminase-dUTP complex structure (cyan, PDB ID: 1XS1, Johansson *et al*., 2005). Nucleotides are presented in standard atom colours. However, carbon is shown in colours corresponding to the colour of the matching enzyme. Mg^{2+} is shown in grey. Of the two tuberculosis structures, the apo-structure more closely resembles the *E. coli* structure. Consequently it is believed that the apo-structure is in an active conformation while the dTTP complex structure is in an inactive conformation. The figure was prepared using PyMOL (DeLano W. L., 2002).

3.2.8 Why dTTP is not hydrolysed by dCTP deaminase:dUTPase

dCTP deaminase:dUTPase is able to hydrolyse dCTP and dUTP, but not dTTP, which acts as an inhibitor. The reason must originate in the difference between dUTP and dTTP and the consequent differences imposed on the structure of the enzyme-nucleotide complex. dTTP only differs from dUTP by the methyl group, which is bound to C5 of the pyrimidine ring.
Figure 3.22 shows the active sites from a superposition of the four structures: *M. tuberculosis* dCTP deaminase:dUTPase-dTTP complex, *M. tuberculosis* apo-dCTP deaminase:dUTPase, *M. jannaschii* E145A dCTP deaminase:dUTPase-α,β-imido dUTP complex (Siggaard *et. al.*, 2008) and *M. tuberculosis* dUTPase-α,β-imido dUTP complex (PDB ID: 1SIX, Chan *et al.*, 2004). From the figure it is seen that the position of the nucleotides in the enzyme-nucleotide structures is almost invariant. Consequently, it must be structural changes in the enzyme associated with the presence of the methyl group of dTTP that causes the enzyme to be inactive towards dTTP. Residues believed to participate in the hydrolysis reaction are the conserved Ser-102 Asp-119 and Gln-148 (see section 3.2.6.2). From figure 3.22 it is seen that these residues superpose well in all four structures. Also, the water molecule suggested to perform the nucleophilic attack is present in all structures with the correct angle for the attack (indicated with solid black arrow in the figure). Based on the suggested mechanism for hydrolysis of dUTP by dUTPase (Chan *et al.*, 2004, section 3.2.6.2), the *M. tuberculosis* dCTP deaminase:dUTPase-dTTP complex should be a competent catalytic complex with the ability to hydrolyse dTTP.
In the *M. tuberculosis* dCTP deaminase:dUTPase-dTTP complex structure and in the *M. jannaschii* E145A dCTP deaminase:dUTPase-α,β-imido dUTP complex structure a second water molecule is found very close to the nucleophilic water molecule (indicated with solid blue arrow in the figure). Also, a minor displacement of serine, which supports the leaving

group, pyrophosphate, during reaction, is observed in the *M. tuberculosis* dCTP deaminase:dUTPase-dTTP complex structure. The presence of this water molecule and the minor displacement of serine do not appear to be factors explaining the lack of reaction.
Therefore an explanation must be found in the structural changes observed in residues 110-117 (see section 3.2.4.2) upon binding of dTTP in *M. tuberculosis* dCTP deaminase:dUTPase. This dramatic change is caused by steric reasons due to the methyl group. Due to binding of dTTP Ala-115 changes its position by 2.8 Å and moves closer to and points the side chain towards the α-phosphoryl hereby pushing a water molecule aside (water 59, subunit A of the apo- *M. tuberculosis* dCTP deaminase:dUTPase structure). A corresponding water molecule (dotted arrow in figure 3.22) is also present in the *M. jannaschii* E145A dCTP deaminase:dUTPase-α,β-imido dUTP complex structure and the *M. tuberculosis* dUTPase-α,β-imido dUTP complex structure. A corresponding molecule is not only present in these three structures as shown in the figure, but also in the vast majority of dUTPase structures deposited in the Protein Data Bank (see table 3.2).

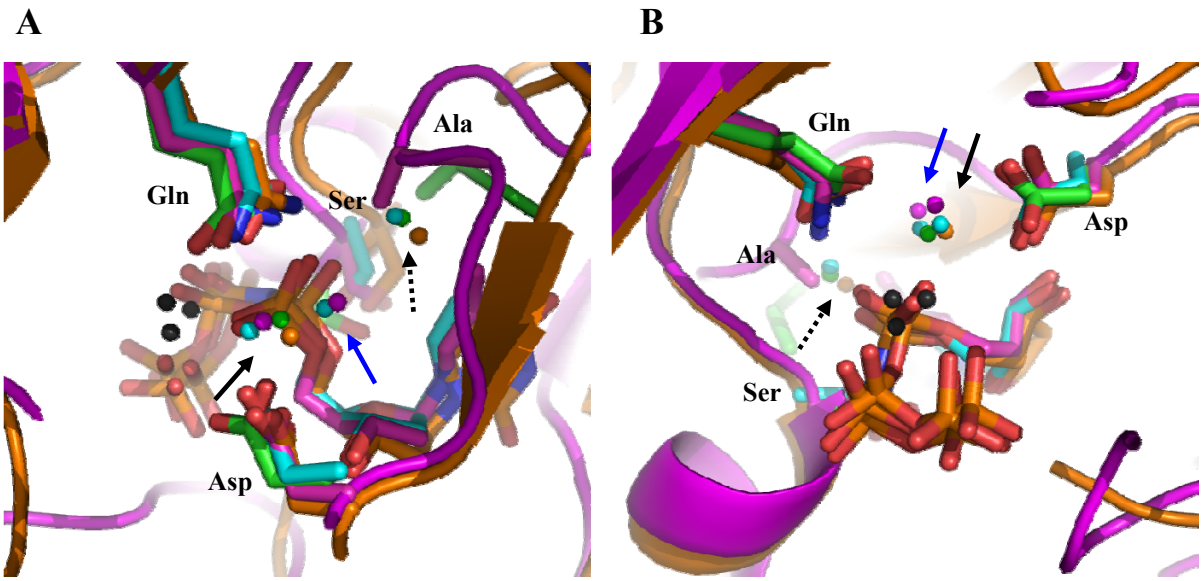


Figure 3.22 Catalytic complexes for hydrolysis. Close up of the active site (ribbon view) from superposition of the *M. tuberculosis* dCTP deaminase:dUTPase apo-structure (green) and dTTP complex structure (pink), the *M. jannaschii* E145A dCTP deaminase:dUTPase-α,β-imido dUTP complex (cyan, PDB ID: 2HXD, Siggaard *et. al.*, 2008) and the *M. tuberculosis* dUTPase-α,β-imido dUTP complex (orange, PDB ID: 1SIX, Chan *et al.*, 2004). From all four structures the side chain of Ser-102 Asp-119 and Gln-148 are shown and from the *M. tuberculosis* dCTP deaminase:dUTPase enzyme the side chain of Ala-115 is also shown. Ribbon views of the backbones of the *M. tuberculosis* enzymes are shown. Nucleotides are presented in standard atom colours. However, carbon is shown in colours corresponding to the colour of the matching enzyme. Important water molecules are also shown in colours corresponding to the colour of the matching enzyme. Mg^{2+} is shown in grey. The active site is shown from two directions (panels A and B). A solid black arrow points to the nucleophilic water molecule while a dotted arrow points to the water molecule which is hydrogen bonded to O2 of α-phosphoryl. A solid blue arrow points to a water molecule next to the nucleophilic water molecule (found in the bifunctional enzymes). The figure prepared with PyMOL (DeLano W. L., 2002).

Table 3.2 Structures examined for including a water molecule corresponding to water 59 from the apo-dCTP deaminase:dUTPase structure. In each case subunit A (or the first stated subunit in the pdb-file) was examined by superposition with the apo- and the dTTP-complex structures of *M. tuberculosis* dCTP deaminase:dUTPase.

Life form / virus	Enzyme	PDB ID	Resolution / Å	Nucleotide	Mutation	Corresponding water (W number)
E. coli	dCTP deaminase	1XS1 [a]	1.80	dUTP	-	Yes (W 263)
		1XS4 [a]	2.53	dCTP	E138A	Yes (W 165)
		1XS6 [a]	2.00	dUTP	E138A	Yes (W 29)
		2J4Q [b]	2.60	dTTP	E138A	No (5 H_2O in model)
		2J4H [b]	2.70	dCTP	H121A	No H_2O in model
M. jannaschii	Bi-functionel	1OGH [c]	1.88	-	-	Yes (W 118, chain B)
		2HXD [d]	2.30	α,β-imido dUTP	E145A	Yes (W 33)
E. coli	dUTPase	1SEH [e]	1.47	dUMP	-	Yes (W 9)
		1SYL [e]	1.95	dUTP	D90N	Yes (W 7)
		1RN8 [e]	1.93	α,β-imido dUTP	-	Yes (W 2)
		1RNJ [e]	1.70	α,β-imido dUTP	D90N	Yes (W 3)
		1EUW [f]	1.05	-	-	Yes (W 228)
		1EU5 [f]	1.45	-	-	Yes (W 229)
Human	dUTPase	1Q5U [g]	2.00	-	-	Yes (W 20)
		1Q5H [g]	2.00	dUDP	-	Yes (W 102)
		2HQU [h]	2.20	α,β-imido dUTP	-	Yes (W 5)
M. tuberculosis	dUTPase	1SNF [i]	1.85	dUMP	-	Yes (W 107)
		1SLH [i]	3.00	dUDP	-	Yes (W 537)
		1MQ7 [i]	1.95	-	-	Yes (W 57)
		1SIX [i]	1.30	α,β-imido dUTP	-	Yes (W 304)
		1SMC [i]	2.10	dUTP	-	Yes (W 204)
		1SM8 [i]	2.90	dUTP (Cr^{3+})	-	No
		1SJN [i]	1.80	α,β-imido dUTP	-	Yes (W 26)
EIAV	dUTPase	1DUN [j]	1.90	-	-	No
		1DUC [j]	2.05	dUDP (Sr^{2+})	-	Yes (W 5)
FIV	dUTPase	1DUT [k]	1.90	-	-	Yes (W 36)
		1F7Q [l]	2.26	dUTP	-	Yes (W 612)
		1F7P [l]	2.30	dUDP	-	Yes (W 632)
		1F7N [l]	2.20	dUMP	-	No
		1F7R [l]	2.50	dUDP	-	Yes (W 225)
		1F7O [l]	2.20	-	-	No
		1F7K [l]	2.20	dUMP	-	No
		1F7D [l]	1.40	-	-	No

[a] Johansson *et al.*, 2005; [b] Johansson *et al.*, 2007; [c] Johansson *et al.*, 2003; [d] Siggaard *et. al.*, 2008; [e] Barabás *et al.*, 2004; [f] Gonzáles et al., 2001; [g] Mol *et al.*, 1996; [h] Varga *et al.*, 2007; [i] Chan *et al.*,2004; [j] Dauter *et al.*, 1999; [k] Prasad *et al.*, 1996; [l] Prasad *et al.*, 2000.

A water molecule in this position makes a hydrogen bond to O2 of the α-phosphoryl and has not previous been considered to play a part in the suggested mechanism of dUTPase. Since the structure of *M. tuberculosis* dCTP deaminase:dUTPase in complex with dTTP represents what is best described with the current knowledge as a catalytic complex for hydrolysis of the nucleotide, this suggests that a water molecule at the position of water 59 (apo-structure) is essential for the reaction to proceed. Due to its position it cannot play the part of an attacking nucleophile. The hydrolysis reaction probably happens as suggested (Chan *et al.*, 2004).

However, the water molecule is necessary for stabilizing the transition state by attracting electron density from α-phosphoryl, hereby leaving the phosphorous atom more susceptible for nucleophilic attack. If the water molecule is not present and the side chain of the hydrophobic residue alanine is pointing towards O2 (the distance is 3.1 Å), the effect is reversed and electrons pushed towards the phosphorous atom. This raises the activation energy of the reaction, rendering it too high for the hydrolysis to take place. This explains why dTTP is not hydrolysed but acts as an inhibitor in enzymes having dCTP deaminase activity.

3.3 Conclusion

M. tuberculosis dCTP deaminase:dUTPase has been crystallized as an apo-enzyme and an enzyme-dTTP complex and data were collected to 2.4 Å and 2.0 Å, respectively. Each crystal structure was solved using molecular replacement and deposited in the Protein Data Bank. The enzyme was also crystallized in the presence of α,β-imido dUTP, data was collected to 3.0 Å and the structure was solved using molecular replacement. The quality of the model was poor and α,β-imido dUTP did not appear in the model.

Major changes in the conformation of the active site and ordering of the C-terminal were apparent upon binding of the inhibitor, dTTP. The inhibition mode of dTTP explaining previously obtained kinetic data has been inferred. Also, it has been explained why dTTP is not hydrolysed by the enzyme, this gives new insight into the reaction mechanism of bifunctional dCTP deaminase:dUTPases as well as monofunctional dUTPases.

Chapter four

4 Concluding remarks and outlook

One of the greater perspectives of the second part of this thesis is to find the evolutionary connection between dCTP deaminases, dUTPases and the bifunctional enzymes based on sequence analysis and to survey the flow of genetic information. In this connection determining functionality based only on sequence is a powerful tool saving numerous of hours spent in the laboratory.

A second perspective of this project is to learn how DNA evolved and how and when dTTP became one of the nucleotides used for incorporation into DNA. dTTP is not synthesized by equivalent paths in all organisms and studying these different paths might help in answering evolutionary questions.

A third perspective is the appearance of a possible drug target in the battle against tuberculosis since the bifunctional enzyme is not found in eukaryotes.

4.1 Key residues and mutational analysis

Several residues have previously been identified as being important for the deamination reaction, the hydrolysis reaction and correct binding of the nucleotide in the active site (section 3.2.6). The monofunctional dUTPases are easy to distinguish from dCTPdeaminase and the bifunctional enzymes since the primary sequence is shorter and they do not contain the conserved residues believed to participate in the deamination reaction. Distinguishing the dCTP deaminases from the bifunctional enzymes is not an easy task and more monofunctional and bifunctional enzymes have to be identified (from kinetic experiments) before certain conclusions may be drawn. Some of the known dCTP deaminases might prove to be bifunctional since they have probably not been tested for dUTPase activity. However, a few conclusions can be drawn from comparing the *E. coli* dCTP deaminase to the bifunctional enzymes from *M. tuberculosis* and *M. jannaschii*. A glutamine is believed to participate in positioning of the nucleophilic water molecule (Chan *et al.*, 2004) for the hydrolysis. In the *E. coli* enzyme an alanine (residue 157) is found in this position. At the same time an arginine side chain is positioned where the nucleophilic water should be (Johansson *et al.*, 2005), probably due to a domino effect caused by a histidine (residue 125). In the position corresponding to this histidine a glycine is present in the bifunctional enzymes and dUTPases examined in this thesis. Besides the residues already believed to participate in the hydrolysis reaction this glycine is speculated to be a key residue in distinguishing between mono- and bifunctionality (section 3.2.6.2).

It would be interesting to include the mutations A157Q and H125G in the *E. coli* dCTP deaminase and examine if this would render the enzyme bifunctional. Mutational studies should also be performed on the bifunctional enzymes to investigate the effects of mutations

on residues assumed to participate in the enzymatic reactions. The effect of C-terminal closure on the reactions should also be investigated e.g. by producing the particular enzyme lacking the C-terminal.

4.2 The transition state of hydrolysis

In section 3.2.8 it was argued that the reason for the impaired hydrolysis of dTTP was the absence of a water molecule, which otherwise made a hydrogen bond to O2 of α-phosphate, and the presence of the Ala-115 side chain in a corresponding position.
One way to gain insight into the hydrolysis reaction, the importance of the water molecule and the transition state would be to mimic the transition state. AlF_3 is widely used in mimicking the transition state in hydrolysis or phosphorylation of nucleotides in enzymatic reactions (Delbaere *et al.*, 2004; Gorrell *et al.*, 2005; Ostermann *et al.*, 2000; Xu *et al.*1997). In dUTPases the α-phosphorous atom is attacked by the activated nucleophilic water molecule, consequently AlF_3 would have to play the part of the α-phosphoryl group. To obtain a transition state model complex one would have to crystallize the enzyme in contact with 2',5'-dideoxyuridine, AlF_3, pyrophosphate and magnesium.
Another way to gain insight into the mechanism of hydrolysis and inhibition would be to perform theoretical calculations concerning the ground states and probable transitions states of the nucleotides when bound in the enzyme. Hereby one could infer the importance of water molecules and amino acid residues present in the vicinity of the nucleotide.

4.3 Crystallization conditions

To gain a better insight into the structures, structural changes and reaction mechanisms it would be beneficial to have atomic resolution structures. Since it appears to be an easy task to crystallize the enzyme but it is difficult to obtain large crystals it may be advantageous to test commercially available additive screens. Also, screening for other crystallization conditions using commercially available screens may reveal conditions for producing larger and better diffracting crystals. Finally, changing the buffer system and making some changes in the purification procedure may also change the crystallization properties of the enzyme.

Appendix one

A.1 Crystallization of dCTP deaminase:dUTPase in contact with inhibitor, dTTP

A.1.1 Plate A1

Enzyme solution: 1.8 mg/ml, 5 mM dTTP, 20 mM $MgCl_2$, 50 mM HEPES pH 6.8.
Temperature: The plate was set up at room temperature and kept at 23ºC.
2 µl of enzyme solution was mixed with 2 µl of reservoir solution.

No.	Precipitant	Salt	Buffer
A1	20 % PEG 400	50 mM $MgCl_2$	100 mM HEPES pH 7.5
B1	20 % PEG 400	100 mM $MgCl_2$	100 mM HEPES pH 7.5
C1	20 % PEG 400	150 mM $MgCl_2$	100 mM HEPES pH 7.5
D1	20 % PEG 400	200 mM $MgCl_2$	100 mM HEPES pH 7.5
A2	25 % PEG 400	50 mM $MgCl_2$	100 mM HEPES pH 7.5
B2	25 % PEG 400	100 mM $MgCl_2$	100 mM HEPES pH 7.5
C2	25 % PEG 400	150 mM $MgCl_2$	100 mM HEPES pH 7.5
D2	25 % PEG 400	200 mM $MgCl_2$	100 mM HEPES pH 7.5
A3	30 % PEG 400	50 mM $MgCl_2$	100 mM HEPES pH 7.5
B3	30 % PEG 400	100 mM $MgCl_2$	100 mM HEPES pH 7.5
C3	30 % PEG 400	150 mM $MgCl_2$	100 mM HEPES pH 7.5
D3	30 % PEG 400	200 mM $MgCl_2$	100 mM HEPES pH 7.5
A4	35 % PEG 400	50 mM $MgCl_2$	100 mM HEPES pH 7.5
B4	35 % PEG 400	100 mM $MgCl_2$	100 mM HEPES pH 7.5
C4	35 % PEG 400	150 mM $MgCl_2$	100 mM HEPES pH 7.5
D4	35 % PEG 400	200 mM $MgCl_2$	100 mM HEPES pH 7.5
A5	40 % PEG 400	50 mM $MgCl_2$	100 mM HEPES pH 7.5
B5	40 % PEG 400	100 mM $MgCl_2$	100 mM HEPES pH 7.5
C5	40 % PEG 400	150 mM $MgCl_2$	100 mM HEPES pH 7.5
D5	40 % PEG 400	200 mM $MgCl_2$	100 mM HEPES pH 7.5

A.1.2 Plates A2 and A3

(Same reservoir solutions, but different temperature)
Enzyme solution: 1.8 mg/ml, 5 mM dTTP, 20 mM $MgCl_2$, 50 mM HEPES pH 6.8.
Reservoirs A1-D5: 2 µl of enzyme solution was mixed with 2 µl of reservoir solution.
Reservoirs A6-D6: 4 µl of enzyme solution was mixed with 2 µl of reservoir solution.
Temperature plate A2: The plate was set up at room temperature and kept at 23ºC.
Temperature plate A3: The plate was set up on ice and kept at 15 ºC.

No.	Precipitant	Salt	Buffer
A1	35 % PEG 400	100 mM $MgCl_2$	100 mM HEPES pH 7.5
B1	35 % PEG 400	150 mM $MgCl_2$	100 mM HEPES pH 7.5
C1	35 % PEG 400	200 mM $MgCl_2$	100 mM HEPES pH 7.5
D1	35 % PEG 400	250 mM $MgCl_2$	100 mM HEPES pH 7.5
A2	40 % PEG 400	100 mM $MgCl_2$	100 mM HEPES pH 7.5
B2	40 % PEG 400	150 mM $MgCl_2$	100 mM HEPES pH 7.5
C2	40 % PEG 400	200 mM $MgCl_2$	100 mM HEPES pH 7.5
D2	40 % PEG 400	250 mM $MgCl_2$	100 mM HEPES pH 7.5
A3	45 % PEG 400	100 mM $MgCl_2$	100 mM HEPES pH 7.5
B3	45 % PEG 400	150 mM $MgCl_2$	100 mM HEPES pH 7.5
C3	45 % PEG 400	200 mM $MgCl_2$	100 mM HEPES pH 7.5
D3	45 % PEG 400	250 mM $MgCl_2$	100 mM HEPES pH 7.5
A4	50 % PEG 400	100 mM $MgCl_2$	100 mM HEPES pH 7.5
B4	50 % PEG 400	150 mM $MgCl_2$	100 mM HEPES pH 7.5
C4	50 % PEG 400	200 mM $MgCl_2$	100 mM HEPES pH 7.5
D4	50 % PEG 400	250 mM $MgCl_2$	100 mM HEPES pH 7.5
A5	55 % PEG 400	100 mM $MgCl_2$	100 mM HEPES pH 7.5
B5	55 % PEG 400	150 mM $MgCl_2$	100 mM HEPES pH 7.5
C5	55 % PEG 400	200 mM $MgCl_2$	100 mM HEPES pH 7.5
D5	55 % PEG 400	250 mM $MgCl_2$	100 mM HEPES pH 7.5
A6	45 % PEG 400	100 mM $MgCl_2$	100 mM HEPES pH 7.5
B6	45 % PEG 400	150 mM $MgCl_2$	100 mM HEPES pH 7.5
C6	45 % PEG 400	200 mM $MgCl_2$	100 mM HEPES pH 7.5
D6	45 % PEG 400	250 mM $MgCl_2$	100 mM HEPES pH 7.5

A.1.3 Plates A4 and A5

(Same reservoir solutions, but different temperature)
Enzyme solution: 1.8 mg/ml, 5 mM dTTP, 20 mM $MgCl_2$, 50 mM HEPES pH 6.8.
Reservoirs A3-D3: 2 µl of enzyme solution was mixed with 2 µl of reservoir solution.
Reservoirs A4-D4: 4 µl of enzyme solution was mixed with 2 µl of reservoir solution.
Temperature plate A4: The plate was set up on ice and kept at 4ºC.
Temperature plate A5: The plate was set up on ice and kept at 15ºC.

No.	Precipitant	Salt	Buffer
A3	30 % PEG 400	200 mM $MgCl_2$	100 mM HEPES pH 7.5
B3	35 % PEG 400	200 mM $MgCl_2$	100 mM HEPES pH 7.5
C3	40 % PEG 400	200 mM $MgCl_2$	100 mM HEPES pH 7.5
D3	35 % PEG 400	250 mM $MgCl_2$	100 mM HEPES pH 7.5
A4	45 % PEG 400	100 mM $MgCl_2$	100 mM HEPES pH 7.5
B4	45 % PEG 400	150 mM $MgCl_2$	100 mM HEPES pH 7.5
C4	45 % PEG 400	200 mM $MgCl_2$	100 mM HEPES pH 7.5
D4	35 % PEG 400	200 mM $MgCl_2$	100 mM HEPES pH 7.5

Appendix two

A.2 Crystallization of apo-dCTP deaminase:dUTPase

A.2.1 Plate B1

Enzyme solution: 1.9 mg/ml, 20 mM $MgCl_2$, 50 mM HEPES pH 6.8.
Temperature: The plate was set up at room temperature and kept at 23ºC.
2 µl of enzyme solution was mixed with 2 µl of reservoir solution.
Reservoir solutions are the same as for plate A1 (appendix A.1).

A.2.2 Plate B2

Enzyme solution: 1.9 mg/ml, 50 mM HEPES pH 6.8.
Temperature: The plate was set up at room temperature and kept at 23ºC.
Reservoirs A1-D4: 2 µl of enzyme solution was mixed with 2 µl of reservoir solution.
Reservoirs A5-D6: 4 µl of enzyme solution was mixed with 2 µl of reservoir solution.

No.	Precipitant	Salt	Buffer
A1	15 % PEG 400		100 mM HEPES pH 7.5
B1	15 % PEG 400	20 mM $MgCl_2$	100 mM HEPES pH 7.5
C1	15 % PEG 400	50 mM $MgCl_2$	100 mM HEPES pH 7.5
D1	15 % PEG 400	70 mM $MgCl_2$	100 mM HEPES pH 7.5
A2	20 % PEG 400		100 mM HEPES pH 7.5
B2	20 % PEG 400	20 mM $MgCl_2$	100 mM HEPES pH 7.5
C2	20 % PEG 400	50 mM $MgCl_2$	100 mM HEPES pH 7.5
D2	20 % PEG 400	70 mM $MgCl_2$	100 mM HEPES pH 7.5
A3	25 % PEG 400		100 mM HEPES pH 7.5
B3	25 % PEG 400	20 mM $MgCl_2$	100 mM HEPES pH 7.5
C3	25 % PEG 400	50 mM $MgCl_2$	100 mM HEPES pH 7.5
D3	25 % PEG 400	70 mM $MgCl_2$	100 mM HEPES pH 7.5
A4	30 % PEG 400		100 mM HEPES pH 7.5
B4	30 % PEG 400	20 mM $MgCl_2$	100 mM HEPES pH 7.5
C4	30 % PEG 400	50 mM $MgCl_2$	100 mM HEPES pH 7.5
D4	30 % PEG 400	70 mM $MgCl_2$	100 mM HEPES pH 7.5
A5	20 % PEG 400		100 mM HEPES pH 7.5
B5	20 % PEG 400	20 mM $MgCl_2$	100 mM HEPES pH 7.5
C5	20 % PEG 400	50 mM $MgCl_2$	100 mM HEPES pH 7.5
D5	20 % PEG 400	70 mM $MgCl_2$	100 mM HEPES pH 7.5
A6	25 % PEG 400		100 mM HEPES pH 7.5
B6	25 % PEG 400	20 mM $MgCl_2$	100 mM HEPES pH 7.5
C6	25 % PEG 400	50 mM $MgCl_2$	100 mM HEPES pH 7.5
D6	25 % PEG 400	70 mM $MgCl_2$	100 mM HEPES pH 7.5

A.2.3 Plates B3 and B4

Enzyme solution: 1.9 mg/ml, 20 mM $MgCl_2$, 50 mM HEPES pH 6.8.
Reservoirs A1-D1: 2 µl of enzyme solution was mixed with 2 µl of reservoir solution.
Reservoirs A2-D2: 4 µl of enzyme solution was mixed with 2 µl of reservoir solution.
Temperature plate B3: The plate was set up on ice and kept at 4ºC.
Temperature plate B4: The plate was set up on ice and kept at 15 ºC.

No.	Precipitant	Salt	Buffer
A1	25 % PEG 400		100 mM HEPES pH 7.5
B1	15 % PEG 400	50 mM $MgCl_2$	100 mM HEPES pH 7.5
C1	20 % PEG 400	50 mM $MgCl_2$	100 mM HEPES pH 7.5
D1	25 % PEG 400	50 mM $MgCl_2$	100 mM HEPES pH 7.5
A2	25 % PEG 400		100 mM HEPES pH 7.5
B2	15 % PEG 400	50 mM $MgCl_2$	100 mM HEPES pH 7.5
C2	20 % PEG 400	50 mM $MgCl_2$	100 mM HEPES pH 7.5
D2	25 % PEG 400	50 mM $MgCl_2$	100 mM HEPES pH 7.5

A.2.4 Plate B5

Enzyme solution 1: 1.9 mg/ml, 50 mM HEPES pH 6.8.
Enzyme solution 2: 4.3 mg/ml, 50 mM HEPES pH 6.8.
Reservoirs A1-D1 and A3-D6: 2 µl of enzyme solution was mixed with 2 µl of reservoir solution.
Reservoirs A2-D2: 4 µl of enzyme solution was mixed with 2 µl of reservoir solution.
Temperature: The plate was set up in a cold room at 4ºC and kept herein.

No.	Enzyme solution	Precipitant	Salt	Buffer
A1	2	20 % PEG 400	50 mM $MgCl_2$	100 mM HEPES pH 7.5
B1	2	25 % PEG 400	50 mM $MgCl_2$	100 mM HEPES pH 7.5
C1	2	30 % PEG 400	50 mM $MgCl_2$	100 mM HEPES pH 7.5
D1	2	35 % PEG 400	50 mM $MgCl_2$	100 mM HEPES pH 7.5
A2	1	20 % PEG 400		100 mM HEPES pH 7.5
B2	1	25 % PEG 400		100 mM HEPES pH 7.5
C2	1	30 % PEG 400		100 mM HEPES pH 7.5
D2	1	35 % PEG 400		100 mM HEPES pH 7.5
A3	1	20 % PEG 400		100 mM HEPES pH 7.5
B3	1	25 % PEG 400		100 mM HEPES pH 7.5
C3	1	30 % PEG 400		100 mM HEPES pH 7.5
D3	1	35 % PEG 400		100 mM HEPES pH 7.5
A4	1	20 % PEG 1000	50 mM $MgCl_2$	100 mM HEPES pH 7.5
B4	1	25 % PEG 1000	50 mM $MgCl_2$	100 mM HEPES pH 7.5
C4	1	30 % PEG 1000	50 mM $MgCl_2$	100 mM HEPES pH 7.5
D4	1	35 % PEG 1000	50 mM $MgCl_2$	100 mM HEPES pH 7.5
A5	1	15 % PEG 4000	50 mM $MgCl_2$	100 mM HEPES pH 7.5
B5	1	20 % PEG 4000	50 mM $MgCl_2$	100 mM HEPES pH 7.5
C5	1	25 % PEG 4000	50 mM $MgCl_2$	100 mM HEPES pH 7.5
D5	1	30 % PEG 4000	50 mM $MgCl_2$	100 mM HEPES pH 7.5
A6	1	10 % PEG 8000	50 mM $MgCl_2$	100 mM HEPES pH 7.5
B6	1	15 % PEG 8000	50 mM $MgCl_2$	100 mM HEPES pH 7.5
C6	1	20 % PEG 8000	50 mM $MgCl_2$	100 mM HEPES pH 7.5
D6	1	25 % PEG 8000	50 mM $MgCl_2$	100 mM HEPES pH 7.5

Appendix three

A.3 Crystallization of dCTP deaminase:dUTPase in contact with substrate analogue, α,β-imido dUTP

A.3.1 Plate C1

Enzyme solution: 1.6 mg/ml, 17 mM $MgCl_2$, 1.7 mM α,β-imido dUTP, 50 mM HEPES pH 6.8.
Temperature: The plate was set up at room temperature and kept at 23ºC.
2.4 µl of enzyme solution was mixed with 2 µl of reservoir solution.
Reservoir solutions are the same as for plate A1 (appendix A.1).

A.3.2 Plates C2 and C3

(Same reservoir solutions, but different temperature)
Enzyme solution: 1.6 mg/ml, 17 mM $MgCl_2$, 1.7 mM α,β-imido dUTP, 50 mM HEPES pH 6.8.
Reservoirs A5-D5: 2.4 µl of enzyme solution was mixed with 2 µl of reservoir solution.
Reservoirs A6-D6: 4 µl of enzyme solution was mixed with 2 µl of reservoir solution.
Temperature plate C2: The plate was set up on ice and kept at 4ºC.
Temperature plate C3: The plate was set up on ice and kept at 15 ºC.

No.	Precipitant	Salt	Buffer
A5	30 % PEG 400	200 mM $MgCl_2$	100 mM HEPES pH 7.5
B5	35 % PEG 400	200 mM $MgCl_2$	100 mM HEPES pH 7.5
C5	40 % PEG 400	200 mM $MgCl_2$	100 mM HEPES pH 7.5
D5	35 % PEG 400	250 mM $MgCl_2$	100 mM HEPES pH 7.5
A6	45 % PEG 400	100 mM $MgCl_2$	100 mM HEPES pH 7.5
B6	45 % PEG 400	150 mM $MgCl_2$	100 mM HEPES pH 7.5
C6	45 % PEG 400	200 mM $MgCl_2$	100 mM HEPES pH 7.5
D6	35 % PEG 400	200 mM $MgCl_2$	100 mM HEPES pH 7.5

A.3.3 Plate C4

Enzyme solution 1: 1.6 mg/ml, 17 mM $MgCl_2$, 1.7 mM α,β-imido dUTP, 50 mM HEPES pH 6.8.
Enzyme solution 2: 1.9 mg/ml, 20 mM $MgCl_2$, 2 mM α,β-imido dUTP, 50 mM HEPES pH 6.8.
2.4 µl of enzyme solution 1 was mixed with 2 µl of reservoir solution in all drops except 1D, 2D and 3D, here 2 µl of enzyme solution 2 was used.
Temperature: The plate was set up in a cold room at 4ºC and kept herein.

No.	Precipitant	Salt	Buffer
A1	25 % PEG 400	150 mM $MgCl_2$	100 mM HEPES pH 7.5
B1	30 % PEG 400	150 mM $MgCl_2$	100 mM HEPES pH 7.5
C1	35 % PEG 400	150 mM $MgCl_2$	100 mM HEPES pH 7.5
D1	25 % PEG 400	200 mM $MgCl_2$	100 mM HEPES pH 7.5
A2	25 % PEG 400	200 mM $MgCl_2$	100 mM HEPES pH 7.5
B2	30 % PEG 400	200 mM $MgCl_2$	100 mM HEPES pH 7.5
C2	35 % PEG 400	200 mM $MgCl_2$	100 mM HEPES pH 7.5
D2	30 % PEG 400	200 mM $MgCl_2$	100 mM HEPES pH 7.5
A3	25 % PEG 400	250 mM $MgCl_2$	100 mM HEPES pH 7.5
B3	30 % PEG 400	250 mM $MgCl_2$	100 mM HEPES pH 7.5
C3	35 % PEG 400	250 mM $MgCl_2$	100 mM HEPES pH 7.5
D3	25 % PEG 400	150 mM $MgCl_2$	100 mM HEPES pH 7.5
A4	20 % PEG 1000	200 mM $MgCl_2$	100 mM HEPES pH 7.5
B4	25 % PEG 1000	200 mM $MgCl_2$	100 mM HEPES pH 7.5
C4	30 % PEG 1000	200 mM $MgCl_2$	100 mM HEPES pH 7.5
D4	35 % PEG 1000	200 mM $MgCl_2$	100 mM HEPES pH 7.5
A5	15 % PEG 4000	200 mM $MgCl_2$	100 mM HEPES pH 7.5
B5	20 % PEG 4000	200 mM $MgCl_2$	100 mM HEPES pH 7.5
C5	25 % PEG 4000	200 mM $MgCl_2$	100 mM HEPES pH 7.5
D5	30 % PEG 4000	200 mM $MgCl_2$	100 mM HEPES pH 7.5
A6	10 % PEG 8000	200 mM $MgCl_2$	100 mM HEPES pH 7.5
B6	15 % PEG 8000	200 mM $MgCl_2$	100 mM HEPES pH 7.5
C6	20 % PEG 8000	200 mM $MgCl_2$	100 mM HEPES pH 7.5
D6	25 % PEG 8000	200 mM $MgCl_2$	100 mM HEPES pH 7.5

A.3.4 Plate C5

Enzyme solution 1: 1.6 mg/ml, 17 mM $MgCl_2$, 1.7 mM α,β-imido dUTP, 50 mM HEPES pH 6.8.

Enzyme solution 2: 1.2 mg/ml, 13 mM $MgCl_2$, 1.3 mM α,β-imido dUTP, 50 mM HEPES pH 6.8.

1 µl of enzyme solution 1 was mixed with 1 µl of reservoir solution.

Temperature: The plate was set up in a cold room at 4ºC and kept herein.

4D and 5D was seeded from Plate C4, number 1A, using the tip of a needle.

No.	Enzyme Solution	Precipitant	Salt	Buffer
A1	1	4 % PEG 8000	200 mM $MgCl_2$	100 mM HEPES pH 7.5
B1	1	6 % PEG 8000	200 mM $MgCl_2$	100 mM HEPES pH 7.5
C1	1	8 % PEG 8000	200 mM $MgCl_2$	100 mM HEPES pH 7.5
D1	1	10 % PEG 8000	200 mM $MgCl_2$	100 mM HEPES pH 7.5
A2	2	10 % PEG 8000	200 mM $MgCl_2$	100 mM HEPES pH 7.5
B2	2	15 % PEG 8000	200 mM $MgCl_2$	100 mM HEPES pH 7.5
A3	1	20 % PEG 400	150 mM $MgCl_2$	100 mM HEPES pH 7.5
B3	1	25 % PEG 400	150 mM $MgCl_2$	100 mM HEPES pH 7.5
D3	2	25 % PEG 400	150 mM $MgCl_2$	100 mM HEPES pH 7.5
A4	1	20 % PEG 400	100 mM $MgCl_2$	100 mM HEPES pH 7.5
B4	1	25 % PEG 400	100 mM $MgCl_2$	100 mM HEPES pH 7.5
D4	2	25 % PEG 400	150 mM $MgCl_2$	100 mM HEPES pH 7.5
A5	2	25 % PEG 400	150 mM $MgCl_2$	100 mM HEPES pH 7.5

B5	1	25 % PEG 400	150 mM $MgSO_4$	100 mM HEPES pH 7.5
C5	1	25 % PEG 400	150 mM $MgCl_2$	100 mM HEPES pH 7.0
D5	1	25 % PEG 400	200 mM $MgCl_2$	100 mM HEPES pH 7.5

A.3.5 Plate C6

Enzyme solution: 1.9 mg/ml, 200 mM $MgCl_2$, 3 mM α,β-imido dUTP, 50 mM HEPES pH 6.8.

2 µl of enzyme solution was mixed with 2 µl of reservoir solution.

Temperature: The plate was set up in a cold room at 4ºC and kept herein.

No.	Precipitant	Salt	Buffer
A4	15 % PEG 400	100 mM $MgCl_2$	100 mM HEPES pH 7.5
B4	18 % PEG 400	100 mM $MgCl_2$	100 mM HEPES pH 7.5
C4	21 % PEG 400	100 mM $MgCl_2$	100 mM HEPES pH 7.5
D4	24 % PEG 400	100 mM $MgCl_2$	100 mM HEPES pH 7.5
A5	4 % PEG 8000	100 mM $MgCl_2$	100 mM HEPES pH 7.5
B5	6 % PEG 8000	100 mM $MgCl_2$	100 mM HEPES pH 7.5
C5	8 % PEG 8000	100 mM $MgCl_2$	100 mM HEPES pH 7.5
D5	10 % PEG 8000	100 mM $MgCl_2$	100 mM HEPES pH 7.5
A6	4 % PEG 8000	200 mM $MgCl_2$	100 mM HEPES pH 7.5
B6	6 % PEG 8000	200 mM $MgCl_2$	100 mM HEPES pH 7.5
C6	8 % PEG 8000	200 mM $MgCl_2$	100 mM HEPES pH 7.5
D6	10 % PEG 8000	200 mM $MgCl_2$	100 mM HEPES pH 7.5

Appendix four

A.4 Structural evidence for a concerted bifunctionality in dCTP deaminase-dUTPase from *Methanocaldococcus jannaschii*

Julie H. B. Siggaard[1], Eva Johansson[1,2,3*], Signe Smedegaard Helt[4], Pernille Harris[4], Sine Larsen[1,2] and Martin Willemoës[5*]

[1]Centre for Crystallographic Studies, Department of Chemistry, University of Copenhagen, Universitetsparken 5, DK-2100 Copenhagen Ø, Denmark.
[2]European Synchrotron Radiation Facility, BP220 F38043 Grenoble Cedex, France.
[3]Present address: E. Johansson, Diabetes Protein Engineering, Novo Nordisk A/S, Novo Nordisk Park, DK-2760 Måløv, Denmark.
[4]Department of Chemistry, Technical University of Denmark, Kemitorvet 207, DK-2800 Kgs. Lyngby, Denmark.
[5]Section for Biochemical Genetics, Department of Biology, University of Copenhagen, Ole Maaløes vej 5, DK-2200 Copenhagen N, Denmark.

Corresponding authors: evjh@novonordisk.com; willemoes@bio.ku.dk

Keywords: bifunctional; dUTPase; dCTP deaminase; induced fit; deoxy-ribonucleotide metabolism; pyrimidine.

Abbreviations used: dCTP deaminase-dUTPase, bifunctional dCTP deaminase-dUTPase; rmsd, root mean square deviation; PP_i, Diphosphate.

Abstract
The crystal structures of two mutant dCTP deaminase-dUTPases from *Methanocaldococcus jannaschii*, E145A and E145Q, have been determined. The mutant enzymes that are impaired in the deaminase reaction were crystallised as a complex with α,β-imido-dUTP or diphosphate, respectively. In the structure of E145Q in complex with diphosphate, the diphosphate occupies the same position as the β- and γ-phosphoryls of the nucleotide analogue in the E145A complex. Firstly, the C-terminal region that is unresolved in previous structures of the apo-form of the enzyme is now in both cases ordered and closes over the active site by interacting with the phosphate backbone of the nucleotide or with the diphosphate. In both complexes a magnesium ion is readily observed to complex with all three phosphoryls in the nucleotide complex or with the diphosphate. A water molecule that is likely to be involved in the nucleotidyl diphosphorylase reaction is observed in the E145A:α,β-imido-dUTP complex. This water molecule is held in place by Asp135, a residue positioned similarly as the catalytic acid in the monofunctional trimeric dUTPase. Based on a comparison of the active site between the bifunctional enzyme and the monofunctional family members, dCTP deaminase and dUTPase, we may suggest that the reaction mechanisms are the same. Furthermore, the conformation of side chains involved in the deaminase reaction is similar in the nucleotide and diphosphate complexes indicating a concerted re-arrangement, or induced fit, of the active site promoted by the interactions between the enzyme and the β- and γ-phosphoryls of the bound nucleotide.

Introduction
The formation of dUMP *de novo*, the precursor of dTTP, occurs in general via two pathways; one by which UDP is reduced to dUDP by ribonucleotide reductase and further phosphorylated to dUTP and hydrolysed to dUMP and diphosphate by dUTPase, and a second, where dUMP is generated by the deamination of a cytosine deoxy-

ribonucleotide. Whereas the dUMP obtained from reduction of UDP is considered a minor supplement, the deamination of a cytosine deoxy-ribonucleotide generally provides up to 50-80% of the dUMP used for dTTP synthesis [1-3]. In eukaryotes [4-6] and Gram-positive bacteria like *Bacillus subtilis* [1] the deamination of dCMP directly to dUMP is catalysed by the zinc containing enzyme dCMP deaminase [7]. In other organisms the deamination occurs at the level of dCTP as in the Gram-negative bacteria *Salmonella typhimurium* [2,8,9] and *Escherichia coli* [10,11]. The product dUTP then gets cleaved by dUTPase to dUMP. Alternatively as it was discovered in the archaea *Methanocaldococcus jannaschii* [12-15] and recently in *Mycobacterium tuberculosis* [16], a bifunctional enzyme (dCTP deaminase-dUTPase) simultaneously deaminates and cleaves dCTP to release dUMP. The enzymes constituting the dCTP dependent dUMP forming pathways; dCTP deaminase and trimeric dUTPase, or the bifunctional enzyme are all three members of a structural family of trimeric enzymes that share a common fold for the subunit [17].
Others and we have previously published the structure of the apo-form of the *M. jannaschii* dCTP deaminase-dUTPase [13,14]. In addition, we have published the structures of the *E. coli* dCTP deaminase [17] and recently the *M. tuberculosis* bifunctional enzyme [16]. We have undertaken an investigation of both the catalytic mechanism of the unusual metal-ion independent deamination reaction carried out by these enzymes[18] and the mechanism by which dTTP regulates the activity [16,19]. The bifunctional enzyme is quite unusual in that it simultaneously modifies a substrate at two different and distant positions within the same molecule. In this work we are addressing the structural properties behind the mechanism by which the bifunctional dCTP deaminase-dUTPase coordinates the deamination and triphosphate cleavage reactions.

Results and Discussion

Crystallization and structure determination
Earlier attempts to crystallize wild type dCTP deaminase-dUTPase with substrate analogues or inhibitors have failed, therefore, two mutant enzymes, E145A and E145Q, impaired in the deaminase reaction[15] were produced. Crystallization experiments of these two mutant enzymes with dCTP, dUTP, dTTP or the intermediate product analogue α,β-imido-dUTP resulted in well-diffracting crystals of E145A in complex with α,β-imido-dUTP and Mg^{2+} and of E145Q in complex with diphosphate and Mg^{2+}. dCTP was used in the latter co-crystallisation experiment, and the observation that one of the two products was bound demonstrates that the E145Q mutant is able to cleave dCTP by way of the dUTPase half-reaction [15]. The E145Q structure revealed 30% of the enzyme in a E145Q:PPi complex and 70% of the enzyme in the apo form.

The overall fold of the *M. Jannaschii* dCTP deaminase-dUTPase trimer
The overall fold of the dCTP deaminase-dUTPase mutant enzymes in complex with either nucleotide or diphosphate was very similar to previously published structures. The C-terminal region appears to fold back over the active site to form two 3_{10}-helices - not only in the presence of nucleotide - but also in the presence of diphosphate. Figure 1A shows a superposition of the structure of the *M. jannaschii* enzyme apo-form [13,16] and the complexes of mutant *M. jannaschii* enzymes described in this work. Also included in this Figure are structures previously published of bifunctional [16] and monofunctional [17] enzyme family members in complex with nucleotides. As can be seen, the formation of the complexes results in an almost identical ordered C-terminal region that closes as a lid over the active site.

The active site
Figure 2 provides a detailed picture of the binding of α,β-imido-dUTP and PP_i to the E145A and E145Q mutant enzymes, respectively. In the structure of the E145A:α,β-imido-dUTP complex, the magnesium ion was octahedrally coordinated to the α-, β- and γ-phosphoryls as well as to three water molecules. There was clear electron density for both the nucleotide as well as the magnesium ion in the structure (Figure 2a). As mentioned, the C-terminal region was found to close over the active site and only the last five amino acid residues were unidentifiable in the electron density maps. The γ-phosphoryl of the bound nucleotide formed hydrogen bonds with Tyr177 and Lys184 from the C-terminal region. The β-phosphoryl group was hydrogen bonded to Arg117 and Ser119, whereas Ser118 and the two water molecules, wat33 and wat87, were hydrogen bond partners for the α-phosphoryl group (Figure 2b). Also, in addition to a hydrogen bond to Thr143, O2 in the uracil moiety of α,β-imido-dUTP was interacting with the C-terminal region through a hydrogen bond to the side chain of Gln188 (Figure 2b). Asp135 (OD2) formed a hydrogen bond to O3' of the deoxy-ribose ring. The other oxygen of the carboxylate group OD1 was only interacting with the water molecules wat87 and wat12 (Figure 2b). There were also noteworthy hydrophobic interactions between the nucleotide and the enzyme. The aromatic side chain

of Phe138 was stacking with the deoxy-ribosyl moiety and Ile142 interacted with the pyrimidine ring of α,β-imido-dUTP (Figure 2a).
The structure refinement revealed that the crystals of E145Q in complex with PP_i contained the protein in two conformations; one similar to the structure determined previously with no ligand and a disordered C-terminal region [13], and one resembling the active site structure of the E145A:α,β-imido-dUTP complex (Figure 2 and 3). To describe this static disorder the side chains of Tyr43, Tyr56 and residues in the segment His128 to Trp133 were modelled in double conformations (Figure 2c and 3). The closure of the active site by the C-terminal region involved a hydrogen bond between the backbone amide of Ser194 and Tyr43 and the interaction between side chains of Lys184 and Tyr177 (Figure 2c) and the diphosphate.
In Figure 3, the side chain position of His128 and Trp133 in the wild type dCTP deaminase-dUTPase without ligand bound (1OGH) is compared to the structures of E145A and E145Q with bound ligand. In the structure of the apo-form of the wild type enzyme His128 was hydrogen bonded to Asp125, but upon binding of nucleotide as represented by the E145A structure, the imidazole ring was turned away from the active site. The position of His128 in the apo-form structure would be likely to cause a steric hindrance for nucleotide binding. Also, the side chain of Arg122 changed its position upon nucleotide or diphosphate binding to the active site. In both the nucleotide and diphosphate complex the side chain of Arg122 was facing towards Phe138 and formed a hydrogen bond to OE1 of Gln188 and the OG1 of Thr143 that also interacted with the nucleotide (Figure 2a and c).

The dCTP deaminase reaction
The catalytic mechanism for the deamination of dCTP catalyzed by dCTP deaminase-dUTPase is likely to be identical to that of *E. coli* dCTP deaminase as suggested by Johansson et al. [17,18]. The position of active site residues involved in the deamination reaction as observed in the E145A:α,β-imido-dUTP complex, and those of *E. coli* dCTP deaminase in complex with dUTP superimpose well. As predicted, Glu145 in dCTP deaminase-dUTPase occupied a position equivalent to Glu138 that earlier was shown to be central to catalysis in the monofunctional *E. coli* dCTP deaminase [17,18]. Interestingly, these residues are found in the same conformation in the E145Q: PP_i complex.

The dUTPase reaction
The E145Q mutant enzyme, although impaired in deaminase activity, has been reported to be a functional dUTPase and as apparent from the crystal complex, can cleave dCTP [15]. The nucleotidyl diphosphorylase reaction is likely to be mediated by Asp135 of which the OD1 may activate wat87 through proton abstraction to make a nucleophilic attack on the α-phosphate of the substrate. The Asp135 of the *M. jannaschii* dCTP deaminase-dUTPase is in a similar position as the catalytic aspartic acid residue identified in trimeric dUTPases [20-22].
A similar orientation of the active site residues was recently observed in the dTTP inhibited form of *M. tuberculosis* dCTP deaminase-dUTPase [16]. The inhibition of the bifunctional *M. tuberculosis* enzyme by dTTP is explained by a rearrangement of residues 109-118 (residues 125-134 in the *M. jannaschii* enzyme). This rearrangement occurs to accommodate the additional methyl group of the thymine moiety of dTTP. As a result, Ala115 of the *M. tuberculosis* enzyme (Ala131 in the *M. jannaschii* enzyme) moves to expel a water molecule corresponding to wat33 in the *M. jannaschii* complex (Figure 2b). This water molecule which would otherwise hydrogen bond to the α-phosphate, is suggested to be crucial for the catalysis of the nucleotidyl diphosphorylase reaction [16]. The structure of the E145A:α,β-imido-dUTP complex lends support to this hypothesis: when superposing the E145A:α,β-imido-dUTP complex with the *M. tuberculosis* dCTP deaminase-dUTPase:dTTP complex, the nucleotides are found in identical positions. The hydrogen bond patterns between the enzyme and the triphosphates are similar and the only notable difference is the position of Ala115 in the *M. tuberculosis* enzyme complex that takes the place of that occupied by wat33 in the *M. jannaschii* enzyme (Figure 4). In the structure of the E145Q:PP_i complex no water molecule corresponding to Wat33 was found although the position of Ala131 made room for it. Although the resolution to 2.6 Å of the E145Q:PP_i complex of course imposed a limititation on the search for water molecules in the structure, the lack of an α-phosphoryl in this complex that otherwise coordinates to Wat33 is likely to explain its absence (Figure 4).

Conclusions
In the structures determined in the present work the nucleotide as well as the diphosphate could easily be positioned based on the electron densities alone (Figure 2a and 2c). The structures of these two complexes revealed large differences when compared to the to the previously solved structure of the apo-form *M. jannaschii*

dCTP deaminase-dUTPase with changes in the positions of amino acyl side chains in the active site and the ordering of the previously unresolved C-terminal region.
As mentioned, each reaction catalysed by dCTP deaminase-dUTPase is likely to occur via the same mechanism as that previously proposed for the monofunctional dCTP deaminase [17,18] and trimeric dUTPase [20,21]. However, the present work provides a new and important insight into the mechanism by which the bifunctional enzyme coordinates the simultaneous deamination and diphosphorylase reaction of dCTP. The imidazole ring of His128 points into the active site in the apo-form of the enzyme, much like what is seen in the dTTP inhibited forms of *E. coli* dCTP deaminase [19] and *M. tuberculosis* dCTP deaminase-dUTPase [16]. The observation that both His 128 and Trp133 move out of the active site to avoid steric hindrance of binding of the nucleobase moiety of the substrate suggests an induced fit mechanism for the binding of nucleotide to the enzyme (Figure 3). Also, the E145Q:PP_i complex was found to exist in two conformations: one resembling the apo-form and one where in addition to an ordered C-terminal region, the residues from His128 to Trp133 had adapted a similar conformation as in the nucleotide complex of E145A (Figure 2c and 3).
Pre-steady state kinetics of the *E. coli* dCTP deaminase revealed a time-lag in the progress curve of product formation that was increased by the presence of dTTP [19]. This lag was interpreted to represent a structural rearrangement of the same region as discussed above (His121 to His124 in the *E. coli* dCTP deaminase) from an inactive conformation to an active conformation in order to be able to bind the substrate [19]. We therefore suggest that the E145Q:PP_i complex demonstrates the likely existence of a concerted structural response of the active site in the bifunctional enzyme that accompanies nucleotide binding and which combines optimal positioning of catalytic residues for both the deaminase and the dUTPase "half-reactions". A mechanism that judged from the structure of the E145Q:PP_i complex appears to depend, at least in part, on the interaction between the enzyme and the β- and γ-phosphates of the substrate.

Materials and Methods

Crystallisation of proteins
The plasmid based on the pET-3a vector (Novagen) used for overproduction of the *M. jannaschii* dCTP deaminase-dUTPase has been described previously [12]. Mutations were introduced using the quick-change site-directed mutagenesis system (Stratagene) according to the supplier's recommendations and using the deoxy-oligonucleotides D145A5-3, GGAAAAATAACCTTGGCGATTGTTGCTTTCGAT; D145A3-5, ATCGAAAG-CAACAATCGCCAAGGTTATTTTTCC; D145N5-3, GGAAAAATAACCTTGCAGATTGTTGCTTTCGAT; D145N3-5, ATCGAAAGCAACAATCTGCAAGGTTATTTTTCC. Protein was produced and purified as previously described [12]. Crystals were obtained by the hanging drop, vapour-diffusion method with 2 µL protein solution, 2 µL mother liquor and 1 mL reservoir volume at room temperature. The PEG/ION screen (Hampton Research) was used for crystallisation of both mutant proteins. Crystals of E145A were grown from a solution of 2 mg mL^{-1} protein, 20 mM Tris-HCl, pH 8.5, 20 mM magnesium chloride and 3 mM α,β-imido-dUTP and reservoir solution of 0.2 M di-potassium hydrogen phosphate and 20 % PEG3350. Crystals of E145Q were grown from a solution of 2 mg mL^{-1} protein, 20 mM Tris-HCl, pH 8.5, 20 mM magnesium chloride and 5 mM dCTP and reservoir solution of 0.2 M tri-potassium citrate and 20 % PEG3350.

Diffraction data collection
Diffraction data from were collected at the ESRF (Grenoble, France), beam line ID23-1. 25 % ethylene glycol mixed with mother liquor was used for cryo protection. The crystals of the E145A variant and E145Q variant diffracted to 2.55 and 2.3 Å resolution, respectively. The diffraction data from crystals of E145Q were auto indexed and integrated using Denzo and scaled with Scalepack [23]. The diffraction data from crystals of E145A were integrated using MOSFLM [24] and scaled using Scala [25]. Both crystals belonged to space group I432 with one protein chain in the asymmetric unit giving a Matthews coefficient of 4.44 $Å^3$/Da, corresponding to a water content of approximately 72 %. Details on data collection, processing and statistics are summarized in Table 1.

Structure determination and refinement
The structures of mutant enzymes were determined using the molecular replacement technique as implemented in the program AMoRe [26] with the previously determined structure of dCTP deaminase-dUTPase in the space group $P2_13$ as the search model (Protein Data Bank code entry: 1OGH, chain A). All water molecules were removed from the model before commencing with molecular replacement. Refinement and model building was carried out using Refmac5 [27] and O [28]. In the E145A structure residues 1 to 199 were modelled and clear electron density for α,β-imido-dUTP was seen. 121 water molecules were furthermore included in the model. In the E145Q structure residues 1 to 199 were modelled and density for PP_i was seen. 71 water molecules were included. Tyr43 and Tyr56 were found in alternate positions and modelled as such. Furthermore, residues 128 to

133 were also found in alternate positions. 70% were in a conformation similar to the structure found in the apo-enzyme, 1OGH, and 30% were in a conformation similar to the one found in the E145A complex structure. The PPi and the C-terminal region were modelled with B-factors of about 70 $Å^2$ as compared to about 30 $Å^2$ to the neighbouring atoms. They were therefore given an occupancy of 0.3, assuming that binding of PPi was the prerequisite necessary for an ordered C-terminal and the rearrangement of residues 128-133. The quality of the structures was checked using WHATIF [29] and PROCHECK [30]. Refinement statistics are summarized in Table 1.

Protein Data Bank accession codes

The atomic coordinates and structure factors have been deposited in the Protein Data Bank with the PDB ID codes 2HXD and 2HXE.

Acknowledgements

We thank Majbritt Thymark for help with the production of enzymes and Flemming Hansen and Susan Arent for assistance during diffraction data collection. We are grateful for the beamtime provided at MAX-LAB (Lund, Sweden) and ESRF (Grenoble, France) and for the Danish Natural Science Research Council contribution to DANSYNC. We also acknowledge the support by the European Community – Research Infrastructure Action under the FP6 programme "Structuring the European Research Area".

References

1. Mollgaard, H. & Neuhard, J. (1978). Deoxycytidylate deaminase from *Bacillus subtilis*. Purification, characterization, and physiological function. *J. Biol. Chem.* **253**, 3536-3542.
2. Neuhard, J. & Thomassen, E. (1971). Deoxycytidine triphosphate deaminase: identification and function in *Salmonella typhimurium*. *J. Bacteriol.* **105**, 657-665.
3. Bianchi, V., Pontis, E. & Reichard, P. (1987). Regulation of pyrimidine deoxyribonucleotide metabolism by substrate cycles in dCMP deaminase-deficient V79 hamster cells. *Mol. Cell. Biol.* **7**, 4218-4224.
4. Nucci, R., Raia, C. A., Vaccaro, C., Rossi, M. & Whitehead, E. P. (1991). Allosteric modifier and substrate binding of donkey deoxycytidylate aminohydrolase (EC 3.5.4.12). *Arch. Biochem. Biophys.* **289**, 19-25.
5. Ellims, P. H., Kao, A. Y. & Chabner, B. A. (1983). Kinetic behaviour and allosteric regulation of human deoxycytidylate deaminase derived from leukemic cells. *Mol. Cell. Biochem.* **57**, 185-190.
6. McIntosh, E. M. & Haynes, R. H. (1984). Isolation of a *Saccharomyces cerevisiae* mutant strain deficient in deoxycytidylate deaminase activity and partial characterization of the enzyme. *J. Bacteriol.* **158**, 644-649.
7. Almog, R., Maley, F., Maley, G. F., Maccoll, R. & Van Roey, P. (2004). Three-dimensional structure of the R115E mutant of T4-bacteriophage 2'-deoxycytidylate deaminase. *Biochemistry* **43**, 13715-13723.
8. O'Donovan, G. A., Edlin, G., Fuchs, J. A., Neuhard, J. & Thomassen, E. (1971). Deoxycytidine triphosphate deaminase: characterization of an *Escherichia coli* mutant deficient in the enzyme. *J. Bacteriol.* **105**, 666-672.
9. Krogan, N. J., Zaharik, M. L., Neuhard, J. & Kelln, R. A. (1998). A combination of three mutations, *dcd*, *pyrH*, and *cdd*, establishes thymidine (Deoxyuridine) auxotrophy in *thyA*$^+$ strains of *Salmonella typhimurium*. *J. Bacteriol.* **180**, 5891-5895.
10. Wang, L. & Weiss, B. (1992). *dcd* (dCTP deaminase) gene of *Escherichia coli*: mapping, cloning, sequencing, and identification as a locus of suppressors of lethal *dut* (dUTPase) mutations. *J. Bacteriol.* **174**, 5647-5653.
11. Weiss, B. & Wang, L. (1994). De novo synthesis of thymidylate via deoxycytidine in *dcd* (dCTP deaminase) mutants of *Escherichia coli*. *J. Bacteriol.* **176**, 2194-2199.
12. Bjornberg, O., Neuhard, J. & Nyman, P. O. (2003). A bifunctional dCTP deaminase-dUTP nucleotidohydrolase from the hyperthermophilic archaeon *Methanocaldococcus jannaschii*. *J. Biol. Chem.* **278**, 20667-20672.
13. Johansson, E., Bjornberg, O., Nyman, P. O. & Larsen, S. (2003). Structure of the bifunctional dCTP deaminase-dUTPase from *Methanocaldococcus jannaschii* and its relation to other homotrimeric dUTPases. *J. Biol. Chem.* **278**, 27916-27922.
14. Huffman, J. L., Li, H., White, R. H. & Tainer, J. A. (2003). Structural basis for recognition and catalysis by the bifunctional dCTP deaminase and dUTPase from *Methanococcus jannaschii*. *J. Mol. Biol.* **331**, 885-896.
15. Li, H., Xu, H., Graham, D. E. & White, R. H. (2003). The *Methanococcus jannaschii* dCTP deaminase is a bifunctional deaminase and diphosphatase. *J. Biol. Chem.* **278**, 11100-11106.
16. Helt, S. S., Thymark, M., Harris, P., Aagaard, C., Dietrich, J., Larsen, S. & Willemoes, M. (2008). Mechanism of dTTP Inhibition of the Bifunctional dCTP Deaminase:dUTPase Encoded by Mycobacterium tuberculosis. *J. Mol. Biol.* **376**, 554-569.
17. Johansson, E., Fano, M., Bynck, J. H., Neuhard, J., Larsen, S., Sigurskjold, B. W., Christensen, U. & Willemoes, M. (2005). Structures of dCTP deaminase from *Escherichia coli* with bound substrate and product: reaction mechanism and determinants of mono- and bifunctionality for a family of enzymes. *J. Biol. Chem.* **280**, 3051-3059.
18. Thymark, M., Johansson, E., Larsen, S. & Willemoes, M. (2007). Mutational analysis of the nucleotide binding site of Escherichia coli dCTP deaminase. *Arch. Biochem. Biophys.* **470**, 20-26.
19. Johansson, E., Thymark, M., Bynck, J. H., Fano, M., Larsen, S. & Willemoes, M. (2007). Regulation of dCTP deaminase from Escherichia coli by nonallosteric dTTP binding to an inactive form of the enzyme. *FEBS J.* **274**, 4188-4198.
20. Chan, S., Segelke, B., Lekin, T., Krupka, H., Cho, U. S., Kim, M. Y., So, M., Kim, C. Y., Naranjo, C. M., Rogers, Y. C., Park, M. S., Waldo, G. S., Pashkov, I., Cascio, D., Perry, J. L. & Sawaya, M. R. (2004). Crystal structure of the *Mycobacterium tuberculosis* dUTPase: insights into the catalytic mechanism. *J. Mol. Biol.* **341**, 503-517.
21. Barabas, O., Pongracz, V., Kovari, J., Wilmanns, M. & Vertessy, B. G. (2004). Structural insights into the catalytic mechanism of phosphate ester hydrolysis by dUTPase. *J. Biol. Chem.* **279**, 42907-42915.
22. Harris, J. M., McIntosh, E. M. & Muscat, G. E. (1999). Structure/function analysis of a dUTPase: catalytic mechanism of a potential chemotherapeutic target. *J. Mol. Biol.* **288**, 275-287.
23. Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 620-633.

24. Leslie, A. G. W. (1992). Recent changes to the MOSFLM package for processing film and image plate data. *Joint CCP4 + ESF-EAMCB Newsletter on Protein Crystallography*, **26**.
25. Evans, P. R. (1993). Data reduction. *Proceedings of CCP4 Study Weekend, 1993, on Data Collection & Processing*, 114-122.
26. Navaza, J. (1994). AMoRe: an automated package for molecular replacement. *Acta Crystallogr. sect. A,* **50**, 157-163.
27. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta Crystallogr. sect. D,* **53**, 240-255.
28. Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. sect. A,* **47**, 110-119.
29. Vriend, G. (1990). WHAT IF: a molecular modeling and drug design program. *J. Mol. Graph.* **8**, 52-56, 29.
30. Laskowski, R. a., Macarthur, M. W., Moss, D. S. & Thornton, J. M. (1993). Procheck - a Program to Check the Stereochemical Quality of Protein Structures. *J. Appl. Crystallogr.* **26**, 283-291.
31. Kraulis, P. J. (1991). Molscript - a Program to Produce Both Detailed and Schematic Plots of Protein Structures. *J. Appl. Crystallogr.* **24**, 946-950.
32. Merritt, E. A. & Bacon, D. J. (1997). Raster3D: Photorealistic molecular graphics. *Methods Enzymol.* **277**, 505-524.

Enzyme:ligand	E145A:α,β-imido- dUTP	E145Q:PP_i
Resolution range (Å)	25 - 2.3 (2.42 - 2.30)[a]	25 - 2.3 (2.69 - 2.62)[a]
R_{merge} (%)[b]	8.9 (43.2)[a]	10.1 (43.4)[a]
Space group	I432	I432
Cell dimensions a,b,c (Å)	171.5	171.6
$<I>/\sigma(I)$	14.4 (3.0)[a]	26.8 (5.4)[a]
Completeness (%)	96.4 (98.9)[a]	100.0 (100.0)[a]
Total no. of reflections	80,955	178,848
No. of unique reflections	18,566	14,787
Monomers in asymmetric unit	1	1
Total no. of reflections	17,636	12,631
No. of reflections (working set)	16,712	11,995
No. of reflections (test set)	924	636
No. of protein atoms	1756	1756
No. of water molecules	122	72
R_{factor} (%)[c]	17.8 (26.4)[a]	18.5 (23.0)[a]
R_{free} (%)[d]	22.1 (35.3)[a]	22.8 (28.0)[a]
Average B_{factor} (Å)	27.1	35.1
Average B_{factor} protein atoms (Å)	30.1	38.3
Average B_{factor} nucleotide or PP_i and Mg^{2+} (Å)	33.1	64.7
Average B_{factor} water atoms (Å)	35.4	36.3
Bond length rmsd[e] from ideal (Å)	0.022	0.021
Bond angle rmsd[e] from ideal (deg.)	1.84	1.74
Ramachandran plot		
Most favoured regions (%)	89.9	92.1
Additional allowed regions (%)	9.0	7.9
Generously allowed regions (%)	1.1	0
Disallowed regions (%)	0	0

Figure legends

Figure 1. Structural comparison of subunits of the bifunctional dCTP deaminase-dUTPase from *M. jannaschii* and *M. tuberculosis* and the monofunctional *E. coli* dCTP deaminase. The colour coding is as follows; *M. jannaschii* apo-form (*yellow*), *M. jannaschii* E145Q:PP_i complex (*cyan*), *M. jannaschii* E145A:α,β-imido-dUTP complex (*green*), *E. coli* dCTP deaminase-dUTP complex (*purple*) (Protein Data Bank entry 1XS1) and the *M. tuberculosis* dCTP deaminase-dUTPase:dTTP complex (*pink*) (Protein Data Bank entry 2QXX). The nucleotides and the diphosphate are in standard colouring. The magnesium ions are in dark grey. The N-and C-terminals of the subunits are as indicated. The apo-form of the *M. jannaschii* enzyme has an unresolved C-terminal region for which reason only the last residue with clear electron density has been modelled. Panels were prepared using PyMol (DeLano Scientific).

Figure 2. Views of the active site of the E145A:α,β-imido-dUTP complex and the E145Q:PP_i complex. (a), Close up stereo picture of the E145A:α,β-imido-dUTP complex. The orange and blue amino acid residues are contributed by two different subunits in the trimer. (b), Schematic view of the active site architecture of the E145A:α,β-imido-dUTP complex. Amino acid residues in black and blue are contributed by two different

subunits in the trimer, Asp135 is coloured in red to indicate its double appearance. (c), Close up stereo picture of the E145Q:PP_i complex. Colouring of amino acid residues is as above. In panel a and c the displayed $2F_{obs} - F_{calc}$ difference electron density maps are contoured at 1.0 σ level at a distance of 1.0 Å around the nucleotide, the diphosphate and the magnesium ions. The Figure was prepared using MOLSCRIPT [31] and Raster3D [32].

Figure 3. Stereo view of the amino acid residues in double conformations in the crystal structure of E145Q:PP_i complex (shown in *yellow* and *blue*) compared to wild-type dCTP deaminase-dUTPase structure with no ligand bound (shown in *red*) and the E145A:α,β-imido-dUTP complex (shown in *green*). The amino acid residues were superimposed using default parameters in the program O [28]. The Figure was prepared using MOLSCRIPT [31] and Raster3D [32].

Figure 4. Two different views, A and B, of an overlay of the active site around the bound ligands in the E145A:α,β-imido-dUTP (*green*) and E145Q:PP_i (*yellow*) complexes and the *M. tuberculosis* bifunctional enzyme in complex with dTTP (*purple*, PDB accession code 2QXX). The E145Q:PP_i complex is represented by two structures, one with no diphosphate bound similar to the apo-form (1OGH) and one with diphosphate bound similar to the E145A:α,β-imido-dUTP complex (*see text*). The spheres coloured as the complexes represent important water molecules. Black spheres represent magnesium ions. The solid arrow points to the proposed catalytic water molecule that attacks the α-phosphoryl in the dUTPase half-reaction [20] (*see text*). The dotted arrow points to the newly identified water molecule proposed to also play an important role in catalysis of the dUTPase half-reaction [16]. The side chain of Ala115 of the *M. tuberculosis* enzyme, indicated in bold, (Ala131 in *M. jannaschii*) that plays a role in the dTTP is compared to the position of the homologues residue, Ala131, in *M. jannaschii*. Panels were prepared using PyMol (DeLano Scientific).

FIGURE 1

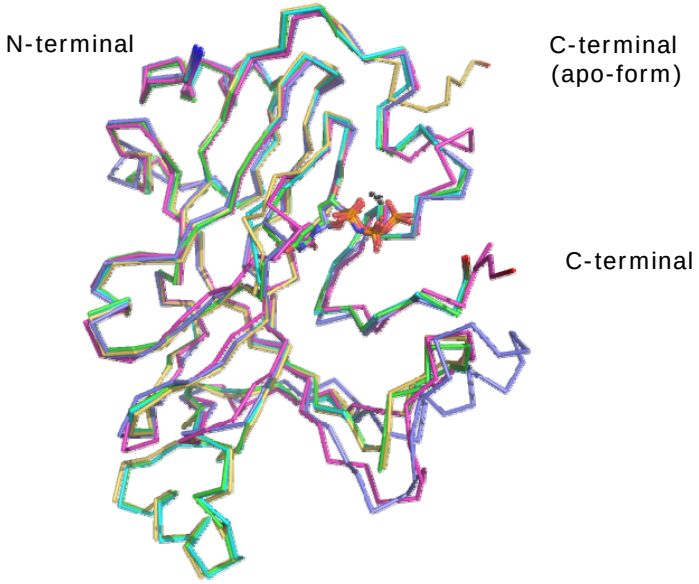


FIGURE 2A

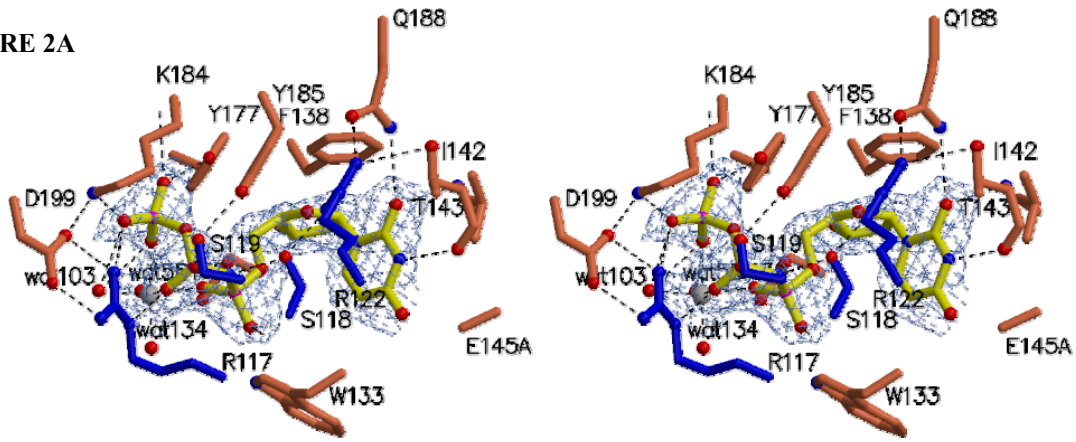


FIGURE 2B

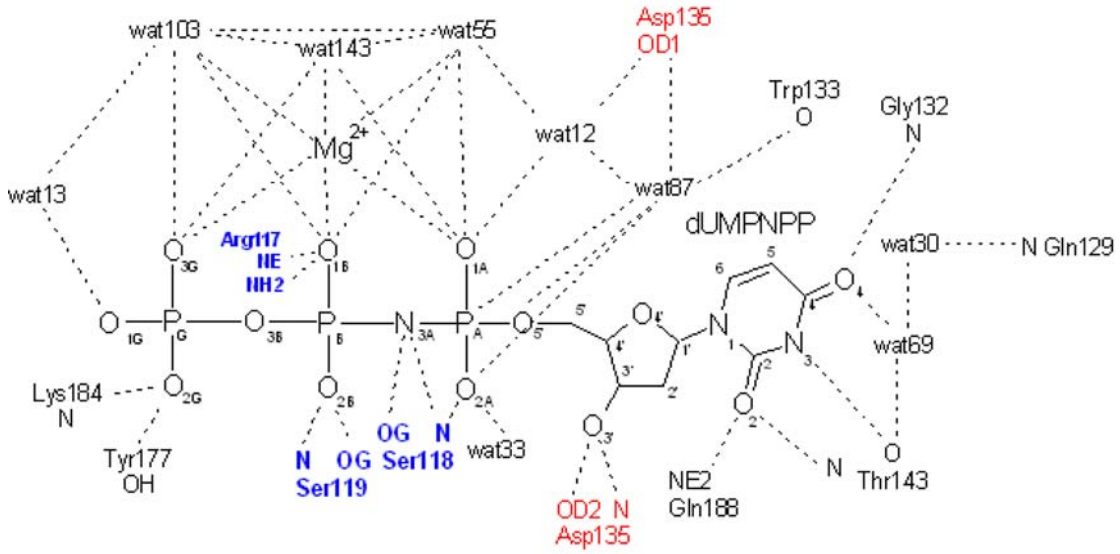


FIGURE 2C

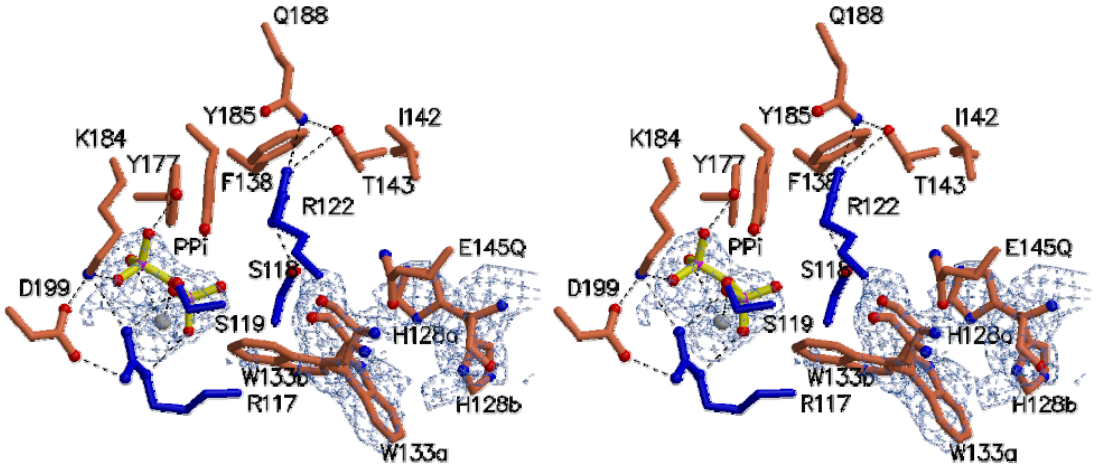


FIGURE 3

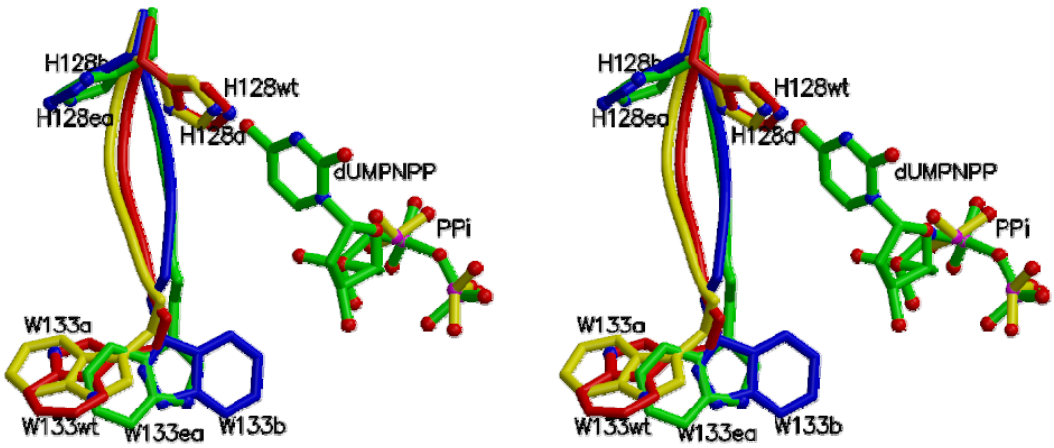


FIGURE 4A

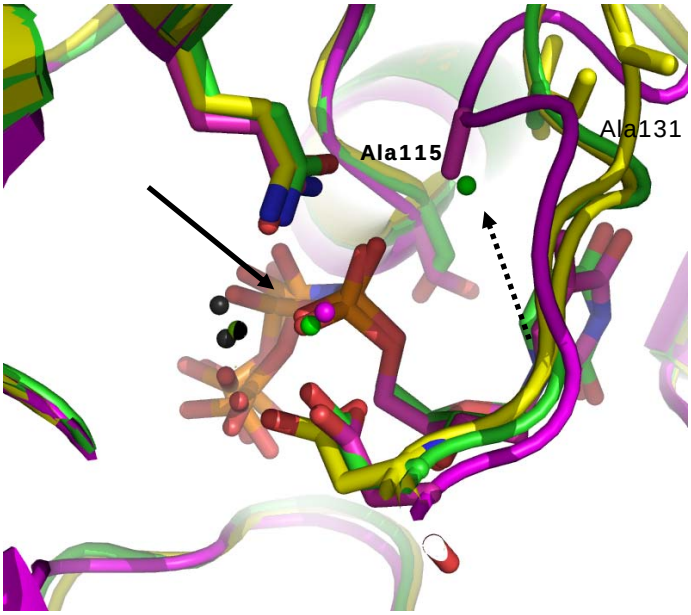
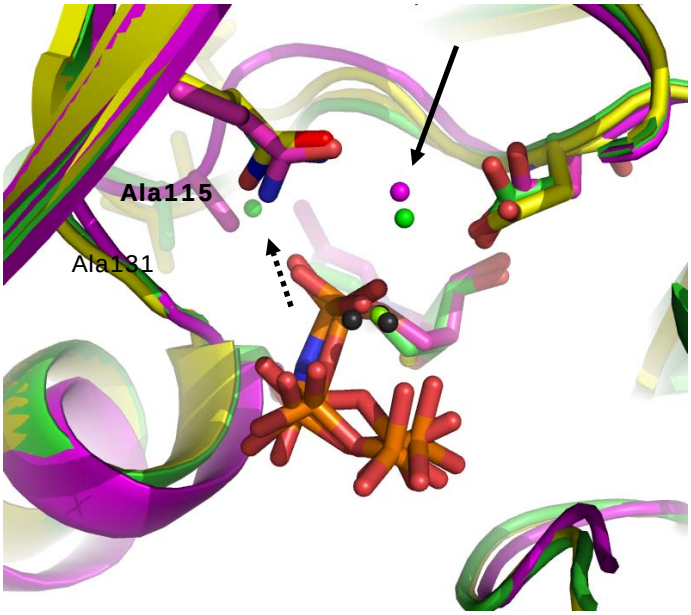


FIGURE 4B



Bibliography

Almog R., Maley F., Maley G. F., Maccoll R. and Van Roey P. Three-dimensional structure of the R115E mutant of T4-bacteriophage 2'-deoxycytidylate deaminase, *Biochemistry* 43 (2004) 13715-13723.

Barabás O., Pongrácz V., Kovári J., Wilmanns M. and Vértessy B. G., Structural insights into the catalytic mechanism of phosphate ester hydrolysis by dUTPase, *J. Biol. Chem.* 279 (2004) 42907-42915.

Beck C. F., Eisenhardt A. R. and Neuhard J., Deoxycytidine triphosphate deaminase of *Salmonella typhimurium*. Purification and characterization, *J. Biol. Chem.* 250 (1975) 609-616.

Berg J. M., Tymoczko J. L. and Stryer L., Biochemistry (5th edition), W. H. Freeman and Company, USA (2002) (ISBN: 0716746840).

Bianchi V., Pontis E. and Reichard P., Regulation of pyrimidine deoxyribonucleotide metabolism by substrate cycles in dCMP deaminase-deficient V79 hamster cells, *Mol. Cell. Biol.* 7 (1987) 4218-4224.

Björnberg O., Neuhard J. and Nyman P. O., A bifunctional dCTP deaminase-dUTP nucleotidohydrolase from the hyperthermophilic archeon *Methanocaldococcus jannaschii*, *J. Biol. Chem.* 278 (2003) 20667-20672.

Chan S., Segelke B., Lekin T., Krupka H., Cho U. S., Kim M.-Y., So M., Kim C.-Y., Naranjo C. M., Rogers Y. C., Park M. S., Waldo G. S., Pashkov I., Cascio D., Perry J. L. and Sawaya M. R., Crystal structure of the *Mycobacterium tuberculosis* dUTPase: insights into the catalytic mechanism, *J. Mol. Biol.* 341 (2004) 503-517.

Climie S., Lutz T., Radul J., Sumner-Smith M., Vandenberg E. and McIntosh E., Expression of trimeric human dUTP pyrophosphatase in *Escherichia coli* and purification of the enzyme, *Protein Expr Purif* 5 (1994) 252-258.

Collaborative Computational Project, Number 4. The *CCP*4 suite: Programs for protein crystallography, *Acta Cryst*. D50 (1994) 760-763.

Dauter Z., Persson R., Rosengren A. M., Nyman P. O., Wilson K. S. and Cedergren-Zeppezauer E. S., Crystal structure of dUTPase from equine infectious anaemia virus; active site metal binding in a substrate analogue complex, *J. Mol. Biol.* 285 (1999) 655-673.

DeLano W.L., The PyMOL Molecular Graphics System (2002) DeLano Scientific, Palo Alto, CA, USA. http://www.pymol.org (accessed December 9th 2007).

Delbaere L. T. J., Sudom A. M., Prasad L., Leduc Y. and Goldie H., Structure/function studies of phosphoryl transfer by phosphoenolpyruvate carboxykinase, *Biochim. Biophys. Acta* 1697 (2004) 271-278.

Emsley P. and Cowtan K., *Coot*: model-building tools for molecular graphics, *Acta Cryst*. D**60** (2004) 2126-2132.

Gadsden M. H., McIntosh E. M., Game J. C., Wilson P. J, and Haynes R. H., dUTP pyrophosphatase is an essential enzyme in *Saccharomyces cerevisiae*, *EMBO J.* 12 (1993) 4425–4431.

González A., Larsson G., Persson R. and Cedergren-Zeppezauer E., Atomic resolution structure of *Escherichia coli* dUTPase determined *ab initio*, *Acta. Cryst.* D57 (2001) 767-774.

Gorrell A., Lawrence S. H. and Ferry J. G., Structural and kinetic analyses of arginine residues in the active site of acetate kinase from *Methanosarcina thermophila*, *J. Biol. Chem.* 280 (2005) 10731-10742.

Grässer F. A., Romeike B. F. M., Neidobitek G., Nicholls J. and Kremmer E., dUTPase in human neoplastic cells as a potential target for therapeutic intervention, *Curr. Protein Pept. Sci.* 2 (2001) 349-360.

Hall T., Ibis Biosciences, A subsidiary of Isis Pharmaceuticals, 1891 Rutherford Road, Carlsbad, CA 92008, http://www.mbio.ncsu.edu/BioEdit/bioedit.html (accessed December 9th 2007).

Helt S. S., 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.* 376 (2008) 554-569.

Johansson E., Björnberg O., Nyman P. O. and Larsen S., Structure of the bifunctional dCTP deaminase-dUTPase from *Methanocaldococcus jannaschii* and its relation to other homotrimeric dUTPases, *J. Biol. Chem.* 278 (2003) 27916-27922.

Johansson E., Fanø M., Bynck J. H., Neuhard J., Larsen S., Sigurskjold B. W., Christensen U. and Willemoës M., Structures of dCTP deaminase from *Escherichia coli* with bound substrate and product: reaction mechanism and determinants of mono- and bifunctionality for a family of enzymes, *J.Biol.Chem.* 280 (2005) 3051-3059.

Johansson E., Thymark M., Bynck J. H., Fanø M., Larsen S., Willemoës M., Regulation of dCTP deaminase from *Eschericia coli* by nonallosteric dTTP binding to an inactive form of the enzyme, *FEBS J.* 274 (2007) 4188-4198.

Kabsch W., Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants, *J. Appl. Cryst.* 26 (1993). 795–800.

Ladner R. D., The role of dUTPase and uracil-DNA repair in cancer chemotherapy, *Curr. Protein Pept. Sci.* 2 (2001) 361-370.

Lamzin V. S., Perrakis A. and Wilson K. S., The ARP/WARP suite for automated construction and refinement of protein models. In *Int. Tables for Crystallography.* Vol. F: Crystallography of biological macromolecules (1st edition), edited by Rossmann M. G. and Arnold E., Kluwer Academic Publishers, The Netherlands (2001) (ISBN: 0792368576).

Larsson G., Nyman P. O. and Kvassman J. O., Kinetic characterization of dUTPase from *Escherichia coli*, *J. Biol. Chem.* 271 (1996) 24010-24016.

Laskowski R. A., MacAthur M. W., Moss D. S. and Thornton J. M., *PROCHECK*: a program to check the stereochemical quality of protein structures, *J. Appl. Cryst.* 26 (1993) 283-291.

Li H., Xu H., Graham D. E. and White R. H., The *methanococcus jannaschii* dCTP deaminase is a bifunctional deaminase and diphosphatase, *J. Biol. Chem.* 278 (2003) 11100-11106.

Madsen D., Johansson P. and Kleywegt G. J, (2002) Indonesia: An integrated sequence analysis system. (In prep.) http://xray.bmc.uu.se/dennis/ (accessed December 9th 2007).

McIntosh E. M. and Hanes R. H., dUTP pyrophosphatase as a potential target for chemotherapeutic drug development, *Acta Biochim. Pol.* 44 (1997) 159-171.

Mol C. D., Harris J. M., McIntosh E. M. and Tainer J. A., Human dUTP pyrophosphatase: uracil recognition by a beta hairpin and active sites formed by three separate subunits,*Structure* 4 (1996) 1077-1092.

Murshudov G. N., Vagin A. A. and Dodson E. J., Refinement of macromolecular structures by the maximum-likelihood method, *Acta Cryst.* D53 (1997) 240-255.

Myllykallio H., Lipowski G., Leduc D., Filee J., Forterre P. and Liebl U., An alternative flavin-dependent mechanism for thymidylate synthesis, *Science* 297 (2002) 105-107.

Møllgaard H. and Neuhard J., Deoxycytidylate deaminase from *Bacillus subtilis*. Purification, characterization, and physiological function, *J. Biol. Chem.* 253 (1978) 3536-3542.

Neuhard J. and Thomassen E., Deoxycytidine triphosphate deaminase: identification and function in *Salmonella typhimurium*. *J. Bacteriol.* 105 (1971) 657-665.

Neuhard J., dCTP deaminase from *Salmonella typhimurium*, *Methods. Enzymol.* 51 (1978) 418-423.

Neuhard J. and Nygaard P., Purines and pyrimidines. In *Escherichia coli and Salmonella typhimurium*. Cellular and molecular biology, edited by Neidhart F. C., Ingraham J. L., Low K. B., Magasanik B., Schaechter M., and Umbarger H. E., ASM Press, Washington, D. C. (1987) (ISBN: 0914826891).

Neuhard J. and Kelln R. A., Biosynthesis and conversions of pyrimidines. In *Escherichia coli and Salmonella*. Cellular and molecular biology (2nd edition), edited by Neidhardt F. C., Curtiss R. III, Ingraham J. L., Lin E. C. C., Low K. B., Magasanik B., Reznikoff W. S., Riley M., Schaechter M. and Umbarger, H. E., ASM Press, Washington, D. C. (1996) (ISBN: 1555810845).

Nord J., Larsson G., Kvassman J. O., Rosengren A. M. and Nyman P. O., dUTPase from the retrovirus equine infectious anemia virus: specificity, turnover and inhibition, *FEBS Lett.* 414 (1997) 271-274.

Ostermann N., Schlichting I., Brundiers R., Konrad M., Reinstein J., Veit T., Goody R. S. and Lavie A., Insights into the phosphoryltransfer mechanism of human thymidylate kinase gained from crystal structures of enzyme complexes along the reaction coordinate, *Structure* 8 (2000) 629-642.

Prasad G. S., Stura E. A., McRee D. E., Laco G. S., Hasselkus-light C., Elder J. H. and Stout C. D., Crystal structure of dUTP pyrophosphatase from feline immunodeficiency virus, *Protein Sci.* 5 (1996) 2429-2437.

Prasad G. S., Stura E. A., Elder J. H. and Stout C. D., Structure of feline immunodeficiency virus dUTP pyrophosphatase and its nucleotide complexes in three crystal forms, *Acta Cryst.* D56 (2000) 1100-1109.

Richards R. G., Sowers L. C., Laszlo J. and Sedwick W. D., The occurrence and consequences of deoxyuridine in DNA, *Advan. Enzyme Reg*. 22 (1984) 157-185.

Shao H., Robek M. D., Threadgill D. S., Mankowski L. S., Cameron C. E., Fuller F. J. and Payne S. L., Characterization and mutational studies of equine infectious anemia virus dUTPase, *Biochim. Biophys. Acta* 1339 (1997) 181-191.

Siggaard J. H. B., Johansson E., Helt S. S., Harris P., Larsen S. and Willemoës M., Structural evidence for a concerted bifunctionality in dCTP deaminase:dUTPase from *Methanocaldococcus jannaschii*, submitted to *J. mol. Biol.* 2008.

Vagin A. and Teplyakov A., *MOLREP*: an automated program for molecular replacement, *J. Appl. Cryst*. 30 (1997) 1022-1025.

Varga B., Barabás O., Kovári J., Tóth J., Hunyadi-Gulyás E., Klement E., Medzihradszky K. F., Tölgyesi F., Fidy J. and Vértessy B. G., Active site closure facilitates juxtaposition of reactant atoms for initiation of catalysis by human dUTPase, *FEBS Lett.* 581 (2007) 4783-4788.

Vriend G., WHAT IF: a molecular modeling and drug design program. *J. Mol. Graphics* 8 (1990) 52-56.

Weiner K. X., Weiner R. S., Maley F., and Maley G. F., Primary structure of human deoxycytidylate deaminase and overexpression of its functional protein in *Escherichia coli*, *J. Biol. Chem.* 268 (1993) 12983-12989.

World Health Organization: http://www.who.int/tb/en/ (accessed December 9th 2007).

Xu Y.-W., Moréra S., Janin J. and Cherfils J., AlF_3 mimics the transition state of protein phosphorylation in the crystal structure of nucleoside diphosphate kinase and MgADP, *Proc. Natl. Acad. Sci.* USA 94 (1997) 3579-3583.