Department of Biosciences and Nutrition
Karolinska Institutet, Stockholm, Sweden

STRUCTURAL AND FUNCTIONAL STUDIES OF GLUTAREDOXINS AND MEMBERS OF THE THIOREDOXIN SUPERFAMILY

Johan Sagemark

Stockholm 2009

Published by Karolinska Institutet. Printed by Universitetsservice US-AB

ISBN 978-91-7409-306-3

ABSTRACT

The active site of *E.coli* glutaredoxin 3 was investigated using a combination of experimental and theoretical techniques. Starting from different conformations, molecular dynamics simulations converged to an active site conformation where the Cys11 thiolate was hydrogen bonded to surrounding amide protons and to the thiol proton of Cys14. The Cys14 $\chi 1$ and $\chi 2$ torsion angles determined by NMR supported the simulation result. The torsion angles of Cys11 could not be determined due to degenerate β-protons. The NMR titration of His15 showed that it was the sidechain was in the ε tautomer form and had a pK_a value of 6.0. The pH-induced unfolding of WT and the C14A and K8A mutants were monitored by UV- and CD-spectroscopy. The Cys11 thiolate becomes protonated when the protein unfolds, or the protein unfolds when Cys11 becomes protonated. The removal of the thiol group in the C14A had a large impact on the outcome of the titrations while the K8A mutant only had a marginal effect.

Human spermatid-specific thioredoxin-1 is a protein that is expressed only in sperm during maturation. Analysis of crystallization trials showed that only one part of the protein was left. CD spectra of the full length protein and the two parts were collected. It was clear from the spectra of the N-terminal domain that it was largely unstructured. In contrast, the C-terminal domain was that of a folded protein and was found to be quite similar to human Trx1.

There are two human dithol glutaredoxins. The redox potentials of these were determined using glutathione redox buffer, direct protein-protein equilibration and thermodynamic linkage. The values were found to be -232 mV and -221 mV for hGrx1 and hGrx2, respectively. Furthermore, a second disulfide bond was discovered. The redox potential was determined to be -317 mV. Hence, it is present as a disulfide in the protein except under very reducing conditions *in vivo*. Phylogenetic analysis showed that there are three distinct groups of glutaredoxins, Grx1, Grx2 and the monothiol Grx5.

A high precision NMR structure of *E.coli* Grx3 was solved. The RMSD of the backbone atoms was 0.26 Å relative to the mean. The Cys11 residue in active site showed two conformations. In the first, the thiolate was the acceptor in a hydrogen bond network where the donors where the surrounding amide protons and the thiol proton of Cys14. The other conformation displayed a hydrogen bond between the Cys11 and Thr10. Double mutant cycle analyses showed that there is a favorable interaction between the side chains of Cys11 and Cys14 (-1.63 kcal/mol). However, the interaction between Thr10 and Cys11 was unfavorable (+0.68 kcal/mol). The structures of the oxidized form and of the Grx3-glutathione complex were recalculated in order to make a valid comparison to the reduced form. The analysis showed that there are small differences at the backbone level between the redox forms. There are some differences in the active site. The most apparent ones are the conformation of Tyr13 and the position of Val52, both important in substrate binding.

LIST OF PUBLICATIONS

I. Foloppe N., **Sagemark J.**, Nordstrand K., Berndt, K.D., Nilsson, L.
Structure, dynamics and electrostatics of the active site of glutaredoxin 3 from *Escherichia coli*: comparison with functionally related proteins.
J. Mol. Biol. **310**, 449-470 (2001)

II. Jiménez A., Johansson C., Ljung J., **Sagemark, J**., Berndt K.D., Ren B., Tibbelin G., Ladenstein R., Kieselbach T., Holmgren A., Gustafsson J.Å., Miranda-Vizuete A.
Human spermatid-specific thioredoxin-1 (Sptrx-1) is a two-domain protein with oxidizing activity.
FEBS Lett. **530**, 79-84 (2002)

III. **Sagemark J**., Elgan T.H., Burglin T.R., Johansson C., Holmgren A., Berndt K.D.
Redox properties and evolution of human glutaredoxins.
Proteins. **68**, 879-892 (2007)

IV. **Sagemark J**., Guntert P., Berndt K.D.
NMR structure of the reduced form of *E.coli* glutaredoxin 3: Relation to other redox forms and mechanistic implications.
Manuscript

CONTENTS

Introduction 1
The thioredoxin superfamily 1
Glutaredoxins 2
Glutaredoxin mechanisms 3
Reaction rates, redox potentials and thiolate stabilization 4
Methods 6
Mutagenesis 6
Protein expression 6
Protein purification 7
Redox potential determination 8
Optical Spectroscopy 9
UV spectroscopy 9
CD spectroscopy 11
NMR spectroscopy 14
Structure determination 14
Titrations 17
Results 19
Paper I 19
Paper II 20
Paper III 20
Paper IV 21
Future perspectives 23
Acknowledgements 25
References 26

LIST OF ABBREVIATIONS

CD	circular dichroism
Cys_{GS}	the cysteine residue in glutathione
DTT	dithiothreitol
DsbA	protein promoting disulfide bond formation in *E.coli*, gene product of *dsbA*
Grx	glutaredoxin
Grx-SG	glutaredoxin-glutathione mixed disulfide complex
GSH	reduced glutathione
GSSG	oxidized glutathione
hGrx	human glutaredoxin
HSQC	heteronuclear single quantum coherence
IPTG	isopropyl-β-D-thiogalactopyranosid
MD	molecular dynamics
NMR	nuclear magnetic resonance
NOE	nuclear overhauser effect
NOESY	NOE spectroscopy
RMSD	root mean square deviation
RNR	ribonucleotide reductase
TOCSY	total correlation spectroscopy
Trx	thioredoxin
UV	ultraviolet

INTRODUCTION

The thioredoxin superfamily members catalyze thiol-disulfide exchange reactions. Depending on the redox potential and substrate specificity of the enzymes in this family, reductions, disulfide bond formations and isomerizations, and de-glutathionylations are catalyzed [1]. Glutaredoxins (Grxs) are found in most living organisms such as virus, bacteria, plants and mammals. In general, Grxs are poor general reductases. Instead they are more specific towards glutathione-mixed disulfides [2]. Glutathione (GSH), γ-Glu-Cys-Gly, is present in millimolar concentrations in the cell. The ratio of reduced to oxidized glutathione is very high under normal conditions, keeping a reducing environment in the cell. During oxidative stress, a number of modifications of thiol groups in proteins can occur. Glutathionylation of these residues and the subsequent reduction by Grxs seem to be important in redox regulation and signaling.

THE THIOREDOXIN SUPERFAMILY

The thioredoxin superfamily members include thioredoxins (Trxs), glutaredoxins (Grxs), the eukaryotic protein disulfide isomerases (PDIs), the bacterial Dsb family, glutathione S-transferases GSTs) and peroxidases. The subfamilies share low sequence identity compared to each other but have a common basic fold, the thioredoxin fold [3]. The name originates from the first structure that was determined, the one of *E.coli* Trx1. The fold is a αβα sandwich: a central mixed four-stranded β-sheet is surrounded by three α-helices (Fig 1). *E.coli* Grx1 [4] and T4 Grx [5] are the smallest proteins in the family and have no additional secondary structure elements than the basic fold. Trx itself has in addition a β-stand and a α-helix at the N-terminus [6]. *E.coli* Grx3 has a small C-terminal α-helix in addition to the basic fold. The human dithiol glutaredoxins 1 and 2 have two additional α-helices, placed at the N- and C-terminus, respectively. DsbA is larger, 21 kDa, due to the insertion of a whole domain after helix 2 [7]. The protein still contains the thioredoxin fold as the extra helices form a separate domain. The active site is situated near the N-terminus of helix 1 of the fold. Two general features in the family is the presence of a C-X-X-C/S sequence and a *cis*-Pro loop in the active site.

Thioredoxins are general disulfide reductants. Its cellular function includes providing reducing equivalents to ribonucleotide reductase (RNR), oxidative stress protection, redox regulation of cellular function and metabolic pathways. The disulfide bond formed in the active site is reduced by thioredoxin reductase, using reducing equivalents from NADPH [1]. There are also a number of more specialized thioredoxins. One of them is the human spermatid-specific thioredoxin 1 (paper II). It is only expressed during spermatogenesis [8]. The oxidizing DsbA, catalyzes the formation of disulfide bonds in the periplasm of *E.coli* [9]. The eukaryotic PDI is located in the lumen of the endoplasmatic reticulum and contains four thioredoxin domains and catalyses disulfide bond oxidation and isomerization [10]. The GSTs contain no dithiols in the active site. The function of this protein is to activate GSH to attack xenobiotic substrates [11]. Peroxidases catalyze the reduction of H_2O_2 and organic hydroperoxides [12] and are the only members of the thioredoxin superfamily that do not have the *cis*-Pro loop.

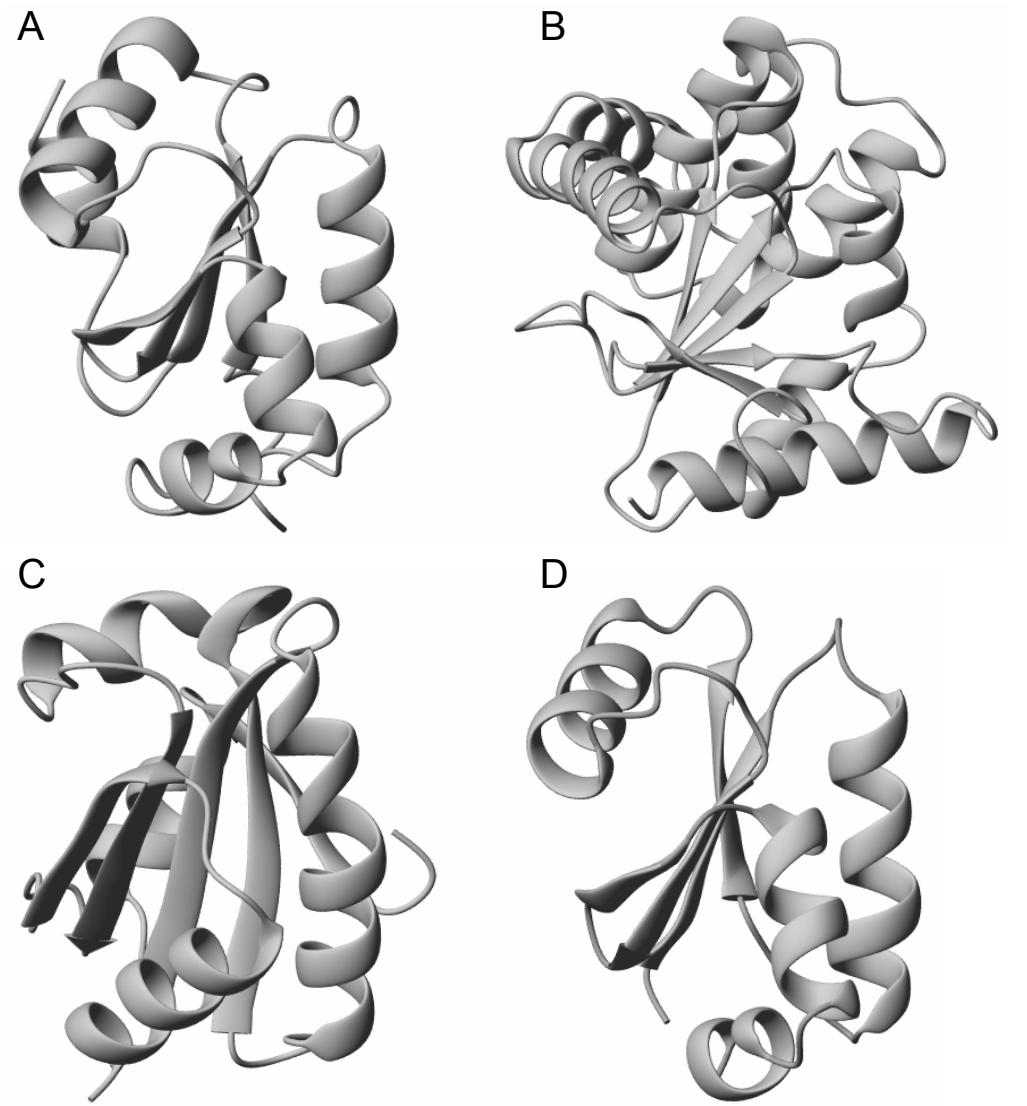

Figure 1. Selected members of the thioredoxin superfamily. E.coli Trx (A), E.coli DsbA (B), Human Grx1 (C) and E.coli Grx3 (D).

GLUTAREDOXINS

E.coli Grx1 was discovered as a ribonucleotide reductase hydrogen donor in cells deficient in Trx [13]. In addition, Grx1 was later found to catalyze the reduction of glutathione mixed disulfide substrates. However, Grx1 could only account for less than 1 % of the de-glutathionylation of substrates in crude extracts from *E.coli* [14]. Trx is inactive in the latter reaction, since it cannot be reduced by GSH [15]. This can partly be attributed to the fact that thioredoxin has a lower redox potential than GSH, -270 compared to -240 mV. Two additional dithiol glutaredoxins were found in *E.coli* [16]. Grx2, with a structure that is more similar to the GSTs than standard Grxs, is responsible for the majority of the GSH-disulfide oxidoreductase activity in the cell [17], while unable to reduce RNR [16]. Grx3 has only 5 % of the Grx1 activity towards RNR [16]. In part, this can be explained by the more reducing redox potential of Grx1, -233 mV compared to -198 mV for Grx3 [18]. Grx1 and Grx3 catalyze GSH-mixed disulfides with a similar rate [19].

In mammals, there are two dithiol glutaredoxins [20]. Grx1 is mainly localized to the cytosol but has also been found in the nucleus [21]. Many of the mammalian Grxs are not active as dithiol reductants, although human Grx1 is able to reduce RNR but not other dithiol substrates [22]. As seen in paper III, human Grx1 was not able to interact with *E.coli* Grx1 in the protein-protein equilibration method of determining redox potentials. In addition to the two cysteines in the active site, hGrx1 contains three additional cysteines. The sequence alignments in paper III show that the N-terminal cysteine is only present occasionally in mammals while the two C-terminal ones are more conserved in the Grx1 clade. The function of these additional cysteines remains unclear. Human Grx2 is present in two alternatively spliced isoforms, one targeted to the nucleus and the other to the mitochondria [23; 24]. In contrast to other glutaredoxins, hGrx2 can be reduced by thioredoxin reductase, in addition to the standard glutathione, glutathione reductase and NADPH system [25]. Two additional cysteines are conserved in the Grx2 clade (paper III). They are located near the N- and C-terminus, respectively. A disulfide bond between Cys28 and Cys113 was shown to be an integral part of the structure of hGrx2. Subsequent structure determination confirmed the presence of this bond which ties the C-terminal helix to the interface of the central β-sheet and the C-terminal of the N-terminal helix [26].

GLUTAREDOXIN MECHANISMS

Two related mechanisms account for the observed activities of glutaredoxins [27] (Fig. 2). Both active site cysteine residues are required for the completion of the dithiol cycle, used for the reduction of ribonucleotide reductase. The nucleophilic N-terminal cysteine attacks the disulfide of the oxidized substrate, producing a covalent, mixed disulfide intermediate. The subsequent formation of an intramolecular disulfide bond occurs by the attack of the active site C-terminal cysteine. Two sequential reductions by glutathione reduce Grx back to dithiol form. In the monothiol mechanism, only the N-

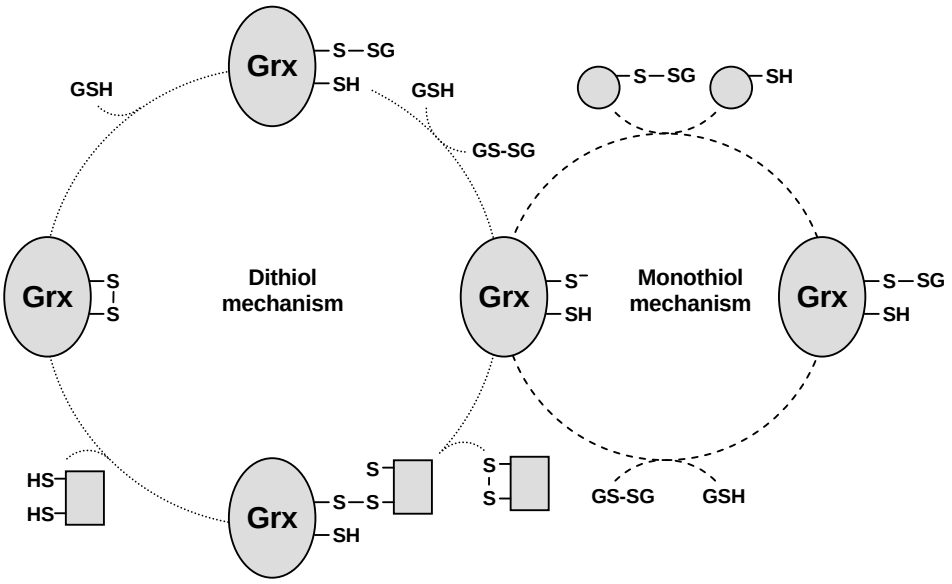

Figure 2. The glutaredoxin mono- and dithiol mechanisms.

terminal cysteine is required. Here, the thiolate specifically attacks the glutathione sulfur of a glutathione-mixed disulfide substrate, releasing the reduced substrate while forming a Grx-glutathione mixed disulfide complex (Grx-SG). A second GSH reduces Grx while the glutathione disulfide formed is subsequently reduced by glutaredoxin reductase, using NADPH as reducing equivalents.

REACTION RATES, REDOX POTENTIALS AND THIOLATE STABILIZATION

The thiol disulfide exchange reaction

$$A^{SH}_{SH} + B^{S}_{S} \underset{k_2}{\overset{k_1}{\rightleftharpoons}} A^{S}_{SH}\text{—}{}^{S}_{HS}B \underset{k_4}{\overset{k_3}{\rightleftharpoons}} A^{S}_{S} + B^{SH}_{SH}$$

is thought to be a S_N2 displacement reaction, similar to the reaction in organic thiols [28]. Here the nucleophilic and leaving group properties of the participating thiols largely determine the reaction rate. Using small thiols, an empirical equation relating the observed rate of the reaction to the pK_a values of the thiol groups has been established [28]. According to the equation, at a pK_a close to the pH of solvent, the dominant property of the attacking thiolate is its nucleophilic behavior. At a pK_a a bit lower than the pH of the solvent, it is instead the leaving group properties of the thiolate that is dominant for the outcome of the reaction rate. The pK_a values of the N-terminal cysteine in thioredoxin superfamily proteins are quite different. *E.coli* Trx has a pK_a of about 7.1 [29], close to the pH of a neutral solution. Therefore, it is optimal for a nucleophilic attack. The other members in the family have more perturbed pK_a values. The N-terminal cysteine in DsbA has a pK_a of around 3.5 [30] and in *E.coli* Grx3 it is below 5.5 [31; 32]. Hence, it is the leaving group abilities of these proteins that are determining the rate of the reaction.

The rates of the reactions determines the redox potential, since the overall equilibrium constant K_{12} is related to the individual rate constants k_1-k_4

$$K_1 = \frac{k_1}{k_2} \qquad K_2 = \frac{k_3}{k_4} \qquad K_{12} = K_1 \cdot K_2$$

Since the rates of the reaction depend on the pK_a value of the thiolate [28], the redox potential is also correlated to the pK_a of the N-terminal cysteine in the active site. The most reducing protein, Trx, has a redox potential of -270 mV [33] and a thiolate pK_a of 7.1 [29]. The most oxidizing protein, DsbA, has a redox potential of -124 mV [9] and a thiolate pK_a of 3.5 [30]. Grx3 has an intermediate redox potential of -233 mV [18] and a thiolate pK_a value that is less than 5.5 [31; 32]. Since the common part of the fold and the placement of the active site is more or less the same in the family, there must be other factors that contribute to the large differences in redox potentials. The dipeptide sequence between the two active site cysteines is different and characteristic within enzymes in the family. Many mutation experiments have shown that the dipeptide sequence is a key factor in determining the redox potential and the pK_a value of the N-terminal cysteine [30; 34; 35; 36; 37]. For example, when replacing the dipeptide in DsbA

(PH) with the dipeptide found in Trx (GP) or Grx (PY), both the pK_a values and the redox potentials obtained are shifted toward the values seen in donor proteins [37]. However, there are additional factors involved since the values are only shifted in the right direction. The two *E.coli* dithiol glutaredoxins 1 and 3 (Fig. 10) have similar folds and identical active site sequences. Still, the difference in redox potential is 35 mV [18]. Most likely, there are structural features on the side chain level and perhaps differences in dynamics that contribute to the activities seen.

As already pointed out several times, the pK_a value of the N-terminal cysteine residue in the active site of the enzymes in the thioredoxin superfamily is perturbed compared with a standard cysteine. Thus, there must be stabilizing factors around the thiolate. Being negatively charged, positive charges from side chains or the helix dipole have been proposed to contribute to the stabilization [38; 39]. Hydrogen bonds from neighboring amide protons or even from the thiol group of the C-terminal active site cysteine have further been proposed [31; 39; 40]. In Grx3, the results from structure determination, double cycle analysis, titrations and molecular dynamics simulations (papers I and IV) all point to the explanation that it is the hydrogen bonds donated by neighboring groups that contribute the most to the thiolate stabilization. Comparing to DsbA and Trx, it seems likely that the general theme is that the number of stabilizing hydrogen bonds largely determine the thiolate pK_a. There are four potential hydrogen bond donors in DsbA while only two in Trx.

In this thesis, the long term goal has been to obtain structural and biophysical information regarding *E.coli* Grx3 in order to

- compare the different redox forms and find structural determinants for the enzymatic mechanism
- quantify the interactions in the active site of the reduced form

In addition, different biophysical techniques have been applied to other members of the thioredoxin superfamily.

METHODS

The following sections briefly describe the most common methods and instrumental techniques used in the laboratory. The emphasis is on practical considerations rather than theory.

MUTAGENESIS

To this point, around 30 mutant proteins, based on six different thioredoxin superfamily proteins, have been used for projects described in this thesis. For this reason, mutagenesis has been an important technique in the laboratory. We have used the Quickchange Site-Directed Mutagenesis kit (Stratagene). The principle is quite simple. The plasmid template is the expression vector, containing the gene to be mutated. Mismatched primers are designed. These are made complimentary to the gene sequence on both sides of the point of mutation. The primers for both strands are made to overlap completely. Using thermal cycling, only the parental plasmid is used as the template for the replications, unlike in PCR. This is due to the complete overlap of the primers used for the opposite strands. At the end of the reaction, the parental DNA is digested by the Dpn I enzyme, specific to methylated and hemimethylated DNA. The mutant DNA is finally transformed into XL-1 Blue cells. The plasmids are subsequently sequenced and transformed into *E.coli* BL-21 expression cells.

PROTEIN EXPRESSION

Using the pET (Novagen) expression vectors in *E.coli* BL-21 (DE3), protein yields of up to 100 mg/L were obtained. While expressing unlabeled protein is generally straightforward, there are some considerations to keep in mind when expressing isotopically labeled proteins for NMR experiments. The spin ½ isotopes of carbon and nitrogen occur with a natural abundance of 1.1 and 0.37 %, respectively. For isotopic labeling, [^{13}C]-glucose and [^{15}N]-$(NH_4)_2SO_4$ are used as the sole sources of carbon and/or nitrogen. To make efficient use of the label, it is a good idea to obtain a sufficient mass of cells by growing in unlabeled media first. However, the cell metabolism is most likely very different when grown in LB compared to grown in minimal media (M9). Both uniformly labeled proteins and proteins with selective cysteine labeling have been produced by the following protocol [41]: Briefly, a cells mass is obtained after an over night growth in LB media. The cell pellet is washed with unlabeled M9 media and then suspended in labeled M9. The cells are allowed to adjust to the changed conditions for an hour before induction.

In a procedure developed for *E.coli* Grx1 [42], the combination of millimolar concentration of intracellular glutathione and the C14S mutant resulted in a Grx1-SG complex in which the protein and ligand were labeled to the same extent.
The technique was successfully applied to form the Grx3-SG complex. The question is how this complex is formed. Reduced glutaredoxins do not bind a single GSH molecule [43]. In addition, the C14S mutant used lacks the active site C-terminal cysteine, preventing the formation of an intramolecular disulfide bond. The only alternative left, assuming that the mono- and dithiol mechanisms are correct, is that the complex is formed by the reduction of a glutathione-mixed disulfide species. However, the

intracellular excess of GSH should subsequently reduce Grx-SG to reduced Grx3. The most reasonable explanation is that oxidative stress during cell lysis results in glutathionylation of various species in the cell [42]. Grx3, with its high affinity for glutathione-mixed disulfides, will form the Grx3-SG complex. The GSH pool is perhaps depleted under these conditions, shifting the equilibrium constant so that the Grx3-SG complex is favored.

PROTEIN PURIFICATION

Beginning with published protocols, considerable optimization has been performed for procedures used in this thesis. In general, the aim has been to use the same buffer throughout the lysis, purification and in the final experiments in order to minimize the time necessary to prepare samples. Most of the work presented here utilizes 50 mM phosphate pH 7.0. The first step in the purification of the thioredoxin superfamily proteins Grxs, Trx and DsbA is actually the lysis. It has been reported that many expressed proteins (up to 30 kDa) are exported from the cell when subjected to freeze-thaw cycles [44]. The mechanism for this is not clear, but likely involves selective poration of the *E. coli* plasma membrane. What is clear is that the cells are not lysed, and consequently the bulk of the proteins stay within the cells. The freeze-thaw method has been used successfully on Trx, Grxs and DsbA. Glutaredoxins, having the same fold and being smaller than Trx, probably is released by the same mechanism as Trx. DsbA, however, is about twice the size compared to Grx and Trx. On the other hand, it is exported to the periplasm of *E.coli*. Here freeze-thaw cycles may simply flush periplasmic proteins into the buffer through existing pores.

Following freeze-thaw treatment and collection of the cell supernatant by centrifugation, proteins are first separated using anion exchange chromatography. The DE52 resin (Whatman) contains the tertiary amine diethylaminoethyl, a weak ion exchanger. The excellent purification results obtained using the DE52 resin is due to the fact that Grxs, Trx and DsbA show very weak binding, eluting early in the gradient while the vast majority of the *E.coli* proteins bind much stronger.

The Grx peaks of interest are then concentrated using an Amicon stirred cell with a 3 kDa cutoff membrane (YM3, Millipore) and further purified by size exclusion chromatography with G-50 Fine (GE Healthcare). Lager proteins are eluted before the desired protein peak while mostly small nucleic acid fragments are eluted after. The impurities remaining in the protein sample are now such a small fraction that they are not seen on a coomassie-stained SDS-gel, in UV-spectra (DNA and RNA content) or in a standard ^{1}H-^{15}N HSQC NMR spectrum.

Following DE52 anion separation, thioredoxin and DsbA are further purified by hydrophobic interaction chromatography (HIC). A particular advantage of this technique as applied here is that no concentration of the eluted peaks is necessary after the DE52 anion step. Instead, the DsbA and Trx containing fractions are pooled and loaded onto the HIC column after addition of 0.7 M and 1 M ammonium sulfate, respectively. A negative gradient of ammonium sulfate is used to elute the proteins. Similar purity as for Grxs is obtained.

REDOX POTENTIAL DETERMINATION

The equilibrium constant K'_{12} of the overall thiol disulfide exchange reaction between two proteins A and B, or between a protein A and a substrate B, is related to the difference in redox potential between two molecules A and B by the Nernst equation

$$\mathrm{A}^{\mathrm{SH}}_{\mathrm{SH}} + \mathrm{B}^{\mathrm{S}}_{\mathrm{S}} \overset{K'_1}{\rightleftharpoons} \mathrm{A}^{\mathrm{S}}_{\mathrm{SH}}\text{—}{}^{\mathrm{S}}_{\mathrm{HS}}\mathrm{B} \overset{K'_2}{\rightleftharpoons} \mathrm{A}^{\mathrm{S}}_{\mathrm{S}} + \mathrm{B}^{\mathrm{SH}}_{\mathrm{SH}}$$

$$K'_1 \cdot K'_2 = K'_{12} = \frac{\left[\mathrm{A}^{\mathrm{S}}_{\mathrm{S}}\right]\left[\mathrm{B}^{\mathrm{SH}}_{\mathrm{SH}}\right]}{\left[\mathrm{A}^{\mathrm{SH}}_{\mathrm{SH}}\right]\left[\mathrm{B}^{\mathrm{S}}_{\mathrm{S}}\right]}$$

$$\Delta E^{\circ'}_{AB} = \frac{RT}{nF} \ln K'_{12}$$

where R is the molar gas constant, T is the temperature, F is the Faraday constant and n represents the two electrons that is transferred in the reaction.

To obtain the redox potential of a protein A, the redox potential of the B species must be known. The most common references are NADPH/$NADP^+$ ($E^{\circ'}$= -315 mV [45]) or GSH/GSSG ($E^{\circ'}$= -240 mV [46]) [47]. There are several approaches to determine the redox potential of proteins [18]. In paper II, two independent methods were used to determine the redox potential of the two human dithiol glutaredoxins. First, the redox potentials were determined with the use of redox buffers. When preparing accurate redox buffers, the ratio of buffer components should not be too far from unity, otherwise errors due to contaminating redox species, multiple dilutions and pipetting of small volumes can be significant. To avoid this problem, the choice of redox buffer therefore set the limit of the practical redox range. Since the redox values were completely unknown to us, an array between -160 and -260 mV was used. The redox potential for glutathione is -240 mV [46] and was therefore the obvious choice in the preparation of these buffers. There was an indication of a second disulfide in hGrx2 (paper II), out of reach of the GSH/GSSG redox buffer range. Therefore, a more reducing redox buffer with equal amounts of reduced and oxidized DTT ($E^{\circ'}$= -312 mV [46]) had to be used. During the approach to redox equilibration, it is important that oxygen and metals are kept out from the samples to ensure no other sources of oxidation or reduction equivalents are present. The use of degassed buffers, the metal chelator EDTA and nitrogen atmosphere are essential to achieve this purpose. Quenching the reaction is also a very important step. It was found that phosphoric acid works well for this purpose [18]. Determination of the redox potentials of the glutaredoxins requires an analytical method allowing the quantification of the individual redox forms once equilibrium is formed. Reversed-phase HPLC has proven to be a powerful technique capable of separating the different active site redox forms of diverse members of the thioredoxin family [18]. The peaks corresponding to the reduced and oxidized forms of the protein are quantified by integration and application of Beer's law. The values are used to determine K_{12} and the redox potential through the Nernst equation.

The second method used to determine redox potential was the so-called direct protein-protein redox equilibrium [18]. The principle is that two redox active proteins are allowed to equilibrate with each other and the resulting oxidized and reduced peaks of the proteins are quantified. To establish an absolute redox potential of an unknown, the standard state redox potential of the partner must be known. Practically, the redox potential of the partner protein should be similar to the one under investigation. If the redox potentials of the interacting proteins are very different, the equilibrium is shifted far in one direction resulting in that one protein will be almost completely reduced while the other will be almost completely oxidized. Hence, obtaining an accurate quantification of all four redox species by integration will be difficult.

OPTICAL SPECTROSCOPY

When light of a particular wavelength is directed to a sample, part of the light can be absorbed by the chromophores in the sample [48]. The physical process involved in absorption is that in a chromophore, electrons are excited from the ground electronic state to a higher electronic state. The electrons return through various vibrational energy levels to the ground electronic state rapidly by loosing the acquired energy as heat when colliding with neighboring molecules. The relationship between light intensity (I) and chromophore concentration (C) forms the familiar definition of absorption (A) referred to as the Lambert-Beer law:

$$\log\left(\frac{I_0}{I}\right) = \varepsilon \cdot C \cdot l = A$$

where I_0 is the intensity of the incoming light, I is the intensity of the non-interacting light passing through the sample, ε is the extinction coefficient in units of $M^{-1}cm^{-1}$, C is the concentration (M), and *l* is the path length of the sample (cm).

A peak in an absorption spectrum can be described by the wavelength, λ_{max} , at which the absorption is at the maximum, and the extinction coefficient at that point, ε_{max}. The absorbance of a chromophore is not only determined by its chemical structure but also by the environment around it. Solvent polarity and solutes are two such environmental factors. Detergents and chaotropic agents like urea or guanidine hydrochloride unfold native proteins, thus causing buried chromophores to be exposed to a significantly more polar solvent. Another common source of local environmental change is the electron redistribution that follows protonation/deprotonation reactions, such as when the pH is changed or in reduction/oxidation. The effect is often a change in λ_{max} and/or ε_{max}. These effects can be used as local probes to follow change, for example protein unfolding or functional group titration. Usually, the wavelength of choice is the one that changes the most.

UV spectroscopy

Ultraviolet (UV) spectroscopy is a common analytical technique in chemical research. UV spectroscopy is particularly central to biochemistry as most molecules in the cellular chemistry contain chromophores in this wavelength region. For proteins, the absorption maxima of the aromatic residues Trp, Tyr and Phe fall conveniently are in the UV wavelength range (260-290 nm) separate from that of commonly used buffers. Since most proteins contain one or more of these residues, UV spectroscopy is often

used to quantify protein concentrations. In addition, UV wavelength spectra are helpful as a routine check for DNA/RNA content in a protein purification process. Since these impurities are not seen on a standard coomassie stained SDS gel, their presence can be verified by a peak at λ_{max} 260 nm.

In addition to the aromatic residues above, cysteine has been shown to have a pH dependant absorption at 230-240 nm [49]. The change in absorption between the thiol group and the thiolate group is about 4000 M^{-1} cm^{-1}, the thiolate being the stronger chromophore (Fig. 3). This phenomenon has been used to explore the pK_a values of members in the thioredoxin superfamily [29; 30]. Practically, there are several considerations. Since the pK_a values are often perturbed in the active sites of these proteins, it is important to know when the protein undergoes pH-induced unfolding. Therefore, an identical titration experiment should be set up using, for example, CD spectroscopy which can provide information on the folded state of the protein. In Grx3, the Cys11 thiolate becomes protonated when the protein starts to unfold, or, the protein unfolds when the thiolate becomes protonated.

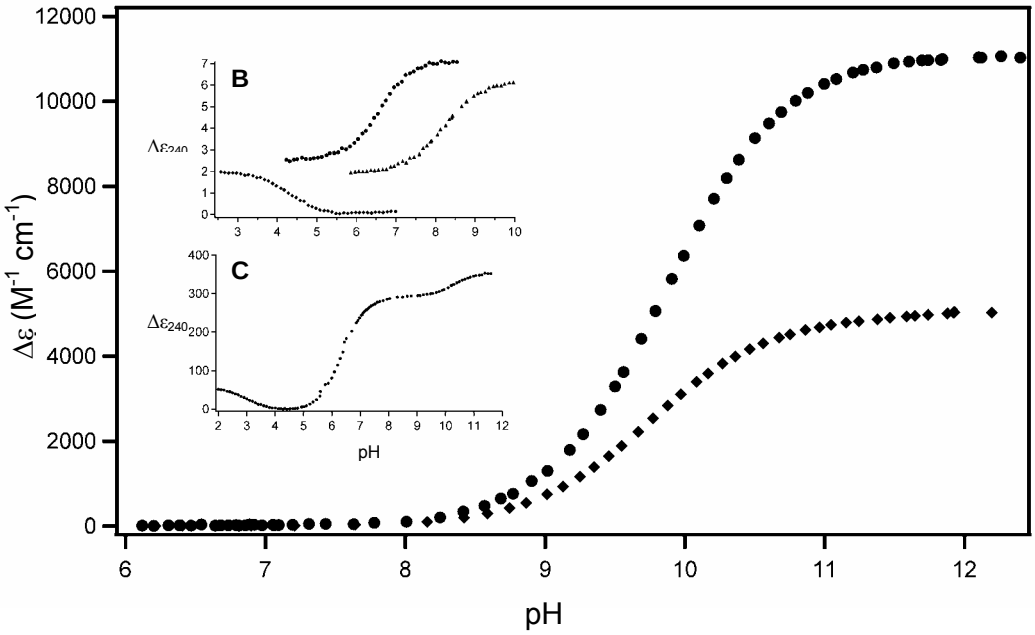

Figure 3. pH titrations of tyrosine and cysteine. UV-spectroscopy at 240 nm is used to monitor the signals. pH titrations of glutamate, histidine and arginine (inset B), pH titration of EDTA (inset C).

A complicating factor when monitoring thiol absorbance is that tyrosine also displays a pH dependant absorption. The λ_{max} is at 275 nm but the tail of the peak is still significant at 240 nm (Fig 3). The difference between protonated and unprotonated Tyr is about 11000 M^{-1} cm^{-1}. The pK_a values of a model Tyr and a model Cys amino acid are almost identical (Fig 3). Therefore, it would be difficult to follow the thiolate-thiol signal in the presence of one or several Tyr residues, unless the pK_a of the thiolate is shifted or the pK_a values of the Tyr residues are known previously.

The thioredoxin superfamily proteins are oxidoreductases, able to have their active site cysteines in both dithiol and disulfide form. The experimental titration of the thiolate

group must therefore be performed under conditions that prevent oxidation of the active site. In addition to using degassed buffer and keeping a nitrogen atmosphere in the UV spectrometer sample compartment, the metal chelating agent EDTA is often added to the sample to reduce metal ion catalyzed oxidation. However, care must taken to insure that the ratio between protein and EDTA is not too low. EDTA contains two tertiary amine groups and four acetate groups. Monitoring the titration of EDTA at 240 nm shows that the molar change in absorption is about 10 % of the molar thiol/thiolate change (Fig. 3, inset C). The pH-dependant change in absorption of other side chains in this region is comparably small. For example, the change in absorption of: a model histidine is 5 M^{-1} cm^{-1}; a model arginine is 4 M^{-1} cm^{-1}; and a model glutamate is 2 M^{-1} cm^{-1} (Fig. 3, inset B).

CD spectroscopy

The high intensity light from a xenon lamp is made circular polarized in a circular dichroism instrument [48]. The electronic vector of linear polarized light can be described as a sinus wave confined to oscillate in a plane. Looking at this wave from one end, the magnitude will change but the direction will be restricted to this plane. Linear polarized light can be considered as a sum of left- and right-circular polarized light with equal magnitudes and opposite rotation velocities. In circular polarized light (left or right), the magnitude of the oscillation is constant while the direction changes in a right- or left-handed manner. The result of one such wave is a helical path traced out by the constant magnitude electronic vector along the propagation of the wave. When right- and left-circular polarized light of equal magnitude (linear polarized light) passes through an optically active sample, the resulting light is still circular polarized but the

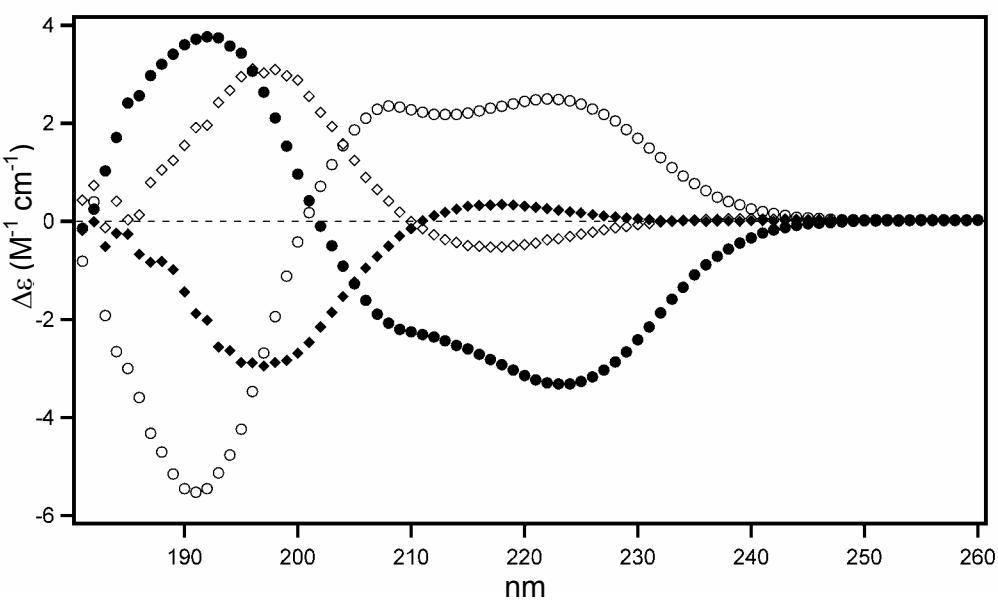

Figure 4. CD spectra of poly-glutamic acid. (●) L-Glu pH 3.3, (○) D-Glu pH 3.3, (♦) L-Glu pH 11.2, (◊) D-Glu pH 11.2. At pH 3.3, poly-Glu is helical and at pH 11.2 it is unstructured.

magnitudes of the left- and right components are no longer equal. Consequently, elliptically polarized light can result from the interaction of linear polarized light with an optically active sample.Two basic criteria must be fulfilled for a sample in order to be useable in a CD experiment. First, the molecule must contain a chromophore that absorbs light a region accessible by the instrument (UV-Vis wavelength range). Hence, if the molecule does

not absorb light, it cannot display a CD effect. Second, the molecule must either contain an optically active chromophore or a symmetric chromophore in an anisotropic environment. Amino acids (except glycin) are chiral due to the α-carbon and the proximity to the amide chromophore results in a CD signal. Aromatic side chains are symmetrical due in part to the free rotation about χ1 and χ2, but can generate CD signals if they are prevented from free rotation in an unsymmetrical environment, for example buried in a protein.

CD spectroscopy is very useful to estimate the gross structural features of a protein without the time consuming procedures involved when obtaining structures using X-ray crystallography or NMR spectroscopy. Model compounds of "pure" helices, sheets and turns can be used to guide an unknown structure (Fig. 4). There are also databases of CD spectra from proteins of known structure from which it is possible to calculate the percentage helices, sheets and others that are present in the protein under investigation.

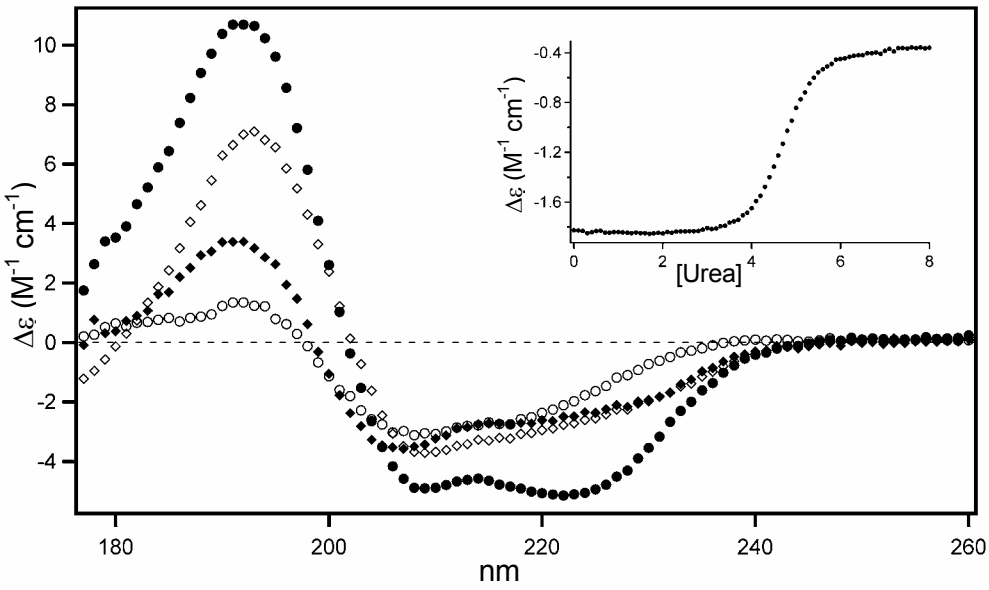

Figure 5. CD spectra of (•) Myoglobin, (◊) DsbA, (♦) Lysozyme, (○) RNase. Urea-induced unfolding of E.coli thioredoxin (inset), monitored at 222 nm.

Information regarding the composition of secondary elements can be obtained from a 180-260 nm CD spectrum (Fig.5). In paper II, analysis of CD spectra recorded of the full length and the N- and C-terminal parts of human spermatid-specific thioredoxin 1 showed that the protein consist of a unstructured N-terminal domain and a thioredoxin-like C-terminal domain. Tertiary structure information can be obtained from analysis of

the aromatic region at 250-300 nm. However, since there are only a few aromatic residues in a protein compared to amide chromophores, the signals are very weak in comparison. Although CD can be used to obtain structure information, the most powerful use of this technique is the great ability to measure conformational change. As stated above, CD spectroscopy can be used in parallel with UV spectroscopy to determine if the protein undergoes a pH-induced unfolding. Another application is the determination of protein conformational stability using urea- or guanidine hydrochloride-induced unfolding (Fig 5, inset). The native protein is subjected to increasing concentrations of denaturant and the CD signal is measured at 222 nm. An important experimental consideration is that even ultrapure urea must be further purified. Since the urea concentration is often increased to over 9 M in a protein unfolding experiment, even small amounts of UV-absorbing contaminants will effectively prohibit that any light reach the detector. For this purpose, 5 % (w/v) AG 501-X8(D) mixed resin (Biorad) was added to the stock solution of urea [50]. The solution was allowed to mix for an hour before filtration.

Analysis of solute unfolding experiments requires the assumption of a reversible, two-state equilibrium between folded and unfolded protein. The following equation describes the dependence of the raw signal, y, on the denaturant concentration [Urea]:

$$y = \frac{\left(y_F + m_F[Urea]\right) + \left(y_U + m_U[Urea]\right) \cdot \exp\left(\frac{m_{trans}[Urea] + \Delta G^{unfolding}}{RT}\right)}{1 + \exp\left(\frac{m_{trans}[Urea] + \Delta G^{unfolding}}{RT}\right)}$$

where y_F and y_U are the CD signals of the folded and unfolded protein at 0 M urea, respectively, m_F and m_U are the urea dependence of the protein CD signals of the pre- and post-transition regions, for folded and unfolded protein, respectively, R is the gas constant (cal K^{-1} mol^{-1}), T is the temperature (K), m_{trans} is the slope of the linear dependence of the Gibbs free energy as a function of urea concentration within the transition region, and $\Delta G^{unfolding}$ is the Gibbs free energy (kcal mol^{-1}) of the protein unfolding reaction. Using these definitions, the conformational stability of a globular protein is a negative number and the $\Delta G^{unfolding}$ is then the negative of this value.

Comparing the conformational stability of mutants is one way to determine if and how strong two residues interact in a protein. This forms the basis of the double mutant cycle analysis [51] and is shown in figure 6. Four proteins are needed. The first protein, P, contains both residues whose interaction is to be tested. The second and third proteins are single mutations of the first (PA and PB), so that one of each residue is replaced, usually by an alanine. The fourth protein is a double mutant (PAB), where both residues have been replaced by alanine. The interaction energy, $\Delta\Delta G$, can be determined by the equation

$$\Delta\Delta G = (\Delta G_{PB} - \Delta G_{PAB}) - (\Delta G_P - \Delta G_{PA}) = (\Delta G_{PA} - \Delta G_{PAB}) - (\Delta G_P - \Delta G_{PB})$$

In this thesis, the method has been applied to probe a number of different interactions in thioredoxin superfamily proteins.

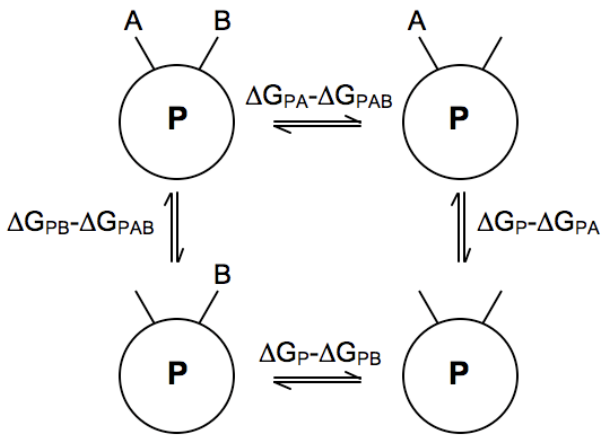

Figure 6. The four proteins needed for a double mutant cycle analysis. The cycle quantifies the interaction between residues A and B in protein P.

NMR SPECTROSCOPY

The atomic nucleus has four important physical properties: mass, electric charge, magnetism and spin [52]. Spin is a form of angular momentum, but it is not produced by rotation. It is an intrinsic property. The spin angular momentum **S** is proportional to the magnetic moment **μ**

$$\boldsymbol{\mu} = \gamma \mathbf{S}$$

where γ is the magnetogyric ratio, characteristic of the isotope.
When a nucleus is subjected to an external magnetic field, due to the fact that it possesses both angular momentum and a magnetic moment, it will be forced into a motion called precession. The frequency of procession is characteristic of the isotope and proportional to the external magnetic field B^0

$$\omega^0 = -\gamma B^0$$

where ω^0 is called the Larmor frequency. Nuclei of a given isotope do not have uniformly constant frequencies as predicted by this equation due to the local magnetic perturbations at the nucleus as a result small local magnetic field differences.
In a sample, due to the combination of the external magnetic field and the fluctuating microscopic magnetic field from electrons and nuclei, a very small net distribution of magnetic moments along the external field is produced. By applying a radio-frequency pulse, the spins can be rotated perpendicular to the external magnetic field and be detected.

Structure determination

A protein structure can be defined by the network of distances between the proton atoms. The process of obtaining a structure by NMR spectroscopy is the collecting of distance constraints between pairs of ^1H nuclei which have been unambiguously

assigned to atoms positions in the protein. The more distance constraints used in a structure calculation, the better the precision of the resulting structure ensemble.
The sequence specific assignment of a protein is obtained in a stepwise manner. The ability to incorporate the spin ½ isotopes ^{13}C and ^{15}N has made it possible to use multiple dimensions to resolve the abundant and often overlapped proton peaks (Fig. 7). In heteronuclear experiments, the magnetization is usually transferred to the protons before detection. The gain in sensitivity is proportional to the ratio of the magnetogyric ratios, and even more if the magnetization also starts on a proton [53].

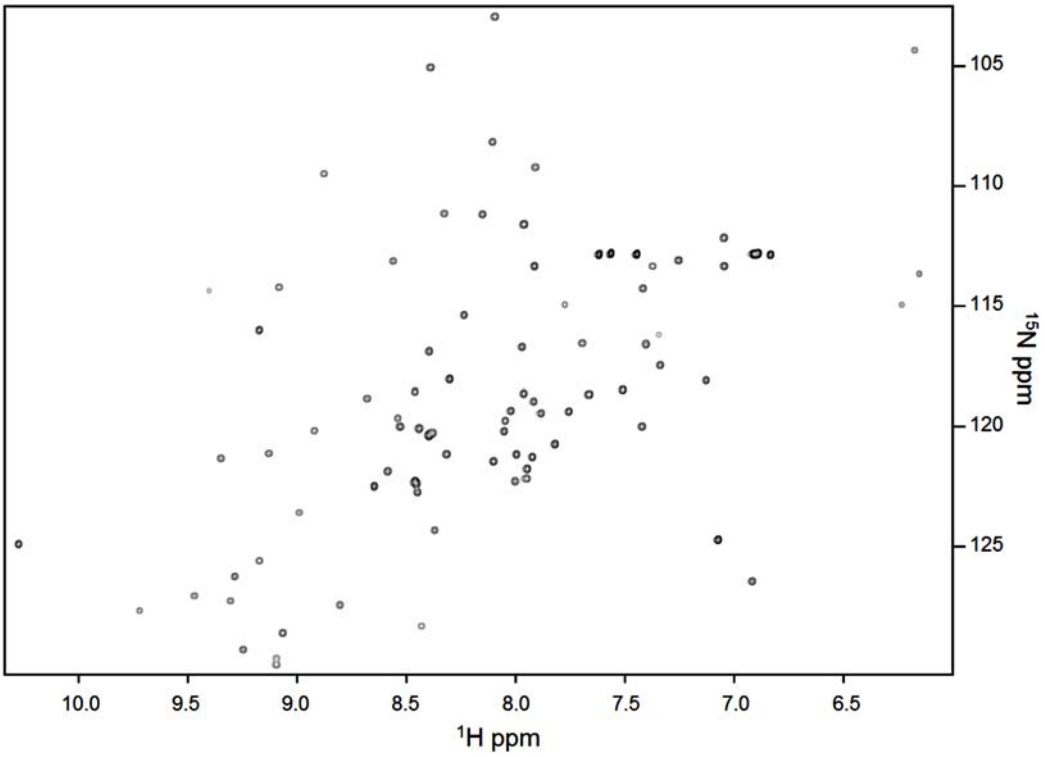

Figure 7. ^{1}H-^{15}N HSQC spectrum of reduced Grx3 recorded on a 800 MHz NMR spectrometer.

In general, the backbone atoms are assigned first. For a small protein such as Grx3, the HNCA and HN(CO)CA combination provide enough information for unambiguous assignment. HNCA correlates the amide of residue (i) with both the intra (i) and the previous (i-1) alpha carbon. In short, the connections

HN_i - N_i – C^{α}_i and HN_i - N_i – C^{α}_{i-1} are seen.

The HN(CO)CA experiment provides information about the above inter-residue correlation only. By combining data from the two spectra, the amino acid chain can be assigned by stepping through the sequence. The process is interrupted at proline residues, which lack the amide proton. In addition, some amides may be missing due to fast solvent exchange. The missing data will have to come from other spectra.
The next step is generally the assignment of the α and β protons. The HBHA(CO)NH experiment correlates the α and β protons of a given amino acid with the amide of the

following residue *via* the carbonyl carbon. The assignment is based on the obtained assignment of the amides in the previous step.

The side chain atoms are assigned using a combination of experiments. The H(CCO)NH experiment (also called HCCH-TOCSY-NNH) correlates the aliphatic protons of an amino acid with the amide of the following residue. The C(CO)NH experiment (also called CCH-TOCSY-NNH) correlates the aliphatic carbons of an amino acid with the amide of the following residue. The HCCH-TOCSY experiment correlates the aliphatic protons of an amino acid with the aliphatic carbons in the same residue. Finally, specific experiments for aromatic side chain, such as CB(CGCD)HD and CB(CGCDCE)HE, are assigned.

Distance constraints are obtained from the integration of assigned NOESY spectra. The Nuclear Overhauser Effect (NOE) is a through space dipole-dipole coupling. In the linear regime, the intensity of a NOE cross-peak is proportional to the inverse of the sixth power of the distance between two protons. Once calibrated using known distances and intensities, the obtained distances are used as upper distance limits in the structure calculation. For a protein of the size of Grx3, a combination of a 2D ^{1}H-^{1}H NOESY and ^{1}H-^{13}C and/or ^{1}H-^{15}N 3D HSQC-NOESY is preferred. The 2D NOESY will contain much better ^{1}H-^{1}H resolution than the 3D NOESY, while the use of a third dimension will help to resolve the overlap seen in the two-dimensional spectrum.

The manual assignment of the NOESY spectra has previously been a time consuming and potentially biased procedure. Due to the number of protons, even in a small protein, many, if not the majority, of the cross-peaks can be assigned to more than one pair of protons. Structure calculation in the CYANA program [54] performs an automatic NOESY assignment, allowing ambiguous assignments. The peaks are ranked and filtered out using several empirical criteria. Two conditions that must be fulfilled include the peak position relative to chemical shift lists and the participation in a network of other assignments. The calculation utilizes torsion angle dynamics. The iterative process is executed in seven cycles (Fig. 8). The obtained structure in one cycle is used to guide the NOE assignment in the next cycle. All NOE cross peaks are

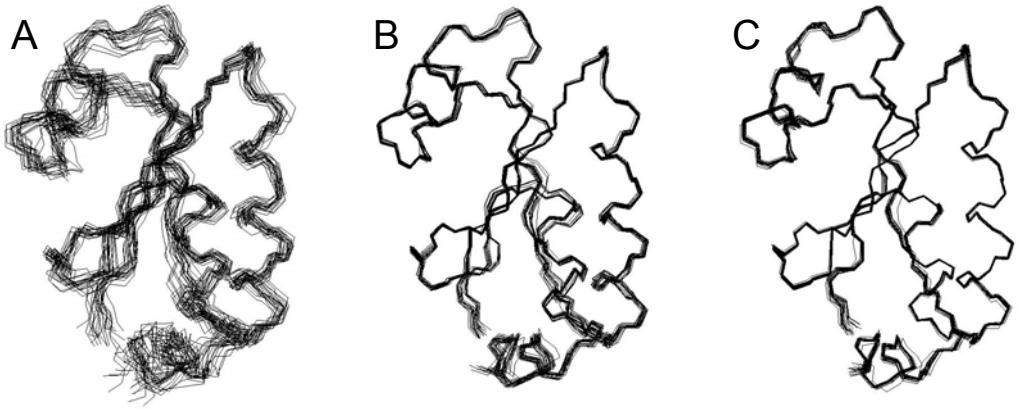

Figure 8. The NMR structure ensemble after the first CYANA cycle (A), the third cycle (B) and the final cycle (C).

assigned to a single pair of hydrogen atoms in the final cycle. The 20 conformers with the lowest target function are subjected to restrained energy minimization in explicit solvent using the AMBER force field [55] using the program Opalp [56].

Titrations

NMR is an excellent tool for the determination of pK_a values and even tautomer forms of amino acids such as histidine side chains in proteins [57]. An example of the power of NMR in this respect is the pK_a determination of the side chain of a model histidine (Fig 9). In practice a ^{15}N-inriched sample is required for histidines in proteins as the imidazole proton peaks are not often resolved. The chemical shifts of the protonated and unprotonated nitrogen atoms in the histidine ring are in the region of 170 and 250 ppm, respectively, making them easily identifiable. Combined with coupling constant data [57], three different peak patterns are possible, one for each neutral tautomer and one for the positively charged form. Hence, one single ^{15}N -HSQC at an appropriate pH can reveal the tautomer form by the pattern of the imidazole peaks. During a titration, the proton and nitrogen chemical shifts are recorded as a function of pH. The pK_a value is obtained by non-linear regression analysis of the titration curve, using the equation

$$\delta = \delta_{HA} - \frac{\delta_{HA} - \delta_A}{1 + 10^{\alpha(pK_a - pH)}}$$

where δ is the observed chemical shift, δ_{HA} is the chemical shift of the protonated form, δ_A is the chemical shift of the deprotonated form. α is the Hill coefficient, reflecting the cooperativity, interpreted as the minimum number of protons involved in the titration.

In addition to histidines, cysteines side chains have also been titrated in this thesis work using NMR. In contrast to histidine, the cysteine side chain does not contain any atoms with overtly unique chemical shifts. To overcome this, specific labeling can be of great help. Both [^{13}Cβ]- and [U^{13}C]- cysteine have been used successfully. The H^{β} and C^{β} chemical shifts are recorded as function of pH. The pK_a is obtained as described above for histidines.

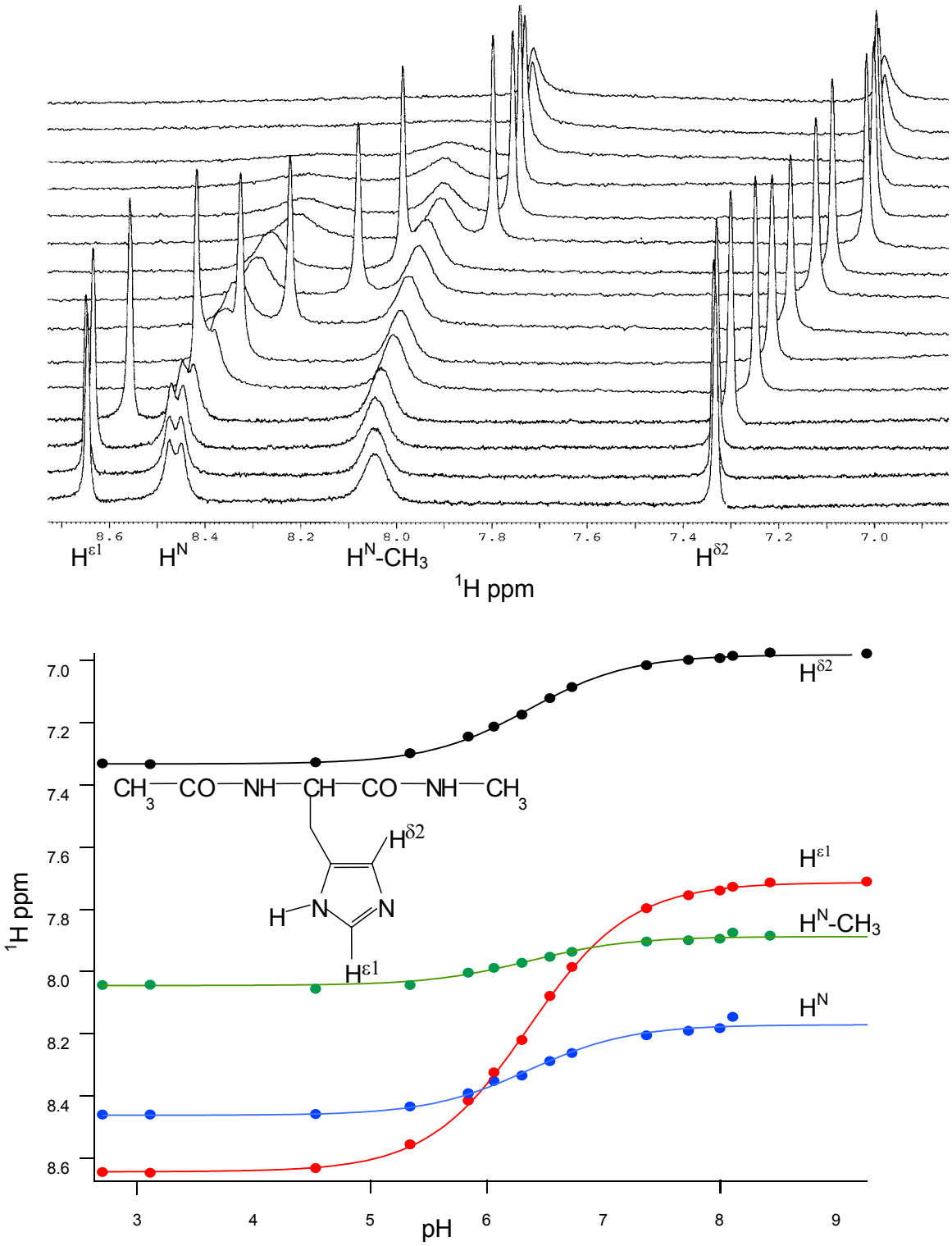

Figure 9. NMR titration of a model histidine. 1D ^1H NMR spectra at different pH (increasing pH from bottom to top)(A). The data were fit simultaneously to a single ionization (pK_a = 6.36) using non-linear regression (B).

RESULTS

Paper I: Structure, dynamics and electrostatics of the active site of glutaredoxin 3 from Escherichia coli: comparison with functionally related proteins

The active site of *E.coli* Grx3 was studied using theoretical and experimental methods. Molecular dynamics simulations started out with two different structures taken from the published NMR ensemble of the glutathione-bound form of Grx3 (PDB ID 3GRX) [58]. The structures were reduced *in silico* before being subjected to 10 ns molecular dynamics (MD) simulation in explicit water using the CHARMM force field. The result showed that the two starting structures had largely converged to an arrangement where the thiolate of the N-terminal cysteine in the active site, Cys11, interacted through hydrogen bonds with the amide protons of Tyr13 and Cys14 and the thiol group of Cys14. NMR coupling constant data showed that the torsion angles of Cys14 in solution were in the same rotomer position as those in the converged simulation providing direct experimental support for the MD result.

MD snapshots every 200 ps from the simulations were used to calculate pK_a values using the UHBD program [59]. Experiments had shown that the pK_a values for Cys11 and Cys14 were below 5.5 and above 10.5, respectively. Using a low dielectric constant of 3.0, it was possible to reproduce those values. Furthermore, estimates of the contributions from surrounding groups and the helix dipole showed that the local environment rather than long range interactions were most responsible for thiolate stabilization.

Histidine 15 is located near the active site. If positively charged, it could contribute to the stabilization of the Cys11 thiolate. The histidine titration was performed by NMR using a series ^{1}H-^{15}N HMQC spectra at varying pH. The pulse sequence delay was set to 22 ms in order to minimize one-bond and maximize two- and three-bond couplings to enhance the intensity of the side chain resonances. It was found that His15 had a pK_a of 6.0 and was in the ε-tautomer form. Hence, the thiolate is not stabilized by His15 electrostatically at neutral pH. The other histidine in the protein, His61, was also found to be in the ε-form and had a pK_a of 5.1. The result for this residue was expected since the $H^{\varepsilon 2}$ proton was observed to hydrogen bond to O' in Gln54.

The impact of neighboring groups on the pK_a value of Cys11 was tested experimentally by making mutants and running cysteine titrations. For obvious reasons, it is impossible to mutate the amide protons of Tyr13 and Cys14. Beside WT (C65Y), the C14A/C65Y and K8A/C65Y mutants were made. The thiol group of Cys14 is expected to interact with Cys11 thiolate though a hydrogen bond. Lys8 is positively charged and is located close to the active site. The pH titration of Cys11 was monitored by UV-spectroscopy at 240 nm. The same setup was also utilized in a second set of titrations, checking pH-induced unfolding at 222 nm in a CD instrument. The result showed that Cys11 titrate when the protein unfolds, or, that the proteins unfold when Cys11 becomes protonated.

It is clear however, that Lys8 has only a marginal effect on the thiolate pK_a compared to WT, while a change in Cys14 shifts the pK_a much more.

Paper II: Human spermatid-specific thioredoxin-1 (Sptrx-1) is a two-domain protein with oxidizing activity

Sptrx1 is a protein that is expressed only in sperm during maturation. Attempts to determine the structure and oligomer state of this molecule using different techniques had been made. Gel filtration had shown that the protein behaved larger than expected and was therefore thought to be either an oligomer or very unsymmetrical monomer. Mass spectrometry provided evidence that at least in the electrospray, Sptrx 1 was a monomer. Crystallization trials produced crystals after one year, but subsequent analysis showed that the N-terminal part had been cleaved off and that only the C-terminal part formed crystals. To help resolve this situation, the two parts that were cleaved from each other were expressed and purified separately.

CD spectra of the full length protein and the two parts were collected. It was clear from initial spectra of the N-terminal domain that the presence of a large negative peak around 190 nm that the N-terminal domain was largely unstructured. In contrast, CD spectra from the C-terminal domain was that of a folded protein and compared to the CD spectrum of human Trx1, and was found to be quite similar and very characteristic of proteins of this family. The signals from the two separate domains were combined in molar ratios and compared to the spectrum of the full length protein. The very close coincidence of the summed spectra to that of the full length protein was taken as evidence that the N-terminal domain is as unstructured in the full length protein as it is separately in solution. Similarly, the folded, Trx-like C-terminal domain was similarly structured in the full length protein as it was expressed as a single domain. This result provides an explanation why crystals were only observed for the C-terminal domain. During the year-long equilibration, the presence of minute amounts of proteases remaining in the sample digested the largely unstructured N-terminal domain, leaving the folded C-terminal domain intact to crystallize.

Paper III: Redox properties and evolution of human glutaredoxins

There are two human glutaredoxins with a dithiol active site. The mainly cytosolic hGrx1 has the standard glutaredoxin active site sequence CPYC while hGrx2, localized in the mitochondria and the nucleus, has the sequence CSYC. The redox potentials were determined using three independent methods. First, since the redox potentials were unknown, an array of glutathione redox buffers was made by combining oxidized and reduced glutathione in suitable amounts to make buffers in the -160 to -260 mV range. The redox potential obtained after equilibration, quenching and HPLC analysis were -232 mV for hGrx1 and -221 mV for hGrx2. To verify the result, a second method was used to determine the redox potential, the so-called protein-protein equilibration method developed in the laboratory [18]. Here, two dithiol active site proteins are allowed to react with each other until equilibrium is reached and then the amounts of each in the oxidized and reduced forms are quantified as in the abovementioned redox buffer

method. Here a toolbox of different proteins of known redox potential is a great advantage. The two human Grxs had identical retention times on the HPLC, preventing the use of them together as redox partners. Therefore, *E.coli* Grx1 was used as a partner protein. The redox potential obtained for hGrx2 was -222 mV. However, the analysis of hGrx1 showed that the initial concentrations of hGrx1 and *E.coli* Grx1 in the different redox states were unchanged after 4 and 24 hours incubation. This specificity has not been reported before. The confirmation of the redox potential for hGrx1 had to come from the thermodynamic cycle method.

The initial studies of hGrx2 with different reducing and oxidizing buffers had shown that three species exist. Two of these were seen in the glutathione buffers above while the third one appeared after DTT treatment. Therefore, a DTT redox buffer of -312 mV was prepared and hGrx2 was allowed to equilibrate with it. Separation and integration of the HPLC peaks resulted in the determination of a redox potential of -317 mV for this second redox pair in hGrx2. Apart from the two active site cysteines, two additional cysteines are present in the protein, Cys28 and Cys113. The strongly reducing redox potential indicated that under all but the most reducing conditions this pair would be in the disulfide redox state – that is to say it would be natively a structural disulfide bond. To further confirm the existence of a structural disulfide bond between Cys28 and Cys113 in hGrx2, a cleavage experiment was designed. There are three methionine residues, located in such a way that chemical digestion with CNBr would produce different fragments depending on if there is a disulfide bond or not present. The digested sample was divided into two samples, one of which was reduced. The electrophoresis gel showed the expected pattern. Mass spectrometry was used to confirm the masses of these fragments. The N- and C-terminus was linked through the disulfide bond while two fragments between them had been cut away which could only have been so if Cys28 and Cys113 were disulfide bonded.

Phylogenetic analysis was used to obtain a more complete picture of the evolution of the human glutaredoxins. Interestingly, while the monothiol hGrx5 was found to be part of a very distinct group, with homologs from bacteria to mammals, hGrx1 and hGrx2 were found to be part of two separate clades with a common ancestor and diverged early in animal evolution.

Paper IV: NMR structure of the reduced form *E.coli* glutaredoxin 3: relation to other redox forms and mechanistic implications

Glutaredoxin 3 is an abundant *Escherichia coli* protein of 9 kDa with a high specificity for glutathionylated mixed disulfides. Results from paper I provided evidence that a network of hydrogen bonds stabilize the active site thiolate. However, experimental data was needed to further address the redox properties and explore the structural differences compared to the oxidized and glutathione-mixed disulfide forms of the protein.

The structure of reduced Grx3 is well defined with a RMSD for the N, C^{α} and C' atoms of the backbone of 0.26 Å and 0.60 Å for the heavy atoms. The topology of the structure has the expected $\alpha\beta\alpha$ sandwich, characteristic of members of the Trx superfamily. The active site cleft is formed by residues in the N-terminus of helix 1, the

cis-Pro loop and the N-terminus of helix 3. The side chain of Cys11 is in two conformations in the ensemble. When $\chi 1$ is *trans*, the thiolate participates in a hydrogen bond network with the neighboring amide protons of Tyr13 and Cys14 and the thiol proton of Cys14. In the *gauche*- rotamer form, Cys11 form a hydrogen bond with the hydroxyl group of Thr10. Whether these two conformations of Cys11 were due to the extensive overlap of the some of the NOE cross peaks or reflected a real dynamic behavior was impossible to know from the structure alone.

In order to gain more insight to the active site, interaction energies ($\Delta\Delta G$) between selected pairs of amino acids were measured using double mutant cycle analysis. The stabilities of the proteins were measured by urea-induced unfolding as monitored by CD spectroscopy. A value of -1.63 kcal/mol for the Cys11-Cys14 pair indicating that there is a favorable interaction between the Cys11 thiolate and the Cys14 thiol. Furthermore, a larger interaction energy was found on replacement of the Cys14 thiol by a Ser14 hydroxyl, confirming a stronger hydrogen bond to Cys11. This is expected due to the stronger electronegativity of oxygen. Finally, the quantification of the Cys11-Thr10 interaction showed that it was unfavorable, +0.68 kcal/mol.

To understand more about the mechanisms that Grx3 uses to reduce different substrates, well defined structures for comparisons are needed to make detailed comparisons. Although structures of oxidized and glutathione-mixed disulfides have been determined previously, new methodology regarding structure calculations prompted us to re-examine these structures. While the oxidized structure was re-calculated using the original data, additional NMR spectra were used for the structure of Grx3-SG. For this purpose a [U-^{13}C;U-^{15}N]- labeled protein-glutathione complex was prepared.

Having three redox states of the protein, a detailed comparison could be made. Superimposing the backbones, it was clear that only small changes occur upon binding or oxidation. Focusing on the active site, the binding cleft contain a mixture of charged and apolar residues that complement glutathione well. Two hydrogen bonds are centrally placed, both are between Cys_{GS} and Val52. While the oxidized form is very similar to Grx3-SG at the backbone of residues 51-53, the reduced form showed a significant displacement, resulting in alternate positions of the amide and carboxyl groups of Thr51 and Val52. The most apparent difference in the structures is the position of the Tyr13 ring. In Grx3-SG, it is situated at the protein surface, interacting with the bound GSH, while it is directed toward the solvent in the other two structures. Plausibly, the Tyr13 ring acts as a landmark for an incoming glutathione moiety.

The thiolate is shielded from the sides by neighboring residues, possibly discriminating against non-glutathione substrates. A change in exposure is seen upon GSH binding. In combination of the movement of the Tyr13 ring toward the protein surface, this creates better access for a second glutathione to reduce the complex. The product, oxidized glutathione, seems to be accompanied by the movement of the Tyr13 ring away from the surface.

FUTURE PERSPECTIVES

DsbA and Thioredoxin

The redox potentials of the members in the thioredoxin superfamily span a large range. At opposite ends are the oxidizing DsbA with -124 mV [60] and the reducing thioredoxin (Trx) with -270 mV [33]. The pK_a of the N-terminal cysteine in the CXXC active site has been shown to be correlated with the redox potential [30; 34; 35; 36; 37]. In DsbA, the pK_a value is around 3.5 [30; 61] while it is around 7.1 [29] in Trx. Through mutagenesis of the dipeptide between the cysteines, it is clear that the sequence has an effect on the N-terminal cysteine pK_a value. Structure analysis and MD simulation has provided important evidence that the number of hydrogen bonds that stabilize the thiolate is the main contributor to the observed perturbed values. The DsbA active site sequence is CPHC. The crystal structure of reduced DsbA indicates that there are four hydrogen bond donors to the thiolate acceptor [39]. In addition to the ones seen in Grx3, a fourth one is donated by the $H^{\delta 2}$ proton in the histidine ring. Furthermore, the ring is assumed to be positively charged, contributing to the thiolate stabilization. The active site sequence in Trx is CGPC. Compared to Grx3 and DsbA, the proline is shifted one step forward in the sequence. One consequence is that one of the hydrogen bonds involving an amide is missing. What is left is the possibility of forming two hydrogen bonds to the thiolate.

We are interested in comparing the structure and dynamics of the active site of representative members of the thioredoxin superfamily. Having explored Grx3, the focus has shifted to DsbA and Trx. We have used a combined approach including molecular dynamics simulations and biophysical experiments to try to resolve some of the questions regarding the active sites of these proteins. The pK_a values for the N-terminal cysteines have been determined by UV-titrations and/or with NMR titrations using specific ^{13}C-cysteine labeling. The pK_a and tautomer state of the histidines in DsbA and Trx P34H has been determined by NMR. Double mutant cycle analyses have been done to quantify the interactions between the active site cysteines and between the N-terminal cysteine and the active site histidine ring in DsbA and Trx P34H. Molecular dynamics simulations have been done for DsbA, Trx P34H and for Trx. For Trx, separate simulations were done with either a thiolate or a thiol group in the N-terminal cysteine. The data show some unexpected results and the dynamic behavior seems to explain at least some of the differences. The analysis of this data is ongoing.

E.coli Grx1 compared to Grx3

While the sequence of the dipeptide between the active site cysteines determines some of the redox properties, additional factors must exist. *E.coli* Grx1 and Grx3 have 33 % sequence identity and an identical active site sequence CPYC. Despite a similar structure (Fig. 10) and an identical active site sequence, Grx1 and Grx3 have a 35 mV difference in redox potential. We have determined the pK_a value of the N-terminal cysteine in Grx1 using NMR. Using the same sample, we have examined reduced Grx1 for the presence of the a C-terminal active site thiol proton to see if this is unique to Grx3 or represents a common, feature in these proteins. Double cycle analyses are under way to quantify the cysteine-cysteine interaction energy and the effect of

positively charged residues in the vicinity of the active thiolate for comparison to the results presented here for Grx3.

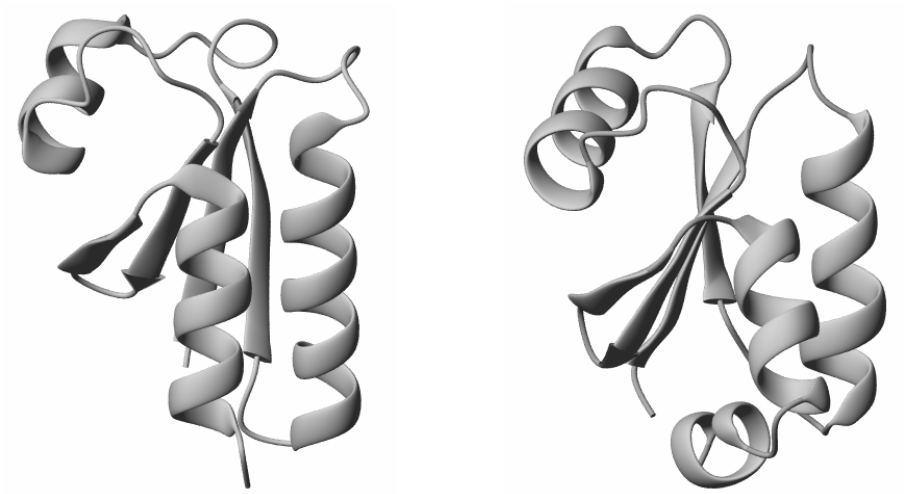

Figure 10. Structures of E.coli Grx1 (left) and E.coli Grx3 (right).

ACKNOWLEDGEMENTS

I would like to thank colleagues and the staff at NOVUM for all these years.

In particular I would like to thank

My supervisor professor Kurt D. Berndt. With his background in physics, chemisty, biology and medicine, he is a real scientist and an excellent teacher. Many of our discussions have turned to the really important things in life, such as byggnadsvård.

Tobias, diskussioner på kontoret angående de mest skiftande ämnena har ibland gjort att vissa dagar inte blivit så effektiva. Resorna till USA och Japan som vi gjort tillsammans har varit riktiga höjdpunkter.

Professor Peter Güntert, for the calculations of the NMR structures and for a nice stay in Yokohama.

Professor Lennart Nilsson som hjälpte oss att komma igång med MD simuleringar.

Past and present colleagues at CSB, in particular Micke, Boel, Linda, Kimberley and Xiofeng.

Catrine Johansson vid MBB i Solna som inledde ett samarbete om humana Grxs med oss.

Karin Fürtenbach som har hjälp till i labbet på slutet och som försett oss med lussekatter lagom till avhandlingsskrivandet.

Släkten: Mamma, pappa, syskon, Eva och Pelle, Milan och Vera.

Familjen. Att få tvillingar, börja bygga upp ett gammalt skiftesverkhus och göra klart avhandlingen under ett och samma år har varit ganska tufft. Nu hoppas jag att det ska bli lite lugnare. Malin, Ella och Arvid, jag älskar er.

REFERENCES

1. Berndt, C., Lillig, C. H. & Holmgren, A. (2008). Thioredoxins and glutaredoxins as facilitators of protein folding. *Biochim Biophys Acta* 1783, 641-50.
2. Lillig, C. H., Berndt, C. & Holmgren, A. (2008). Glutaredoxin systems. *Biochim Biophys Acta* 1780, 1304-17.
3. Martin, J. L. (1995). Thioredoxin--a fold for all reasons. *Structure* 3, 245-50.
4. Sodano, P., Xia, T. H., Bushweller, J. H., Björnberg, O., Holmgren, A., Billeter, M. & Wüthrich, K. (1991). Sequence-specific 1H n.m.r. assignments and determination of the three-dimensional structure of reduced Escherichia coli glutaredoxin. *Journal of Molecular Biology* 221, 1311-24.
5. Eklund, H., Ingelman, M., Soderberg, B. O., Uhlin, T., Nordlund, P., Nikkola, M., Sonnerstam, U., Joelson, T. & Petratos, K. (1992). Structure of oxidized bacteriophage T4 glutaredoxin (thioredoxin). Refinement of native and mutant proteins. *J Mol Biol* 228, 596-618.
6. Holmgren, A., Söderberg, B. O., Eklund, H. & Brändén, C. I. (1975). Three-dimensional structure of Escherichia coli thioredoxin-S2 to 2.8 A resolution. *Proceedings of the National Academy of Sciences of the United States of America* 72, 2305-9.
7. Martin, J. L., Bardwell, J. C. & Kuriyan, J. (1993). Crystal structure of the DsbA protein required for disulphide bond formation in vivo. *Nature* 365, 464-8.
8. Miranda-Vizuete, A., Ljung, J., Damdimopoulos, A. E., Gustafsson, J. A., Oko, R., Pelto-Huikko, M. & Spyrou, G. (2001). Characterization of Sptrx, a novel member of the thioredoxin family specifically expressed in human spermatozoa. *J Biol Chem* 276, 31567-74.
9. Wunderlich, M. & Glockshuber, R. (1993). Redox properties of protein disulfide isomerase (DsbA) from Escherichia coli. *Protein Sci* 2, 717-26.
10. Gilbert, H. F. (1998). Protein disulfide isomerase. *Methods in Enzymology* 290, 26-50.
11. Armstrong, R. N. (1994). Glutathione S-transferases: structure and mechanism of an archetypical detoxication enzyme. *Adv Enzymol Relat Areas Mol Biol* 69, 1-44.
12. Rhee, S. G., Kang, S. W., Chang, T. S., Jeong, W. & Kim, K. (2001). Peroxiredoxin, a novel family of peroxidases. *IUBMB Life* 52, 35-41.
13. Holmgren, A. (1976). Hydrogen donor system for Escherichia coli ribonucleoside-diphosphate reductase dependent upon glutathione. *Proceedings of the National Academy of Sciences of the United States of America* 73, 2275-9.
14. Holmgren, A. (1979). Glutathione-dependent synthesis of deoxyribonucleotides. Purification and characterization of glutaredoxin from Escherichia coli. *Journal of Biological Chemistry* 254, 3664-71.
15. Holmgren, A. (1979). Glutathione-dependent synthesis of deoxyribonucleotides. Characterization of the enzymatic mechanism of Escherichia coli glutaredoxin. *Journal of Biological Chemistry* 254, 3672-8.
16. Åslund, F., Ehn, B., Miranda-Vizuete, A., Pueyo, C. & Holmgren, A. (1994). Two additional glutaredoxins exist in Escherichia coli: glutaredoxin 3 is a hydrogen donor for ribonucleotide reductase in a thioredoxin/glutaredoxin 1 double mutant. *Proceedings of the National Academy of Sciences of the United States of America* 91, 9813-7.
17. Potamitou, A., Holmgren, A. & Vlamis-Gardikas, A. (2002). Protein levels of Escherichia coli thioredoxins and glutaredoxins and their relation to null mutants, growth phase, and function. *J Biol Chem* 277, 18561-7.
18. Åslund, F., Berndt, K. D. & Holmgren, A. (1997). Redox potentials of glutaredoxins and other thiol-disulfide oxidoreductases of the thioredoxin

superfamily determined by direct protein-protein redox equilibria. *Journal of Biological Chemistry* 272, 30780-6.

19. Lundström-Ljung, J., Vlamis-Gardikas, A., Åslund, F. & Holmgren, A. (1999). Reactivity of glutaredoxins 1, 2 and 3 from Escherichia coli and protein disulfide isomerase towards glutathionyl-mixed disulfides in ribonuclease A. *FEBS Letters* 443, 85-8.
20. Holmgren, A., Johansson, C., Berndt, C., Lonn, M. E., Hudemann, C. & Lillig, C. H. (2005). Thiol redox control via thioredoxin and glutaredoxin systems. *Biochem Soc Trans* 33, 1375-7.
21. Lysell, J., Stjernholm Vladic, Y., Ciarlo, N., Holmgren, A. & Sahlin, L. (2003). Immunohistochemical determination of thioredoxin and glutaredoxin distribution in the human cervix, and possible relation to cervical ripening. *Gynecol Endocrinol* 17, 303-10.
22. Padilla, C. A., Martinez-Galisteo, E., Barcena, J. A., Spyrou, G. & Holmgren, A. (1995). Purification from placenta, amino acid sequence, structure comparisons and cDNA cloning of human glutaredoxin. *Eur J Biochem* 227, 27-34.
23. Gladyshev, V. N., Liu, A., Novoselov, S. V., Krysan, K., Sun, Q. A., Kryukov, V. M., Kryukov, G. V. & Lou, M. F. (2001). Identification and characterization of a new mammalian glutaredoxin (thioltransferase), Grx2. *J Biol Chem* 276, 30374-80.
24. Lundberg, M., Johansson, C., Chandra, J., Enoksson, M., Jacobsson, G., Ljung, J., Johansson, M. & Holmgren, A. (2001). Cloning and expression of a novel human glutaredoxin (Grx2) with mitochondrial and nuclear isoforms. *J Biol Chem* 276, 26269-75.
25. Johansson, C., Lillig, C. H. & Holmgren, A. (2004). Human mitochondrial glutaredoxin reduces S-glutathionylated proteins with high affinity accepting electrons from either glutathione or thioredoxin reductase. *J Biol Chem* 279, 7537-43.
26. Johansson, C., Kavanagh, K. L., Gileadi, O. & Oppermann, U. (2007). Reversible sequestration of active site cysteines in a 2Fe-2S-bridged dimer provides a mechanism for glutaredoxin 2 regulation in human mitochondria. *J Biol Chem* 282, 3077-82.
27. Bushweller, J. H., Åslund, F., Wüthrich, K. & Holmgren, A. (1992). Structural and functional characterization of the mutant *Escherichia coli* glutaredoxin (C14S) and its mixed disulfide with glutathione. *Biochemistry* 31, 9288-93.
28. Szajewski, R. P. & Whitesides, G. M. (1980). Rate constants and equilibrium constants for thiol-disulfide interchange reactions involving oxidized glutathione. *J. Am.Chem.Soc* 102, 2011-2026.
29. Dyson, H. J., Jeng, M. F., Tennant, L. L., Slaby, I., Lindell, M., Cui, D. S., Kuprin, S. & Holmgren, A. (1997). Effects of buried charged groups on cysteine thiol ionization and reactivity in Escherichia coli thioredoxin: structural and functional characterization of mutants of Asp 26 and Lys 57. *Biochemistry* 36, 2622-36.
30. Grauschopf, U., Winther, J. R., Korber, P., Zander, T., Dallinger, P. & Bardwell, J. C. (1995). Why is DsbA such an oxidizing disulfide catalyst? *Cell* 83, 947-55.
31. Nordstrand, K., Åslund, F., Meunier, S., Holmgren, A., Otting, G. & Berndt, K. D. (1999). Direct NMR observation of the Cys-14 thiol proton of reduced Escherichia coli glutaredoxin-3 supports the presence of an active site thiol-thiolate hydrogen bond. *FEBS Letters* 449, 196-200.
32. Foloppe, N., Sagemark, J., Nordstrand, K., Berndt, K. D. & Nilsson, L. (2001). Structure, dynamics and electrostatics of the active site of glutaredoxin 3 from Escherichia coli: comparison with functionally related proteins. *J Mol Biol* 310, 449-70.
33. Lundström, J. & Holmgren, A. (1993). Determination of the reduction-oxidation potential of the thioredoxin-like domains of protein disulfide-isomerase from the equilibrium with glutathione and thioredoxin. *Biochemistry* 32, 6649-55.

34. Joelson, T., Sjöberg, B. M. & Eklund, H. (1990). Modifications of the active center of T4 thioredoxin by site-directed mutagenesis. *Journal of Biological Chemistry* 265, 3183-8.
35. Krause, G., Lundström, J., Barea, J. L., Pueyo de la Cuesta, C. & Holmgren, A. (1991). Mimicking the active site of protein disulfide-isomerase by substitution of proline 34 in Escherichia coli thioredoxin. *Journal of Biological Chemistry* 266, 9494-500.
36. Mössner, E., Huber-Wunderlich, M. & Glockshuber, R. (1998). Characterization of Escherichia coli thioredoxin variants mimicking the active-sites of other thiol/disulfide oxidoreductases. *Protein Science* 7, 1233-44.
37. Huber-Wunderlich, M. & Glockshuber, R. (1998). A single dipeptide sequence modulates the redox properties of a whole enzyme family. *Folding & Design* 3, 161-71.
38. Kortemme, T. & Creighton, T. E. (1995). Ionisation of cysteine residues at the termini of model alpha-helical peptides. Relevance to unusual thiol pKa values in proteins of the thioredoxin family. *J Mol Biol* 253, 799-812.
39. Guddat, L. W., Bardwell, J. C. & Martin, J. L. (1998). Crystal structures of reduced and oxidized DsbA: investigation of domain motion and thiolate stabilization. *Structure* 6, 757-67.
40. Jeng, M. F., Holmgren, A. & Dyson, H. J. (1995). Proton sharing between cysteine thiols in Escherichia coli thioredoxin: implications for the mechanism of protein disulfide reduction. *Biochemistry* 34, 10101-5.
41. Marley, J., Lu, M. & Bracken, C. (2001). A method for efficient isotopic labeling of recombinant proteins. *J Biomol NMR* 20, 71-5.
42. Bushweller, J. H., Holmgren, A. & Wuthrich, K. (1993). Biosynthetic 15N and 13C isotope labelling of glutathione in the mixed disulfide with Escherichia coli glutaredoxin documented by sequence-specific NMR assignments. *Eur J Biochem* 218, 327-34.
43. Yang, Y. F. & Wells, W. W. (1991). Catalytic mechanism of thioltransferase. *J Biol Chem* 266, 12766-71.
44. Johnson, B. H. & Hecht, M. H. (1994). Recombinant Proteins Can Be Isolated from Escherichia-Coli-Cells by Repeated Cycles of Freezing and Thawing. *Bio-Technology* 12, 1357-1360.
45. Clark, W. M. (1960). *Oxidation potentials of organic systems*, Williams & Wilkins, Baltimore.
46. Gilbert, H. F. (1995). Thiol/disulfide exchange equilibria and disulfide bond stability. *Methods Enzymol* 251, 8-28.
47. Gilbert, H. F. (1990). Molecular and cellular aspects of thiol-disulfide exchange. *Adv Enzymol Relat Areas Mol Biol* 63, 69-172.
48. Berndt, K. D., Flocco Henningsson, M., Hebert, H., Koeck, P. & Nilsson, L. (1999). *Structural biochemistry and biophysics*, Karolinska institutet.
49. Benesch, R. E. & Benesch, R. (1955). The Acid Strength of the -Sh Group in Cysteine and Related Compounds. *Journal of the American Chemical Society* 77, 5877-5881.
50. Street, T. O., Courtemanche, N. & Barrick, D. (2008). Protein folding and stability using denaturants. *Methods Cell Biol* 84, 295-325.
51. Fersht, A. R. (1999). *Structure and mechanism in protein science*, Freeman.
52. Levitt, M. H. (2001). *Spin dynamics*, Wiley.
53. Cavanagh, J., Fairbrother, W. J., Palmer III, A. G. & Skelton, N. J. (1996). *Protein NMR spectroscopy: principles and practice.*, Academic press, San Diego.
54. Güntert, P. (2004). Automated NMR structure calculation with CYANA. *Methods Mol Biol* 278, 353-78.
55. Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W. & Kollman, P. A. (1995). A

Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *Journal of the American Chemical Society* 117, 5179-5197.

56. Luginbühl, P., Güntert, P., Billeter, M. & Wüthrich, K. (1996). The new program OPAL for molecular dynamics simulations and energy refinements of biological macromolecules. *J. Biomol. NMR* 8, 136-146.
57. Blomberg, F., Maurer, W. & Ruterjans, H. (1977). Nuclear magnetic resonance investigation of 15N-labeled histidine in aqueous solution. *Journal of the American Chemical Society* 99, 1849-59.
58. Nordstrand, K., Åslund, F., Holmgren, A., Otting, G. & Berndt, K. D. (1999). NMR structure of Escherichia coli glutaredoxin 3-glutathione mixed disulfide complex: implications for the enzymatic mechanism. *Journal of Molecular Biology* 286, 541-52.
59. Madura, J. D., Briggs, J. M., Wade, R. C., Davis, M. E., Luty, B. A., Ilin, A., Antosiewicz, J., Gilson, M. K., Bagheri, B., Scott, L. R. & McCammon, J. A. (1995). Electrostatics and Diffusion of Molecules in Solution - Simulations with the University-of-Houston Brownian Dynamics Program. *Computer Physics Communications* 91, 57-95.
60. Zapun, A., Bardwell, J. C. & Creighton, T. E. (1993). The reactive and destabilizing disulfide bond of DsbA, a protein required for protein disulfide bond formation in vivo. *Biochemistry* 32, 5083-92.
61. Nelson, J. W. & Creighton, T. E. (1994). Reactivity and ionization of the active site cysteine residues of DsbA, a protein required for disulfide bond formation in vivo. *Biochemistry* 33, 5974-83.