Olena Dobrovolska

Structure-functional Characterization of Mammalian Redox Proteins

Methionine sulfoxide reductase B1 (MsrB1),
Glutaredoxin domain (Grx) of TGR, and Thioredoxin (Trx)

Thesis for the degree of Philosophiae Doctor

Trondheim, January 2013

Norwegian University of Science and Technology
Faculty of Natural Sciences and Technology
Department of Biotechnology

NTNU – Trondheim
Norwegian University of
Science and Technology

Acknowledgements

I cordially thank to all those people who helped me to complete this thesis.

First, I want to thank to the Department of Biotechnology and NT faculty of NTNU for providing the financial support which allowed me to undertake this research.

I express a special gratitude to my supervisor, Professor Alexander Dikiy, for his continious support, inspiring guidance and faith in me. His kindness and patience as well as his academic experience are invaluable for me.

I am also extremely indebted to my co-supervisor Dr. Elena Shumilina. She has provided advices many times during the research process, and has been helpful giving constructive comments for my progress report presentations.

I thank to our student Henrik Waldal Holen for performing the assignment of MsrB1-Co, to Rebecca Del Conte for calculating the magnetic susceptibility tenzor, and to Ann-Sissel Ulset for performing HPLC experiments.

I want to thank to all the teachers which courses I have taken during my PhD studies at NTNU, UiO, and University of Gothenburg. I gratefully acknowledge support and training from BioStruct, the Norwegian national graduate school in Structural biology.

I also express my sincere thanks to Dr. Finn Aachmann from whom I received help during the first weeks of my PhD study.

Finally, I thank my family and friends for their support and help throughout my study.

Summary

This thesis was written at the Department of Biotechnology of Norwegian University of Science and Technology (NTNU) and made as a completion of the three-year PhD program. It summarises the studies of the thiol redox active mammalian enzymes operating within thioredoxin defence system. Aiming to characterize their structural and functional aspects high resolution NMR spectroscopy technique was mainly used along with other complementary techniques. This study covers four research projects dedicated to the following proteins: methionine sulfoxide reductase B1 (MsrB1), thioredoxin (Trx), methionine sulfoxide reductase B (MsrBs), and Grx domain of mouse TGR (Grx).

The reduction mechanism of MsrB1, focusing on structural aspects of the intermolecular protein complex formation between MsrB1 and its functional partner Trx was the subject of the first study.

Analysis and systematization of the currently available structural data about the key members of the cellular antioxidant defence system, MsrB enzymes, were presented in the second study.

Thioredoxin glutathione reductase (TGR) is a member of the mammalian thioredoxin reductase family that has a monothiol glutaredoxin (Grx) domain attached to the thioredoxin reductase module. Grx structure determination and characterization was performed in the third project.

MsrB1 contains zinc ion coordinated by four cysteines. Recombinantly expressed in *E.coli* cells in cobalt-supplemented medium MsrB1 was demonstrated to uptake cobalt ion. Structural studies of cobalt-containing MsrB1 were described in the fourth project.

TABLE OF CONTENTS

I. Introduction

I.1. Thioredoxin and Glutaredoxin systems

All aerobic life proceses require oxygen. However, aerobic respiration inevitably produces reactive oxygen species (ROS) as byproducts, which may damage many cellular components icluding nucleic acids, proteins and lipids. The level of ROS production is normally controlled by the highly powerful antioxidant systems of the organism. The imbalance of the ROS production and the antioxidant defense, causes oxidative stress, which through a series of events, deregulates the cellular functions, leading to a various diseases and pathological conditions [1,2,3].

To prevent the ROS-induced disorders and maintain the redox control of cellular processes the organisms utilize two main enzymatic antioxidant systems, thioredoxin (Trx) and glutaredoxin (Grx) systems [4]. The key members of two systems are represented in Figure I.1.1.
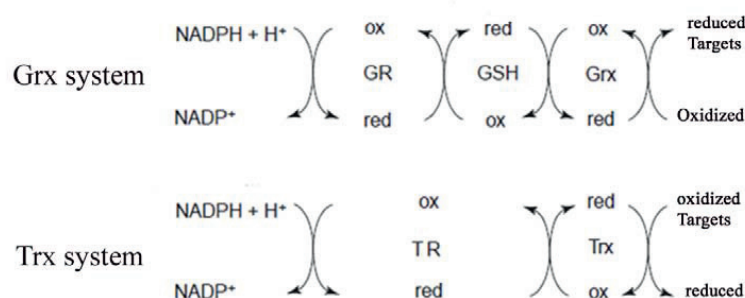
Figure I.1.1. Scheme depicting the key players in defense against the oxidative stress in Trx and Grx systems.

The Trx system is composed of thioredoxin reductase (TR), thioredoxin (Trx), and Trx peroxidase, while the Grx system constitute glutathione reductase (GR), a γ-Glu-Cys-Gly tripeptide (GSH), glutaredoxin (Grx), and glutathione peroxidase (GPx). Trxs and Grxs belong to a family of thiol-disulfide oxidoreductases, and characterized by the Trx-like fold [5] and a common dithiol-disulfide active site motif Cys-X-X-Cys [5,6]. Trxs catalyze a reversible reduction of disulfides utilizing its both cysteines in the active site. The N-terminal active site cysteine has a low pK_a value at around 7 [7] and acts as nucleophile attacking the target disulfide, forming a covalent-fixed intermediate, which is in turn reduced by the C-terminal resolving cysteine, leading to the reduction of target protein and disulfide formation in the active site of Trx. This disulfide is reduced by TR using electrons from NADPH (Figure I.1.2). Grxs catalyse reactions via two distinct but functionally connected mechanisms. During

the first mechanism, so called 'dithiol mechanism', the disulfide, formed between the two thiol groups upon reduction of protein disulfide (as in Trx), is reduced by one molecule of GSH, forming of a mixed disulfide between N-terminal cysteine of Grx and GSH. The obtained mixed intermediate disulfide is further reduced by the second molecule of GSH. For reduction of protein mixed disulfides (PSSG) Grx utilizes only the N-terminal cysteine, i.e. acting via 'monothiol mechanism'. Glutathione disulfide (GSSG) molecule, formed in both reactions, is further regenerated by GR at the expense of NADPH.
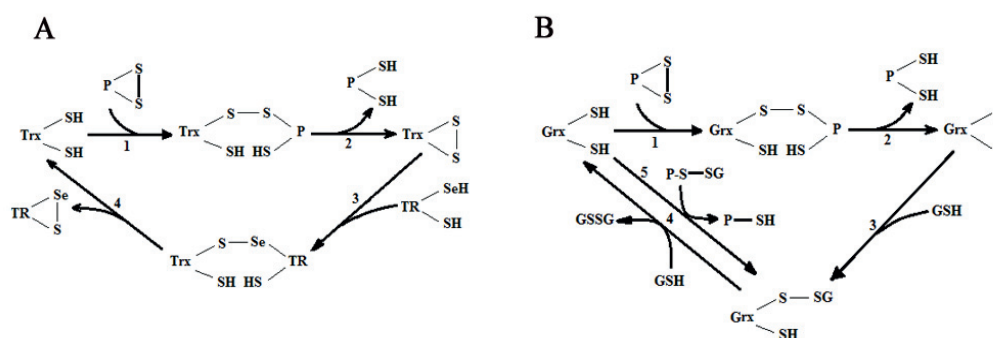
Figure I.1.2. Redox cycles and reaction mechanisms of Trxs (A) and Grxs (B). Upon reduction of protein disulfide the N-terminal active site cysteine of Trx forms a covalent-mixed disulfide intermediate (A,1), which further reduced by its C-terminal active site thiolate (A,2), forming intramolecular disulfide in Trx. Oxidized Trx is reduced by TR using electrons from NADPH (A, 3,4). Grxs reduce protein disulfides in a similar manner to Trxs (B, 1). Oxidized Grxs are further reduced by GSH, forming Grx-SG intermediate (B,3). Protein deglutathionylation reaction also lead to Grx-SG formation (B,5), which is reduced by the second GSH molecule (B,4). Picture adapted from [8].

NADPH-dependent glutathione reductase (GR) and thioredoxin reductase (TR) are the main antioxidant enzymes providing the transfer of reducing equivalents via reversible thiol oxidoreduction to the oxidized protein targets in the corresponding systems. Mammalian TRs and GRs are homodimers and possess structural and functional homology, although TRs has an additional C-terminal Gly-Cys-SeCys-Gly active site, a feature, which also distinguish them from their bacterial counterparts [9,10,11,12,13].

I.2. Selenoproteins

Selenocysteine is one of the naturally occurring amino acids in proteins. In this residue, that is a cysteine analog – sulfur is replaced by selenium. These two elements have a similar electronegativity (2,58 for sulfur and 2,55 for selenium [14]), however, Se is a stronger nucleophile [15,16,17]. Furthermore, the pK_a of selenocysteine is more acidic than that of cysteine (pK_a (Sec) = 5.2–5.6 while pK_a value of Cys is around 8.3) [16,18], which

makes Sec deprotonated (anionic form) at physiological pH and more reactive, than cysteine. Due to the higher selenium nucleophilicity, selenoproteins are typically more active than their cysteine (Cys) analogs [19]. This high catalytic activity has been regarded as a key reason why Sec is used in biological systems [20,21,22,23].

In general, selenoproteins are represented by a small number of protein families, and are found in many of bacterial, archaeal and eukaryal species. In eukaryotes Sec is encoded during translation by the UGA codon, normally a translation termination codon. This mechanism is complex, and requires the Selenocysteine Insertion Sequence (SECIS) in the 3'UTR region of mRNA and several protein factors [24,25]. The most abundant selenoproteins in mammals are the glutathione peroxidases, thioredoxin reductases, methionine sulfoxide reductase B1 (MsrB1), selenoproteins M (SelM) and W (SelW). The biological advantage of and evolutionary pressure for a catalytic Sec is a subject of debate because of the high cellular 'costs' of Sec incorporation [24].

I.3.   Metals in proteins

Metal ions are essential for living cells and are crucial to the structure and function of many proteins [26]. Occurrence of the metal ions as enzymatic cofactors is shown in Figure I.3.
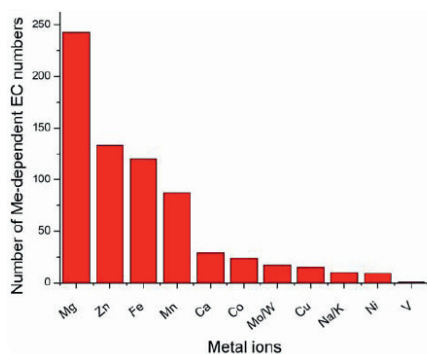
Figure I.3.1. Distribution of metals as cofactors. Occurrence of various metal ions as cofactors in enzymes with known structure, the bars indicating the number of distinct enzymes (EC number) which depend on the given metal ion for catalysis. Figure reproduced from [26].

In proteins metals serve different functions, such as [27]:

1. Structural – bound metal forms a structural motif and/or assists correct protein folding;

2. Metal storage – uptake, binding and release of metals in soluble form;

3. Electron transfer – bound metal facilitates uptake, storage and release of electrons;

4. Dioxygen binding – bound metal facilitates uptake, storage and release of molecular oxygen;

5. Catalytic - bound metal facilitates binding, activation and turnover.

In proteins metals are bound either to endogenous ligands (backbone or a side chain atoms of the same polypeptide) or exogenous ligands (provided by other molecules bound to the polypeptide) [27]. A list of endogenous metal ligands is represented in Table I.3.2.

Table I.3.2. A list of endogenous ligands. Endogenous ligands divided into three groups by which atom is donating electrons to the metal ion. In the left column the ligand groups are listed, the examples of the metal groups are listed in the right column. X represents any residue, unless specific residues are listed in brackets. Figure modified from [27].

Many protein-bound metals are divalent metals. The natural order of stability for divalent metals sets out a resulting trend ($Ca^{2+} < Mg^{2+} < Mn^{2+} < Fe^{2+} < Co^{2+} < Zn^{2+} < Ni^{2+} < Cu^{2+}$), often called the Irving-Williams series [28]. The trend follows a decrease in ionic radii, which leads to stronger metal-ligand bonds. Zn^{2+} is larger than Cu^{2+} and, thus, is an exeption. It is assumed that proteins insert metals by sponteneous self-assembly [29].

I.3.1 Transition metals – coordination geometry and ligand field splitting

Many of the biological metals are transition metals, which have d-orbital electrons. d-orbital electrons allow many of these metals to appear in a variety of oxidation states. Many of the transition metals allow *d*-orbital hybridization in complex with ligands, and, thus,

coordination of more ligands and a higher variety of the coordination geometries [29]. Common geometries of ligand-metal complexes with coordination nnumber 3, 4, 5 and 6 are represented in figures I.3.1. and I.3.2.
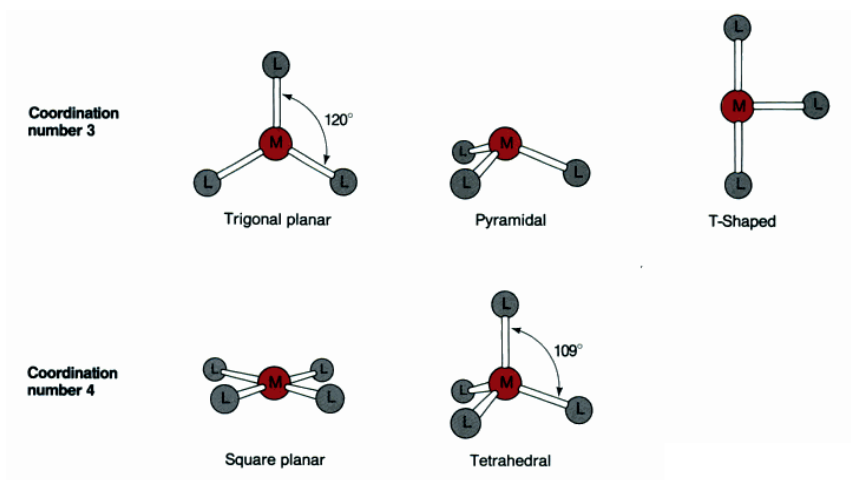
Figure I.3.1. Common coordination geometries for coordination numbers 3 and 4. Taken from [29].
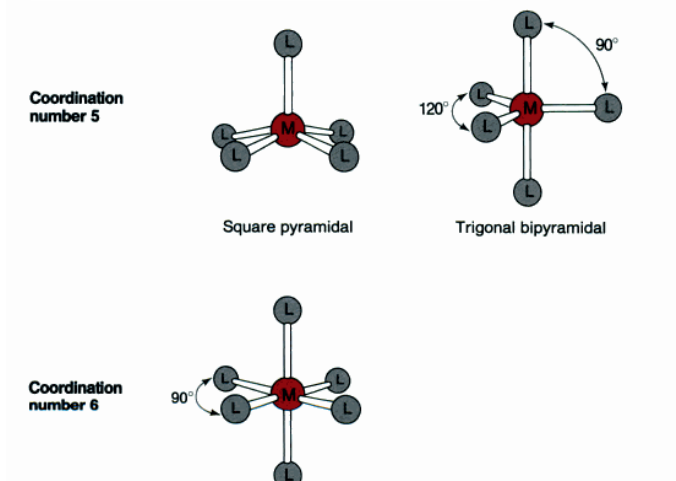
Figure I.3.2. Common coordination geometries for coordination numbers 5 and 6. Taken from [29].

For a free transition metal ion the d-orbitals are degenerate (equal energy on each level). Upon binding with the ligands the d-orbitals of a metal ion become non-degenerate, and the energy gap between the orbitals depends on the geometry of the metal-ligand complex. This is reffered to as ligand-field splitting. The number of unpaired electrons depends on the geometry and the number of the d-orbital electrons (Figure I.3.3) [29]. Unpaired electrons are responsible for paramagnetism, a phenomenon with important consequences for various spectroscopic methods.
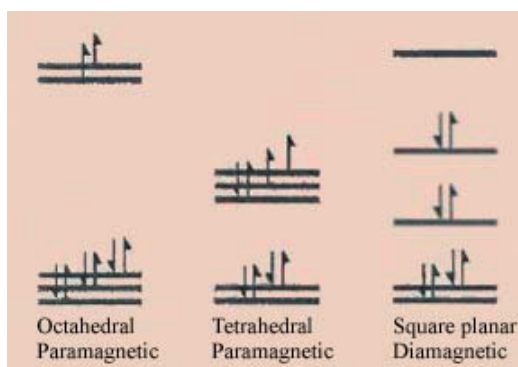
Figure I.3.3. Ligand–field splitting diagrams, orbital occupancies, and magnetic properties for d^8 Ni (II) complexes having octahedral, tetrahedral, and square planar geometries.

Among all metal ions, Zn^{2+} is one of the most important trace metal ions in living organisms and an essential cofactor in many metabolic enzymes and regulatory proteins [30]. Zinc-binding sites in proteins play mainly either catalytic or a structural role and possess 4- or 6-fold ligand coordination (5-fold coordination is rather rare) [31]. Catalytic Zn^{2+} binds preferentially to water molecules and histidines in catalytic sites of enzymes, such as human carbonic anhydrase II [32], carboxypeptidase [33], thermolysin [34] etc. Structural Zn sites have four protein ligands (preferentially Cys) and no bound water molecule. Aspartate transcarbomylase [35], cytochrome c oxidase [36], horse alcohol dehydrogenase [37], many protein kinases [38], tRNA synthases [39], 'zinc finger' proteins [40] are some examples of enzymes with structural zinc sites. Because of its d^{10} close-shell electronic configuration, zinc is regarded as spectroscopically 'silent' metal ion. To spectroscopically characterize proteins' zinc-sites the substitution with several paramagnetic metal ions (Co^{2+}, Ni^{2+}, Mn^{2+}, Cu^{2+}), preferentially with Co^{2+}, is often used. Generally, the substitution of Zn^{2+} by Co^{2+} in the proteins is performed by either displacement of the original metal ion (preparation of the apoprotein) followed by addition of Co^{2+}, or by biosynthetic incorporation, which involves growing the host organism in a Co^{2+}- supplemented and Zn^{2+} - deficient medium [41]. The latter method is sometimes considered as *in vivo* substitution of Co^{2+} for Zn^{2+} [39,42,43,44,45]. Co^{2+} is a paramagnetic metal ion and has d^7 electronic configuration (incomplete outer shell), which allows the *d-d* transitions that absorb within UV-visible region. Coordination chemistry and ionic radii similarity makes cobalt an excellent structural and functional model for zinc [46]. The ability of Co^{2+} ion to substitute Zn^{2+} in proteins was reported elsewhere and has been used to study active sites of some naturally occurring zinc enzymes [47,48,49]. For many systems with a catalytic zinc-site, Co-substituted enzymes exhibited similar or higher catalytic activity [29].

II. Redox proteins in Thioredoxin system

II.1. Methionine sulfoxide reductase B1 (MsrB1)

Although all amino acids are susceptible to oxidation, only two sulfur-containing amino acids, cysteine and methionine, could be reversibly oxidised [50]. Cysteines are implicated in catalytic cycle of many enzymes, they can contribute to the protein structure stabilization *via* disulfide bond formation or metal ion coordination, and participate in redox signaling processes [51]. Methionines, representing one of the major targets of ROS, by its reversible oxidation reduce the intracellular level of ROS, restore the enzymatic function of oxidized protein and inhibit or induce certain cellular events in the organism. Oxidation of methionine results in formation of two enantiomers, S-MetO and R-MetO. Thioredoxin system catalyzes the reduction of free and protein–bound MetO *via* thiol-based oxidoreductases methionine sulfoxide reductases (Msrs). Msrs constitute two distinct family of enzymes, MsrA and MsrB, each specific for reduction of S-MetO and R-MetO isomers, respectively (Figure II.1.1).
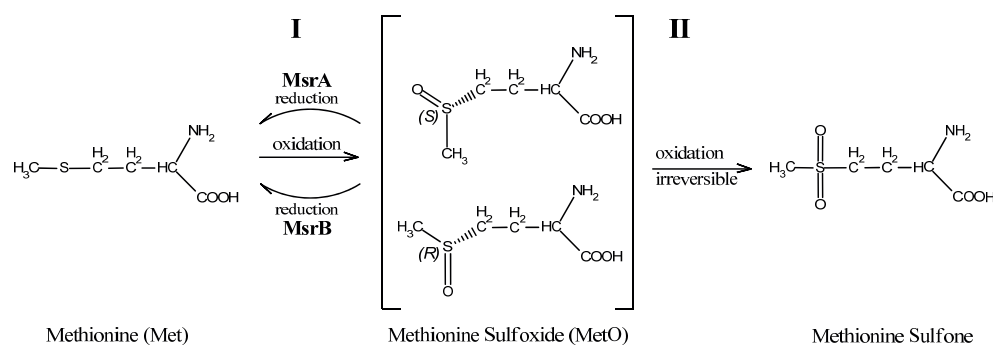
Figure II.1.1. The mechanism of methionine oxidation/reduction. Step I: reversible oxidation of methione gives two enantiomers: S- and R-methionine sulfoxide. Step II: irreversible oxidation of methionine sulfoxide to methionine sulfone. Picture taken from [52].

Mammalians contain three different subclasses of MsrB: MsrB1, MsrB2 and MsrB3. While mammalian MsrB1 contains selenocysteine (Sec) in the active site and resolving Cys, MsrB2 and MsrB3 contain only a catalytic cysteine residue [53,54]. All of them contain zinc ion, coordinated by four cysteines.

Among all members of MsrB family selenoprotein MsrB1 is the most active in the reduction of MetO. The three-dimensional structure of mammalian MsrB1, containing Cys95Sec, has recently been determined by high resolution NMR spectroscopy [55] and is represented on Figure II.1.2. It consists of two antiparallel β-sheets. Both N-terminal and C-terminal regions are flexible. The first β-sheet is three stranded, forming the back side of the

structure, whereas the second sheet has five strands forming the front side. The active site is situated in the second β-sheet. The four cysteines Cys23, Cys26, Cys71, and Cys74, situated outside the protein active site, coordinate zinc ion, stabilizing the structure of MsrB1.
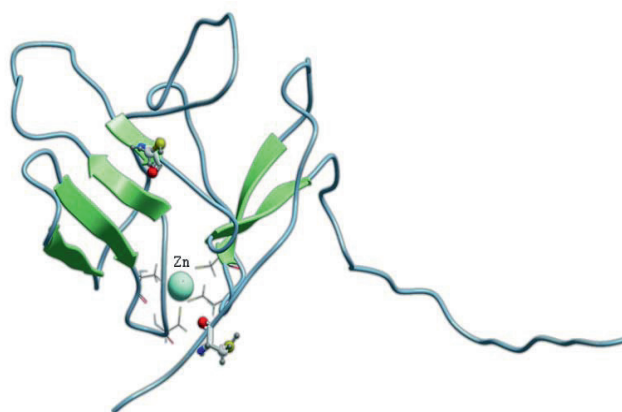
Figure II.1.2. Structure of MsrB1 (pdb code 2kv1) [55].

II.1.1 MsrB1-Thioredoxin interaction

The mechanism of a catalytic MetO reduction by Msrs employs a sulfenic acid chemistry and S-S/Se-S formation [56,57,58,59,60]. First, a catalytic Sec95 attacks a sufoxide moiety of MetO, resulting in formation of selenenic acid intermediate and concomitant release of methionine. Second, a resolving Cys4, situated on the mobile N-terminal region, flips in towards the catalytic Sec95 and attacks the selenenic acid intermediate forming a selenylsulfide bond Cys4-Sec95. Hydrophobic interactions between aromatic residues in the N-terminal region and the active site play an important role during the catalysis, assisting the movement of the N-terminal region upon formation of the intramolecular disulfide [55,61]. The intramolecular disulfide bond in MsrB1 is further reduced by thioredoxin (Trx) [62,63]. The proposed thioredoxin-dependent catalytic cycle mechanism for MsrB1 is illistrated on Figure II.1.1.1.
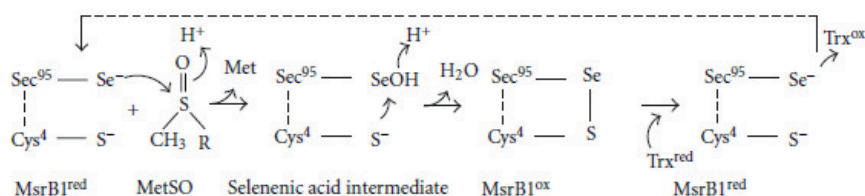
Figure II.1.1.1. The biochemical cycle of selenoprotein MsrB1. Picture taken from [52].

During the intermediate step of the thioredoxin-catalyzed reaction the formation of the intermolecular protein complex between thioredoxin and its substrate takes place. Some studies suggested that specific structural recognitions exist between them [64,65]. Elucidation of the structural aspects of the MsrB1-Trx complex formation would provide a better understanding of the full catalytic mechanism.

II.2. Mammalian thioredoxin glutathione reductase (TGR)

Mammalian TRs belong to a family of pyridine nucleotide disulfide oxidoreductases, and are represented by three different subclasses: TR1 (cytosolic TR), TR3 (mitochondrial TR) and TGR (thioredoxin glutathione reductase). TR1 and TR3 were shown to be essential for mouse embryogenesis, and ubiquitously expressed in various tissues and cell types [66,67]. TGR is expressed predominantly in testes in cytosol of spermatids at the time of mitochondrial sheath formation [68]. Among TRs TGR is unusual because it combines the elements of both Trx and Grx systems: TGR has an N-terminal Grx domain, containing a monothiol CPHS catalytic motif, naturally fused to a canonical TR module, sequentially related to TR1 (Figure II.2.1) [69,70,71,72]. TGR exhibits broad substrate specificity: it can reduce various elements of both thioredoxin and glutaredoxin systems. In fact, it was demonstrated that TGR can catalyze reactions associated with Grx (deglutathionylation), GR (NADPH-dependent reduction of GSSG) and TR (NADPH-dependent reduction of Trx) activities [71].
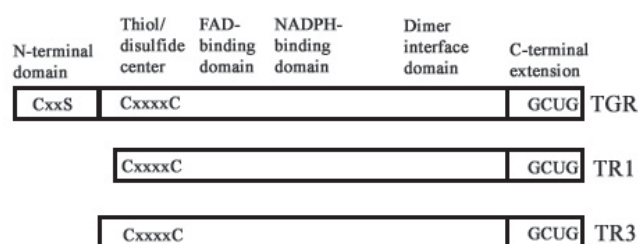
Figure II.2.1. Schematic representation of the domain organization of mammalian TRs. Picture adapted from [73]. All three proteins contain active center disulfides (CxxxC motifs), FAD- and NADPH- binding domains, a C-terminal extension of GCUG-tetrapeptide, dimer interface domain and other features of the pyridine nucleotide disulfide oxidoreductase family. In addition, TGR contains N-terminal Grx domain.

Based on the available structures of pyridine nucleotide disulfide oxidoreductases, molecular modeling of TGR homodimer was performed (Figure II.2.2) by Sun et al. [71]. In the obtained model, Cys-Sec motif of one subunit transfers electrons from the thiol-disulfide center (TR module) to Grx domain of the second subunit. The proposed reaction mechanism suggested the following electron flow from NADPH through several redox centers within

TGR: NADPH → FAD → thiol-disulfide center → the C-terminal Sec-containing tetrapeptide → the active site Cys residue within the Grx domain → downstream substrate [71].
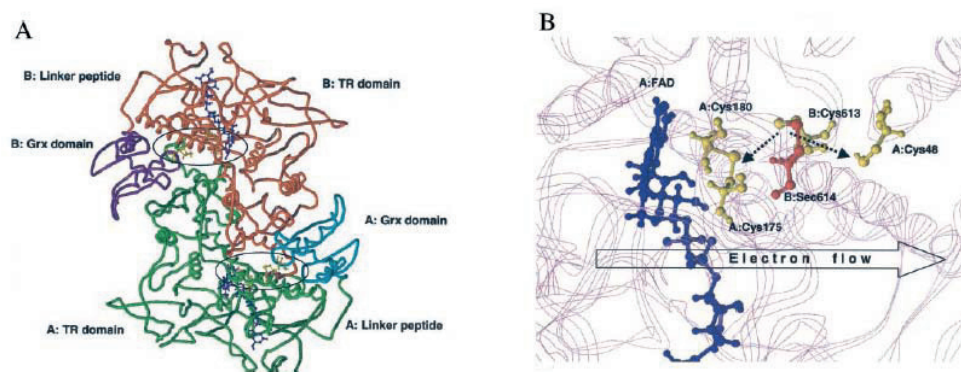
Figure II.2.2. Molecular modeling and reaction mechanism of TGR. Panel A) Molecular model of TGR. The TR domain of subunit A is shown in green and of subunit B in red. The Grx domain of subunit A is shown in cyan and of subunit B in purple. Active centers are circled. Two FAD molecules are shown in blue, and Cys48, Cys175, Cys180, Cys613, and Sec614 in both subunits are shown in yellow. Panel B) Molecular model of the enzyme active center. FAD molecule is shown in blue. Cys48, Cys175, Cys180 of subunit A and Cys613 of subunit B are shown in yellow. Sec614 of subunit B in shown in red. Predicted interactions of the Cys613/Sec614 dipeptide with the disulfide center Cys175/Cys180 and the active center of the Grx domain Cys48 are shown by dashed arrows. In the initial step of catalysis by TGR, NADPH reduces FAD. The predicted direction of electron flow in subsequent steps of catalysis (A: FAD→A: Cys175/Cys180→B:Cys613/Sec614→A: Cys48).

In support to the obtained TGR model, the enzymatic studies of the native purified protein demonstrated a key role of C-terminal Sec in TGR function [74]. Further studies argued that Grx domain of TGR is responsible for Grx and GR activities of the enzyme, and its involvment in formation and isomerization of disulfide bonds in spermatids [75].

The first NMR analysis of Grx domain of mouse TGR was performed by Shumilina E. [76], and as a result ^{1}H, ^{13}C and ^{15}N NMR assignment was reported. Through this study it was found that N-terminal part of Grx, containing 22 amino acids, is unstructured. The overall resemblance of ^{15}N-^{1}H HSQC patterns between the full-length Grx and its shortened form (lacking N-terminal residues) indicated on their structural similarity. Also, the full-length protein had a higher stability than the short Grx, and was also missing the HSQC signals, corresponding to residues C105 and D106, D74 and A77 in the corresponding spectra. These findings may indicate on slight structural differences between the two proteins due to the presence/absence of the N-terminal part. The obtainment of the three-dimensional structure of Grx would shed light on this phenomenon and, thus, may provide novel aspects of Grx function.

III. Spectroscopic methods in structural proteomics

All spectroscopic methods are based on exposing a sample to electromagnetic irradiation of a given frequency or frequency range, upon which scattering or absorbance of this irradiation by the sample leads to a change in the recorded output intensity [5]. Different spectroscopic techniques operate over different and limited frequency ranges within this broad spectrum Figure III.1
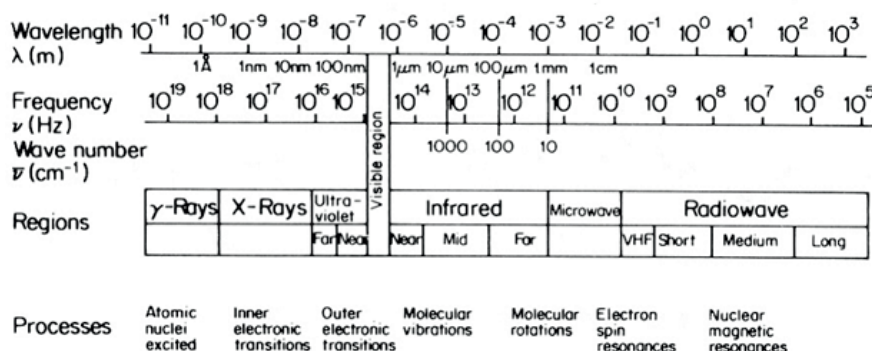
Figure III.1: The electromagnetic radiation spectrum. The different frequency ranges of the electromagnetic radiation spectrum with subclassification of frequency range. Molecular, nucleic or electronic influence of the different frequency ranges are also indicated.

III.1. Nuclear Magnetic Resonance on Proteins

NMR spectroscopy and X-ray crystallography are the only two experimental techniques which provide a complete protein structure elucidation. Below, the basic principles of NMR spectroscopy and the experiments most relevant to this thesis are described. The following paragraph is adapted from several sources ([77,78,79,80], 'Protein NMR - A Practical Guide' website available at *http://www.protein-nmr.org.uk/spectra.html)* so the references will not be further mentioned.

NMR spectroscopy is based on the observation of the resonance frequency of a spin. Spin is an intrinsic property that certain isotopes possess while others do not. The preferable nuclei for liquid NMR are the spins with the quantum number of one half. The most frequently found atoms in proteins, that possess spin of one half, are hydrogen (^{1}H isotope), carbon (^{13}C isotope), nitrogen (^{15}N isotope), and phosphorous (^{31}P isotope). Hydrogens, having only one proton in the atom nucleus (referred to as a protons) are the most abundant in proteins (natural occurrence 99,98%), whereas ^{13}C and ^{15}N make up only 1.1% and 0.4%, respectively, of all naturally occuring carbon and nitrogen atoms.

The spin needs to experience a magnetic field and be excited to be observable. After excitation, the spin will precess with a resonance frequency (Figure III.1.1, A), and with the passage of time will return to equilibrium. The return of magnetization to equilibrium is called relaxation, whith the rate depending on the local mobility of the nuclei. When nuclei with a spin I=1/2 are placed in a magnetic field there are two possible spin states, either the spin is oriented up or the spin is down (Figure III.1.1, A), i.e. the two energy levels generated – α and β (Figure III.1.1).
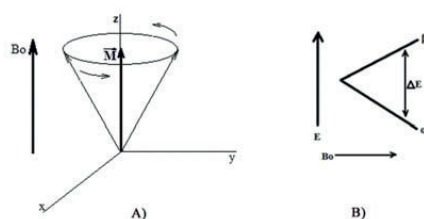
Figure III.1.1. A) Spin precessing around the magnetic field direction (upward and downward). The precession frequency is the same as the resonance frequency of the nuclei. Mz is the sum magnetization vector. B) Energy level diagram for a spin with I= ½ in a magnetic field B_0.

These energy states are separated by an amount ΔE , which is field dependent:

$$\Delta E = \left(\frac{h\gamma}{2\pi} \right) B_0 ,$$

where h is the Planck's constant, γ the gyromagnetic ratio ('spin intrinsic property') and B_0 is the magnitude of the applyied static magnetic field. At equilibrium the magnetization is precessing around the field direction in a non coherent manner, which results in non observable magnetization. Magnetization becomes coherent, i.e observable, by tilting it away from Z-axis into the XY plane applying radio frequency pulse.
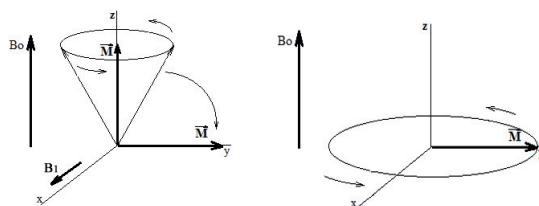
Figure III.1.2. The bulk magnetization after perturbation by a pulse. The net magnetization is symbolised by the thick vector and called My.
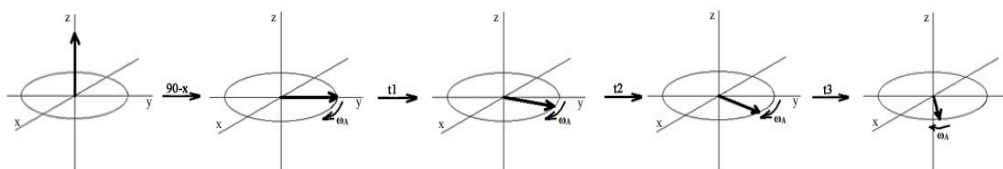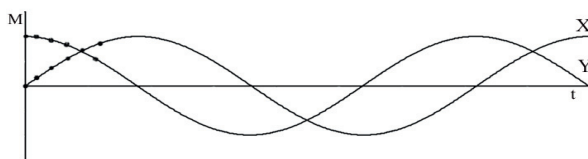
Figure III.1.3. The evolution of the net magnetization and the size of the net magnetization vector in X and Y directions as a function of time. The points on the curves of the lower parts of the figure illustrate how the digitized net magnetization changes with time.

Figure III.1.3 illustrates what happens when a pulse renders magnetization observable and this magnetization starts to evolve. In Figure the evolution of magnetization give a frequency oscillation of the NMR signal.

The observation frequency may be expressed in terms of the magnetogyric ratio and the applied field (Larmor law):

$$\nu = \left(\frac{\gamma}{2\pi} \right) B_0 \,.$$

The resonance frequency of the spin depends on the magnetic field, the nuclei and on the chemical surroundings of the nuclei. The latter property of the frequency introduces a possibility to observe an individual chemical shift of all protons in a protein molecule.

During absorption, the magnetic dipole moment of a nucleus deviates from its equilibrium position along the direction of the magnetic field (z-axis) towards the xy-plane where it is precessing. This tipping of the magnetic dipole moment is recorded by an oscillator coil in the instrument (reciever) producing a Free Induction Decay (FID) curve, which after Furie transformation transforms the time domain into the frequency domain, thus, acquiring an NMR spectrum.

III.1.1   Spin-spin interactions through electrons

Two nuclei that have spins and are connected by chemical bond will be J coupled. Scalar coupling is the isotropic part of the J coupling, while dipolar coupling is the anisotropic component of the latter. In practice anisotropy of J coupling is rarely observed, therefore, term scalar coupling and J coupling is often used as synonyms. The effect of J coupled nuclei will affect the spectrum. This is because the J coupling introduces a local field on the J coupled nuclei that either is against or along the local field observed at the nuclei. This field difference will in turn lead to a splitting of the peak, which complicates the spectrum. Small J couplings may not be observable in spectra of molecules with long correlation times (low mobility) and thus with wide peaks. Splitting may thus not be observed in biomolecular spectra. They are however used to transfer magnetization between nuclei.

III.1.2   Nuclear Overhauser Effect

When irradiation with RF of one spin (S) causes the perturbation of other nucleus spins close in space nuclei (I) *via* dipole-dipole interactions*,* the Nuclear Overhauser Effect arises, resulting in signal enhancement of other nuclei. The NOE intensity can be related to the distance r between the irradiated and observed spin by an equation:

$$NOE \propto \frac{1}{\langle r^6 \rangle} f(\tau_c),$$

Where $f(\tau_c)$ is a function of correlation time, which accounts for the influence of the motional averaging process on the observed NOE.

The energy diagram of two-spin coupled system is shown in the following figure Figure III.1.2.
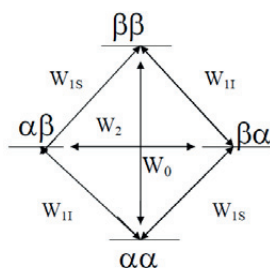
Figure III.1.2. The energy and transitions of two – spin system.

W_{1I} and W_{1S} are single quantum transition probability rates for saturated and observed spins, respectively. W_2 and W_0 are probability rates for double and zero quantum transitions. The W_{1I} and W_{1S} rates are responsible for the spin-lattice relaxation process. After saturation of the I spin with RF, a population difference will be created across the energy levels that is different from that in the unperturbed state. This creates the Nuclear Overhauser Effect at unsaturated spin S. The magnitude of NOE can be calculated according to the following formula:

$$\varepsilon_{NOE} = 1 + \frac{\gamma_I}{\gamma_S} \left(\frac{(W_2 - W_0)}{W_0 + 2W_{1S} + W_2} \right)$$

γ_I and γ_S are the gyromagnetic ratios for spins I and S respectively.

As can bee seen in the formulae, the magnitude of NOE and its sign depends strongly on W_2 and W_0 transitions. To induce the spin transition a molecule should rotate with a

frequency similar to the difference between the energy levels. The rotation of the molecule, and, thus, the correlation time vary depending on the molecular type and solution.

For small molecules NOE is positive, for large molecules (like proteins) NOE is negative.

III.1.3   NMR experiments with proteins

NMR spectroscopy technique is an important technique for structure determination of biomolecules. Structure determination is a process where structures are calculated by using computer algorithms, starting from the primary sequence of the molecules and a set of measured interatomic interactions as well as dihedral angles that are structure determining. Prior to determining the NMR constraints, the NMR signals must be assigned (the atom in the protein responsible for the signal is determined). In principle, there are two different approaches to NMR signal assignment. The first one is applicable to small and medium sized molecules where only protons are used to resolve the spectra. When the protein is large and the spectral resolution is decreased, the approach is different. Such proteins have larger peak widths and extensive signal overlap, which introduces ambiguity in chemical shift assignment making it difficult to analyze. The signal overlap is solved by increasing the dimensionality of the experiments, i.e. by additional protein labeling, which allows signal separation based on heteroatom, and, therefore, simplify protein assignment. Labeling means that specific spin half isotopes are brought into the molecules either by synthesizing from labeled amino acids or by growing proteins from cells systems capable of living off labeled media. The different labeling schemes and the typical NMR choice of NMR techniques are shown in Table 1.

Table 1. Suggested protein labeling as a function of molecular size and suggested NMR experiments.

Molecular mass, kDa	Residue number	Labeling	NMR techniques for structure determination
0-6	0-50	none	^{1}H, NOESY, TOCSY
3-12	30-100	^{15}N	HSQC, 3D ^{15}N-NOESY-HSQC, 3D ^{15}N TOCSY-HSQC
6-20	60-250	^{15}N, ^{13}C	HNCA, HNCO, HNCOCA, HNCACB
15-25	150-200	^{15}N, ^{13}C	TROSY type: HNCA, HNCO, HNCOCA, HNCACB

15-40	150-350	^{2}H, ^{15}N, ^{13}C	TROSY type: HNCA, HNCO, HNCOCA, HNCACB
25>	250>	^{2}H, ^{15}N, ^{13}C	TROSY type: HNCA, HNCO, HNCOCA, HNCACB

Homonuclear spectra

COSY (Correlation Spectroscopy), TOCSY (Total Correlation Spectroscopy) and NOESY (Nuclear Overhauser Enhancement Spectroscopy) 2D proton-proton NMR spectroscopic techniques are the most often used in the assignment process and subsequent structure calculation of smal and medium sized proteins. Figure III.1.3.1 shows a typical NMR pulse sequence used on unlabeled samples.

^{1}H   t1   MIX   t2

Figure III.1.3.1. Schema illustrating the two time periods giving in 2D spectrum. t2 is the observed signal, while t1 is the incremented indirect dimension. MIX is the sequence used to ensure that the signal is transferred between spins. Horizontal black bar is the preparation pulse; the red bar illustrates the signal in different evolution dimensions.

Both COSY and TOCSY experiments use the J^2 and J^3 couplings to transfer magnetization between nuclei, which turns up as a cross-peak in the 2D spectrum between nuclei the magnetization comes from to the nuclei the magnetization goes to. The COSY technique is only efficient at transferring magnetization from a proton connected to one carbon to a proton connected to the neighboring carbon. In TOCSY experiments the magnetization may be transferred between a chain of J^3 connected protons, which means that they are able to connect J couplings within an amino acid, but not between the amino acids in a row. In NOESY spectra by through space cross relaxation. NOESY signals depend on atoms being close in space, but not nessesarily connected through covalent bonds. The upper distance for magnetization transfer in a 2D proton-proton NOESY is typically 5Å.

Heteronuclear spectra

The ^{15}N and ^{13}C labeling increases spectral resolution comparing to the 2D spectra, as it introduces the third dimension, allowing to assign the chemical shift values of the labeled proteins unambiguously. All experiments on ^{15}N and ^{13}C labeled proteins use HSQC (Heteronuclear Single Quantum Coherence) type of magnetization transfer (INEPT). In

HSQC experiments the magnetization is transferred via J^1 couplings to the wanted hetero nuclei, which may be either ^{15}N or ^{13}C and back. Observation is done on protons and the ^{15}N or ^{13}C are observed indirectly. All experiments below depend on labeling with both isotopes unless otherwise stated. In the accompanying illustrations, atoms that are contribute to the respective spectra by cross peaks are colored grey, whereas atoms that are involved in transferring signals but are not represented with cross peaks are colored blue.

^{15}N-^1H **HSQC (requires ^{15}N-labeling)** shows cross peaks for each N-HN pair of a protein. The spectrum is mostly represented by cross peaks from backbone N-HN, but also from side chains of certain amino acids. Since Pro residues do not have N-HN pairs, they are not visible in the spectra (Figure III.1.3.2).

Figure III.1.3.2. ^{15}N-^1H **HSQC:** Schematic representation of amino acids of a protein and the atoms involved in the effects recorded in an HSQC-experiment.

All experiments on ^{15}N and ^{13}C labeled proteins use HSQC type of magnetization transfer (INEPT transfer step).

HNCA (3D) correlates the N-HN pair with Cαi and Cαi-1, thus giving two cross peaks in the third dimension for each N-HN pair (Figure III.1.3.3). The cross peak for Cαi is usually more intense.
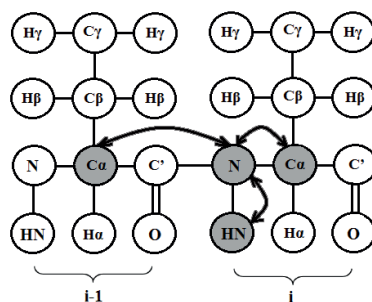
Figure III.1.3.3. **HNCA:** Schematic representation of amino acids of a protein and the atoms involved in the effects recorded in an HNCA-experiment.

CBCANH gives individual cross peaks for Cαi, Cβi, Cαi-1 and Cβi-1 for each N-HN pair. **CBCA(CO)NH** gives individual cross peaks only for Cαi-1 and Cβi-1. Comparison of the two spectra makes it possible to distinguish which cross peaks belong to which amino acid (Figure III.1.3.4).
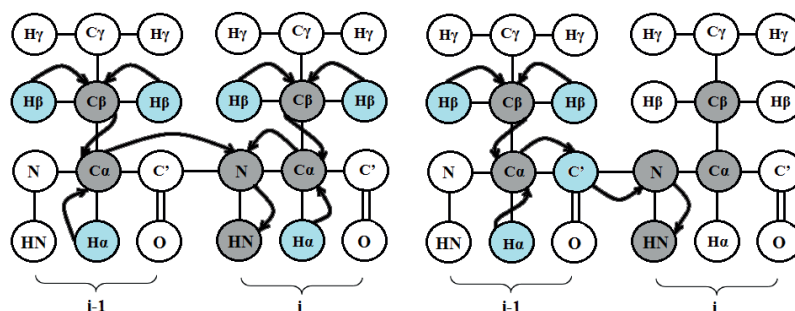
Figure III.1.3.4. **CBCANH and CBCA(CO)NH.** Schematic representation of amino acids of a protein and the atoms involved in the effects recorded in CBCANH- and CBCA(CO)NH-experiments (left and right, respectively).

HN(CA)CO gives individual cross peaks for COi and COi-1 for each N- HN pair. **HNCO** gives individual cross peaks only for COi-1. Comparison of the two spectra makes it possible to distinguish which cross peaks belong to which amino acid (Figure III.1.3.5).
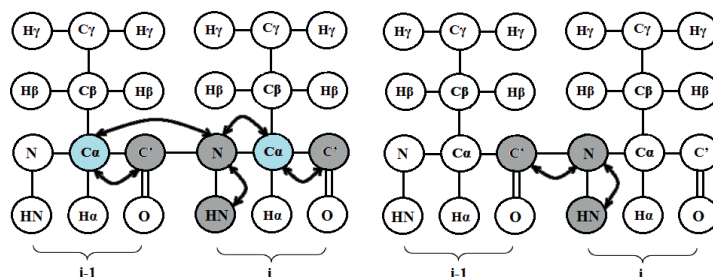
Figure III.1.3.5. **HN(CA)CO and HNCO.** Schematic representation of amino acids of a protein and the atoms involved in the effects recorded in HN(CA)CO- and HNCO-experiments (left and right, respectively).

HBHANH gives individual cross peaks for Hαi, Hβi, Hαi-1 and Hβi-1 for each N- HN pair. **HAHB(CO)NH** gives individual cross peaks only for Hαi-1 and Hβi-1. Comparison of the two spectra makes it possible to distinguish which cross peaks belong to which amino acid (Figure III.1.3.6).
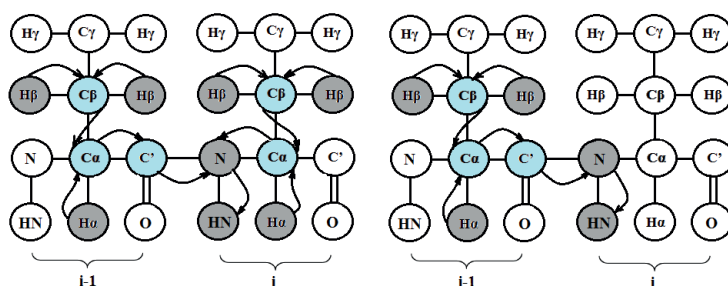
Figure III.1.3.6. Schematic representation of amino acids of a protein and the atoms involved in the effects recorded in **HBHANH**- and **HAHB(CO)NH**-experiments (left and right, respectively).

HCCH-COSY (requires ^{13}C-labeling) establishes correlations of ^{1}H and ^{13}C resonances within H-C-C-H molecular fragments using one bond ^{1}H-^{13}C and ^{13}C-^{13}C j couplings (Figure III.1.3.7).
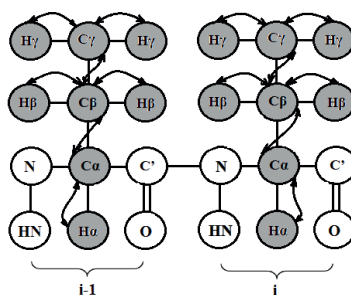
Figure III.1.3.7. Schematic representation of amino acids of a protein and the atoms involved in the effects recorded in a COSY-experiment.

HCCH-TOCSY (requires ^{13}C-labeling) correlates and gives cross peaks for all carbon atoms of a given amino acid (except CO) and all hydrogen atoms of the same amino acid (except HN) (Figure III.1.3.8)
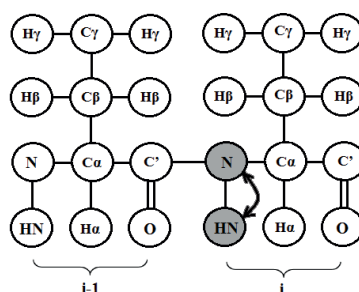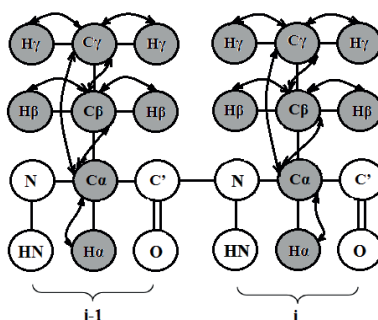

Figure III.1.3.8. Schematic representation of amino acids of a protein and the atoms involved in the effects recorded in a TOCSY-experiment.

III.1.4   NMR on proteins containing paramagnetic metal ions

Paramagnetic metal ions (pMe) contain at least one unpaired electron. The unpaired electron interacts with the nuclei of the atoms directly linked to the metal ion (contact interaction) and atoms within a radius of 10-30 Å (dipolar interaction) depending on the metal nature. Examples of pMes found naturally in proteins are Fe^{2+}, Fe^{3+}, Cu^{2+}, Mn^{2+}, Mn^{3+}, Ni^{2+} and Co^{2+}. Not native to a protein pMes could be also introduced via in vitro biochemical/chemical methods or creating recombinant proteins with pMe binding sites. A paramagnetic metal ion present in protein influences the magnetic properties of surrounding nuclei in both time-dependent and time-averaged manner. Time dependent fluctuations arising from the unpaired electrons at the metal center provide relaxation mechanisms, leading to the shorter T_1 (longitudinal) and T_2 (transverse) relaxation times, which in the case of T_2 leads to signal broadening. Time-average magnetic field significantly affects resonance frequencies of nuclei close to the paramagnetic center, which results in shifting of many signals outside the diamagnetic region. Figure III.1.4. illistrates an example of a 1D proton NMR spectrum of a protein, containing paramagnetic metal ion (MsrB1-Co, manuscript). Thus, the presense of pMe in proteins complicates their characterization in comparison with proteins containing diamagnetic metals or no metal ions, making, NMR characterization of paramagnetic protein a challenging endeavour.
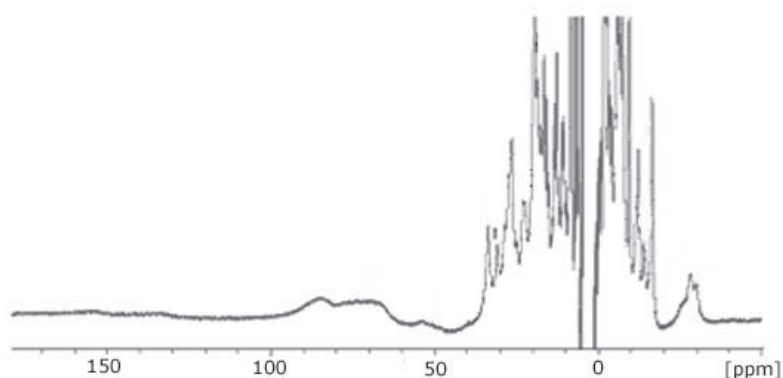
Figure III.1.4. A 1D proton NMR spectrum of MsrB1-Co (the manuscript), showing both high chemical shifts and signal broadening.

A paramagnetic shift

In a protein with pMe, the observed chemical shift for a given nucleus can be expressed as

$$\delta_{\text{obs}} = \delta_{\text{dia}} + \delta_{\text{pm}}$$

where δ_{obs} is the observed chemical shift, δ_{dia} is the diamagnetic, and δ_{pm} is the paramagnetic contribution to the observed chemical shift, respectively. The paramagnetic contribution δ_{pm} arises from the hyperfine interaction of unpaired electron(s) with the nucleus. The hyperfine interactions comprise two terms. First, the Fermi contact term (A_{iso}) is caused by delocalization of the spin density from the metal orbitals onto the ligands orbitals [81]. The chemical shift caused by Fermi contact contribution can be expressed as:

$$\delta_{F_c}^{j} = \frac{A_{iso}^{j}/h}{3\beta\gamma_j/2\pi}\left(\frac{\chi_{xx}}{g_{xx}} + \frac{\chi_{yy}}{g_{yy}} + \frac{\chi_{zz}}{g_{zz}}\right),\qquad\qquad [82]$$

where χ_{ii} and g_{ii} (ii= xx, yy, zz) are the principal components of the paramagnetic susceptibility tenzor χ and the g-tenzor. β is the Bohr magneton, γ_j is the gyromagnetic ratio of the nucleus j, and A_{iso}^{j} is the Fermi contact contribution between nucleus j and the unpaired electron spin.

The second term of hyperfine interaction, which can cause an additional paramagnetic shift, is the dipolar part. This shift has an isotropic component, which causes the pseudocontact shift δ_{pc} if the magnetic susceptibility tenzor of the metal ion is anisotropic or if the metal ion has a spin S > ½ and possesses a significant zero-field splitting. It affects the chemical shifts of both ligated and not ligated nuclei. The δ_{pc} for the spin S=1/2 could be expressed as:

$$\delta_{pc} = \frac{1}{12\pi N}\left(\Delta\chi_{ax}\left(3\cos^2\vartheta - 1\right)R^{-3} + 1.5\Delta\chi_{rh}(\sin^2\theta\cos 2\phi)R^{-3}\right),$$

$$\Delta\chi_{ax} = \chi_{zz} - 0.5(\chi_{xx} + \chi_{yy}),$$

$$\Delta\chi_{rh} = \chi_{xx} - \chi_{yy},$$

where N is Avogadro's number, $\Delta\chi_{ax}$ and $\Delta\chi_{rh}$ are axial and rhombic magnetic susceptibilities. R, ϑ and ϕ are the spherical polar coordinates of the nucleus j realtive to the principal axes of the χ-tenzor.

Both terms, the Fermi contact contribution and dipolar part vary both depend on the spin state and magnetic anisotropy of the metal ion.

III.1.5   Protein structure determination by NMR

The procedure of protein structure determination with NMR spectroscopy icludes several steps:

1. Obtaining the needed NMR spectra on the target protein.

2. Spectrum assignment, which involves determining the amino acid, its sequence position and determining which specific nucleus is responsible for a NMR peak.

3. Analysis of the structural data. Dihedral angle determination.

4. Structure calculation based on the obtained available data.

Typically all these steps are intertwined, as different spectra and spectral conditions may give different (additional) and complementary information, which may sufficiently help to solve the structure.

Structure determination of proteins is based on the distance information derived from NOE, dihedral angles based on the three-bond scalar coupling between nuclei (J^3 couplings), and from the angular information obtained with residual dipolar couplings. The NOE is caused by cross-relaxation of two spins and depends on the distance between these nuclei to the power of six. Thus, the observation of a cross-peak in a NOESY spectrum implies a close distance between the two nuclei (within 5Å). NOE distance constraints are the most important in structure determination of biological molecules by NMR.

For the purposes of current work the NOE constraints have been derrived by combining ^{15}N-^{1}H-HSQC with the 3D ^{15}N-NOESY, ^{13}C-^{1}H-HSQC with the 3D ^{13}C-NOESY, and aro-^{13}C-^{1}H-HSQC with the 3D aro-^{13}C-NOESY, which were subsequently used in structure calculation. In addition to the distance constraints, the procedure of protein structure determination was supplemented with the torsion angle constrains for the backbone were derived using program TALOS (Tosrion angle Likelyhood Obtained from shift and sequence similarity).

Structure calculation and minimization

Several computer programs have been developed for structure determination of biomolecules, one of the most used programs today is CYANA [83]. The caluculation starts from an unstructured linear protein/peptide and use torsion angle dynamics approaches to calculate structures that fulfill the constraints. The programs generate a set of random mobilities for the different residues used in the structure calculations, and the final output structure depends on the input constraints. CYANA uses a simulated annealing procedure

where the structures receive a high kinetic energy by being heated up, and, thus, able to overcome the energy barriers of the potential hypersurface. The system will gain a global minimum (the correct fold) and slowly cool down to a normal temperature. Typically about 100 structures are calculated and the 20 with the lowest energy structures will then be investigated further to check how the structures are compared to the structural constraints used for the structure calculation.

Structure calculation is often an iterative process where the results from one calculation are used to evaluate assignments and reassess the structural constraints. The quality of the structure improves while the diversity between the structures should decrease, as the work on the structure proceeds. The obtained throguh calculation final 20 structures are evaluated using constraint violations, visual inspection, and different software e.g. PROCHECK-NMR [84].

The structure calculations for NMR structures use a simplified force field, where only the most dominant part of the conformational energy is used. This is time saving for the structure calculation, however, the resulting structures may be unfavourable in terms of full 'physical' energy function. Thus, the final structures from CYANA are further energy minimized with a full force field. In the current work, the energy minimizatioin was performed by AMBER program [85].

III.2. UV-visible spectroscopy

A wide range of biomolecules absorb light in the ultraviolet and visible frequency range often at a characteristic wavelengths. As UV-visible light of a given intensity passes through a sample, a fraction of this light is absorbed by the sample. This fraction is proportional to the concentration (c) of the species responsible for absorption, the light path length through the sample (l) and the molar extinction coefficient (ε) [86]. This relationship, called the Lambert-Beer law, is the basis of UV-vis spectroscopy and is expressed as

$$Abs = \log \frac{I_0}{I} = \varepsilon c l$$

where I_0 is the intensity of the irradiating light, I is the intensity of the light leaving the sample. $\log \frac{I_0}{I}$ is called the absorbance, and is often simplified by the designation Abs (or A). The molar extinction coefficient varies with the absorbing species, the wavelength of irradiating light and solvent [87].

IV. Aims and scope of this study

The main objective of the study was to establish structural and functional pecularities of mammalian thiol redox active proteins of thioredoxin system. High-resolution NMR spectroscopy was main experimental technique through this work. Three proteins were in scope of this study: methionine sulfoxide reductase B1 (MsrB1), its functional partner thioredoxin (Trx), and Grx domain of mouse TGR (Grx).

Studies of the first two proteins were aiming to determine the amino acids involved in protein complex formation between MsrB1 and Trx. This required an oxidized form of MsrB1 along with its assignment, and a mutant form of Trx at the place of resolving cysteine (Ser35Cys), an important factor, providing generation of a stable intermolecular complex required for NMR studies. To determine the Trx amino acids participating in protein complex formation with oxidized MsrB1, the availability of the backbone assignment of Ser35Cys Trx was crucial. Once the stable protein complex was generated and analysed, further aim was to characterize the driving forces of protein complex formation, which required generation of its model based on protein-protein docking and molecular dynamics simulation.

Along with our research, knowledge available in literature regarding MsrB proteins accumulated in recent years required systematization. Aiming to summarize what is known to date about MsrBs, a review describing evolution, sequential, structural and coordination peculiarities of MsrBs was written.

Studies of Grx domain of mouse TGR were aiming to determine its three-dimensional structure, as it is a critical step in understanding of enzyme's function and catalytic mechanism. Once the structure of Grx was determined, the secondary goals were to characterize the role of certain structural elements in Grx function and in entire TGR enzyme.

Recombinant expression of MsrB1 in medium containing Co^{2+} resulted in biosynthesis of paramagnetic cobalt-containing MsrB1. Characterization of the obtained protein, in particular, the influence of cobalt coordination on enzyme's structure and function, required its NMR assignment. Upon completion of the assignment, further goal was to derive pseudocontat shift values and calculate magnetic susceptibility tenzor, a crucial constraints in structure determination of paramagnetic protein.

The obtained research results are presented in the following papers.

List of appendix papers

Paper I:

Dobrovolska O, Rychkov G, Shumilina E, Nerinovski K, Schmidt A, et al. (2012) **Structural Insights into Interaction between Mammalian Methionine Sulfoxide Reductase B1 and Thioredoxin**. Journal of Biomedicine and Biotechnology 2012: 586539.

Paper II:

Shumilina E, Dobrovolska O, Dikiy A, (2012) **Evolution of Structural and Coordination Features within Methionine Sulfoxide Reductase B Family** (in press)

Paper III:

Dobrovolska O., Shumilina E, Vadim N. Gladyshev and Alexander Dikiy, **Structural analysis of glutaredoxin domain of *Mus musculus* thioredoxin glutathione reductase,** (submitted to PLOS ONE).

Paper IV:

Shumilina E, Dobrovolska O, Rebecca Del Conte, Henrik Waldal Holen and Alexander Dikiy, (2012) **Competitive Zinc for Cobalt Substitution in Mammalian MsrB1 Overexpressed in *E.coli*: Structural and Functional Insights** (manuscript).

V. Summary of Results and Discussions

V.1. Structural characterization of the Trx-MsrB1 complex (Paper I)

V.1.1 NMR spectroscopy studies

NMR studies of stable interprotein complex MsrB1-Trx required production of mutants of both oxidized MsrB1 and reduced Trx in sufficient amounts. Since expression of selenocysteine–containing MsrB1 is impossible in bacterial cells, Sec95 of MsrB1 was replaced with Cys. To generate a stable interprotein and catalytically not active complex, thioredoxin, containing Ser35Cys mutation was obtained. To monitor amino acids participating in interprotein complex formation from both proteins, NMR titration was performed. However, prior this investigation, a backbone assignment of reduced mutant Trx, and assignment of ^{15}N-^{1}H HSQC spectrum of oxidized MsrB1, were carried out. The resulted backbone assignment of mutant Trx was deposited in BMRB under the accession number 16850. Figure V.1.1.1 illustrates a fragment of ^{15}N-^{1}H HSQC spectra of oxidized and reduced MsrB1.
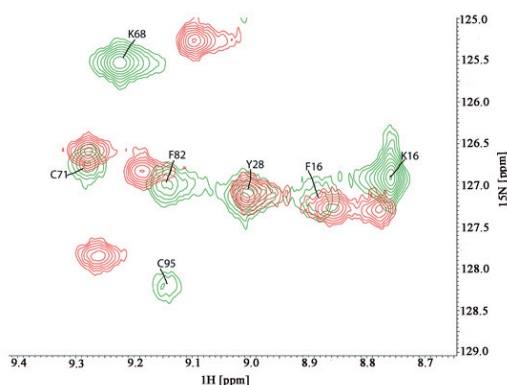
Figure V.1.1.1. Fragment of ^{15}N-^{1}H HSQC spectra of oxidized and reduced MsrB1. Green cross-peaks belong to reduced MsrB1; oxidized form of MsrB1 corresponds to red cross-peaks.

^{15}N-labeled oxidized MsrB1 was titrated with unlabeled thioredoxin and subsequent ^{15}N-^{1}H HSQC spectra were recorded. And vice versa, ^{15}N-enriched reduced Trx was titrated with unlabeled oxidized MsrB1, followed by recording the corresponding ^{15}N-^{1}H HSQC spectra. The residues, experiencing perturbations, i.e. signal broadening or/and chemical shift change were mapped for both oxidized MsrB1 and reduced Trx. Figure V.1.1.2 represents a fragment of ^{15}N-^{1}H HSQC experiments performed for Trx.
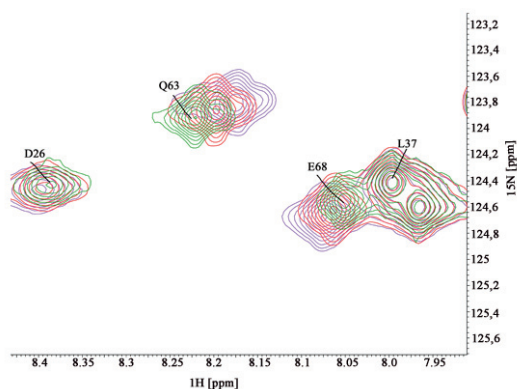
Figure V.1.1.2. ^{15}N labelled TrxC35S titration with non-labelled oxidized MsrB1. Green cross-peaks belong to free Trx; red – ratio Trx: MsrB1^{ox}=1:1; magenta – ratio Trx: MsrB1^{ox}=1:3.

The analysis of the NMR titration data is represented on the Figures V.1.1.3 and V.1.1.4.
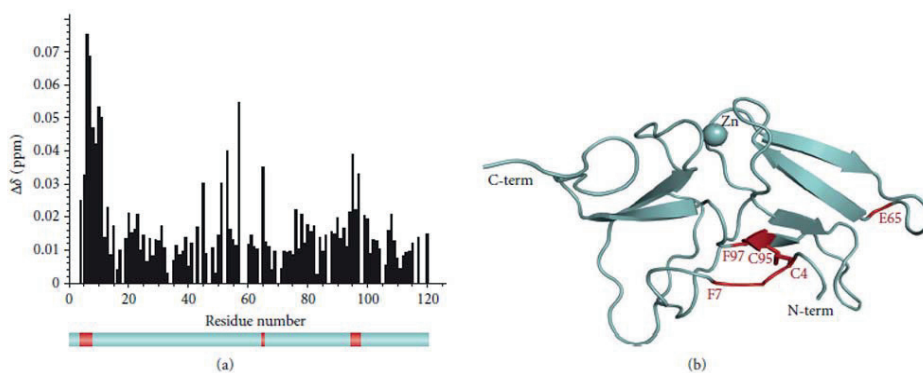
Figure V.1.1.3: (a) Chemical shift difference observed for each residue of MsrB1 upon NMR titration by Trx. Lower panel shows in red the residues belonging to the protein active site revealing maximal changes upon titration. (b) The 3D structure of oxidized mammalian MsrB1. "Hot points" used for molecular docking are evidenced in red.
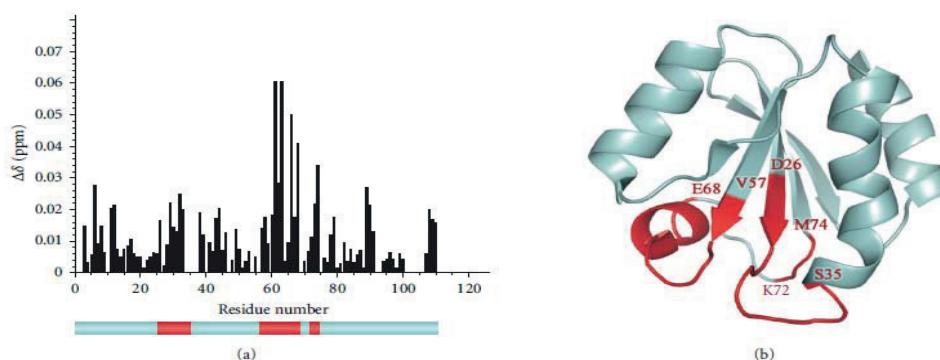
Figure V.1.1.4: (a) Chemical shift difference observed for each residue of Trx upon NMRtitration by MsrB1. Lower panel shows in red the residues revealing maximal changes upon titration. (b) The 3D structure of oxidized mammalian Trx. "Hot points" used for molecular docking are evidenced in red.

Cys4-Phe7, Glu65, and Cys95-Phe97 of MsrB1, which belong to the protein's active site and its neignbouring amino acids, were found to be involved in interaction with Trx. Throughout the Trx titration by MsrB1 Asp26-Ser35, Val57-Glu68, and Lys72-Met74 were found to participate in protein complex formation with MsrB1^{ox}. As in the case with MsrB1, the reported residues of Trx belong to its active site and are situated on the external loops close in space. The revealed segments are indicated on the Figures V.1.1.3 and V.1.1.4 with red colour.

V.1.2 Biocomputing of Trx-MsrB1 complex

The averaged structures of mutants of oxidized MsrB1 and Trx were generated using molecular dynamics simulations (Figures V.1.1.3 (b) and V.1.1.4 (b)). To identify possible options of mutual arrangement of MsrB1 and Trx protein-protein docking was performed. As criteria, a distance of 10 Å between Cys95-Cys4 disulfide of MsrB1 and S$^{\gamma}$ atom Cys32 of thioredoxin was used. Based on the obtained NMR data the 'hot points of docking' were found, and more than 1000 spatial structure orientations were generated, which were further sorted into four groups (Table 2). The representative structure of each group was subjected to 5 ns molecular dynamics simulations. The analysis of the averaged structures of the complexes obtained after 5 ns of MD simulations reveal that only for complex D the distance SγCys32 Trx–SγCys4 MsrB1 remains the smallest among all possible complexes. Therefore, the MD trajectory for this complex was extended to 20ns (Figures V.1.2 (a) and V.1.2 (c)).

Table 2. Structural parameters for MsrB1-Trx complexes (four groups) revealed by protein docking. The distances between sulphur atom of Cys32 of Trx and the nearest sulphur atom of MsrB1 disulfide bond as well as three pairs of the nearest aminoacids approaching in space for each group (to outline the spatial orientation of the proteins within a complex) are presented.

	Group A			Group B			Group C			Group D		
Distance, Å (Trx Cys32) S/S–S bond (MsrB1)	7,1			7,8			7,8			5,3		
Nearest neighbouring contacts between residues												
MsrB1 residues	W43	H39	P87	S2	W43	P87	W43	N62	P87	F6	P42	R93
Trx residues	M37	E95	S90	W31	A92	E70	M37	K72	K96	K72	D60	K36

Figure V.1.2. The snapshots (A, B, and C) of the protein complex formation along the molecular dynamics trajectory. The panels (a), (b), and (c) show zoomed views (interactions interfaces) of the snapshots (A) and (C), respectively.

According to MD calculations, the N-terminal tail of MsrB1 upon oxidation gets closer to the β-strand of the protein formed by Arg93-Ile96 (Figures V.1.2 (A) and V.1.2 (a)). Further, elongation of the mentioned MsrB1's β-strand occurs additionally involving residues I96-S98 accompanied by formation of three additional hydrogen bonds with residues S5-F7 (Figures V.1.2 (B) and V.1.2 (b)). This newly constituted strand belongs to the N-terminus of MsrB1, which now is stabilized by the intramolecular disulfide bond formed upon oxidation. At the final stage of the interaction, Trx adjusts its unstructured loop to the newly formed N-terminal β-strand of MsrB1, leading to formation of interprotein β-sheet, composed of six β-strands, where five of them belong to MsrB1 and the sixth to Trx (residues K72-M74) (Figures V.1.2 (C) and V.1.2 (c)). Thus, in developing of two new β-strands (one intramolecular MsrB1 and one intermolecular), nine aminoacids participate.

Upon the formation of the protein intermolecular complex, six new hydrogen bonds evolve (three of them are interproteins), thus stabilizing the intermolecular complex. Analysis of MD trajectories for MsrB1-Trx complex (prior formation of the intermolecular disulfide bond) indicates an existence of two conformations of C4-C95 MsrB1 bond. The following energies have been obtained (kcal/mol): 3.1 and 3.2 (vw), −0.5 and −1.7 (to), 2.6 and 1.5

(total) for "strained" and "relaxed" S– S bond, respectively. During the MD simulations this MsrB1 intramolecular S–S bond prefers its "strained" conformation. This bond thus becomes totally destabilized by 2.6 kcal/mol for "strained" or by 1.5 kcal/mol for "relaxed" conformation. However, the presence of six hydrogen bonds inside the developed interprotein β-sheet obviously overcompensates the straining of Cys4-Cys95 bond in MsrB1, as energy profit from one hydrogen bond formation in β-sheet can be estimated by −0.78 kcal/mol. Therefore, the formation of the interprotein MsrB1-Trx complex becomes energetically preferable.

V.2. Evolution of Structural and Coordination Features within Methionine Sulfoxide Reductase B Family (paper II)

Different classes of Msrs

There are two main classes of Msrs: MsrA – which can reduce protein-bound or free S-epimer of methionine sulfoxide [88,89,90,91,92,93,94] and MsrB – which reduce both the protein-bound R-epimer of MetO, and, with less efficiency, free R-MetO [54,95,96,97]. Since each of the Msr classes has its own, different from each other three-dimensional fold, and significantly distinct, well conserved amino acid sequence it is possible to refer to the convergent evolution for MetO reduction in living organisms [98,99,100,101]. MsrA protein sequences, are characterized by the invariable signature motif "GCFWG/C'' [50,102,103] and MsrB - "RXCXN" (where X indicates any amino acid) or in case of MsrB1 - "RXUXF" [53,99,104]. Structures of some archaea, bacterial and eukaryotic MsrB proteins have previously been determined using either X-ray or NMR techniques (Figure V.2.1).

Both MsrA and MsrB were found in the most genomes [53], however, their distribution within the life kingdoms is quite different. Without exception, all studied eukaryotes and cyanobacteria contain msrA and msrB genes [101]. In prokaryotes, MsrA and MsrB genes can form two separate transcription units [100], with not adjacent loci on the chromosome. Alternatively, genes can either be transcriptionally [105] or translationally fused and form in the last case the two-domain protein - MsrAB [106]. In prokaryotes, the Msrs distribution is rather variable: in some thermophile or anaerobic archaea no Msrs are present. Some organisms possess both Msrs, while others only MsrA but no MsrB. However, there are no MsrB-containing organisms which would also not contain MsrA [101]. A possible explanation for the different MsrA/MsrB distribution in diverse organisms might be the existence of an enzymatic stereospecific preference for S-epimers during Met oxidation in

cells, or, alternatively, a more damaging effect of S-MetO for the cell. Otherwise, MsrA and/or MsrB might have other unknown biochemical functions along with protein repair, which would account for their unequal distribution in genomes [101].

The number of msrA and msrB ortholog copies also greatly varies in the different organisms. For example, *E.coli* contains one copy of msrA and msrB, *Arabidopsis* harbors five msrAs and nine msrBs, *Rhizobium meliloti* has three msrA and three msrB genes, while mammals have one msrA and three msrB genes, i.e. MsrB1, MsrB2 and MsrB3 [53,54,93,107,108,109]. Gene duplication is a frequent event in genome evolution across all three domains of life. The organism's necessity to have alleles expressed under different conditions might be a possible explanation of this fact [100,110,111,112].

MsrB subcellular distribution in eukaryotic cells

Although MsrB1 does not have clearly predictable nuclear TS, it is rich in positively charged residues, which could serve to transfer the protein into the nucleus. Both MsrB2 and MsrB3B form of MsrB3 are targeted to mitochondria. Normally, mitochondrial TS are represented by about 20–60 amino acid residues with abundant positive charges and frequent hydroxylated residues [113]. Targeting sequences are predicted to form amphipathic α-helices in membranes or in membrane-like environments, whereas in aqueous solution they show little structural organization [114,115,116,117,118]. Mouse MsrB2 contains a typical mitochondrial signal with the high proportion of arginine residues at the N-terminus (MARLLRALRGLPLLQAPGRLARG) [54].

MsrB3 should be considered separately. In humans and other species (e.g. zebrafish) MsrB3 were found to be translated in both ER (MsrB3A) and mitochondrial (MsrB3B) forms due to alternative splicing that produces contrasting ER and mitochondrial signals [54]. It should be noted, that the mitochondrial location of the human MsrB3B, was not immediately clear because it had contrasting N- and C-terminal signals [119]: on the N-terminus – mitochondrial and on the C-terminus – KDEL-like ER signal [54]. In contrast, rodents have only one form of MsrB3 found only in the ER [119]. At the same time, mouse MsrB3 has both ER and mitochondrial signal peptides at the N-terminus: the mitochondrial signal is located between the ER signal and the common MsrB domain. The role of the mouse MsrB3 mitochondrial signal, which is functional if placed as an N-terminal sequence, remains unclear and requires further research [119]. MsrB3A and rodent MsrB3 are targeted to ER with C-terminus (KDEL-like) ER retention sequence (KAEL for human and RAEL for mouse MsrB) [54,119].

Mitochondria, the major source of ROS in the cell, have two different MsrBs: MsrB2 and MsrB3B. MsrB2 is the most active at lower concentrations of Met and inhibited by high concentrations of the substrate, whereas MsrB3B is most active at concentrations of methionine sulfoxide more than 1mM [54]. In addition, these two MsrBs show differential tissue expression [120]. The above argumentation might probably explain the occurrence of two MsrBs in mitochondria.

Selenocysteine in Msrs

Selenoproteins are typically more active than their cysteine (Cys) mutants due to the higher selenium nucleophilicity [19]. This high catalytic activity has been regarded as a key reason why Sec is used in biological systems [20,21,22,23]. However, the nearest protein environment of Cys or Sec, as well as the overall structure of the protein can significantly influence the redox potential of both residues, and the small changes in the amino acid composition in the region close to the active site can lead to comparable catalytic efficiency of the sulfur homolog of a selenium-dependent enzyme [121]. Thus, only a higher activity of selenoproteins cannot be unique reason to use Sec instead of Cys. Sec utilization by living organisms is a compromise: on one hand, Sec utilization provides higher activity, a broader range of substrates in microenvironmental conditions (e.g. pH) where an enzyme is active. On the other hand, there is a limitation imposed by the electron donors (or acceptors, depending on the reaction), dependence on selenium, and the availability and complexity of the Sec insertion system [121].

Seleno-containing MsrBs are found in invertebrates and vertebrates, but not in bacteria or plants [53,122]. Mammals have one selenoprotein MsrB (MsrB1) and two Cys-containing homologs (MsrB2 and MsrB3). MsrB1 contains selenocysteine in the place of the catalytic cysteine residue normally present in other MsrBs [53,54]. MsrB1 is the most active in the reduction of MetO among all members of the MsrB family. The mechanism of a catalytic MetO reduction by Msrs employs a sulfenic acid chemistry and S-S/Se-S formation [56,57,58,59,60].

MsrB structural description

Generally, the structure cores of all catalytic and structurally characterized MsrBs have a highly similar β-fold consisting of two β-sheets, one with three strands and another with five strands (Figure V.2.1). However, there is a dramatic variability with regards to length and secondary structure composition of N- and C-terminal regions of these proteins.

This can be clearly observed from superimposition of MsrB structures [123]: while the β-core is well conserved, the terminal parts of the proteins show significant differences.
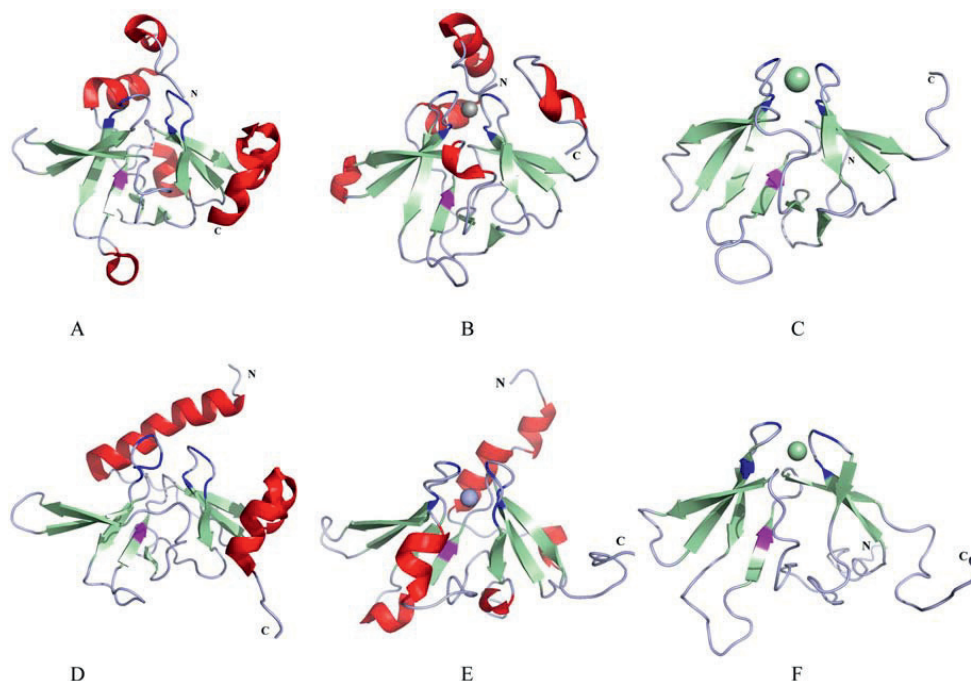
Figure V.2.1. Three-dimensional structures of proteins belonging to MsrB family.
A – MsrB *N. Gonorrhoeae*, 3HCH, [99]; B – MsrB *X. campestris*, 3HCI, [124]; C – MsrB1 *H. sapience*, 3MAO, not pubblished; D – MsrB *B. subtilis*, 1XM0, not pubblished; E – MsrB2 *M. musculus*, 2L1U, [125]; F – MsrB1 *M. musculus*, 2KV1,[126].

Comparative structural analysis revealed the differences between selenoprotein and non-selenoprotein MsrBs, which corresponds to different mechanisms employed by these enzymes. The presence of α-helical structures in the N-terminal region of bacterial MsrBs suggests that large structural alterations during catalytic act are unlikely in these enzymes due to rigidity imposed by α-helices. Instead, the absence of any secondary structure in the N-terminus of Sec-containing MsrB1 does not impose any rigidity constraints. As a result, the catalytic reaction in Sec-containing MsrB1 occurs through the formation of internal selenide-sulfide bridge between the catalytic Sec and the resolving Cys situated on mobile N-terminus, while the reaction between methionine sulfoxide and other MsrBs occurs through formation of sulfenic acid intermediate, and any disulfide bridge is formed within this reaction [127]. Such peculiarities of catalytic reactions between Sec- and Cys-containing proteins are the result of evolutionary changes of catalytic Cys to Sec accompanied by other adaptations such as distinct resolving Cys and flexible N-terminus.

Another aspect of structural comparison of MsrBs (Figure V.2.1) is that mammalian MsrB2 is more similar to bacterial MsrBs rather than to mammalian MsrB1 as it contains three α-helices on the exterior of the protein in the N-terminal region. Indeed, the proposed catalytic mechanism for mammalian MsrB2 protein is the same as that proposed for bacterial MsrBs [125]. This observation indicates that the functional similarities within MsrB family are determined not by the organism, but rather by the nature of the catalytic residue and the protein secondary structure.

Zinc ion in MsrB

Sequence analysis of MsrBs homologs shows the existence of proteins with two different set of metal binding sites: one with the "classic" four Cys ligands for zinc binding and another with different "not classic" amino acids at the correspondent coordinating positions [128]. It was suggested the existence of a strictly Zn-dependent family of proteins and a family that normally has Zn in metal binding site with four cysteines, but it can be expressed also as zinc-independent forms where cysteine ligands are substituted by other residues [129]. Thus, for some protein family the same protein may exist either with zinc or without.

The majority of MsrB proteins, including MsrB1, MsrB2, MsrB3 and the most part of bacterial MsrB, have four cysteines that bind Zn ion [103,130]. Some of the bacterial MsrB subclasses, however, have only Zn-independent forms (e.g. MsrB domain of MsrAB from *Neisseriaceae*, *Bacillus* sp. and *St. pneumoniae*), some have only Zn-containing forms (MsrB from *B. pseudomallei*, *X. campestris*) and others - both forms (MsrB from *S. meliloti; V. cholera*) (unpublished result from [128]). Several structures of MsrB proteins are presently reported (Figure V.2.2). Some of them are characterized by the presence of Zn ion, while other lacks it. From the structural comparison (Figure V.2.2), it is evident that both Zn-containing and Zn-independent forms have identical structure of metal binding site among themselves and to the corresponding region of no metal homologs. All of them have two unstructured loops between four β-sheets. This region seems very well conserved in all the reported structures. In the cases where Zn ion is absent, it is possible that some other bonds and not the Zn ion preserve the fold of the region. Some Zn-binding site studies were carried out by different laboratories [103,130]. In these studies it was shown that in MsrB from *Drosophila* and *E.coli* mutation of Cys at the positions 23, 26, 71 and 74 (numbering corresponds to MsrB1 from Mus musculus) on GSGS, DSSS and DSSA results in Zn losing by protein. The reversible mutation of DSSS pattern to CCCC results in a more stable, Zn-containing protein. It was suggested, that the role of the metal in Zn-containing MsrBs is to

stabilize the core structure and thus, to adopt the active site conformation for efficient reductase activity. However, as it was stressed, this assumption is in apparent contradiction with the fact that the bacterial MsrB, lacking the two CXXC signatures is as active as the Zn-containing MsrB [130]. It was suggested that a more careful analysis of stabilizing elements for MsrBs that do not contain CXXC motif should be performed.
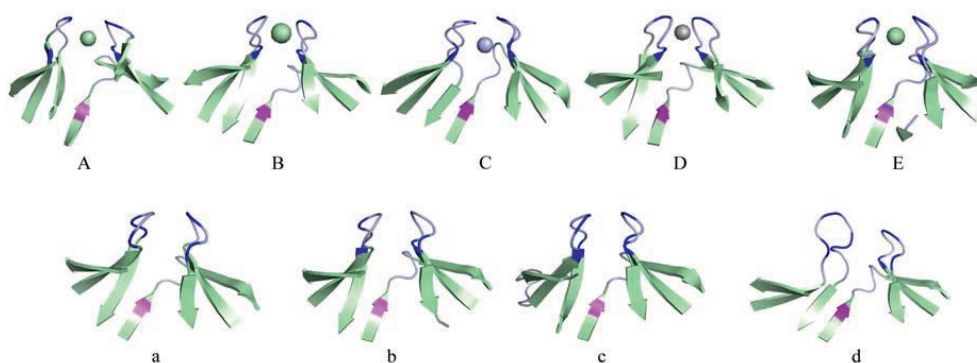
Figure V.2.2. Three-dimensional structure of Zn-binding site and thecorresponding regions of MsrBs family. For the explanation see text. A - MsrB1, *M. musculus*, 2KV1, [123]; B - MsrB1, *H. sapience*, 3MAO, not published; C - MsrB2, *M. musculus*, 2L1U, [125]; D - MsrB, *X. campestris*, 3HCI, [124]; E - *B. pseudomallei*, 3CEZ/3CXK, not published; a – MsrB, *N. meningitidis*, 3HCH, [124]; b - MsrB , *N. gonorhoeae*, 1L1D, [99]; c-MsrB, *S. pneumoniae* 3E0O, [131]; d – MsrB, *B.subtilis*, 1XM0 not published. Blue color: cysteine and "not classic" ligands in metal binding site; magenta: catalytic cysteine.
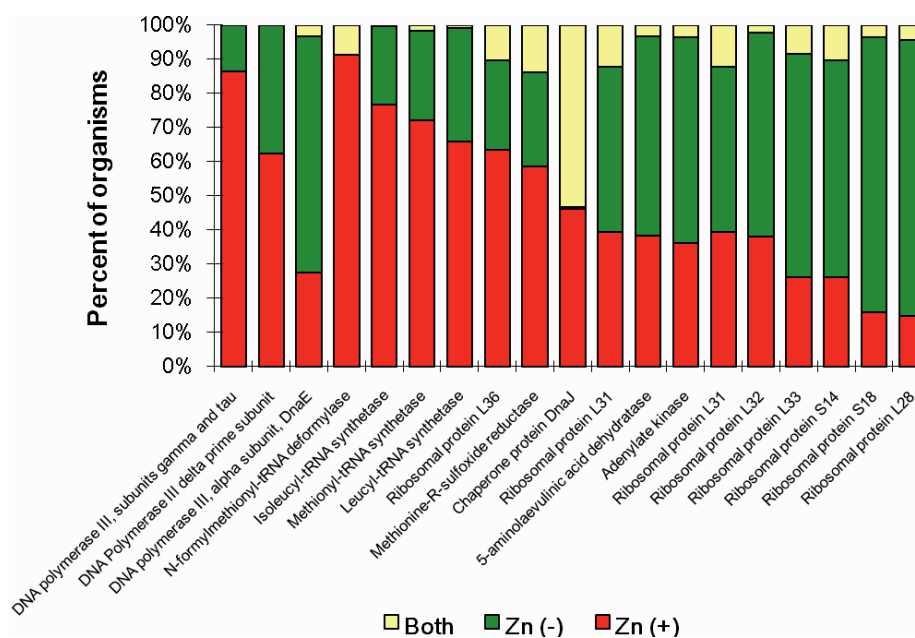
Figure V.2.3. Occurrence of Zn-dependent and Zn-independent forms of representative Zn protein families in bacteria. Zn (+), organisms only containing Zn-dependent form; Zn (-), organisms only containing Zn-independent form; Both, organisms containing both forms. The figure is taken from [128].

V.3. Structure determination of Grx domain of mouse TGR (paper III)

The determination of Grx structure, for which full assignment was already performed and deposited (BMRB accession number 17636), required the obtainment of NOE distance constraints. For this purpose the C^{13} and N^{15}-labeled Grx was recombinantly expressed and purified, and subsequently the C^{13}-^{1}H-HSQC with 3D C^{13}-NOESY, ^{15}N-^{1}H HSQC with 3D ^{15}N-NOESY, and aro C^{13}-^{1}H-HSQC with 3D aro-C^{13}-NOESY NMR spectra were aqcuired, and analysed. To obtain the latter two spectra, the sample buffer was exchanged to D_2O, which allowed to monitor signals belonging to aromatic residues buried in protein core. NOE cross-peaks of the first 22 amino acid residues of Grx were beyond detection (it was previously established that this region is mobile), therefore, the acquired spectra were assigned within the range of Ala23 to Asp124, thus, excluding the N-terminal part. The closest structural analog, Grx of human TGR (pdb code 3H8Q), was used as a reference. As a result, total 894 NOE geometrical constraints were derived, of which 733 were both long- and medium-range distance constraints. In addition to distance constraints 182 dihedral angle constraints for Grx were obtained from TALOS. Both distance and dihedral angle constraints were further imported into CYANA2.1 for structure calculation. The final 20 structures were further energy minimized in AMBER 9. The 20 structure conformers of Grx obtained in CYANA were deposited in PDB under the code 2lv3. Figure V.3 illustrates the obtained superimposition of 20 minimized conformers with the lowest target function and the ribbon representation of the minimized conformer closest to the mean structure of Grx with the secondary structure elements.
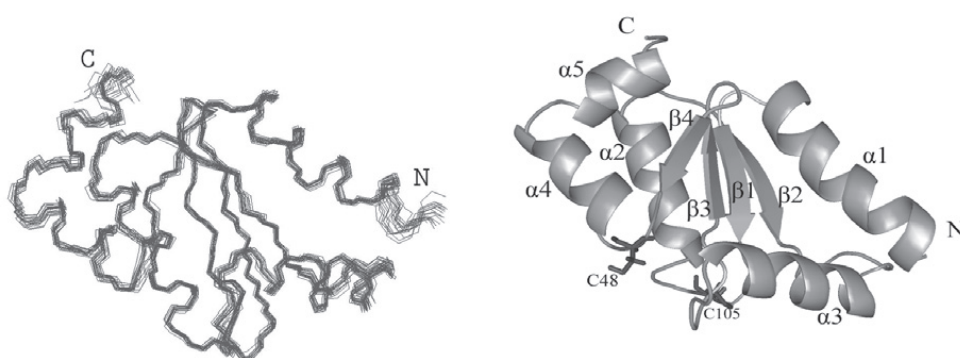
Figure V.3 Solution structure of the reduced Grx domain of mouse TGR. Left: overview of backbone superimposition of 20 conformers with the lowest target function. Right: overview of the ribbon representation of the minimized conformer closest to the mean structure, the active site cysteine (C48) and C-terminal C105 are marked.

Grx domain is a compact Trx-like spherical molecule with a central core of four-stranded β-sheets flanked on either side by five α-helices arranged in the order α_1-β_1-α_2-β_2-α_3-β_3-β_4-α_4-α_5 (Figure V.3). The N-terminal region begins with an α_1 (residues Arg 24 - Glu 36), followed by β_1 consisting of residues Val 40 to Ser 44. The active site Cys 48 - Ser 51 (-CPHS- motif) is situated on the unstructured loop between β_1 and α_2 (residues Arg 53 – Ser 59). The strand β_2 comprises residues Asn 66 to Glu 69; following a loop, α_3 consists of residues Gly 76 to Ser 87, followed by β_3 (Asn 94 – Val 97) and β_4 (Val 100 – Gly 103). The C-terminal region includes α_4 (residues Arg 107 – Asn 114) and α_5 (residues Leu 116 – Leu 120), connected through a hinge section. Strands β_1 and β_2 are parallel, and strand β_3 is antiparallel with β_1 and β_4. Helices α_1 and α_3 pack on one side of the β-sheet, whereas α_2, α_4 and α_5 are on the other. Packing of the sandwich-like architecture is mainly maintained by hydrophobic interactions between the sheet and helices. The determined Grx structure shares significant structural similarity with the modelled Grx domain of mouse TGR.

V.3.1 Analysis of the N-terminal part of Grx

The N-terminal region was left beyond the structure calculation because of lack of corresponding NOE signals. However, bioinformatics studies of this region predicted that it features a mitochondrial targeting sequence. Mitochondrial origin of this sequence was further supported by our NMR data. Being insoluble in water solutions, mitochondrial signals in membraine-like environment are predicted to form amphipathic α-helixes. To examine the N-terminal part of Grx, the protein was treated with negatively charged detergent SDS, a substance which is widely used in NMR for preparation of membrane-like media. Subsequently recorded ^{15}N-^{1}H HSQC spectra clearly demonstrated the appearance of a number of new cross-peaks (Figure V.3.1.1), which indicates on formation of protein's secondary structure. Aiming to improve the quality, the temperature for acquiring the spectra was slightly increased up to 30°C and 42°C. The number of the new cross-peaks corresponded to the expected for the N-terminal part. This observation proves structurization of the positively charged N-terminus in negatively charged environment.
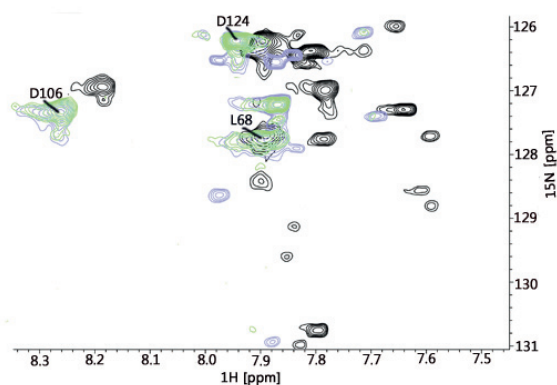
Figure V.3.1.1. A fragment of ^{15}N-^{1}H HSQC spectra of the reduced Grx domain of mouse TGR. Green shows Grx HSQC spectrum at 30°C, light blue shows Grx HSQC spectrum in the presence of SDS at 30°C, and black corresponds to Grx HSQC spectrum in the presence of SDS at 42°C.

Interestingly, the full-length protein assignment was lacking several amino acids (C105, D106, D74, A77), because of the broadening of corresponding NMR signals. However, when the N-terminal part was cut, these signals were assigned. Based on this finding it was porposed that the N-terminus could have an impact on these residues through electrostatic interactions. Figure V.3.1.2 shows electrostatic potential calculated for the obtained structure of Grx. As can be seen, the four missing amino acids (C105, D106, D74, A77) are positioned in the negatively charged region, which could be suggested to interact with the positively charged N-terminus, and therefore lead to the broadening of NMR signal. This hypothesis could explain why these signals were observed in the short form of Grx.
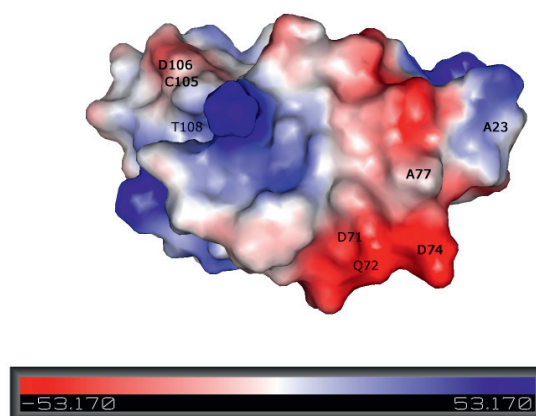
Figure V.3.1.2. Surface charge distribution of the reduced Grx domain of mouse TGR.

Another interesting observation was found during the D$_2$O exchange experiments, performed for the full-length Grx and its short form. During the first 30 minutes in D$_2$O, Grx

exchanged ten residues more with respect to sGrx (D33, G37, N38, V40, S59, V63, D71, Q72, E85, T108). However, after 3 hours of incubation in D_2O, both Grx and sGrx reveal an similar pattern of exchanged/not exchanged residues. Altogether these experiments showed that although the final rate of water exchange is the same both for Grx and sGrx, the short term dynamics of the water exchange is different for Grx and sGrx. The observed differences mostly regard residues belonging to the negative patch involved in the suggested interaction with the positively charged N-terminus (see above). The fact that these residues exchanged within the first 30 min in Grx, while in sGrx they exchanged only 3 hours later indicates that the N-terminus in some way promotes faster rates of water diffusion into Grx protein.

V.3.2 Characterization of GSH binding site

Sequence alignment of the domains with other Grx and Grx domains revealed a characteristic GSH-binding site, which was further characterized with the help of NMR. In particular, the NMR titration experiments were performed, where ^{15}N-labeled Grx was titrated with unlabelled both oxidized and reduced glutathione. Residues experiencing chemichal shift change and/or signal broadening upon titration with the substrate were detected and analysed. As a result, the NMR titration demonstrated that Grx binds oxidized and reduced glutathione *via* the same residues. Fragments of ^{15}N HSQC experiments and the analysed data are represented on figure V.3.2.
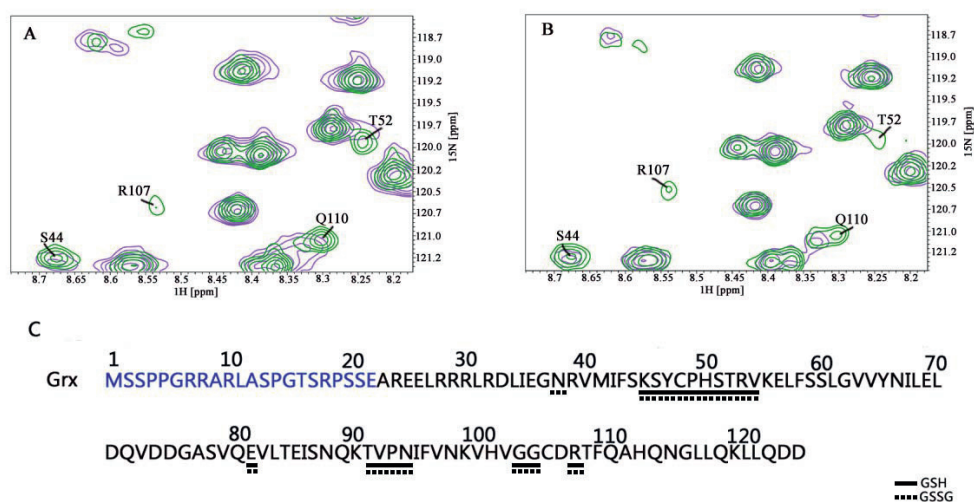
Figure V.3.2. Fragments of ^{15}N-^{1}H HSQC spectra of ^{15}N-labeled reduced Grx domain of mouse TGR titrated with unlabelled GSH and GSSG (panels A and B, respectively). Green corresponds to free Grx and magenta to Grx incubated with GSH/GSSG. Only the residues for which alteration of NMR parameters upon titration was observed are marked. Panel C: qualitative representation of the data. Solid and dashed horizontal lines below the

Grx amino acid sequence highlight the residues interacting with GSH and GSSG, respectively, and, as can be seen, significantly overlap. The N-term of Grx is marked in blue.

V.4. Competitive Zinc for Cobalt Substitution in Mammalian MsrB1 Overexpressed in *E.coli*: Structural and Functional Insights (Paper IV)

MsrB1 recombinantly expressed in *E.coli* cells in M9 media, supplemented with cobalt chloride, resulted in light blue colour of the protein. The presence of cobalt ion within MsrB1 was proven by its 1D proton NMR spectra, which contained several paramagneticaly shifted signals, and UV-visible spectra (Figure V.4A), which exhibited absorption bands at 345, 635, 665, 695 and 725 nm, characteristic for CysS→Co^{2+} LMCT band, and *d-d* transition bands of Co^{2+} respectively diagnostic of tetrahedral coordination geometry.
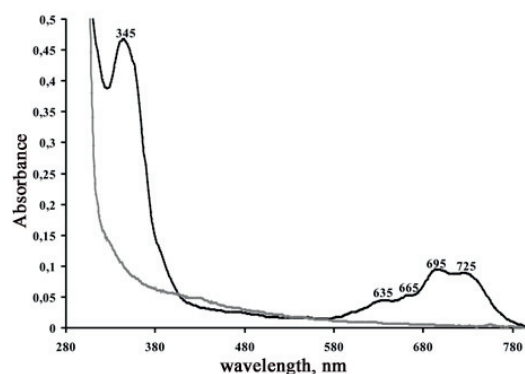
Figure V.4A. The electron UV-visible absorption spectra of reduced MsrB1 protein at pH=5.5; black – MsrB1-Co, grey – MsrB1-Zn.

The fact that cobalt ion is coordinated to MsrB1 protein indicates that Co^{2+} is present in MsrB1 either at the same position as Zn^{2+} ion replacing it, or cobalt binds to another protein site, and the resulted MsrB1 protein containes both metals. A pattern of ^{1}H-^{15}N HSQC spectrum MsrB1, expressed in zinc-containing medium was different from the one, expressed in cobalt supplemented medium. However, ^{1}H-^{15}N HSQC pattern of their equimolar mixture was identical to ^{1}H-^{15}N HSQC of MsrB1 expressed in medium containing both zinc and cobalt. This observation allowed to conlcude that when both metals are present in the medium the expression of two forms MsrB1-Zn and MsrB1-Co takes place.

Activity measurements performed for both MsrB1-Zn and MsrB1-Co (reduction of MetO) showed similar results (Figure V.4B).
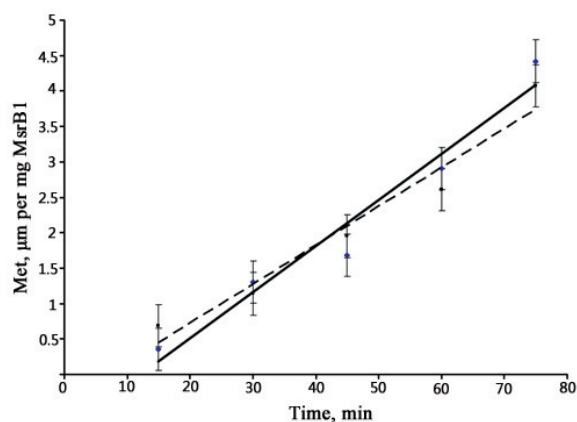
Figure V.4B. Enzyme activity of MsrB1-Co (dashed) and MsrB1-Zn (solid).

The fact that the two protein forms have similar level of activity may indicate the same fold. To structurally characterize MsrB1-Co NMR spectroscopy was further used.

V.4.1 NMR studies of MsrB1-Co

To observe all paramagnetically influenced NMR signals of MsrB1-Co protein, ^{1}H NMR spectrum was recorded using proton dedicated probe at 600 MHz spectrometer, which possesses higher sensitivity and allows to increase the number and intensities of proton signals in ^{1}H NMR spectrum. The full ^{1}H NMR spectrum of MsrB1-Co derivative is reported in Figure V.3.1.1. As can be observed from the figure, the paramagneticaly shifted signals can be found as far as of 350 ppm and they are as broad as 16000 Hz (signal at 347 ppm). As it is known, paramagnetic Co^{2+} ion has electronic configuration of *d7* and can be found in its either low (S=1/2) or high spin (S=3/2). The pattern observed in ^{1}H NMR spectrum of MsrB1-Co (Figure V.4.1) according to the available literature data [132] indicate that in the case of MsrB1-Co high spin cobalt ion is coordinated to the protein.
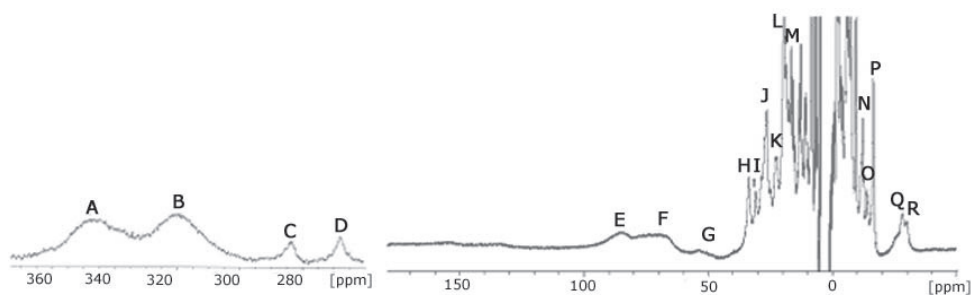
Figure V.4.1. ^{1}D NMR spectrum of MsrB1-Co, acquired at 600MHz, 298K.

V.4.1.1.   Heteronuclear NMR assignment of MsrB1-Co

To structurally characterize MsrB1-Co heteronuclear assignment of the protein was performed. Of the polypeptide backbone, 64% amino acids were assigned (79 of the total 124), side chains assignments of MsrB1-Co covered 62 out of 79 residues for which backbone assignment were accomplished, which constitute 50% of the entire protein.

Through the analysis of D_2O exchange experiments it was identified 18 residues with HN with very slow solvent exchangability rate: Phe 12, Val 20, Tyr 21, Ala 45, Phe 46, Ile 50, His 51, Val 55, Lys 61, Leu 67, Phe 82, Phe 94, Phe 97, Ser 100, Leu 101, Val 104, Ala 112 and Leu 117. Lack of exchangability with deuterium is generally an indication that the HN are involved in a strong hydrogen bond, typical for residues which are part of a tightly packed hydrophobic core of the protein [133]. Consistent with this, all the identified residues have hydrophobic side chains.

In the downfield region of the ^{1}H NMR spectrum there are eight signals which are the broadest (in the spectrum) and shifted from 80 to 350 ppm signals signals A-H (Figure V.4.1). Since the four cysteines making up the zinc-binding site in the native structure were unassigned (S1), it was supposed that the mentioned signals belong to eight β-CH$_2$ protons of four cysteines coordinating cobalt ion as they are the closest to paramagnetic center, and therefore, should have the maximal both contact and dipolar paramagnetic contribution to chemical shifts and relaxation values among all MsrB1-Co protons.

The undertaken attempts to find pairwise assignment within these eight extremely broad signals through ^{1}D NOE spectra did not result in observation of any connecting constraints. However, some NOE's among the observed paramagnetic signals were observed as shown in Figure V.4.1.1. The signals A-H tentatively assigned as β-CH$_2$ protons of four coordinating cysteines, while much sharper signals, observed upon saturation of signals C and D (Figure V.4.1.1B), tentatively assigned as Hα proton signals of the same Cys residue to which belong the saturated signals. In addition, the fact that upon saturation of signals C and D different NOEs are observed indicates that signals C and D belong to different cysteines coordinating Co ion.
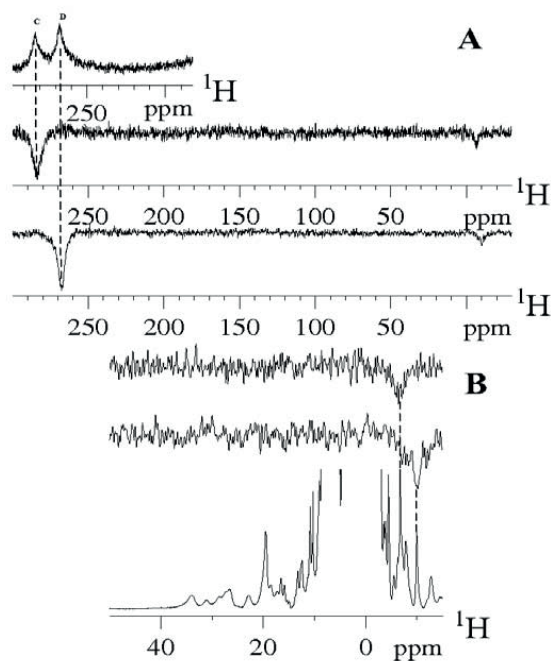
Figure V.4.1.1. A) ^{1}D NOE spectra saturating signals C and D. B) Signals, observed upon saturation of C and D signals, and assigned as Hα protons of the same Cys residue to which belong the saturated signals.

V.4.1.2.   Comparative analysis of MsrB1-Zn and MsrB1-Co

Comparison of ^{1}H and ^{15}N resonanses of pure zinc– and pure cobalt- MsrB1 derivatives showed that generally almost all NH and N resonances within the protein changed their chemical shifts passing from MsrB1-Zn to MsrB1-Co. This may mean that MsrB1 either structurally changes upon Zn for Co substitution or the latter paramagnetic ion strongly affects all chemical shifts within MsrB1 protein while the structure of the protein by itself remains invariable. Figure V.4.1.2. illustrates the chemical shift differences observed for amidic protons as a function of their distance from cobalt ion (supposing that the cobalt ion takes the same position as Zn ion in MsrB1 protein).
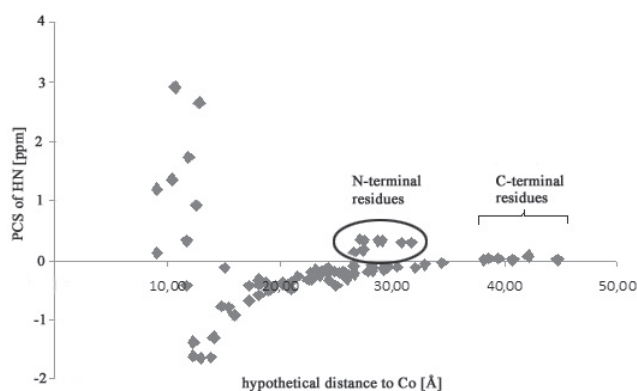
Figure V.4.1.2. A plot of PCS versus the distance from NH group of each MsrB1 residue to zinc ion (hypothetically to Co ion).

This behavior is in agreement with the expected pseudocontact paramagnetic contribution to the observed chemical shift which is reversed to the distance between the nucleus and paramagnetic center in power of three [132]. However, several amino acids (T49, I50, H51, D53, S54, V55, F97, S99, S100, and L101) diverge from this rule. These residues constitute two unstructured loops residing on the opposite site of the two structural CxxC-motifs, demonstrated to be responsible for cobalt coordination. Also, N-terminal residues, which have slightly positive PSCs, were found to deviate from the rule. This may indicate that N-terminus spends significant time in proximity of the metal binding site. Indeed, the superimposition of all the 20 conformers of MsrB1 revealed fixed core and many allowed conformations for N- and C-termina (Figure V.4.1.3, [126]), including the proximity to the metal binding site.
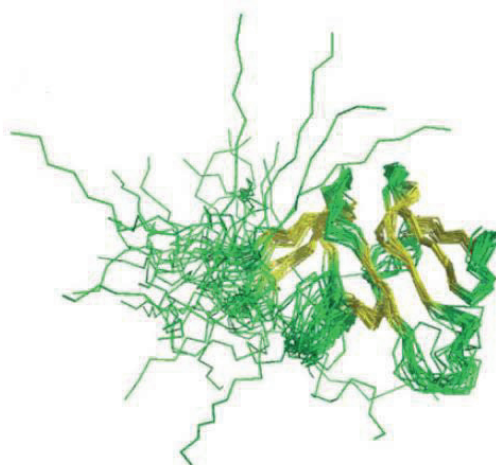
Figure V.4.1.3. The 20 calculated conformers for MsrB1 with lowest target function overlayed. β-strands indicated in yellow and loops and flexible chains - in green.

Also, as the catalytic activity assumes the formation of a disulfide bridge between Cys95 and Cys4, it is expected that the latter, residing on the N-terminus, would tend towards proximity of the catalytic Cys, positioned near the metal binding site.

V.4.1.3.   Obtainment of the structure of MsrB1-Co

For the obtainment of the structure of MsrB1-Co the minimized MsrB1-Zn (2kv1) and the experimental pseudo contact shift values were used. The obtained structure of MsrB1-Co was found to be eventually very similar to that of MsrB1-Zn. The active site presents little reorientation of the side chain of the coordinating cysteins and a movement of about 1 Å for the metal ion. Figure V.4.1.3.1 illustrates the superimposition of MsrB1-Zn (grey) and MsrB1-Co (cyan) active sites.
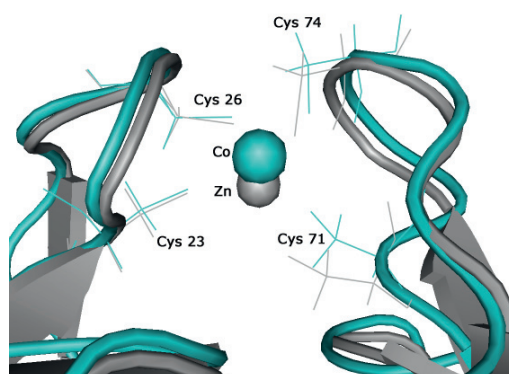
Figure V.4.1.3.1. Structural comparison of metal-binding site of MsrB1-Zn (2kv1, grey) and MsrB1-Co (cyan), obtained with the minimization through AMBER program on the base of the experimental pseudocontact shift values.

The magnetic susceptibility tensor calculated for MsrB1-Co protein (χ_{ax}=15.19 x10^{-32} m^3 and χ_{rh}=-0.97 x10^{-32}m^3) showed a good agreement between the experimental and calculated pceudocontact shift values (Figure V.4.1.3.2), which is an indication of a good quality structure of MsrB1-Co. The identified tensor presents a prevalent axial component which is in accordance with a disordered tetrahedral high-spin Co^{2+} coordination reported for other systems like Co-rubredoxin and Co-desulforedoxin [134].
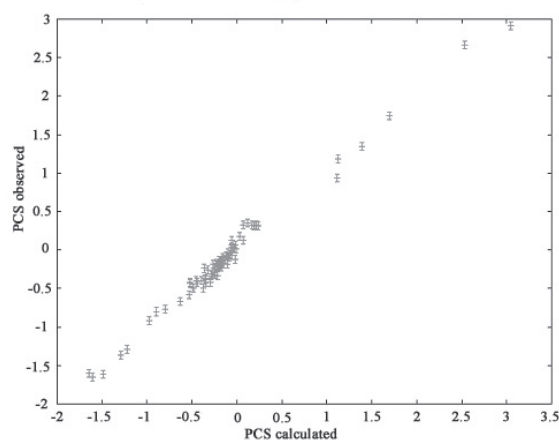
Figure V.4.1.3.2. Plot of the observed vs calculated PCSs in the minimized MsrB1-Co structure.

The performed study describes a phenomenon of Co for Zn substitution in mammalian protein, MsrB1, under expression in prokaryote cells (*E.coli*). The uptake of the correct metal by the individual polypeptide chains in bacterial cells is achieved by the metal ion concentration gradient in the cytosol and periplasm, decreasing and increasing with the Irving-Williams series. The zinc sensory protein (which regulates zinc efflux) has dissociation constant of magnitude 10^{-12} M, which means that the zinc concentration above 10^{-12} M keeps the zinc efflux protein active. Conversely, the dissociation constant of the cobalt sensory protein is in the magnitude of 10^{-7} M [135]. Thus, the intracellular concentration of cobalt ion is 100,000 times higher than for zinc to begin cobalt efflux, which compensates for the lower affinity to ligands. In addition, the ionic radius of zinc in tetrahedral coordination is closer to cobalt than to any other metal in the Irving-Williams series (0.60 and 0.58 A respectively) [136]. Perhaps that the Co^{2+} uptake by MsrB1 upon expression in *E.coli* is due to both intracellular conditions created by *E.coli* and the zinc ionic radii similarity. However, the question whether the same phenomenon is relevant for mammalian cells remains and requires further studies.

VI. Further work

VI.1. Structural analysis of the mammalian TGR function

The determined structure of Grx allowed characterization of its glutathione binding site. It was also found that Grx has the same binding site for both oxidized and reduced glutathione. However, structural characterization of the GR activity of TGR, its catalytic mechanism and the level of Grx involvement in this process remains to be estimated. The knowledge of the structure of the TR module, and, thus, the whole TGR enzyme, along with applied biocomputing tools would help to understand the full catalytic mechanism of this enzyme.

VI.2. Characterization of bacterial MsrBs using Cobalt as a probe

Since the presense of zinc ion in MsrB proteins was predicted by bioinformatics based on the two Zn-binding CXXC motifs, many of the bacterial MsrBs, that contain different well-conserved amino acids at the positions of the four cysteines, were reported to not bind zinc ion. However, the structures of metal binding site in zinc-containing protein and the corresponding region in MsrBs without metal are very similar. The reported MsrB structures lacking zinc ion less tightly coordinate it by non-cysteine residues and simply loose zinc ion upon purification. We demonstrated that MsrB1 easily substitutes zinc for cobalt upon expression in cobalt-supplemented medium; the same method could be applied to study other MsrB proteins and, therefore, experimentally confirm their ability to bind zinc by non-cysteine ligands.

VI.3. Further characterization of MsrB1-Co

Our finding and characterization of MsrB1-Co questions the role of Zn ion in MsrB enzymes. MsrB1-Co has shown to reduce methionine sulfoxide at the similar level as the native protein. Therefore, it would be interesting to test whether the protein complex formation between MsrB1-Co and Trx will take place.

VII. References

1. Rhee SG (1999) Redox signaling: hydrogen peroxide as intracellular messenger. Exp Mol Med 31: 53-59.
2. Goswami A, Dikshit P, Mishra A, Mulherkar S, Nukina N, et al. (2006) Oxidative stress promotes mutant huntingtin aggregation and mutant huntingtin-dependent cell death by mimicking proteasomal malfunction. Biochem Biophys Res Commun 342: 184-190.
3. Halliwell B (2006) Oxidative stress and neurodegeneration: where are we now? J Neurochem 97: 1634-1658.
4. Holmgren A (1989) Thioredoxin and glutaredoxin systems. J Biol Chem 264: 13963-13966.
5. Martin JL (1995) Thioredoxin--a fold for all reasons. Structure 3: 245-250.
6. Holmgren A (1985) Thioredoxin. Annu Rev Biochem 54: 237-271.
7. Kallis GB, Holmgren A (1980) Differential reactivity of the functional sulfhydryl groups of cysteine-32 and cysteine-35 present in the reduced form of thioredoxin from Escherichia coli. J Biol Chem 255: 10261-10265.
8. Berndt C, Lillig CH, Holmgren A (2007) Thiol-based mechanisms of the thioredoxin and glutaredoxin systems: implications for diseases in the cardiovascular system. Am J Physiol Heart Circ Physiol 292: H1227-1236.
9. Luthman M, Holmgren A (1982) Rat liver thioredoxin and thioredoxin reductase: purification and characterization. Biochemistry 21: 6628-6633.
10. Arscott LD, Gromer S, Schirmer RH, Becker K, Williams CH, Jr. (1997) The mechanism of thioredoxin reductase from human placenta is similar to the mechanisms of lipoamide dehydrogenase and glutathione reductase and is distinct from the mechanism of thioredoxin reductase from Escherichia coli. Proc Natl Acad Sci U S A 94: 3621-3626.
11. Tamura T, Stadtman TC (1996) A new selenoprotein from human lung adenocarcinoma cells: purification, properties, and thioredoxin reductase activity. Proc Natl Acad Sci U S A 93: 1006-1011.
12. Gladyshev VN, Jeang KT, Stadtman TC (1996) Selenocysteine, identified as the penultimate C-terminal residue in human T-cell thioredoxin reductase, corresponds to TGA in the human placental gene. Proc Natl Acad Sci U S A 93: 6146-6151.
13. Zhong L, Arner ES, Holmgren A (2000) Structure and mechanism of mammalian thioredoxin reductase: the active site is a redox-active selenolthiol/selenenylsulfide formed from the conserved cysteine-selenocysteine sequence. Proc Natl Acad Sci U S A 97: 5854-5859.
14. Muttenthaler M, Alewood PF (2008) Selenopeptide chemistry. Journal of Peptide Science 14: 1223-1239.
15. Hondal RJ, Nilsson BL, Raines RT (2001) Selenocysteine in native chemical ligation and expressed protein ligation. J Am Chem Soc 123: 5140-5141.
16. Huber RE, Criddle RS (1967) Comparison of the chemical properties of selenocysteine and selenocystine with their sulfur analogs. Arch Biochem Biophys 122: 164-173.
17. Pearson RG, Sobel H, Songstad J (1968) Nucleophilic Reactivity Constants toward Methyl Iodide and Trans-[Pt(Py)2cl2]. Journal of the American Chemical Society 90: 319-&.
18. Muttenthaler M, Alewood PF (2008) Selenopeptide chemistry. J Pept Sci 14: 1223-1239.
19. Metanis N, Keinan E, Dawson PE (2006) Synthetic seleno-glutaredoxin 3 analogues are highly reducing oxidoreductases with enhanced catalytic efficiency. J Am Chem Soc 128: 16684-16691.

20. Bock A, Forchhammer K, Heider J, Baron C (1991) Selenoprotein synthesis: an expansion of the genetic code. Trends Biochem Sci 16: 463-467.
21. Bell IM, Fisher ML, Wu ZP, Hilvert D (1993) Kinetic studies on the peroxidase activity of selenosubtilisin. Biochemistry 32: 3754-3762.
22. Stadtman TC (1996) Selenocysteine. Annu Rev Biochem 65: 83-100.
23. Hatfield DL, Gladyshev VN (2002) How selenium has altered our understanding of the genetic code. Mol Cell Biol 22: 3565-3576.
24. Arner ES (2010) Selenoproteins-What unique properties can arise with selenocysteine in place of cysteine? Experimental Cell Research 316: 1296-1303.
25. Allmang C, Wurth L, Krol A (2009) The selenium to selenoprotein pathway in eukaryotes: more molecular partners than anticipated. Biochimica Et Biophysica Acta 1790: 1415-1423.
26. Andreini C, Bertini I, Cavallaro G, Holliday GL, Thornton JM (2008) Metal ions in biological catalysis: from enzyme databases to general principles. J Biol Inorg Chem 13: 1205-1218.
27. Holm RH, Kennepohl P, Solomon EI (1997) ChemInform Abstract: Structural and Functional Aspects of Metal Sites in Biology. ChemInform 28: no-no.
28. Irving H, Williams RJP (1953) The Stability of Transition-Metal Complexes. Journal of the Chemical Society: 3192-3210.
29. Lippard SJ, Berg JM (1994) Principles of bioinorganic chemistry. Mill Valley, Calif.: University Science Books. XVII, 411 s. : ill. p.
30. Lee Y, Lim C (2008) Physical basis of structural and catalytic Zn-binding sites in proteins. Journal of molecular biology 379: 545-553.
31. Purcell KF, Kotz JC (1977) Inorganic chemistry. Philadelphia: Saunders. xix, 1116 p. p.
32. Eriksson AE, Jones TA, Liljas A (1988) Refined structure of human carbonic anhydrase II at 2.0 Å resolution. Proteins: Structure, Function, and Bioinformatics 4: 274-282.
33. Rees D, Lewis M, Lipscomb W (1983) Refined crystal structure of carboxypeptidase a at 1· 54 Å resolution*. Journal of molecular biology 168: 367-387.
34. Matthews BW (1988) Structural Basis of the Action of Thermolysin and Related Zinc Peptidases. Accounts of Chemical Research 21: 333-340.
35. Kantrowitz ER, Lipscomb WN (1988) Escherichia coli aspartate transcarbamylase: the relation between structure and function. Science 241: 669.
36. Capaldi RA (1990) Structure and function of cytochrome c oxidase. Annual review of biochemistry 59: 569-596.
37. Jeloková J, Karlsson C, Estonius M, Jörnvall H, Höög JO (1994) Features of structural zinc in mammalian alcohol dehydrogenase. European Journal of Biochemistry 225: 1015-1019.
38. Gavel OY, Bursakov SA, Rocco GD, Trinc o J, Pickering IJ, et al. (2008) A new type of metal-binding site in cobalt-and zinc-containing adenylate kinases isolated from sulfate-reducers Desulfovibrio gigas and Desulfovibrio desulfuricans ATCC 27774. Journal of inorganic biochemistry 102: 1380-1395.
39. Xu B, Krudy G, Rosevear P (1993) Identification of the metal ligands and characterization of a putative zinc finger in methionyl-tRNA synthetase. Journal of Biological Chemistry 268: 16259.
40. O'Connor T, Graves R, De Murcia G, Castaing B, Laval J (1993) Fpg protein of Escherichia coli is a zinc finger protein whose cysteine residues have a structural and/or functional role. Journal of Biological Chemistry 268: 9063.
41. Maret W, Vallee B (1993) Cobalt as probe and label of proteins. Methods in enzymology 226: 52.
42. Speckhard DC, Wu FYH, Wu CW (1977) Role of the intrinsic metal in RNA polymerase from Escherichia coli. In vivo substitution of tightly bound zinc with cobalt. Biochemistry 16: 5228-5234.

43. Salowe SP, Marcy AI, Cuca GC, Smith CK, Kopka IE, et al. (1992) Characterization of zinc-binding sites in human stromelysin-1: Stoichiometry of the catalytic domain and identification of a cysteine ligand in the proenzyme. Biochemistry 31: 4535-4540.

44. Curdel A, Iwatsubo M (1968) Biosynthetic incorporation of cobalt into yeast alcohol dehydrogenase. FEBS letters 1: 133.

45. Xu B, Trawick B, Krudy GA, Phillips RM, Zhou L, et al. (1994) Probing the metal binding sites of Escherichia coli isoleucyl-tRNA synthetase. Biochemistry 33: 398-402.

46. Hanson G, Berliner L (2008) Metals in Biology: Springer.

47. Maret W, Andersson I, Dietrich H, Schneider-Bernlohr H, Einarsson R, et al. (1979) Site-specific substituted cobalt(II) horse liver alcohol dehydrogenases. Preparation and characterization in solution, crystalline and immobilized state. Eur J Biochem 98: 501-512.

48. Shi Y, Beger RD, Berg JM (1993) Metal binding properties of single amino acid deletion mutants of zinc finger peptides: studies using cobalt (II) as a spectroscopic probe. Biophysical journal 64: 749-753.

49. Holland DR, Hausrath AC, Juers D, Matthews BW (1995) Structural-Analysis of Zinc Substitutions in the Active-Site of Thermolysin. Protein science 4: 1955-1965.

50. Vogt W (1995) Oxidation of methionyl residues in proteins: tools, targets, and reversal. Free Radic Biol Med 18: 93-105.

51. Netto LE, de Oliveira MA, Monteiro G, Demasi AP, Cussiol JR, et al. (2007) Reactive cysteine in proteins: protein folding, antioxidant defense, redox signaling and more. Comp Biochem Physiol C Toxicol Pharmacol 146: 180-193.

52. Dobrovolska O, Rychkov G, Shumilina E, Nerinovski K, Schmidt A, et al. (2012) Structural Insights into Interaction between Mammalian Methionine Sulfoxide Reductase B1 and Thioredoxin. Journal of Biomedicine and Biotechnology 2012: 586539.

53. Kryukov GV, Kumar RA, Koc A, Sun Z, Gladyshev VN (2002) Selenoprotein R is a zinc-containing stereo-specific methionine sulfoxide reductase. Proc Natl Acad Sci U S A 99: 4245-4250.

54. Kim HY, Gladyshev VN (2004) Methionine sulfoxide reduction in mammals: characterization of methionine-R-sulfoxide reductases. Mol Biol Cell 15: 1055-1064.

55. Aachmann F, Sal L, Kim H, Marino S, Gladyshev V, et al. (2010) Insights into function, catalytic mechanism and fold evolution of selenoprotein methionine sulfoxide reductase B1 through structural analysis. Journal of Biological Chemistry.

56. Kim HY, Gladyshev VN (2007) Methionine sulfoxide reductases: selenoprotein forms and roles in antioxidant protein repair in mammals. Biochem J 407: 321-329.

57. Boschi-Muller S, Azza S, Sanglier-Cianferani S, Talfournier F, Van Dorsselear A, et al. (2000) A sulfenic acid enzyme intermediate is involved in the catalytic mechanism of peptide methionine sulfoxide reductase from Escherichia coli. J Biol Chem 275: 35908-35913.

58. Boschi-Muller S, Olry A, Antoine M, Branlant G (2005) The enzymology and biochemistry of methionine sulfoxide reductases. Biochim Biophys Acta 1703: 231-238.

59. Olry A, Boschi-Muller S, Branlant G (2004) Kinetic characterization of the catalytic mechanism of methionine sulfoxide reductase B from Neisseria meningitidis. Biochemistry 43: 11616-11622.

60. Kim HY, Kim JR (2008) Thioredoxin as a reducing agent for mammalian methionine sulfoxide reductases B lacking resolving cysteine. Biochem Biophys Res Commun 371: 490-494.

61. Kim HY, Gladyshev VN (2005) Different catalytic mechanisms in mammalian selenocysteine- and cysteine-containing methionine-R-sulfoxide reductases. Plos Biology 3: e375.
62. Boschi-Muller S, Olry A, Antoine M, Branlant G (2005) The enzymology and biochemistry of methionine sulfoxide reductases. Biochimica Et Biophysica Acta 1703: 231-238.
63. Kim HY, Kim JR (2008) Thioredoxin as a reducing agent for mammalian methionine sulfoxide reductases B lacking resolving cysteine. Biochemical and biophysical research communications 371: 490-494.
64. Boschi-Muller S, Gand A, Branlant G (2008) The methionine sulfoxide reductases: catalysis and substrate specificities. Archives of biochemistry and biophysics 474: 266-273.
65. Qin J, Clore GM, Kennedy WP, Kuszewski J, Gronenborn AM (1996) The solution structure of human thioredoxin complexed with its target from Ref-1 reveals peptide chain reversal. Structure 4: 613-620.
66. Conrad M, Jakupoglu C, Moreno SG, Lippl S, Banjac A, et al. (2004) Essential role for mitochondrial thioredoxin reductase in hematopoiesis, heart development, and heart function. Molecular and Cellular Biology 24: 9414-9423.
67. Jakupoglu C, Przemeck GK, Schneider M, Moreno SG, Mayr N, et al. (2005) Cytoplasmic thioredoxin reductase is essential for embryogenesis but dispensable for cardiac development. Molecular and Cellular Biology 25: 1980-1988.
68. Su D, Novoselov SV, Sun QA, Moustafa ME, Zhou Y, et al. (2005) Mammalian selenoprotein thioredoxin-glutathione reductase - Roles in disulfide bond formation and sperm maturation. Journal of Biological Chemistry 280: 26491-26498.
69. Sun QA, Wu YL, Zappacosta F, Jeang KT, Lee BJ, et al. (1999) Redox regulation of cell signaling by selenocysteine in mammalian thioredoxin reductases. Journal of Biological Chemistry 274: 24522-24530.
70. Su D, Novoselov SV, Sun QA, Moustafa ME, Zhou Y, et al. (2005) Mammalian selenoprotein thioredoxin-glutathione reductase. Roles in disulfide bond formation and sperm maturation. J Biol Chem 280: 26491-26498.
71. Sun QA, Kirnarsky L, Sherman S, Gladyshev VN (2001) Selenoprotein oxidoreductase with specificity for thioredoxin and glutathione systems. Proc Natl Acad Sci U S A 98: 3673-3678.
72. Sun QA, Zappacosta F, Factor VM, Wirth PJ, Hatfield DL, et al. (2001) Heterogeneity within animal thioredoxin reductases. Evidence for alternative first exon splicing. J Biol Chem 276: 3106-3114.
73. Sun QA, Kimarsky L, Sherman S, Galdyshev V (2001) Selenoprotein oxidoreductase with specificity for thioredoxin and glutathione systems. Faseb Journal 15: A968-A968.
74. Sun QA, Su D, Novoselov SV, Carlson BA, Hatfield DL, et al. (2005) Reaction mechanism and regulation of mammalian thioredoxin/glutathione reductase. Biochemistry 44: 14528-14537.
75. Sun QA, Su D, Novoselov SV, Carlson BA, Hatfield DL, et al. (2005) Reaction mechanism and regulation of mammalian thioredoxin/glutathione reductase. Biochemistry 44: 14528-14537.
76. Shumilina E, Solda A, Gerashchenko M, Gladyshev VN, Dikiy A (2012) (1)H, (1)(3)C, and (1)(5)N NMR resonance assignments of reduced full length and shortened forms of the Grx domain of Mus musculus TGR. Biomol NMR Assign 6: 103-107.
77. Wuthrich K (1990) Protein structure determination in solution by NMR spectroscopy. J Biol Chem 265: 22059-22062.
78. Wuthrich K (1991) NMR of proteins and nucleic acids.
79. Ernst RR, Bodenhausen G, Wokaun A (1987) Principles of nuclear magnetic resonance in one and two dimensions. Oxford: Clarendon Press. xxiv, 610 s. : ill. p.

80. Malliavin TE (2006) Quantitative analysis of biomolecular NMR spectra: A prerequisite for the determination of the structure and dynamics of biomolecules. Current Organic Chemistry 10: 555-568.

81. Bertini I, Luchinat C (1986) NMR of paramagnetic molecules in biological systems: Benjamin/Cummings Pub. Co.

82. Kurland RJ, McGarvey BR (1970) Isotropic NMR shifts in transition metal complexes: The calculation of the fermi contact and pseudocontact terms. Journal of Magnetic Resonance (1969) 2: 286-301.

83. Guntert P, Braun W, Wuthrich K (1991) Efficient computation of three-dimensional protein structures in solution from nuclear magnetic resonance data using the program DIANA and the supporting programs CALIBA, HABAS and GLOMSA. J Mol Biol 217: 517-530.

84. Laskowski RA, Macarthur MW, Moss DS, Thornton JM (1993) Procheck - a Program to Check the Stereochemical Quality of Protein Structures. Journal of Applied Crystallography 26: 283-291.

85. Case DA, Cheatham TE, 3rd, Darden T, Gohlke H, Luo R, et al. (2005) The Amber biomolecular simulation programs. J Comput Chem 26: 1668-1688.

86. Ingle JD, Crouch SR (1988) Spectrochemical analysis. Englewood Cliffs, N.J.: Prentice Hall. XV, 590 s. : ill. p.

87. Nelson DL, Lehninger AL, Cox MM (2008) Lehninger principles of biochemistry. New York: Freeman. XXIX, 1158, [1105] s. : ill. p.

88. Sharov VS, Ferrington DA, Squier TC, Schoneich C (1999) Diastereoselective reduction of protein-bound methionine sulfoxide by methionine sulfoxide reductase. FEBS Lett 455: 247-250.

89. Moskovitz J, Poston JM, Berlett BS, Nosworthy NJ, Szczepanowski R, et al. (2000) Identification and characterization of a putative active site for peptide methionine sulfoxide reductase (MsrA) and its substrate stereospecificity. J Biol Chem 275: 14167-14172.

90. Lee BC, Lee YK, Lee HJ, Stadtman ER, Lee KH, et al. (2005) Cloning and characterization of antioxidant enzyme methionine sulfoxide-S-reductase from Caenorhabditis elegans. Arch Biochem Biophys 434: 275-281.

91. Boschi-Muller S, Gand A, Branlant G (2008) The methionine sulfoxide reductases: Catalysis and substrate specificities. Arch Biochem Biophys 474: 266-273.

92. Caldwell P, Luk DC, Weissbach H, Brot N (1978) Oxidation of the methionine residues of Escherichia coli ribosomal protein L12 decreases the protein's biological activity. Proc Natl Acad Sci U S A 75: 5349-5352.

93. Moskovitz J, Weissbach H, Brot N (1996) Cloning the expression of a mammalian gene involved in the reduction of methionine sulfoxide residues in proteins. Proc Natl Acad Sci U S A 93: 2095-2099.

94. Kuschel L, Hansel A, Schonherr R, Weissbach H, Brot N, et al. (1999) Molecular cloning and functional expression of a human peptide methionine sulfoxide reductase (hMsrA). FEBS Lett 456: 17-21.

95. Huang W, Escribano J, Sarfarazi M, Coca-Prados M (1999) Identification, expression and chromosome localization of a human gene encoding a novel protein with similarity to the pilB family of transcriptional factors (pilin) and to bacterial peptide methionine sulfoxide reductases. Gene 233: 233-240.

96. Bar-Noy S, Moskovitz J (2002) Mouse methionine sulfoxide reductase B: effect of selenocysteine incorporation on its activity and expression of the seleno-containing enzyme in bacterial and mammalian cells. Biochem Biophys Res Commun 297: 956-961.

97. Hansel A, Jung S, Hoshi T, Heinemann SH (2003) A second human methionine sulfoxide reductase (hMSRB2) reducing methionine-R-sulfoxide displays a tissue expression pattern distinct from hMSRB1. Redox Rep 8: 384-388.

98. Kauffmann B, Favier F, Olry A, Boschi-Muller S, Carpentier P, et al. (2002) Crystallization and preliminary X-ray diffraction studies of the peptide methionine sulfoxide reductase B domain of Neisseria meningitidis PILB. Acta Crystallogr D Biol Crystallogr 58: 1467-1469.

99. Lowther WT, Weissbach H, Etienne F, Brot N, Matthews BW (2002) The mirrored methionine sulfoxide reductases of Neisseria gonorrhoeae pilB. Nat Struct Biol 9: 348-352.

100. Ezraty B, Aussel L, Barras F (2005) Methionine sulfoxide reductases in prokaryotes. Biochim Biophys Acta 1703: 221-229.

101. Zhang XH, Weissbach H (2008) Origin and evolution of the protein-repairing enzymes methionine sulphoxide reductases. Biol Rev Camb Philos Soc 83: 249-257.

102. Lowther WT, Brot N, Weissbach H, Matthews BW (2000) Structure and mechanism of peptide methionine sulfoxide reductase, an "anti-oxidation" enzyme. Biochemistry 39: 13307-13312.

103. Kumar RA, Koc A, Cerny RL, Gladyshev VN (2002) Reaction mechanism, evolutionary analysis, and role of zinc in Drosophila methionine-R-sulfoxide reductase. J Biol Chem 277: 37527-37535.

104. Brot N, Weissbach L, Werth J, Weissbach H (1981) Enzymatic reduction of protein-bound methionine sulfoxide. Proc Natl Acad Sci U S A 78: 2155-2158.

105. Singh VK, Moskovitz J (2003) Multiple methionine sulfoxide reductase genes in Staphylococcus aureus: expression of activity and roles in tolerance of oxidative stress. Microbiology 149: 2739-2747.

106. Olry A, Boschi-Muller S, Marraud M, Sanglier-Cianferani S, Van Dorsselear A, et al. (2002) Characterization of the methionine sulfoxide reductase activities of PILB, a probable virulence factor from Neisseria meningitidis. J Biol Chem 277: 12016-12022.

107. Hansel A, Heinemann SH, Hoshi T (2005) Heterogeneity and function of mammalian MSRs: enzymes for repair, protection and regulation. Biochim Biophys Acta 1703: 239-247.

108. Vougier S, Mary J, Friguet B (2003) Subcellular localization of methionine sulphoxide reductase A (MsrA): evidence for mitochondrial and cytosolic isoforms in rat liver cells. Biochem J 373: 531-537.

109. Lee BC, Dikiy A, Kim HY, Gladyshev VN (2009) Functions and evolution of selenoprotein methionine sulfoxide reductases. Biochim Biophys Acta 1790: 1471-1477.

110. Ohno S (1970) Evolution by gene duplication. Berlin, New York,: Springer-Verlag. xv, 160 p. p.

111. Zhang J (2003) Evolution by gene duplication: an update. Trends in Ecology & Evolution 18: 292-298.

112. Zhang J (2012) Genetic Redundancies and Their Evolutionary Maintenance
Evolutionary Systems Biology. In: Soyer OS, editor: Springer New York. pp. 279-300.

113. Neupert W (1997) Protein import into mitochondria. Annu Rev Biochem 66: 863-917.

114. von Heijne G (1986) Towards a comparative anatomy of N-terminal topogenic protein sequences. J Mol Biol 189: 239-242.

115. von Heijne G (1986) Mitochondrial targeting sequences may form amphiphilic helices. EMBO J 5: 1335-1342.

116. Roise D, Schatz G (1988) Mitochondrial presequences. J Biol Chem 263: 4509-4511.

117. Lemire BD, Fankhauser C, Baker A, Schatz G (1989) The mitochondrial targeting function of randomly generated peptide sequences correlates with predicted helical amphiphilicity. J Biol Chem 264: 20206-20215.

118. von Heijne G (1990) Protein targeting signals. Curr Opin Cell Biol 2: 604-608.

119. Kim HY, Gladyshev VN (2004) Characterization of mouse endoplasmic reticulum methionine-R-sulfoxide reductase. Biochem Biophys Res Commun 320: 1277-1283.

120. Jung S, Hansel A, Kasperczyk H, Hoshi T, Heinemann SH (2002) Activity, tissue distribution and site-directed mutagenesis of a human peptide methionine sulfoxide reductase of type B: hCBS1. FEBS Lett 527: 91-94.

121. Gromer S, Johansson L, Bauer H, Arscott LD, Rauch S, et al. (2003) Active sites of thioredoxin reductases: why selenoproteins? Proc Natl Acad Sci U S A 100: 12618-12623.

122. Fomenko DE, Xing W, Adair BM, Thomas DJ, Gladyshev VN (2007) High-throughput identification of catalytic redox-active cysteine residues. Science 315: 387-389.

123. Aachmann FL, Sal LS, Kim HY, Marino SM, Gladyshev VN, et al. (2010) Insights into function, catalytic mechanism, and fold evolution of selenoprotein methionine sulfoxide reductase B1 through structural analysis. J Biol Chem 285: 33315-33323.

124. Ranaivoson FM, Neiers F, Kauffmann B, Boschi-Muller S, Branlant G, et al. (2009) Methionine sulfoxide reductase B displays a high level of flexibility. J Mol Biol 394: 83-93.

125. Aachmann FL, Kwak GH, Del Conte R, Kim HY, Gladyshev VN, et al. (2011) Structural and biochemical analysis of mammalian methionine sulfoxide reductase B2. Proteins 79: 3123-3131.

126. Aachmann FL, Sal LS, Kim HY, Marino SM, Gladyshev VN, et al. (2010) Insights into Function, Catalytic Mechanism, and Fold Evolution of Selenoprotein Methionine Sulfoxide Reductase B1 through Structural Analysis. Journal of Biological Chemistry 285: 33315-33323.

127. Kim HY, Gladyshev VN (2005) Different catalytic mechanisms in mammalian selenocysteine- and cysteine-containing methionine-R-sulfoxide reductases. PLoS Biol 3: e375.

128. Zhang Y, Gladyshev VN (2011) Comparative genomics of trace element dependence in biology. J Biol Chem 286: 23623-23629.

129. Makarova KS, Ponomarev VA, Koonin EV (2001) Two C or not two C: recurrent disruption of Zn-ribbons, gene duplication, lineage-specific gene loss, and horizontal gene transfer in evolution of bacterial ribosomal proteins. Genome Biology 2.

130. Olry A, Boschi-Muller S, Yu H, Burnel D, Branlant G (2005) Insights into the role of the metal binding site in methionine-R-sulfoxide reductases B. Protein Sci 14: 2828-2837.

131. Kim YK, Shin YJ, Lee WH, Kim HY, Hwang KY (2009) Structural and kinetic analysis of an MsrA-MsrB fusion protein from Streptococcus pneumoniae. Mol Microbiol 72: 699-709.

132. BERTINI I, LUCHINAT C (1984) High spin cobalt (II) as a probe for the investigation of metalloproteins. Advances in inorganic biochemistry: 71-111.

133. Katta V, Chait BT (1993) Hydrogen-Deuterium Exchange Electrospray-Ionization Mass-Spectrometry - a Method for Probing Protein Conformational-Changes in Solution. Journal of the American Chemical Society 115: 6317-6321.

134. Moura I, Teixeira M, LeGall J, Moura JJ (1991) Spectroscopic studies of cobalt and nickel substituted rubredoxin and desulforedoxin. J Inorg Biochem 44: 127-139.

135. Iwig JS, Leitch S, Herbst RW, Maroney MJ, Chivers PT (2008) Ni(II) and Co(II) sensing by Escherichia coli RcnR. J Am Chem Soc 130: 7592-7606.

136. Holm RH, Kennepohl P, Solomon EI (1996) Structural and Functional Aspects of Metal Sites in Biology. Chemical Reviews 96: 2239-2314.

Paper I

Research Article

Structural Insights into Interaction between Mammalian Methionine Sulfoxide Reductase B1 and Thioredoxin

Olena Dobrovolska,[1] Georgy Rychkov,[2, 3, 4] Elena Shumilina,[1] Kirill Nerinovski,[3, 5] Alexander Schmidt,[2, 3, 4] Konstantin Shabalin,[3, 4] Alexander Yakimov,[2, 3, 4] and Alexander Dikiy[1]

[1] *Department of Biotechnology, Norwegian University of Science and Technology, 7491 Trondheim, Norway*
[2] *Biophysics Department, St. Petersburg State Polytechnical University, St. Petersburg 195251, Russia*
[3] *Center of Nanobiotechnology, St. Petersburg State Polytechnical University, St. Petersburg 195251, Russia*
[4] *Department of Molecular and Radiation Biophysics, Petersburg Nuclear Physics Institute, The RAS, Gatchina 188300, Russia*
[5] *Department of Quantum Magnetic Phenomena, St. Petersburg State University, St. Petersburg 198504, Russia*

Correspondence should be addressed to Alexander Dikiy, alex.dikiy@biotech.ntnu.no

Maintenance of the cellular redox balance has vital importance for correcting organism functioning. Methionine sulfoxide reductases (Msrs) are among the key members of the cellular antioxidant defence system. To work properly, methionine sulfoxide reductases need to be reduced by their biological partner, thioredoxin (Trx). This process, according to the available kinetic data, represents the slowest step in the Msrs catalytic cycle. In the present paper, we investigated structural aspects of the intermolecular complex formation between mammalian MsrB1 and Trx. NMR spectroscopy and biocomputing were the two mostly used through the research approaches. The formation of NMR detectable MsrB1/Trx complex was monitored and studied in attempt to understand MsrB1 reduction mechanism. Using NMR data, molecular mechanics, protein docking, and molecular dynamics simulations, it was found that intermediate MsrB1/Trx complex is stabilized by interprotein β-layer. The complex formation accompanied by distortion of disulfide bond within MsrB1 facilitates the reduction of oxidized MsrB1 as it is evidenced by the obtained data.

1. Introduction

Oxygen is vital for all aerobic biological processes. However, about 5% of it is converted into reactive oxygen species (ROS) [1]. Methionine residues in proteins are susceptible to oxidation by reactive oxygen and nitrogen species leading to formation of methionine sulfoxide (MetSO). This modification can result in loss of proteins' function [2]. Organisms developed a complex antioxidant defence system that includes enzymatic and nonenzymatic antioxidants [3–5]. Methionine sulfoxide reductases (Msrs) are redox repairing enzymes which reduce MetSO back to methionine (Met). A number of published reports describe the role of methionine sulfoxide reductases in antioxidant defence and the regulation of protein function [6–10]. Methionine sulfoxide

reductases reduce both free and protein-bound MetSO, back to Met in the presence of thioredoxin (Trx). Two distinct families constitute Msrs: MsrA, which specifically reduces S-epimer, and MsrB, which is stereospecific for R-epimer of methionine sulfoxide [11–14].

Mammalian MsrB enzymes constitute three different subclasses: MsrB1, MsrB2, and MsrB3 [15–20]. All of them contain Zn^{2+} ion, coordinated by two motifs CxxC (two cysteines separated by two residues), which stabilize their structure. Whereas mammalian MsrB1 contains a resolving cysteine (Cys4) and catalytic selenocysteine (Sec95) in its active site, the other two subclasses, MsrB2 and MsrB3, in analogy with bacterial MsrBs, contain only catalytic cysteine [21–24]. NMR solution [25] and X-ray crystal (PDB entry
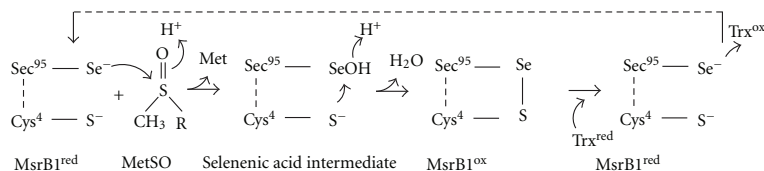
Scheme 1: The biochemical cycle of selenoprotein MsrB1.
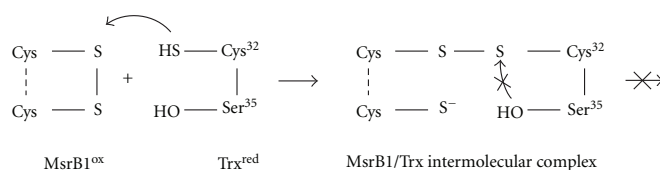
Scheme 2: MsrB1^{ox}-Trx^{red} complex formation. Here, Cys35 of Trx is replaced with catalytically (redox) inactive Ser residue.

3MAO) structures of mammalian MsrB1 have recently been determined.

Reduction of methionine sulfoxide by selenocysteine containing MsrB1 is a multistep process [24, 25]. First, a catalytic Sec95 attacks a sulfoxide moiety of the substrate, MetSO, resulting in the formation of selenenic acid intermediate and concomitant release of methionine. Second, a resolving Cys4 attacks the selenenic acid intermediate to form an intramolecular selenide-sulfide bond, and finally, a fully reduced enzyme is regenerated by reduction of the latter bond with thioredoxin, a natural electron donor (Scheme 1). It has been established that thioredoxin is the natural reducing agent for the members of Msr family, though DTT can be used as the reductant *in vitro* [14, 24, 26]. Reduction of oxidized MsrB2 and MsrB3 as well as of the most bacterial MsrBs occurs through a similar mechanism without formation of the intramolecular disulfide bridge. However, it was reported that bacterial MsrB from *Neisseria meningitides* upon oxidation form intramolecular disulfide bridge [27]. It was shown that its reduction by Trx is overall rate-limiting step of MsrB cycle and it is described as an apparent irreversible process.

Thioredoxin is a ubiquitous protein that plays an important role in maintaining redox balance in cells [28]. In most of its reactions, Trx reduces substrate disulfide bond. The proposed mechanism of Trx-catalysed substrate disulfide reduction [29–31] includes nucleophilic attack by thiolate of Cys32 supported by hydrophobic interactions resulting in transient mixed disulfide formation followed by nucleophilic attack of the resolving Cys35 generating Trx-S$_2$ and reduced protein.

Summarising, reduction of MsrBs (mammalian and bacterial) disulfide bond by Trx, can be described by the following steps. First, thioredoxin and MsrB form an intermolecular complex. In the second step, the just-formed intermolecular bond gets resolved by the thioredoxin resolving cysteine 35, followed by the formation of reduced MsrB and oxidized Trx. Subsequently, the S–S bridge of oxidized thioredoxin is further reduced by NADPH and the specific

FAD-containing enzyme thioredoxin reductase. Some studies suggest that specific structural recognitions exist between oxidized substrate disulfide bond and reduced Trx [14, 32]. The knowledge of the structure aspects of the MsrB/Trx complex formation would help to understand full catalytic mechanism of methionine sulfoxide reductases.

Here, we have investigated the structural aspects of the interaction between mammalian oxidized MsrB1 and reduced thioredoxin. Mutants of MsrB1 and Trx were produced in order to get an intermolecular complex suitable for its investigation using NMR spectroscopy. Firstly, since Sec-containing MsrB1 is not possible to express in bacterial cells, Sec95 of MsrB1 was replaced by Cys giving Sec95Cys mutant (hereafter MsrB1). Further, thioredoxin containing a Cys35Ser mutation (hereafter Trx) was obtained as this modification removes the second resolving Cys35 leading to the formation of a detectable intermediate MsrB1-Trx (Scheme 2). The resultant MsrB1-Trx complex is stable and could, therefore, be structurally characterized by NMR spectroscopy and with the aid of computational approach.

2. Materials and Methods

2.1. Microorganisms and Plasmids. The genes of C-terminal his-tagged mouse MsrB1 (Sec95Cys) and human Trx (Cys35Ser) mutants, cloned into pET21 expression vectors were kindly provided by Professor V. Gladyshev's group. *Escherichia coli* strains ER2566 (New England Biolabs) and BL21 (DE3) (Novagen) cells were respectively transformed with the constructs using the standard protocol [33]. The transformed cells were spread on several LB agar plates containing 100 mg/L of ampicillin and were further stored at 4°C.

2.2. Protein Expression, Purification, and Sample Preparation. E. coli BL21(DE3) cells carrying plasmid pET21-Trx and E. coli ER2566 cells containing pET21-MsrB1 were grown in LB-media with 100 mg/L ampicillin. To uniformly label proteins either with ^{15}N or with ^{15}N-^{13}C, cells were grown in

M9 minimal media containing 1 g/L 99%-enriched $^{15}NH_4Cl$, 4 g/L $^{13}C_6$-glucose and 100 mg/L of ampicillin. MsrB1 and Trx were expressed, by growing host cells at 37°C until an OD_{600} reached 0.8, followed by induction of protein synthesis with 1 mM IPTG and subsequent incubation for 3 h at 37°C. Cells were harvested by centrifugation and disrupted by sonication in 50 mM phosphate buffer, pH 8.0, containing 400 mM NaCl, 0.01% Tween 20 (Sigma Aldrich); 5 mM β-mercaptoethanol as well as complete protease inhibitor cocktail (Roche) (1 tab for 50 mL).

Cell extract was clarified by centrifugation and filtering through 0,45 μm filter. Clear supernatant was loaded onto the preequilibrated in binding buffer (50 mM Na phosphate buffer, pH 8.0, 400 mM NaCl, 5 mM β-mercaptoethanol, and 5 mM imidazole) Ni-NTA column (Novagen). The column was step-washed with the increasing concentration of imidazole (5, 10, and 20 mM) in binding buffer followed by elution of protein with 250 mM imidazole in binding buffer. The purity of the samples was examined with SDS-PAGE gel (Invitrogen). The yield of the proteins per liter of growth medium was 20 mg and 10 mg for MsrB1 and Trx, respectively.

The obtained reduced MsrB1 was further subjected to oxidation by excess of dabsyl-Met-R-SO (kindly provided by Professor H-Y. Kim) for 3 h at 25°C in 20 mM phosphate buffer, 20 mM NaCl, and pH 7.5. The final NMR MsrB1^{ox} and Trx^{red} samples contained 1,5–2 mM of protein in 10 mM NaCl, 10 mM phosphate, pH 5.5, 90%H_2O, 10% D_2O buffer. The sample for the backbone assignment of reduced double enriched Trx (^{15}N- and ^{13}C) also contained 5 mM DTT.

2.3. NMR Spectroscopy. All NMR spectra were recorded at 298 K either on Bruker Avance 600 MHz spectrometer, equipped with 5 mm z-gradients TXI (H/C/N) cryoprobe at the NMR centre of NT faculty, NTNU or Varian DirectDrive NMR System 700 MHz spectrometer, equipped with 5 mm z-gradients salt tolerant H/C/N probe at the SPbSPU. Proton chemical shifts were referenced to external 3-(trimethylsilyl)-propane-sulfonic acid sodium salt (DSS), while ^{15}N and ^{13}C chemical shifts were referenced indirectly to a liquid ammonia and DSS, respectively, based on the absolute frequency ratios [34].

The comparison between reduced and oxidized forms of MsrB1 was performed by analyzing 2D ^{15}N-^1H HSQC spectra of both reduced and oxidized MsrB1. The spectra were acquired using pulse sequence from the standard pulse sequence library. N-H coupling constant was set to 90 ms, the relaxation delay in HSQC experiments was of 1 s. 2048 complex points were collected in F2 dimension, while 256 were collected in F1 dimension. 32 scans per each transient have been recorded.

^1H, ^{13}C, and ^{15}N backbone resonance assignments for Trx protein were achieved using ^{15}N HSQC, HNCA, CBCA(CO)NH, CBCANH, HBHANH, HBHA(CO)NH, and ^{15}N NOESY NMR spectra from the standard pulse sequence library.

The NMR data were processed with the BRUKER XWin-NMR version 3.5 and Varian VNMRJ version 2.2C software.

Spectral analysis was performed using CARA version 1.8.4.2 [35].

In order to map the interacting sites of both Trx and MsrB1 proteins, that is, to determine which aminoacids are involved in the formation of the interprotein complex, NMR titration of each protein was performed. ^{15}N-^1H HSQC of ^{15}N labeled oxidized MsrB1 was recorded followed by acquisition of the series of HSQC spectra of oxidized MsrB1 in the presence of increasing amounts of ^{15}N-unlabeled Trx. Likewise, the ^{15}N-^1H HSQC spectrum of ^{15}N-enriched pure Trx was recorded first, followed by the set of HSQC spectra of Trx containing an increasing amount of unlabeled oxidized MsrB1. All titration experiments were carried out three times at 298 K at two different pH values 5.5 and 6.5 in order to have statistically significant results.

2.4. Molecular Modelling. Structural calculations for MsrB1 and Trx and their complex were performed using molecular modelling techniques, including molecular mechanics and protein docking (Molsoft ICM Pro 3.6 program package [36], ECEPP/3 force field [37]) as well as molecular dynamics simulations (GROMACS program package version 4.0.7 [38], G53a6 force field [39]). VMD [40] and ICM Pro 3.6 program packages were used to visualize and analyze molecular dynamics (MD) trajectories and the resulting averaged spatial structures of the proteins and their complexes. The initial sets of atomic coordinates of the proteins (reduced forms) were taken from PDB (2kv1 and 3trx for MsrB1 and Trx, resp.).

MD simulations were carried out using a standard protocol including the following steps: energy minimization, water box equilibration, and productive run. Both structures of MsrB1 protein and four complexes of its oxidized form with reduced Trx were placed in dodecahedral water box. The box dimensions were chosen in such a way that water shell around the protein structures was no less than 12 Å. Water molecules were represented by SPC model [41]. The intrinsic MsrB1 positive charge was neutralized with two Cl^- ions, and the negative charge of the protein complex was neutralized by three Na^+ ions. The energy minimization procedure was executed with steepest descent method. During 300 ps equilibration of the water molecules surrounding protein, spatial position of its atoms was constrained. The equilibration procedure was followed by two productive runs with 5 ns and 20 ns in order to separate proteins and their complexes, respectively. Distances between Zn^{2+} ion and Sγ atoms of the coordinating cysteines of MsrB1 (Cys23, Cys26, Cys71, and Cys74) were restricted with help of harmonic potential having a minimum at a value of 2.35 Å. The averaged molecular structures were obtained by the cluster analysis [42] after each productive run of MD.

MD integration step was equal to 2 fs. LINCS algorithm [43] was used to constrain the covalent bond distances and the valence angle values. The temperature of the system was held at 300 K with velocity-rescale algorithm [44]. To hold system pressure at 1 atm value, Berendsen barostate [45] with the time constant for coupling, $\tau = 0.5$ ps was used. Long-range electrostatics was calculated with PME method [46]. The cut-off distance of nonbonded interactions was

set to 10 Å for van der Waals interactions and to 14 Å for electrostatic interactions.

To obtain the disulfide bond between Cys95 and Cys4 residues in oxidized MsrB1, MD simulations in explicit water box with the described above protocol were used, except that the distance between Sγ atoms of two cysteines was restrained at 2.8 Å with the force exceeding van der Waals repulsion of these atoms. The disulfide bond itself was formed and its geometry was optimized by molecular mechanic procedures in ECEPP/3 force field implemented in ICM Pro 3.6 program package. To construct the mutant form Cys35Ser of Trx, the side chain of the mutated aminoacid together with its neighbours situated within 5 Å radius, were subjected to the energy minimization procedure in ECEPP/3 potential. The average structures of oxidized MsrB1 and reduced Trx have been subsequently subjected to the protein-protein docking in ICM Pro 3.6. This algorithm uses optimal docking area parameter representing the protein surface regions with the maximum dehydration energy arising upon formation of the tight protein-protein contact. To reduce conformational sampling calculations, only the "hot points" that were situated near MsrB1 and Trx active sites were used within the study. Interaction energies were calculated using potential grids with 0.5 Å cells, and the truncated van der Waals potential with maximum 1.0 kcal/mol was used.

The conformations of MsrB1-Trx complex, obtained from docking, having a representative distance between Sγ Cys95 of MsrB1 and Sγ C32 of Trx less than 10 Å, were taken for further analysis and sorted into four groups depending on the mutual orientations of the two proteins. Then, the representative conformations from each of four groups were refined by the above-described molecular dynamics simulation methods.

3. Results and Discussion

3.1. Assignment of Oxidized MsrB1. Upon oxidation, MsrB1 undergoes structural changes caused by the intramolecular Cys95/Cys4 disulfide bond formation [25]. In order to assign ^{1}H and ^{15}N NMR spectra of the oxidized MsrB1, the HSQC spectra of MsrB1 protein in both oxidized and reduced states were compared [25, 47].

This analysis revealed close similarity between the HSQC spectra of both redoxed forms: the majority of the cross-peaks assigned in the case of reduced MsrB1 protein remained either at the same position in ^{1}H-^{15}N HSQC spectra or their positions were slightly altered. This overall spectral correspondence between reduced and oxidized forms of MsrB1 allowed us to unambiguously identify and assign the cross-peaks related to the most of the aminoacids of the oxidized MsrB1. Nevertheless, the signal broadening and the shift of some of the cross-peaks in the oxidized protein corresponding to the residues either belonging to the protein active site or situated in its vicinity were observed (Cys4, Phe82, Cys95, and Ile96). The performed comparison provided us with the assignment of the ^{1}H and ^{15}N chemical shifts of oxidized MsrB1 protein and indicated that no major structural changes occur in MsrB1 protein upon oxidation.

3.2. Assignment of Reduced Trx. The protein's backbone assignment was performed using a standard procedure. The availability of the assignment of native human Trx (313 K, pH = 5.5) [48] assisted in the obtainment of the present assignment. As a result, 95% of all Trx aminoacids were identified and assigned in the present study. The backbone assignments of Trx's ^{1}H, ^{13}C, and ^{15}N nuclei were deposited in BioMagResBank under accession number BMRB-16850. The obtained within the study assignments for both oxidized MsrB1 and reduced Trx provided a basis for our further investigations on interaction between Trx and MsrB1.

3.3. Monitoring of Interaction of Oxidized MsrB1-Trx. NMR spectroscopy is a commonly used technique for mapping the interacting site of a protein upon complex formation with its ligands. Within the study, it was performed the NMR titration of oxidized MsrB1 with reduced Trx and vice versa. To determine whether the interaction of these proteins can be observed on NMR time scale and, in the positive case, to determine which aminoacids from those proteins are involved in the interaction, the following approach was used. ^{15}N-enriched protein (either MsrB1 or Trx) was titrated with nonenriched partner, and their interaction was monitored by comparative analysis of ^{1}H-^{15}N HSQC spectra of a single protein and the protein in the presence of its partner. Since only one protein was enriched with ^{15}N NMR active isotope in each single experiment, the presence in solution of another protein did not hamper the observation of [1]H-^{15}N spectrum under analysis, and it was possible to clearly identify aminoacids changing their NMR parameters (either chemical shifts (shifting of the cross-peak) or relaxation rates (broadening of the cross-peak)), which indicated that these residues are involved in interaction with another protein. This analysis allowed us to determine which aminoacids from each protein are directly involved in the complex formation.

3.4. The NMR Titration of Oxidized MsrB1 with Reduced Trx. The titration of MsrB1 revealed its aminoacids involved in binding with Trx. The integral chemical shift differences calculated for all MsrB1 residues using the relationship $\Delta^1H + \Delta^{15}N/7$ ($\Delta^1H + \Delta^{15}N/5$ for glycine residues) [49] and plotted as a function of the aminoacid number, are shown in Figure 1(a). These data show that the residues belonging to the MsrB1 active site as well as their neighbouring residues directly participate in interaction with reduced Trx (C4-F7, E65, and C95-F97). It was also observed that the MsrB1 HSQC cross peaks corresponding to the residues revealing the maximal shifts upon interaction with Trx became broader (increase of transverse relaxation rates), thus supporting our finding in pinpointing of the interaction relevant aminoacids. Altogether, our studies indicate that these MsrB1 residues are involved in protein complex formation with Trx.

3.5. The NMR Titration of Reduced Trx with Oxidized MsrB1. The integral ^{1}H-^{15}N HSQC chemical shift differences for all Trx residues plotted as a function of the aminoacid number were obtained as described above and are shown in

(a)

(b)
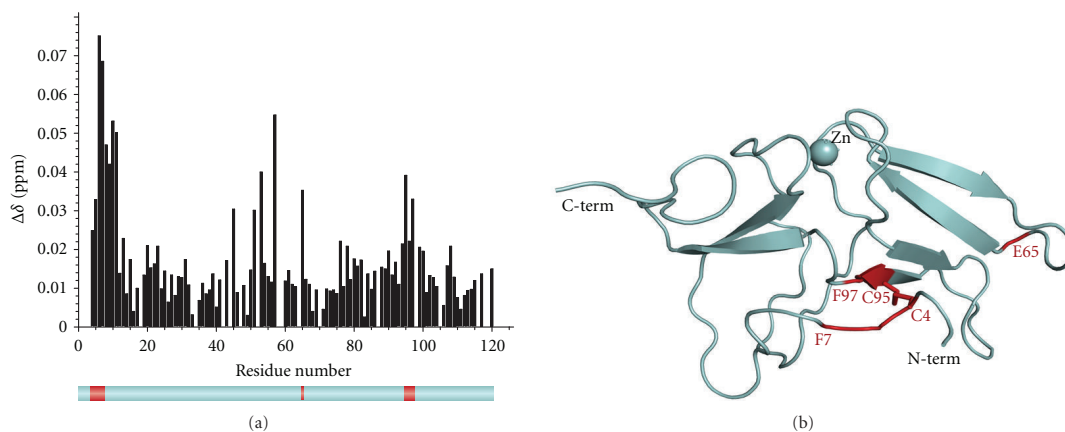
FIGURE 1: (a) Chemical shift difference observed for each residue of MsrB1 upon NMR titration by Trx. Lower panel shows in red the residues belonging to the protein active site revealing maximal changes upon titration. (b) The 3D structure of oxidized mammalian MsrB1. "Hot points" used for molecular docking are evidenced in red.

(a)

(b)
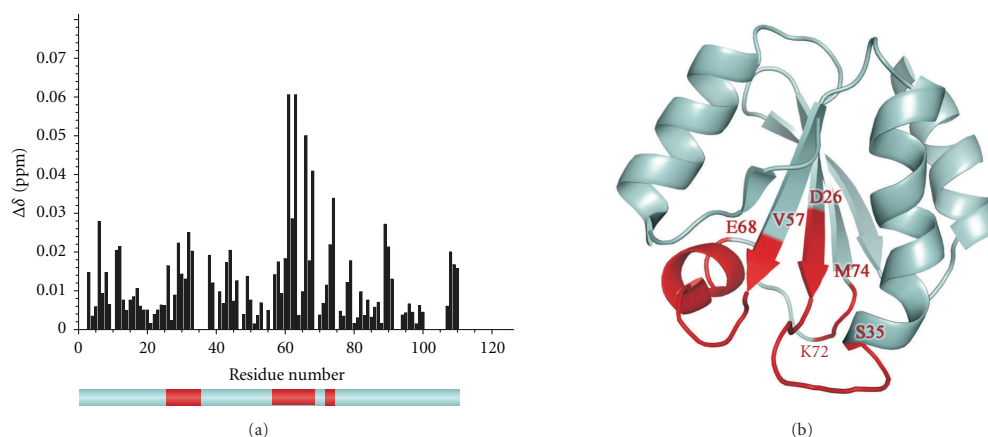
FIGURE 2: (a) Chemical shift difference observed for each residue of Trx upon NMR titration by MsrB1. Lower panel shows in red the residues revealing maximal changes upon titration. (b) The 3D structure of oxidized mammalian Trx. "Hot points" used for molecular docking are evidenced in red.

Figure 2(a). The titration of Trx has revealed three aminoacid areas which are involved in interaction with MsrB1: D26-S35, V57-E68, and K72-M74. These segments are evidenced in red on the lower panel of Figure 2(a). Our results indicate that: (i) Trx active site (as expected) is involved in the interaction; (ii) the monitored segments are situated mostly on the Trx external loops; (iii) these segments are rather close in space, thus supporting and validating our findings.

3.6. Structural Modelling of Oxidized MsrB1 and Reduced Trx. The average structures of oxidized MsrB1 and reduced Trx were obtained from the molecular dynamics simulations starting from coordinates of the reduced proteins taken from PDB (2kv1 and 3trx for MsrB1 and Trx, resp.). The calculated structures are shown in Figures 1(b) and 2(b).

3.7. Docking of Oxidized MsrB1 and Trx. Upon formation of an intermolecular protein complex, catalytic Cys32 (active residue) and Cys35 (resolving residue) of Trx should become close in space to the disulfide bond, connecting Cys4 and Cys95 residues of oxidized MsrB1. Such arrangement makes possible that Cys32 of Trx attacks the MsrB1 disulfide bond. A protein docking procedure was carried out in order to identify possible options for mutual arrangement of these proteins. Based on the above-mentioned distance restrains and experimental NMR data, the "hot points of docking", were found (evidenced for both proteins in Figures 1(b) and 2(b) in red), and more than 1000 spatial structure orientations were generated. Our experimental data on residues involved in complex formation and application of 10 Å distance restrains between the Sγ atoms of Cys32 of Trx

TABLE 1: Structural parameters for MsrB1-Trx complexes (four groups) revealed by protein docking. The distances between sulphur atom of Cys32 of Trx and the nearest sulphur atom of MsrB1 disulfide bond as well as three pairs of the nearest aminoacids approaching in space for each group (to outline the spatial orientation of the proteins within a complex) are presented.

	Group A			Group B			Group C			Group D		
Distance, Å (Trx Cys32) S/S–S bond (MsrB1)	7,1			7,8			7,8			5,3		
Nearest neighbouring contacts between residues												
MsrB1 residues	W43	H39	P87	S2	W43	P87	W43	N62	P87	F6	P42	R93
Trx residues	M37	E95	S90	W31	A92	E70	M37	K72	K96	K72	D60	K36

(A)   (B)   (C)

(a)   (b)   (c)

FIGURE 3: The snapshots (A, B, and C) of the protein complex formation along the molecular dynamics trajectory. The panels (a), (b), and (c) show zoomed views (interactions interfaces) of the snapshots (A) and (C), respectively.

and Cys4-Cys95 disulfide bond of MsrB1 were used to sort these structures into four groups. The neighbouring contacts between residues and the distances between sulphur atom of Cys32 (Trx) and the nearest sulphur atom of oxidized MsrB1 disulfide bond, differentiating each group, are shown in Table 1.

Among the four represented complexes, the fourth (D) is the most plausible since the distance in this structure between Cys32 of Trx, and two sulphur atoms of disulfide bond in MsrB1 were found to be minimal (5.3 Å). Further, each of the four representative structures has been subjected to 5 ns molecular dynamics simulations.

3.8. Molecular Dynamics Procedure. The aim of the molecular dynamics simulation was to refine the possible structure

of the Trx-MsrB1 complex taking into account the obtained NMR data, the effects of water environment, and flexibility of proteins backbone and side chains. The analysis of the averaged structures of the complexes obtained after 5 ns of MD simulations reveal that only for complex D the distance $S\gamma Cys32$ Trx–$S\gamma Cys4$ MsrB1 remains the smallest among all possible complexes. Therefore, the MD trajectory for this complex was extended to 20 ns (Figures 3(a) and 3(c)). It is necessary to emphasize that even if these data suggest that the $S\gamma Cys32$ Trx is closer to $S\gamma Cys4$ of oxidized MsrB1, the analogous distance to Cys95 is rather comparable. This observation precludes us to determine which Cys residue from intramolecular MsrB1 disulfide bridge will be subsequently targeted to form intermolecular disulfide bond.

As it was calculated by our MD, the N-terminal tail of MsrB1 upon oxidation gets closer to the β-strand of the protein formed by Arg93-Ile96 (Figures 3(A) and 3(a)). Further, elongation of the mentioned MsrB1's β-strand occurs additionally involving residues I96-S98 accompanied by formation of three additional hydrogen bonds with residues S5-F7 (Figures 3(B) and 3(b)). This newly constituted strand belongs to the N-terminus of MsrB1, which now is stabilized by the intramolecular disulfide bond formed upon oxidation. At the final stage of the interaction, Trx adjusts its unstructured loop to the newly formed N-terminal β-strand of MsrB1, leading to formation of interprotein β-sheet, composed of six β-strands, where five of them belong to MsrB1 and the sixth to Trx (residues K72-M74) (Figures 3(C) and 3(c)). Thus, in developing of two new β-strands (one intramolecular MsrB1 and one intermolecular), nine aminoacids participate. This observation is in good agreement with our NMR titration results (Figures 1 and 2), thus further validating our calculations. In addition, β-strand of MsrB1 formed by G77-F82 extends periodically to residues G75-G77 allowing residues G75-L78 to form hydrogen bonds with the second antiparallel β-strand of MsrB1 constituted by A66-G72. The distance between SγCys32 of thioredoxin and SγCys95 of MsrB1, a reference distance, oscillates during the interaction from the smallest value of about 3.2 Å (first 100 ps of the trajectory) to 9.5 Å and then down to 4.8 Å. Such approaching may cause a catalytic act, upon which the disulfide bond becomes resolved followed by reduction of MsrB1.

Upon the formation of the protein intermolecular complex, six new hydrogen bonds evolve (three of them are interproteins), thus stabilizing the intermolecular complex. Analysis of MD trajectories for MsrB1-Trx complex (prior formation of the intermolecular disulfide bond) indicates an existence of two conformations of C4-C95 MsrB1 bond. The disulfide bond in both conformations is strained in different extent. The degree of strain could be estimated by comparison of the energies for the van der Waals (vw) and torsion (to) terms of disulfide bond for oxidized MsrB1 in the step preceding the complex formation and in the free state. In accordance with ECEPP/3 force field, the following energies have been obtained (kcal/mol): 3.1 and 3.2 (vw), −0.5 and −1.7 (to), 2.6 and 1.5 (total) for "strained" and "relaxed" S–S bond, respectively. It is worth to mention that during the MD simulations this MsrB1 intramolecular S–S bond prefers its "strained" conformation. This bond thus becomes totally destabilized by 2.6 kcal/mol for "strained" or by 1.5 kcal/mol for "relaxed" conformation. However, the presence of six hydrogen bonds inside the developed interprotein β-sheet obviously overcompensates the straining of Cys4-Cys95 bond in MsrB1, as energy profit from one hydrogen bond formation in β-sheet can be estimated by −0.78 kcal/mol [50]. Therefore, the formation of the interprotein MsrB1-Trx complex becomes energetically preferable.

4. Conclusions

In this study, both the experimental and the computational approaches were combined to investigate MsrB1-Trx protein complex formation. Through the analysis of the NMR data, the aminoacids involved in protein complex formation have been determined. The average 3D structures of oxidized MsrB1 and reduced Trx were generated and subsequently subjected to protein docking and molecular dynamics. Complementary use of the mentioned approaches indicates that the formation of the interprotein β-sheet, stabilized by new hydrogen bonds, distorting the MsrB1 intramolecular disulfide bond. This disulfide bond straining favours the formation of intermolecular disulfide bridge of MsrB1-Trx complex. The study evidences structural and energetic driving forces of MsrB1-Trx complex formation.

References

[1] U. Bandyopadhyay, D. Das, and R. K. Banerjee, "Reactive oxygen species: oxidative damage and pathogenesis," *Current Science*, vol. 77, no. 5, pp. 658–666, 1999.

[2] W. Vogt, "Oxidation of methionyl residues in proteins: tools, targets, and reversal," *Free Radical Biology and Medicine*, vol. 18, no. 1, pp. 93–105, 1995.

[3] E. R. Stadtman, H. Van Remmen, A. Richardson, N. B. Wehr, and R. L. Levine, "Methionine oxidation and aging," *Biochimica et Biophysica Acta*, vol. 1703, no. 2, pp. 135–140, 2005.

[4] B. Friguet, "Oxidized protein degradation and repair in ageing and oxidative stress," *FEBS Letters*, vol. 580, no. 12, pp. 2910–2916, 2006.

[5] I. Petropoulos and B. Friguet, "Maintenance of proteins and aging: the role of oxidized protein repair," *Free Radical Research*, vol. 40, no. 12, pp. 1269–1276, 2006.

[6] J. Moskovitz, E. Flescher, B. S. Berlett, J. Azare, J. M. Poston, and E. R. Stadtman, "Overexpression of peptide-methionine sulfoxide reductase in saccharomyces cerevisiae and human T cells provides them with high resistance to oxidative stress," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 95, no. 24, pp. 14071–14075, 1998.

[7] H. Ruan, X. D. Tang, M. L. Chen et al., "High-quality life extension by the enzyme peptide methionine sulfoxide reductase," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 99, no. 5, pp. 2748–2753, 2002.

[8] J. Moskovitz, "Roles of methionine suldfoxide reductases in antioxidant defense, protein regulation and survival," *Current Pharmaceutical Design*, vol. 11, no. 11, pp. 1451–1457, 2005.

[9] H. Weissbach, L. Resnick, and N. Brot, "Methionine sulfoxide reductases: history and cellular role in protecting against oxidative damage," *Biochimica et Biophysica Acta*, vol. 1703, no. 2, pp. 203–212, 2005.

[10] H. Y. Kim and V. N. Gladyshev, "Methionine sulfoxide reductases: selenoprotein forms and roles in antioxidant protein repair in mammals," *Biochemical Journal*, vol. 407, no. 3, pp. 321–329, 2007.

[11] R. Grimaud, B. Ezraty, J. K. Mitchell et al., "Repair of oxidized proteins: identification of a new methionine sulfoxide reductase," *Journal of Biological Chemistry*, vol. 276, no. 52, pp. 48915–48920, 2001.

[12] F. Neiers, A. Kriznik, S. Boschi-Muller, and G. Branlant, "Evidence for a new sub-class of methionine sulfoxide reductases B with an alternative thioredoxin recognition signature," *Journal of Biological Chemistry*, vol. 279, no. 41, pp. 42462–42468, 2004.

[13] H. Weissbach, F. Etienne, T. Hoshi et al., "Peptide methionine sulfoxide reductase: structure, mechanism of action, and biological function," *Archives of Biochemistry and Biophysics*, vol. 397, no. 2, pp. 172–178, 2002.

[14] S. Boschi-Muller, A. Gand, and G. Branlant, "The methionine sulfoxide reductases: catalysis and substrate specificities," *Archives of Biochemistry and Biophysics*, vol. 474, no. 2, pp. 266–273, 2008.

[15] A. Lescure, D. Gautheret, P. Carbon, and A. Krol, "Novel selenoproteins identified in silico and in vivo by using a conserved RNA structural motif," *Journal of Biological Chemistry*, vol. 274, no. 53, pp. 38147–38154, 1999.

[16] W. Huang, J. Escribano, M. Sarfarazi, and M. Coca-Prados, "Identification, expression and chromosome localization of a human gene encoding a novel protein with similarity to the pilB family of transcriptional factors (pilin) and to bacterial peptide methionine sulfoxide reductases," *Gene*, vol. 233, no. 1-2, pp. 233–240, 1999.

[17] S. Jung, A. Hansel, H. Kasperczyk, T. Hoshi, and S. H. Heinemann, "Activity, tissue distribution and site-directed mutagenesis of a human peptide methionine sulfoxide reductase of type B: hCBS1," *FEBS Letters*, vol. 527, no. 1–3, pp. 91–94, 2002.

[18] A. Hansel, S. H. Heinemann, and T. Hoshi, "Heterogeneity and function of mammalian MSRs: enzymes for repair, protection and regulation," *Biochimica et Biophysica Acta*, vol. 1703, no. 2, pp. 239–247, 2005.

[19] H. Y. Kim and V. N. Gladyshev, "Methionine sulfoxide reduction in mammals: characterization of methionine-R-Sulfoxide reductases," *Molecular Biology of the Cell*, vol. 15, no. 3, pp. 1055–1064, 2004.

[20] H. Y. Kim and V. N. Gladyshev, "Characterization of mouse endoplasmic reticulum methionine-R-sulfoxide reductase," *Biochemical and Biophysical Research Communications*, vol. 320, no. 4, pp. 1277–1283, 2004.

[21] S. Boschi-Muller, A. Olry, M. Antoine, and G. Branlant, "The enzymology and biochemistry of methionine sulfoxide reductases," *Biochimica et Biophysica Acta*, vol. 1703, no. 2, pp. 231–238, 2005.

[22] J. Moskovitz, "Methionine sulfoxide reductases: ubiquitous enzymes involved in antioxidant defense, protein regulation, and prevention of aging-associated diseases," *Biochimica et Biophysica Acta*, vol. 1703, no. 2, pp. 213–219, 2005.

[23] G. V. Kryukov, R. A. Kumar, A. Koc, Z. Sun, and V. N. Gladyshev, "Selenoprotein R is a zinc-containing stereo-specific methionine sulfoxide reductase," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 99, no. 7, pp. 4245–4250, 2002.

[24] H. Y. Kim and V. N. Gladyshev, "Different catalytic mechanisms in mammalian selenocysteine- and cysteine-containing methionine-R-sulfoxide reductases," *PLoS Biology*, vol. 3, no. 12, article e375, pp. 1–10, 2005.

[25] F. L. Aachmann, L. S. Sal, H. Y. Kim, S. M. Marino, V. N. Gladyshev, and A. Dikiy, "Insights into function, catalytic mechanism, and fold evolution of selenoprotein methionine sulfoxide reductase B1 through structural analysis," *Journal of Biological Chemistry*, vol. 285, no. 43, pp. 33315–33323, 2010.

[26] H. Y. Kim and J. R. Kim, "Thioredoxin as a reducing agent for mammalian methionine sulfoxide reductases B lacking resolving cysteine," *Biochemical and Biophysical Research Communications*, vol. 371, no. 3, pp. 490–494, 2008.

[27] M. Antoine, S. Boschi-Muller, and G. Branlant, "Kinetic characterization of the chemical steps Involved in the catalytic mechanism of methionine sulfoxide reductase A from Neisseria meningitidis," *Journal of Biological Chemistry*, vol. 278, no. 46, pp. 45352–45357, 2003.

[28] A. Holmgren, "Thioredoxin and glutaredoxin systems," *Journal of Biological Chemistry*, vol. 264, no. 24, pp. 13963–13966, 1989.

[29] A. Holmgren, "Thioredoxin," *Annual Review of Biochemistry*, vol. 54, pp. 237–271, 1985.

[30] G. B. Kallis and A. Holmgren, "Differential reactivity of the functional sulfhydryl groups of cysteine-32 and cysteine-35 present in the reduced form of thioredoxin from Escherichia coli," *Journal of Biological Chemistry*, vol. 255, no. 21, pp. 10261–10265, 1980.

[31] A. Holmgren, "Thioredoxin structure and mechanism: conformational changes on oxidation of the active-site sulfhydryls to a disulfide," *Structure*, vol. 3, no. 3, pp. 239–243, 1995.

[32] J. Qin, G. M. Clore, W. M. P. Kennedy, J. R. Huth, and A. M. Gronenborn, "Solution structure of human thioredoxin in a mixed disulfide intermediate complex with its target peptide from the transcription factor NFκB," *Structure*, vol. 3, no. 3, pp. 289–297, 1995.

[33] F. M. Ausubel, *Short Protocols in Molecular Biology: A Compendium of Methods from Current Protocols in Molecular Biology*, vol. 1, John Wiley & Sons, Hoboken, NJ, USA, 5th edition, 2002.

[34] H. Zhang, S. Neal, and D. S. Wishart, "RefDB: a database of uniformly referenced protein chemical shifts," *Journal of Biomolecular NMR*, vol. 25, no. 3, pp. 173–195, 2003.

[35] R. L. J. Keller, *Optimizing the Process of Nuclear Magnetic Resonance Spectrum Analysis and Computer Aided Resonance Assignment*, Zürich, 2005.

[36] R. Abagyan, M. Totrov, and D. Kuznetsov, "ICM—a new method for protein modeling and design: applications to docking and structure prediction from the distorted native conformation," *Journal of Computational Chemistry*, vol. 15, no. 5, pp. 488–506, 1994.

[37] G. Nemethy, K. D. Gibson, K. A. Palmer et al., "Energy parameters in polypeptides. 10. Improved geometrical parameters and nonbonded interactions for use in the ECEPP/3 algorithm, with application to proline-containing peptides," *Journal of Physical Chemistry*, vol. 96, no. 15, pp. 6472–6484, 1992.

[38] B. Hess, C. Kutzner, D. Van Der Spoel, and E. Lindahl, "GRGMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation," *Journal of Chemical Theory and Computation*, vol. 4, no. 3, pp. 435–447, 2008.

[39] C. Oostenbrink, A. Villa, A. E. Mark, and W. F. Van Gunsteren, "A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6," *Journal of Computational Chemistry*, vol. 25, no. 13, pp. 1656–1676, 2004.

[40] W. Humphrey, A. Dalke, and K. Schulten, "VMD: visual molecular dynamics," *Journal of Molecular Graphics*, vol. 14, no. 1, pp. 33–38, 1996.

[41] C. D. Berweger, W. F. van Gunsteren, and F. Müller-Plathe, "Force field parametrization by weak coupling. Re-engineering SPC water," *Chemical Physics Letters*, vol. 232, no. 5-6, pp. 429–436, 1995.

[42] X. Daura, K. Gademann, B. Jaun, D. Seebach, W. F. Van Gunsteren, and A. E. Mark, "Peptide folding: when simulation meets experiment," *Angewandte Chemie International Edition*, vol. 38, no. 1-2, pp. 236–240, 1999.

[43] B. Hess, H. Bekker, H. J. C. Berendsen, and J. G. E. M. Fraaije, "LINCS: a Linear constraint solver for molecular simulations," *Journal of Computational Chemistry*, vol. 18, no. 12, pp. 1463–1472, 1997.

[44] G. Bussi, D. Donadio, and M. Parrinello, "Canonical sampling through velocity rescaling," *Journal of Chemical Physics*, vol. 126, no. 1, Article ID 014101, 2007.

[45] H. J. C. Berendsen, J. P. M. Postma, W. F. Van Gunsteren, A. Dinola, and J. R. Haak, "Molecular dynamics with coupling to an external bath," *The Journal of Chemical Physics*, vol. 81, no. 8, pp. 3684–3690, 1984.

[46] T. Darden, D. York, and L. Pedersen, "Particle mesh Ewald: an N·log(N) method for Ewald sums in large systems," *The Journal of Chemical Physics*, vol. 98, no. 12, pp. 10089–10092, 1993.

[47] L. S. Sal, F. L. Aachmann, H. Y. Kim, V. N. Gladyshev, and A. Dikiy, "NMR assignments of 1H, 13C and 15N spectra of methionine sulfoxide reductase B1 from Mus musculus," *Biomolecular NMR assignments*, vol. 1, no. 1, pp. 131–133, 2007.

[48] J. D. Forman-Kay, G. M. Clore, P. C. Driscoll, P. Wingfield, F. M. Richards, and A. M. Gronenborn, "A proton nuclear magnetic resonance assignment and secondary structure determination of recombinant human thioredoxin," *Biochemistry*, vol. 28, no. 17, pp. 7088–7097, 1989.

[49] R. A. Williamson, M. D. Carr, T. A. Frenkiel, J. Feeney, and R. B. Freedman, "Mapping the binding site for matrix metalloproteinase on the N-terminal domain of the tissue inhibitor of metalloproteinases-2 by NMR chemical shift perturbation," *Biochemistry*, vol. 36, no. 45, pp. 13882–13889, 1997.

[50] V. Munoz and L. Serrano, "Elucidating the folding problem of helical peptides using empirical parameters. II. Helix macrodipole effects and rational modification of the helical content of natural peptides," *Journal of Molecular Biology*, vol. 245, no. 3, pp. 275–296, 1995.

Paper II

Trondheim, November 9th 2012

TO WHOM IT MAY CONCERN

This letter is to certify that Elena Shumilina, Olena Dobrovolska and Alexander Dikiy have contributed a manuscript entitled "Evolution of Structural and Coordination Features within Methionine Sulfoxide Reductase B Family". This manuscript will constitute a chapter in the book "The Structural Basis of Biological Energy Generation" in the Advances in Respiration and Photosynthesis series to be published by Springer.

The manuscript is accepted for publication pending minor revisions and should be published in the near future.

Sincerely,

Editor "The Structural Basis of Biological Energy Generation"
in Advances in Photosynthesis and Respiration series

Martin F. Hohmann-Marriott, PhD
Associate Professor
phone: +47 735 93309
email: martin.hohmann-marriott@ntnu.no

Department of Biotechnology
Norwegian University of Science and Technology, NTNU
N-7491 Trondheim, Norway

Evolution of Structural and Coordination Features within Methionine Sulfoxide Reductase B Family

Elena Shumilina[1], Olena Dobrovolska[1] and Alexander Dikiy[1,]*

[1]Department of Biotechnology, Norwegian University of Science and Technology, N-7491 Trondheim, Norway;

Keywords: Methionin sulfoxide, reductase, redox stress, redox reaction, structure-functional relationships, selenoproteins, Zn-binding site.

Abbreviations: MsrB - methionine sulfoxide reductase B; Met - methionine; ROS - reactive oxygen species

[*]Corresponding author: Tel.:+4773597863, Fax: +4773591283; E-mail: alex.dikiy@biotech.ntnu.no

Table of Contents

In this review, we aim to summarize the data about evolution, sequential, structural and coordination peculiarities of the MsrBs family representing important redox proteins without taking into consideration their catalytic mechanisms.

I. Methionine residue in proteins: its oxidation and reduction

All amino acids found in proteins are susceptible to oxidation. However, only two of them, sulfur-containing cysteine and methionine, can be reversibly oxidized. Along with tryptophan these residues are the most susceptible to oxidation (*1*). The process of methionine (Met) oxidation consists of two steps (Fig.1). During the first step, Met is reversibly oxidized to methionine sulfoxide (MetO). Methionine sulfone is the product of a deeper, irreversible oxidation of MetO during the second step. Methionine oxidation can be caused either by physiological agents, like superoxide, hydrogen peroxide, hydroxyl radicals (termed as reactive oxygen species – ROS), hypochlorous acid, chloramines and some of the metal ions (as copper and iron). Physiological oxidants are also capable to carry out Met oxidation. These include either reactans able to oxidize all amino acids (*2, 3*), or hydrogen peroxide, chloramine T, N-chloro-succinimide and dimethyl sulfoxide, which oxidize Met rather selectively (*1, 4*). It was found that the ability of Met to be oxidized strongly depends on its solvent exposure (*5*). Indeed, surface-exposed residues are more readily oxidized, than the ones hidden inside the protein's structure.

Living organisms developed a defence system against unwanted protein oxidation which includes the reduction of reversibly oxidized amino acids (Cys and Met) as well as the removal of the irreversible oxidation products from the cell. Methionine sulfoxide reductases (Msrs), being part of this system, can reduce both free or protein-bound methionine sulfoxide back to methionine (Fig.1) (*2, 6*). Msr genes are found among the most conserved in almost all kindoms of life (*7, 8*).

The reversibility of Met oxidation can be seen from different points of view. First, surface-located Met residues, being exposed to the ROS attack, reduce the free radicals cellular level, thus, acting as endogenous antioxidants (*4, 5*). Second, oxidation of some Met residues may leads to loss of enzyme activity, thus Msrs reduction of such residues allows the protein recover their finctions (*9*). Finally, the process of periodic oxidation/reduction of Met in proteins plays a regulatory role inhibiting or inducing some physiological events. For example, functional role of the methionine oxidation in regulation of the Shaker voltage dependent K^+ channels was examined (*10, 11*). The results show that Met oxidation and

reduction facilitated by MsrA regulate the channel inactivation time course. Another example is the calcium-binding protein – calmodulin. Oxidation of one of the calmodulin Met residues results in a 30-fold decrease in the ability of the protein to bind calcium (*12*). Additional studies have shown that this decrease can be reversed by the addition of MsrA (*13*). It was suggest that the reversible oxidation of specific Met residues in calmodulin may regulate its activity and calcium cell homeostasis (*9, 10, 14*).

II. Different classes of Msrs

There are two main classes of Msrs: MsrA – which can reduce protein-bound or free S-epimer of methionine sulfoxide (*15-21*) and MsrB – which reduce both the protein-bound R-epimer of MetO, and, with less efficiency, free R-MetO (Fig.1) (*22-25*). Since each of the Msr classes has its own, different from each other three-dimensional fold, and significantly distinct, well conserved amino acid sequence it is possible to refer to the convergent evolution for MetO reduction in living organisms (*7, 26-28*). MsrA protein sequences, are characterized by the invariable signature motif "GCFWG/C'' (*1, 29, 30*) and MsrB - "RXCXN" (where X indicates any amino acid) or in case of MsrB1 - "RXUXF" (Fig.2, underlined with the dash line) (*6, 8, 27*). Structures of some archaea, bacterial and eukaryotic MsrB proteins have previously been determined using either X-ray or NMR techniques (Table 1, Fig.3).

The overall structural comparison of the MsrB family was reported earlier (*31, 32*). It was emphasized that the central cores of all proteins in the family are well conserved. However, while bacterial MsrB and mammalian MsrB2 have both α-helix and β-sheets as structural elements, MsrB1 represents a more flexible structure containing only β-sheets (Fig. 3). The detailed structural comparison of MsrB proteins is reported in the chapter "V. MsrB structural description".

Both MsrA and MsrB were found in the most genomes (*8*), however, their distribution within the life kingdoms is quite different. Without exception, all studied eukaryotes and cyanobacteria contain msrA and msrB genes (*28*). In prokaryotes, MsrA and MsrB genes can form two separate transcription units (*7*), with not adjacent loci on the chromosome. Alternatively, genes can either be transcriptionally (*33*) or translationally fused and form in the last case the two-domain protein - MsrAB (*34*). In prokaryotes, the Msrs distribution is rather variable: in some thermophile or anaerobic archaea no Msrs are present. Some organisms possess both Msrs, while others only MsrA but no MsrB. However, there are no MsrB-containing organisms which would also not contain MsrA (*28*). A possible explanation for the different MsrA/MsrB distribution in diverse organisms might be the existence of an

enzymatic stereospecific preference for S-epimers during Met oxidation in cells, or, alternatively, a more damaging effect of S-MetO for the cell. Otherwise, MsrA and/or MsrB might have other unknown biochemical functions along with protein repair, which would account for their unequal distribution in genomes (28).

The number of msrA and msrB ortholog copies also greatly varies in the different organisms. For example, *E.coli* contains one copy of msrA and msrB, *Arabidopsis* harbors five msrAs and nine msrBs, *Rhizobium meliloti* has three msrA and three msrB genes, while mammals have one msrA and three msrB genes, i.e. MsrB1, MsrB2 and MsrB3 (8, 14, 20, 25, 35, 36). Gene duplication is a frequent event in genome evolution across all three domains of life. The organism's necessity to have alleles expressed under different conditions might be a possible explanation of this fact (7, 37-39).

III. MsrB subcellular distribution in eukaryotic cells

In eukaryotic cells, specific metabolic reactions (defense against ROS attack (40)) take place within different intracellular compartments. For this reason most proteins are allocated to only a single cellular location and, thus, need some structural information for their correct targeting (41). Mitochondrial and endoplasmic reticulum (ER) targeting sequences (mTS and erTS, respectively) are usually placed on the N-terminus of proteins, nuclear TS (nTS) are often distributed internally, and peroxisomal TS are located at the C-termini. It has been proposed that the N-terminus of the polypeptide that bears mTS attaches to specific mitochondrial membrane receptors and initiates its translocation before the complete protein is folded (41, 42). Subsequently, the signaling sequence is cleaved from the polypeptide and excreted in the endoplasmic membrane for digestion.

Mammals have three MsrBs: MsrB1, MsrB2 and MsrB3 proteins that are localized in different cellular compartments (Table 2) (25). MsrB1 is a cytosolic and nuclear protein, MsrB2 is translocated into mitochondria while MsrB3 occurs in the ER and mitochondria (25). Alternative splicing is a conserved mechanism to regulate subcellular distribution of methionine sulfoxide reductases in mammals and other animals (43). The distribution of MsrB3 will be discussed further.

MsrB1, the most abundant and active MsrB protein in mammals, is present in cytosol and nucleus (25). nTS consists of one or two short sequences that are rich in positively charged amino acids (e.g. Lys and Arg). Nuclear localization signals can be located anywhere in the protein, but usually form patches or loops on the protein surface. Although MsrB1 does not have clearly predictable nuclear TS, it is rich in positively charged residues, which could serve to transfer the protein into the nucleus.

Both MsrB2 and MsrB3B form of MsrB3 are targeted to mitochondria. Normally, mitochondrial TS are represented by about 20–60 amino acid residues with abundant positive charges and frequent hydroxylated residues (*41*). Targeting sequences are predicted to form amphipathic α-helices in membranes or in membrane-like environments, whereas in aqueous solution they show little structural organization (*44-48*). Mouse MsrB2 contains a typical mitochondrial signal with the high proportion of arginine residues at the N-terminus (MARLLRALRGLPLLQAPGRLARG) (*25*).

MsrB3 should be considered separately. In humans and other species (e.g. zebrafish) MsrB3 were found to be translated in both ER (MsrB3A) and mitochondrial (MsrB3B) forms due to alternative splicing that produces contrasting ER and mitochondrial signals (*25*). It should be noted, that the mitochondrial location of the human MsrB3B, was not immediately clear because it had contrasting N- and C-terminal signals (*49*): on the N-terminus – mitochondrial and on the C-terminus – KDEL-like ER signal (Fig. 2, ER signal underlined with black line) (*25*). In contrast, rodents (rat and mouse) have only one form of MsrB3 found only in the ER (*49*). At the same time, mouse MsrB3 has both ER and mitochondrial signal peptides at the N-terminus: the mitochondrial signal is located between the ER signal and the common MsrB domain (Table 1). The role of the mouse MsrB3 mitochondrial signal, which is functional if placed as an N-terminal sequence, remains unclear and requires further research (*49*). MsrB3A and rodent MsrB3 are targeted to ER with C-terminus (KDEL-like) ER retention sequence (KAEL for human and RAEL for mouse MsrB) (*25, 49*).

It follows from the above discussion, that oxidized methionines can be repaired in different cellular compartments in mammals. Mitochondria, the major source of ROS in the cell, have two different MsrBs: MsrB2 and MsrB3B. MsrB2 is the most active at lower concentrations of Met and inhibited by high concentrations of the substrate, whereas MsrB3B is most active at concentrations of methionine sulfoxide more than 1mM (*25*). In addition, these two MsrBs show differential tissue expression (*50*). The above argumentation can probably explain the occurrence of two MsrBs in mitochondria.

IV. Selenocysteine in Msrs

Selenocysteine is one of the naturally occurring amino acids in proteins. In this residue, that is a cysteine analog – sulfur is replaced by selenium. These two elements have a similar electronegativity (2,58 for sulfur and 2,55 for selenium (*51*)), however, Se is a stronger nucleophile (*52-54*). Furthermore, the pK_a of selenocysteine is more acidic than that of cysteine (pK_a *(*Sec*)* = 5.2–5.6 while pK_a value of Cys is around 8.3) (*51, 53*). This means that at physiological pH the Sec residue will be deprotonated (anionic form) and more reactive,

while the cysteine residue would still remain protonated. Several oxidoreductases including glutathione peroxidase, thioredoxin reductase, methionine sulfoxide reductase, have selenocysteine (Sec) in their active sites (*55*). Due to the higher selenium nucleophilicity, selenoproteins are typically more active than their cysteine (Cys) mutants (*56*). This high catalytic activity has been regarded as a key reason why Sec is used in biological systems (*57-60*). However, the nearest protein environment of Cys or Sec, as well as the overall structure of the protein can significantly influence the redox potential of both residues, and the small changes in the amino acid composition in the region close to the active site can lead to comparable catalytic efficiency of the sulfur homolog of a selenium-dependent enzyme (*61*). Taking in account this possibility to have comparable with Se-containing proteins activity for Cys-analogs, only a higher activity of selenoproteins cannot be unique reason to use Sec instead of Cys. Thus, Sec utilization by living organisms is the compromise. On the one hand, Sec utilization provides higher activity, a broader range of substrates and of microenvironmental conditions (e.g. pH) in which enzyme is active. On the other hand, it exists the limitation imposed by electron donors (or acceptors depending on the reaction), dependence on selenium, and the availability and complexity of the Sec insertion system (*61*). It is worth mentioning, that the S to Se replacement may also completely change the protein's function (*55*). For example, a Sec-containing form of subtilisin or glutathione S-transferase converts these proteins into a peroxidase (*58, 62*)

Seleno-containing MsrBs are found in invertebrates and vertebrates, but not in bacteria or plants (*8, 63*). In contrast, selenoprotein forms of MsrA are also found in bacteria and lower eukaryotes, including a unicellular green alga (*63-65*). Mammals have one selenoprotein MsrB (MsrB1) and two Cys-containing homologs (MsrB2 and MsrB3). The MsrB1 contains selenocysteine in the place of the catalytic cysteine residue normally present in other MsrBs (Fig.2, Cys95 and Sec95 shown in red) (*8, 25*). MsrB1 is the most active in the reduction of MetO among all members of the MsrB family. The mechanism of a catalytic MetO reduction by Msrs employs a sulfenic acid chemistry and S-S/Se-S formation (*66-70*). Upon reduction of methionine sulfoxide MsrBs get oxidized forming an intramolecular disulfide/selenide-sulfide bond. The second step of the MsrBs biological cycle includes its reduction or, in other words, the reduction of the S-S/Se-S bond. Thioredoxin (Trx) is generally thought to be a natural reductant for MsrBs (*68, 70-72*). The Se-S bond is characterized by a lower potential and can be more challenging for reduction than the S-S bond (*55*). MsrB1 has highly conserved regions that distinguish it from the cysteine-containing MsrBs and can facilitate the MsrB1[ox] reduction. Fig. 2 illustrates a multiple-sequence alignment of MsrBs, and reveals

a set of conservative regions in MsrB1 with respect to other MsrBs. Recently, the simulation of the Trx-MsrB1 interaction dynamic was published (*73*). NMR titration and molecular dynamics simulations established which MsrB1 amino acids are involved in the initial interaction with Trx (Fig.2, underlined with blue lines and black rectangle). These include the S5-F7, P42, A66-G72, G75-F82 and R93-S98 amino acids. The sequence alignment shows that all these amino acids groups belong to well conserved MsrB1 regions. In addition, their composition is significantly different from non-selenium MsrBs. The spatial arrangement of these groups is shown in Fig.4 (in orange color). It therefore appears that in the course of evolution the amino acid composition of the Se-containing MsrB1 changed, due to a facilitated Se-S reduction by Trx. Indeed, the selenoprotein MsrB1 is a better substrate for Trx than Cys-containing MsrB2 and MsrB3 proteins (*55*).

V. MsrB structural description

All structurally characterized MsrB (both mammalian and bacterial) have a highly similar β-fold consisting of two β-sheets, one with three strands and another with five strands (Table 1, Fig. 3). However, the mammalian MsrB2 protein and bacterial MsrBs additionally have several α-helices at the N-terminal part of the protein. A common feature of these α-helices is that they are not rigidly fixed to the β-strand core structure. It was suggested that these helixes might assist for interaction of MsrBs with different substrates (*31*). The latter structural feature is absent in MsrB1 protein. From the sequence alignment of different MsrBs it appears that the selenoprotein MsrB1 is the smallest and the most compact among MsrBs (*31, 74*). MsrB1 is characterized, as it was shown by recent NMR studies by the presence of mobile N- and C-terminus tails (*31*). The mobility of N-terminus is extremely important because this tail is involved in catalytic activity of the MsrB. Indeed, Cys4, situated at the N-terminal tail, plays role of a resolving cysteine at the methionine sulfoxide reduction. An interesting aspect of the evolution of the MsrB superfamily is the dramatic variability with regard to length and secondary structure composition of N- and C-terminal regions in these proteins. This can be clearly observed from superimposition of MsrB structures (*31*): while the β-core is well conserved, the terminal parts of the proteins show significant differences. This feature is further supported by bacterial MsrBs: although their core structures are very similar, their N- and C-terminal regions are not superimposable. Therefore, it appears that the evolution of MsrB1 tightly couple the shorter N-terminal region which is not characterized by any secondary structure, with the use of N-terminus resolving Cys together with the catalytic Sec.

Comparative structural analysis of mouse MsrB1 and bacterial MsrBs indicates structural differences between selenoprotein and non-selenoprotein MsrBs that result in the different mechanisms of the redox reaction catalyzed by MsrBs. These differences may be even deduced from the analysis of methionine sulfoxide reductase B secondary structure. The presence of the α-helical structures in the N-terminal region of bacterial MsrBs suggests that large structural alterations during catalytic act are unlikely in these enzymes due to rigidity imposed by α-helices. Instead, the absence of any secondary structure in the N-terminus of Sec-containing MsrB1 does not impose any rigidity constraints. As the result, the catalytic reaction in Sec-containing MsrB1 occurs through the formation of internal selenide-sulfide bridge between the catalytic Sec and the resolving Cys situated on mobile N-terminus, while the reaction between methionine sulfoxide and other MsrBs occurs through formation of sulfenic acid intermediate and any disulfide bridge is formed within this reaction (*55*). Such peculiarities of catalytic reactions between Sec- and Cys-containing proteins are the result of evolutionary changes of catalytic Cys to Sec accompanied by other adaptations such as distinct resolving Cys and flexible N-terminus.

It should be also noted from structural comparison of MsrBs proteins (Fig.3) that mammalian MsrB2 is more similar to bacterial MsrBs rather than to mammalian MsrB1 as it contains three α-helices on the exterior of the protein in the N-terminal region. Indeed, the proposed catalytic mechanism for mammalian MsrB2 protein is the same as that proposed for bacterial MsrBs (*32*). This observation indicates that the functional similarities within MsrB family are determined not by organism, but rather by the nature of the catalytic residue and the protein secondary structure.

VI. Zinc ion in MsrB

Zinc is one of the most important biological metals (*75-78*). The ability to determine the primary protein structure through the translation of DNA sequences now allows the prediction of zinc binding sites and, thereby, the enzyme function (*79*). It is possible to distinguish three types of zinc binding sites: *catalytic, co-catalytic* and *structural*. For the purpose of this review, we will discuss here only the latter two classes. Structural zinc sites have four protein ligands and no bound water molecule. There is at least one short spacer between coordinating zinc ion residues, generally containing two amino acids (*80, 81*). Cysteines are the most common, but not unique ligands in such sites. The second most prevalent ligand is His. Cys and Cys/His- containing sites are also called 'zinc fingers'. However, also Asp and Glu residues are able to bind zinc, even if they perform this task less

frequently (*80, 82, 83*). Multiple Glu/Asp metal binding sites can be too flexible for zinc and lead to a weak binding constants due to fast dissociation rates of the zinc ion from such sites. The usual presence of Ca^{2+} and Mg^{2+} in such acidic ligand sites does correlate with the weak binding constants of metal ions for protein binding sites (*80*). There is an interesting example of ligands substitution in adenylate kinases belonging to different organisms. The adenylate kinase from *Bacillus stearothermophilus* contains a zinc-binding site composed of four Cys ligands (*84*). The adenylate kinase from *Bacillus subtilis* also contains a structural zinc site (*85*). In the latter case, the fourth Cys has been replaced by an Asp residue in this structural zinc site.

In co-catalytic sites, metals may be important to the overall fold of the protein as well as catalytic function. Asp and His predominate as ligands in this type of zinc sites where the frequency of coordinating residues occurrence is maximal for aspartate and histidine, and is lower for glutamate (*80*). These sites can also contain unusual zinc ligands such as amide carbonyls provided by Asn, Gln and the peptide backbone; hydroxyl groups from Ser, Thr and Tyr and the amine nitrogen of Lys or the N-terminal amino acid of the protein (*80*). The ligands are often part of a *β*-sheet or are provided by amino acids one or two residues before or after a *β*-sheet. A related interesting case was found for the co-catalytic Zn-binding site of β-lactamase. The crystallization of β-lactamase from the *Bacillus cereus* at a pH 5.6 in 0.1 M $ZnSO_4$ in a citrate/cacodylate buffer leads to a protein with one Zn ion (*86*). However, once the same protein was crystallized in different, more basic condition (at pH 7.0 in the presence of 0.5 mM $ZnSO_4$ in Tris buffer, (*87*)) the resulting protein contained two Zn ions. Thus, the prediction of a Zn binding site based only on the presence of Cys/His ligands cannot give complete pictures of the zinc distribution in proteins.

Sequence analysis of MsrBs homologs shows the existence of proteins with two different set of metal binding sites: one with the "classic" four Cys ligands for zinc binding and another with different "not classic" amino acids at the correspondent coordinating positions (*88*). It was suggested the existence of a strictly Zn-dependent family of proteins and a family that normally has Zn in metal binding site with four cysteines, but it can be expressed also as zinc-independent forms where cysteine ligands are substituted by other residues (*89*). Thus, for some protein family the same protein may exist either with zinc or without. A systematic genome analysis showed that approximately 20% of the zinc protein families in the Protein Data Bank (PDB) have a significant number of zinc-independent forms (zinc-dependent form was described as a homolog which preserved three or four known ligands Cys or His, and a zinc-independent form was defined by having fewer than three known ligands) (Fig. 6) (*88*).

In another study of the Zur (a repressor of zinc transport regulation in bacteria) (*90*) it was suggested that non-zinc-binding paralogs were expressed under zinc-restricted conditions to partially replace the zinc-binding proteins, thereby freeing up some zinc to be used by the essential zinc-binding proteins.

The majority of MsrB proteins have four cysteines that bind Zn ion (*30, 91*). They include MsrB1, MsrB2, MsrB3 and the most part of bacterial MsrB (Fig.2). Some of the bacterial MsrB subclasses, however, have only Zn-independent forms (e.g. MsrB domain of MsrAB from *Neisseriaceae*, *Bacillus* sp. and *St. pneumoniae*), some have only Zn-containing forms (MsrB from *B. pseudomallei*, *X. campestris*) and others - both forms (MsrB from *S. meliloti; V. cholera*) (unpublished result from (*88*)). For example, *S. meliloti* has both Zn-dependent and independent forms. In Zn-independent form, four cysteines are substituted by Asp, Ser and Gly (Fig.2, dash rectangles). While Asp and Ser can hypothetically bind Zn, no Gly has been mentioned as Zn ligand. However, in the immediate vicinity of Gly highly conserved His, Asp, Glu and Ser are present and it could be that the latter residues also weekly coordinate zinc ion. The same "not classic" ligands are present in MsrB domain of bacterial fused MsrAB.

Several structures of MsrB proteins are presently reported (Table 1, Fig. 5). Some of them are characterized by the presence of Zn ion, while other lacks it. From the structural comparison (Fig. 5), it is evident that both Zn-containing and Zn-independent forms have identical structure of metal binding site among themselves and to the corresponding region of no metal homologs. All of them have two unstructured loops between four β-sheets. This region seems very well conserved in all the reported structures. It should be noted, that the unpublished structure of MsrB1 from *Homo sapiens* has iron ion in the metal binding site, however all the protein's structural features to the rest of the proteins reported in Fig. 5 remained unaltered. In the cases where Zn ion is absent, it is possible that some other bonds and not the Zn ion preserve the fold of the region. Some Zn-binding site studies were carried out by different laboratories (*30, 91*). In these studies it was shown that in MsrB from *Drosophila* and *E.coli* mutation of Cys at the positions 23, 26, 71 and 74 (numbering corresponds to MsrB1 from *Mus musculus*) on GSGS, DSSS and DSSA results in Zn losing by protein. The reversible mutation of DSSS pattern to CCCC results in a more stable, Zn-containing protein. It was suggested, that the role of the metal in Zn-containing MsrBs is to stabilize the core structure and thus, to adopt the active site conformation for efficient reductase activity. However, as it was stressed, this assumption is in apparent contradiction with the fact that the bacterial MsrB, lacking the two CXXC signatures is as active as the Zn-containing MsrB (*91*). It was

suggested that a more careful analysis of stabilizing elements for MsrBs that do not contain CXXC motif should be performed.

The analysis of the crystallization conditions for the published structures of MsrBs where Zn ion is absent (no "classic" Zn-binding CXXC motif) suggests that the crystallization occurs in mild acidic environments (*27, 92, 93*). Some of the "not classical" Zn ligands can be protonated at these conditions. As mentioned above, Zn binding can be influenced by the solution pH value. In other words, when the Zn ligands are protonated, they are not able to coordinate metal ion. In addition, the reported absence of Zn ion for some experimentally characterized MsrB may simply be due to the fact that metal ion is lost during the protein's isolation and purification since the coordination bonds between Zn and oxygen containing ligands are weaker with respect to usual cysteine ligand. Thus, zinc ion in such cases could escape detection. The discussed above structural similarities within metal coordinating loops among all MsrBs would indirectly support this suggestion.

Recently in our laboratory it was demonstrated (*94*) that Cys containing MsrB1 expressed in *E.coli* in the cobalt-containing M9 media can replace Zn by Co. Paramagnetic Co ion insertion can be easily detected by both UV-Vis (Co-containing protein has a characteristic light blue color) and NMR spectroscopies (by the appearance of paramagnetically shifted signals). It is known (*95, 96*), that Co ion can be used to substitute Zn ion since the both metal ions have similar radii and coordination properties. Thus, the proposed by us technique represents an easy method to check and confirm whether Zn binds only by four cysteines and no other ligands at different, more basic pH can coordinate metal ion.

Summarizing, most MsrBs have four cysteines that bind zinc. It has been assumed that the metal binding site plays a structural role in the MsrB family. Some of the organisms have also MsrBs proteins, which possess "not classical" cysteine ligands. In such cases, the determined structures also report a conserved overall folding that might suggest that the proteins still weakly bind metal ion.

VII. Conclusion

Living organisms develop efficient and complex defense system against oxidative stress. The MsrBs proteins play an important role in this system. They can reduce R-isomer of methionine sulfoxide back to methionine. Such reversibility of Met oxidation provides the organism the possibility to reduce intracellular level of the ROS, restore the enzymatic function of oxidized protein and inhibit or induce certain cellular events.

MsrBs were found in about all kingdoms of life. The absence of MsrB in some thermophilic and anaerobic bacteria may indicate that either the MetO reduction at high temperature does

not require a catalyst, or MetO is not produced in significant amounts. Whereas prokaryotes have only one type of MsrB, mammals possess three: MsrB1 (the most active mammalian MsrB1 containing catalytic selenocysteine), MsrB2 and MsrB3. The use of Sec in proteins can represent a compromise between a more active enzyme that can work with different substrates at a wider pH range, and the complex genetic mechanism of Se insertion with the more challenging Se-S bond reduction. The different eukaryotic MsrBs have peculiar cellular distribution to be effectively involved in various redox and other catalytic pathways occurring in specific compartments. Mostly MsrBs compartmentalization is achieved by alternative splicing and the proteins are imported into the organelles by specific signals.

Generally, the structure of all catalytic cores of structurally characterized MsrBs is rather similar containing two β-sheets. In addition, bacterial MsrBs and mammalian MsrB2 are characterized by the presence of N-terminal α-helixes, while mammalian MsrB1 has only β-sheets and very flexible N- and C-termini. This flexibility plays an important role in the MsrB1 catalysis. The structural difference between mammalian MsrB1 and other MsrB proteins is due to evolutionary changes occurred with MsrB class of proteins that also presumes different catalytic mechanisms for MsrB1 and other cysteine homologs belonging to this class.

Mostly, MsrBs are Zn-containing proteins. The existence of the metal was predicted by bioinformatics based on the presence of two Zn-binding CXXC motifs. However, some of the bacterial MsrBs have different well-conserved amino acids at the positions of the four cysteines, even if the structures of metal binding site in Zn-containing protein and the corresponding region in MsrBs without metal are very similar. Theoretically, there is a possibility that Zn can be coordinated by non-cysteine ligands at more alkaline pH and with less affinity. Nevertheless, this speculation requires experimental confirmation.

Thus, in the course of evolution the members of MsrB family developed their amino acids sequence and their fold for efficient functioning in different organisms, different cell compartments at different environmental conditions.

Acknowledgements

Tables

Table 1. Structural data of MsrB proteins

Superkingdom / Phylum	Species	Method	PDB entries	Year	Ref.
MsrB					
Proteobacteria	Neisseria gonorrhoeae	X-ray	1L1D	2002	(*27*)
Proteobacteria	Burkholderia pseudomallei	X-ray	3CEZ /3CXK	2008	(*97*)
Archaea/ Euryarchaeota	Methanothermobacter thermautotrophicus	NMR	2K8D	2008	UP
Bacteria	Streptococcus pneumoniae	X-ray	3E0M/3E0O	2009	(*92*)
Proteobacteria	Xanthomonas campestris	X-ray	3HCI	2009	(*93*)
Proteobacteria	Neisseria meningitidis	X-ray	3HCG	2009	(*93*)
Bacteria	Bacillus subtilis	NMR	2KZN/1XM0	2012	(*98*)
MsrB1					
Mammalia	Mus musculus, MsrB1	NMR	2KAO/2KV1	2010	(*31*)
Mammalia	Homo sapience	X-ray	3MAO	2010	UP
MsrB2					
Mammalia	Mus musculus, MsrB2	NMR	2L1U	2011	(*32*)
MsrB3 – not reported					

UP-unpublished; Ref. - references

Table 2. Subcellular distribution of mammalian MsrBs

MsrB1	Cytosol, Nucleus	MsrB1
MsrB2	Mitochondria	mTS-MsrB2
Rodent MsrB3	ER	erTS-mTS-MsrB3-RAEL
MsrB3A	ER	MsrB3A-KAEL
MsrB3B	Mitochondria	mTS-MsrB3B-KAEL

Figure legend

Fig. 1. Mechanism of methionine oxidation. Step I: reversible oxidation of methione gives two enantiomers: S- and R-methionine sulfoxide. Step II: irreversible oxidation of methionine sulfoxide to methionine sulfone.

Fig. 2. Multiple sequence alignment of proteins belonging to MsrBs family.

Numbering is consensus sequence with mouse MsrB1. In abbreviation MsrAB only MsrB domain of protein is shown. Conserved catalytic site of MsrBs stressed with black dash line. erTS of MsrB3 underline in black. Amino acids from MsrB1 involved in the initial interacting stage with Trx stressed with blue line (*73*). The organisms abbreviation is as follows: NG - *Neisseria Gonorrhoeae*; NM – *Neisseria meningitides*; BC - *Bacillus cereus*; HI - *Haemophilus influena*; HP - *Helicobacter pylori*; SP - *Streptococcus pneumoniae*; BS - *Bacillus subtilis*; CO – Clostridium, *Alkaliphilus oremlandii*; SM147 - *Sinorhizobium meliloti* SM11, 147 amino acids; SM164 - *Sinorhizobium meliloti* SM11, 164 amino acids; SM135 - *Sinorhizobium meliloti* SM11, 135 amino acids; EC - *Escherichia coli*; BP - *Burkholderia pseudomallei*; AV - *Azotobacter vinelandii*; MT - *Methanothermobacter thermautotrophicus*; HA - *Herminiimonas arsenicoxydans*; VO - *Vibrio ordalii*; MM - *Mus Musculus*; HS - *Homo sapience*; DR - *Danio rerio* (zebrafish); AT - *Arabidopsis thaliana*; RN - *Rattus norvegicus*; PA - *Pongo abelii*; OS - *Oryza sativa* Japonica Group; BT - *Bos taurus*.

Fig. 3. Three-dimensional structures of proteins belonging to MsrB family.

A – MsrB *N. Gonorrhoeae*, 3HCH, (*27*); B – MsrB *X. campestris*, 3HCI, (*93*); C – MsrB1 *H. sapience*, 3MAO, not pubblished; D – MsrB *B. subtilis*, 1XM0, not pubblished; E – MsrB2 *M. musculus*, 2L1U, (*32*); F – MsrB1 *M. musculus*, 2KV1, (*31*).

Fig. 4. Three-dimensional structure of mouse MsrB1, 2KV1,(*31*). In orange: conserved regions of the Se-containing MsrB1 involved in the initial stage of Trx interaction (the data taken from (*73*)).

Fig. 5. Three-dimensional structure of Zn-binding site and thecorresponding regions of MsrBs family. For the explanation see text. A - MsrB1, *M. musculus*, 2KV1, (*31*); B - MsrB1, *H. sapience*, 3MAO, not published; C - MsrB2, *M. musculus*, 2L1U, (*32*); D - MsrB, *X. campestris*, 3HCI, (*93*); E - *B. pseudomallei,* 3CEZ/3CXK, not published; a – MsrB, *N. meningitidis*, 3HCH, (*93*); b - MsrB , *N. gonorhoeae,* 1L1D, (*27*); c- MsrB, *S. pneumoniae* 3E0O, (*92*); d – MsrB, *B.subtilis,* 1XM0 not published. Blue color: cysteine and "not classic" ligands in metal binding site; magenta: catalytic cysteine.

Fig. 6. Occurrence of Zn-dependent and Zn-independent forms of representative Zn protein families in bacteria. Zn (+), organisms only containing Zn-dependent form; Zn (-), organisms only containing Zn-independent form; Both, organisms containing both forms. The figure is taken from (*88*).
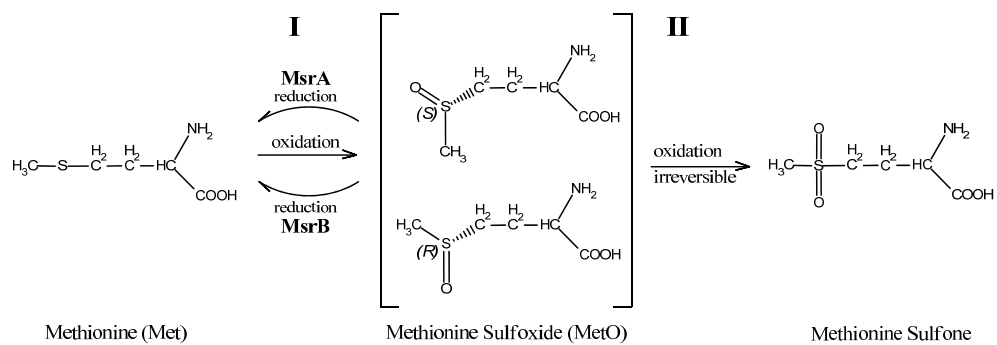
Figure 1.

Methionine (Met)   Methionine Sulfoxide (MetO)   Methionine Sulfone

Figure 2.

```
          ...  .|....|....|....|  ..|...  .|....|....|....|....|.      ..|....|...  .  |....|....|....|..  ..|....  |..  ..|....|.  ....|....  ....|....|
            5       15       25       35       45                    65       75       85       95      105      125
MsrAB NG   nsa teyafsheydhlfkpgiyv dvvsge plfssadkydsgcgwpsftrpid-------aksvtehddfsfnm rrtevraraadshlghvf p-dgprdkggl ryeingasl kfiplegmda agygalkgev k
MsrAB NM   nsa teyafsheydhlfkpgiyv dvvsge plfssadkydsgcgwpsftrpid-------aksvtehddfsynm rtevraadshlghvf p-dgprdkggl ryeinsasl kfiplegmda agygalkskv k
MsrAB BC   eng tepfrneywnhkeeglyv divsge plftsldkfdsgcgwpsftkpvm-------sasvkekmdvshnm trtevrakegdahlghvf p-dgpg-pngl ryeinsasl rfiplekelek egygdfliilf gnkk
MsrAB HI   nkh tersfrneywdnfqpgiyv dvttge pvfssndkfesgcgwpsftkpii-------kdvvhyetdnsfnm qrtevlaragmahlghvf d-dgpkdkggl ryeinsasl kfiplaemek agygyliqsi kk
MsrAB HP   nkh tekpfeneyynkeeegiyv dittge plfssadkydsgcgwpsfskpin-------kdvvkyeddeslnr krievlsrigkahlghvf n-dgpkelggl ryeinsasl rfiplkdmek egygefipyi kkgelkkyiq dkkth
MsrAB SP   eaa teapftnaydqtfeegiyv dittge plffakdkfasgcgwpsfsrpis-------kelihyykdlshgm erievrsrsgshlghvf t-dgprelggl ryeinsasl rfvakdemek agygyllpyi nk

MsrB  BS   nng teppfqneywdhkeeglyv divsgk plftskdkfdsqcgwpsftkpi--------eeeveekldtshgm irtevrartadshlghvf n-dgpg-pngl ryeinsaal rfvpkhklke egyesylhlf nklehh
MsrB  CO   nna terpftseynahdeeglyv dlvsge plfskhkfdagcgwpsftepie-------penvielrdnshgm itevrdkhdshlghvf p-dgpkeaggl rflikkedlek egygeylkif e
MsrB  SM147 neg terp-tgeytdnkrpgiyv dlvspe plfasadkfdsgcgwpsftkpiv-------panvnelrdnshgm itevrsahdshlghvf p-dgpqdrggl rflpreemea egygggyfnqv edi
MsrB  SM164 eeg teppssaldne-krngvfh sagoel pvyssdakydsgtgwpsfwqalp--------gaigtreddtlfv trtechdrdsghlghif d-dgpp-ptgk rheinglal sfvpeta
MsrB  SM135 ehg terp-tgpfldnkrhgtyk gyoqgr alfrsdtkfdsgcgwpsyfapvd-------pdaiaeyvdrshlm vrtelrgadsdahlghvf p-dgpg-ptgl ryelngtam tfeed

MsrB  EC   nhg teppftgrllhnkrdgvyh elieda plfhsqtkydsgcgwpsfyepvs-------eesiryikdlshgm qrieirogsedahlghvf p-dgpq-ptge ryevnnsasl rftdgengee ing
MsrB  BP   haa teppftgeytdtedagiyh cvvcgt alfesgakyhsgcgwpsyfkpid-------gevidekmdythgm trvevrenqeqahlghvf e-dgprdktgl ryeinsaal nfeak
MsrB  AV   lgg terafsgeyyatsvpgvyh cacoda plfdsdakydagcgwpsyfqpva-------asaltsrddyshgm hrievlearedahlghvf p-dgpp-ptgl ryeinsasl rlkprdeg
MsrB  MT   kag teppftgkyhdlhddgiyr eicogt dlfdsetkfdsgtgwpsfydvvs-------ehniklredrslgm vrcevlcaredahlghvf d-dgpr-ptgk ryemnsaal kfiprdqig
MsrB  HA   haa teapfsgrywdhhehgiyt cvccnt plfasdakfdsgcgwpsyftaln-------pdnvkekidraygm irtevienvedahlghvf n-dgpp-ptgl ryeinsasl rfdatpdlne s
MsrB  VO   qrg teapfsgkllhnrdtgvya etccss plfvsnnkydsgcgwpsfdapik--------dtairyledlthgm irteireavedshlghvf p-dgpq-ttge rfevnsvsl ifnkk

MsrB2 MM   ekg teapfsgmylnnketgmyh cvccds plfssekkycsgtgwpsfseaygskgsdesht gilrrldtslgc prmevvekqeeahlghvf p-dgpk-ptgq rfeinsval kfkpskp
MsrB2 HS   ekg teppfsgiylnnkeagmyh cvccds plfssekkycsgtgwpsfseahgtsgsdesht gilrrldtslgs artevvekqeeahlghvf d-dgpg-pngq rfeinsval kfkprkh
MsrB2 DR   ekg tevpfsgiylnhnevgmyh cvccds plfsseakydagagwpsfceahgtwekdeshan ivrrpdnslgs tgtevickhedahlghvf e-dgpd-ptgq rfeinsval nfkprek
MsrB2 AT   qkg teypgtgeynkvfddgiyc eagogt plyksttkfdsgcgwpaffdglp---------gaitrtpdpdgr- -rieiteaaegghlghvf kgegfptptde rhevnsisl kftpenptl
MsrB2 RN   ekg teapfsgmylknketgmyh cvccds plfssekkycsgtgwpsfseahgtkgsdesht gilrrldtslgc prmevvekqeeahlghvf p-dgpdp-tgq rfeinsval kfkpskp

MsrB3 MM   ekg tesafegeythhkdpgiyk cvvcgt plfksetkfdsgsgwpafhdvis-------seaieftddfsygm hrvettsesqgaqahlghlf d-dgprp-tgk ryeinsasl sftpadssea egsgikesgs paaadrael
MsrB3 HS   ekg tesafegeythhkdpgiyk cvvcgt plfksetkfdsgsgwpsfhdvis-------seaitftddfsygm hrvettsesqgaqahlghvf d-dgprp-tgk ryeinsasl sftpadssgt aeggsgvasp aqadkael
MsrB3 PA   ekg tesafegeythhkdpgiyk cvvcgt plfksetkfdsgsgwpsfhdvis-------seaitftddfsygm hrvettsesqgaqahlghvf d-dgprp-tgk ryeinsasl sftpadssga aegesgvasp aqadkael
MsrB3 DR   qkg tesaftgkyvdnkeegtyh cvvcga plfmsekkfdsgsgwpsffdlin-------kdaiaqtddfsygm hrvettsesqgahlghlf d-dgprp-tkk ryeinsasl sfqpknsate eveasgaagq ssasknee1
MsrB3 AT   qkg teypgtgeyvnfdkegvyg cvgcna plyksttkfnagcgwpaffegyp---------gaitrttdpdgr- -rieineatcgghlghvf kgegfatptde rhevnsivl kftpaass1
MsrB3 OS   lkg teypgtgeydklfaegvye cagogt plyksstkfnsgcgwpafyegfp---------gaiartpdpdgr- -rieiteaaegghlghvf kgegfntptde rhevnsisl kfipasedsk 1

MsrB1 MM   msf esfipgevfqnhfepgiyv eakesy elfsshskyahsspwpaftetih-------ppisvtkcpeknr-p ealkvsegkegnglghef lndgpkr-gqs rfuifsssl kfvpkgkeaa asqgh
MsrB1 HS   msf esfipgevfqnhfepgiyv eakogy elfssrskyahsspwpaftetih-------assvakrpehnr-a ealkvsegkegnglghef lndgpkp-gqs rfuifsssl kfvpkgkets asqgh
MsrB1 PA   msf esfipgevfqnhfepgiyv eakogy elfssrskyahsspwpaftetih-------assvakrpehnr-a ealkvsegkegnglghef lndgpkp-gqs rfuifsssl kfvpkgkets asqgh
MsrB1 DR   msf esfipgeiykdhfesgmyv cagogy elfssrskyehsspwpaftetih-------efpsvs+qeerw--- gaykvregkegnglghef vndgpkh-gls rfuifsssl kfipkvkneq q
MsrB1 BT   msf esfihgeifqnhfepgiyv eakogy elfssrskyahsspwpaftetih-------afpsvakrpehnr-p ealkvsegregnglghef lndgpkr-gqs rfuifsssl kfipkaeets asqgq
```

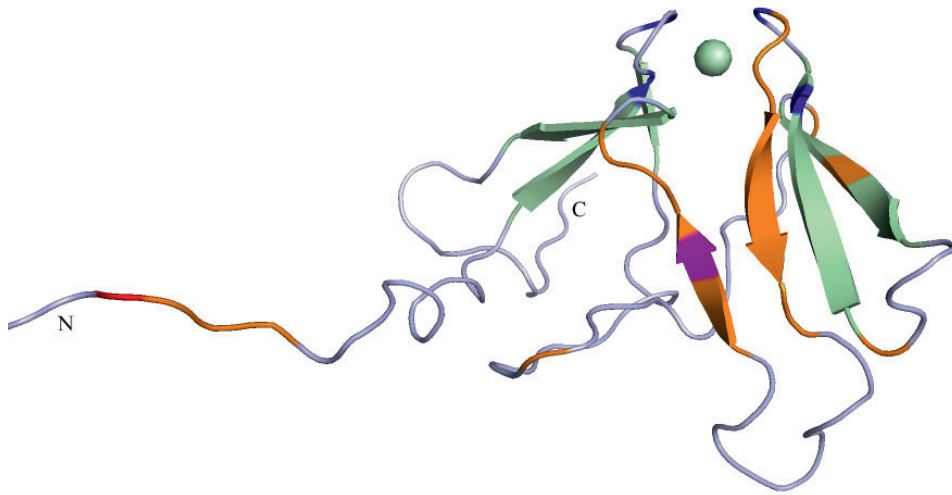**Figure 3.**

A

B

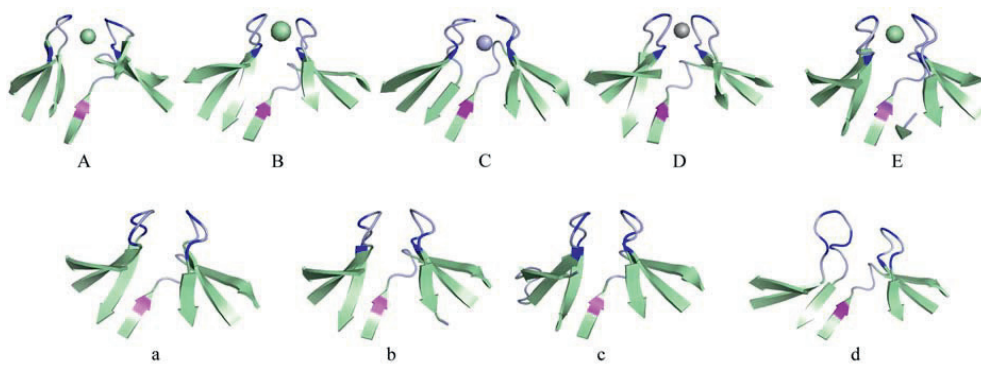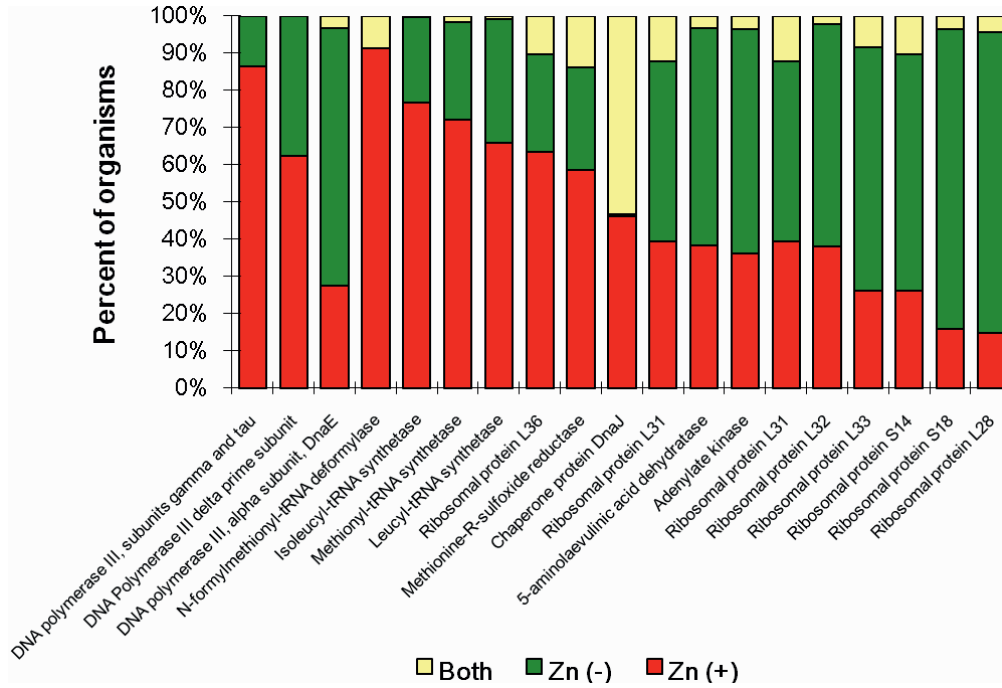C

D

E

F

Figure 4.

N

C

Figure 5.

Figure 6.

REFERENCES

1. Vogt, W. (1995) Oxidation of methionyl residues in proteins: tools, targets, and reversal, *Free Radic Biol Med 18*, 93-105.
2. Brot, N., and Weissbach, H. (1983) Biochemistry and physiological role of methionine sulfoxide residues in proteins, *Arch Biochem Biophys 223*, 271-281.
3. Shechter, Y. (1986) Selective oxidation and reduction of methionine residues in peptides and proteins by oxygen exchange between sulfoxide and sulfide, *J Biol Chem 261*, 66-70.
4. Levine, R. L., Mosoni, L., Berlett, B. S., and Stadtman, E. R. (1996) Methionine residues as endogenous antioxidants in proteins, *Proc Natl Acad Sci U S A 93*, 15036-15040.
5. Levine, R. L., Berlett, B. S., Moskovitz, J., Mosoni, L., and Stadtman, E. R. (1999) Methionine residues may protect proteins from critical oxidative damage, *Mech Ageing Dev 107*, 323-332.
6. Brot, N., Weissbach, L., Werth, J., and Weissbach, H. (1981) Enzymatic reduction of protein-bound methionine sulfoxide, *Proc Natl Acad Sci U S A 78*, 2155-2158.
7. Ezraty, B., Aussel, L., and Barras, F. (2005) Methionine sulfoxide reductases in prokaryotes, *Biochim Biophys Acta 1703*, 221-229.
8. Kryukov, G. V., Kumar, R. A., Koc, A., Sun, Z., and Gladyshev, V. N. (2002) Selenoprotein R is a zinc-containing stereo-specific methionine sulfoxide reductase, *Proc Natl Acad Sci U S A 99*, 4245-4250.
9. Brot, N., and Weissbach, H. (2000) Peptide methionine sulfoxide reductase: biochemistry and physiological role, *Biopolymers 55*, 288-296.
10. Ciorba, M. A., Heinemann, S. H., Weissbach, H., Brot, N., and Hoshi, T. (1997) Modulation of potassium channel function by methionine oxidation and reduction, *Proc Natl Acad Sci U S A 94*, 9932-9937.
11. Ciorba, M. A., Heinemann, S. H., Weissbach, H., Brot, N., and Hoshi, T. (1999) Regulation of voltage-dependent K+ channels by methionine oxidation: effect of nitric oxide and vitamin C, *FEBS Lett 442*, 48-52.
12. Yao, Y., Yin, D., Jas, G. S., Kuczer, K., Williams, T. D., Schoneich, C., and Squier, T. C. (1996) Oxidative modification of a carboxyl-terminal vicinal methionine in calmodulin by hydrogen peroxide inhibits calmodulin-dependent activation of the plasma membrane Ca-ATPase, *Biochemistry 35*, 2767-2787.
13. Sun, H., Gao, J., Ferrington, D. A., Biesiada, H., Williams, T. D., and Squier, T. C. (1999) Repair of oxidized calmodulin by methionine sulfoxide reductase restores ability to activate the plasma membrane Ca-ATPase, *Biochemistry 38*, 105-112.
14. Hansel, A., Heinemann, S. H., and Hoshi, T. (2005) Heterogeneity and function of mammalian MSRs: enzymes for repair, protection and regulation, *Biochim Biophys Acta 1703*, 239-247.
15. Sharov, V. S., Ferrington, D. A., Squier, T. C., and Schoneich, C. (1999) Diastereoselective reduction of protein-bound methionine sulfoxide by methionine sulfoxide reductase, *FEBS Lett 455*, 247-250.
16. Moskovitz, J., Poston, J. M., Berlett, B. S., Nosworthy, N. J., Szczepanowski, R., and Stadtman, E. R. (2000) Identification and characterization of a putative active site for peptide methionine sulfoxide reductase (MsrA) and its substrate stereospecificity, *J Biol Chem 275*, 14167-14172.

17. Lee, B. C., Lee, Y. K., Lee, H. J., Stadtman, E. R., Lee, K. H., and Chung, N. (2005) Cloning and characterization of antioxidant enzyme methionine sulfoxide-S-reductase from Caenorhabditis elegans, *Arch Biochem Biophys 434*, 275-281.

18. Boschi-Muller, S., Gand, A., and Branlant, G. (2008) The methionine sulfoxide reductases: Catalysis and substrate specificities, *Arch Biochem Biophys 474*, 266-273.

19. Caldwell, P., Luk, D. C., Weissbach, H., and Brot, N. (1978) Oxidation of the methionine residues of Escherichia coli ribosomal protein L12 decreases the protein's biological activity, *Proc Natl Acad Sci U S A 75*, 5349-5352.

20. Moskovitz, J., Weissbach, H., and Brot, N. (1996) Cloning the expression of a mammalian gene involved in the reduction of methionine sulfoxide residues in proteins, *Proc Natl Acad Sci U S A 93*, 2095-2099.

21. Kuschel, L., Hansel, A., Schonherr, R., Weissbach, H., Brot, N., Hoshi, T., and Heinemann, S. H. (1999) Molecular cloning and functional expression of a human peptide methionine sulfoxide reductase (hMsrA), *FEBS Lett 456*, 17-21.

22. Huang, W., Escribano, J., Sarfarazi, M., and Coca-Prados, M. (1999) Identification, expression and chromosome localization of a human gene encoding a novel protein with similarity to the pilB family of transcriptional factors (pilin) and to bacterial peptide methionine sulfoxide reductases, *Gene 233*, 233-240.

23. Bar-Noy, S., and Moskovitz, J. (2002) Mouse methionine sulfoxide reductase B: effect of selenocysteine incorporation on its activity and expression of the seleno-containing enzyme in bacterial and mammalian cells, *Biochem Biophys Res Commun 297*, 956-961.

24. Hansel, A., Jung, S., Hoshi, T., and Heinemann, S. H. (2003) A second human methionine sulfoxide reductase (hMSRB2) reducing methionine-R-sulfoxide displays a tissue expression pattern distinct from hMSRB1, *Redox Rep 8*, 384-388.

25. Kim, H. Y., and Gladyshev, V. N. (2004) Methionine sulfoxide reduction in mammals: characterization of methionine-R-sulfoxide reductases, *Mol Biol Cell 15*, 1055-1064.

26. Kauffmann, B., Favier, F., Olry, A., Boschi-Muller, S., Carpentier, P., Branlant, G., and Aubry, A. (2002) Crystallization and preliminary X-ray diffraction studies of the peptide methionine sulfoxide reductase B domain of Neisseria meningitidis PILB, *Acta Crystallogr D Biol Crystallogr 58*, 1467-1469.

27. Lowther, W. T., Weissbach, H., Etienne, F., Brot, N., and Matthews, B. W. (2002) The mirrored methionine sulfoxide reductases of Neisseria gonorrhoeae pilB, *Nat Struct Biol 9*, 348-352.

28. Zhang, X. H., and Weissbach, H. (2008) Origin and evolution of the protein-repairing enzymes methionine sulphoxide reductases, *Biol Rev Camb Philos Soc 83*, 249-257.

29. Lowther, W. T., Brot, N., Weissbach, H., and Matthews, B. W. (2000) Structure and mechanism of peptide methionine sulfoxide reductase, an "anti-oxidation" enzyme, *Biochemistry 39*, 13307-13312.

30. Kumar, R. A., Koc, A., Cerny, R. L., and Gladyshev, V. N. (2002) Reaction mechanism, evolutionary analysis, and role of zinc in Drosophila methionine-R-sulfoxide reductase, *J Biol Chem 277*, 37527-37535.

31. Aachmann, F. L., Sal, L. S., Kim, H. Y., Marino, S. M., Gladyshev, V. N., and Dikiy, A. (2010) Insights into function, catalytic mechanism, and fold evolution of selenoprotein methionine sulfoxide reductase B1 through structural analysis, *J Biol Chem 285*, 33315-33323.

32. Aachmann, F. L., Kwak, G. H., Del Conte, R., Kim, H. Y., Gladyshev, V. N., and Dikiy, A. (2011) Structural and biochemical analysis of mammalian methionine sulfoxide reductase B2, *Proteins 79*, 3123-3131.

33. Singh, V. K., and Moskovitz, J. (2003) Multiple methionine sulfoxide reductase genes in Staphylococcus aureus: expression of activity and roles in tolerance of oxidative stress, *Microbiology 149*, 2739-2747.

34. Olry, A., Boschi-Muller, S., Marraud, M., Sanglier-Cianferani, S., Van Dorsselear, A., and Branlant, G. (2002) Characterization of the methionine sulfoxide reductase activities of PILB, a probable virulence factor from Neisseria meningitidis, *J Biol Chem 277*, 12016-12022.

35. Vougier, S., Mary, J., and Friguet, B. (2003) Subcellular localization of methionine sulphoxide reductase A (MsrA): evidence for mitochondrial and cytosolic isoforms in rat liver cells, *Biochem J 373*, 531-537.

36. Lee, B. C., Dikiy, A., Kim, H. Y., and Gladyshev, V. N. (2009) Functions and evolution of selenoprotein methionine sulfoxide reductases, *Biochim Biophys Acta 1790*, 1471-1477.

37. Ohno, S. (1970) *Evolution by gene duplication*, Springer-Verlag, Berlin, New York,.

38. Zhang, J. (2003) Evolution by gene duplication: an update, *Trends in Ecology & Evolution 18*, 292-298.

39. Zhang, J. (2012) Genetic Redundancies and Their Evolutionary Maintenance

Evolutionary Systems Biology, (Soyer, O. S., Ed.), pp 279-300, Springer New York.

40. Go, Y. M., and Jones, D. P. (2008) Redox compartmentalization in eukaryotic cells, *Biochim Biophys Acta 1780*, 1273-1290.

41. Neupert, W. (1997) Protein import into mitochondria, *Annu Rev Biochem 66*, 863-917.

42. Truscott, K. N., Pfanner, N., and Voos, W. (2001) Transport of proteins into mitochondria, *Rev Physiol Biochem Pharmacol 143*, 81-136.

43. Kim, H. Y., and Gladyshev, V. N. (2006) Alternative first exon splicing regulates subcellular distribution of methionine sulfoxide reductases, *BMC Mol Biol 7*, 11.

44. von Heijne, G. (1986) Towards a comparative anatomy of N-terminal topogenic protein sequences, *J Mol Biol 189*, 239-242.

45. von Heijne, G. (1986) Mitochondrial targeting sequences may form amphiphilic helices, *EMBO J 5*, 1335-1342.

46. Roise, D., and Schatz, G. (1988) Mitochondrial presequences, *J Biol Chem 263*, 4509-4511.

47. Lemire, B. D., Fankhauser, C., Baker, A., and Schatz, G. (1989) The mitochondrial targeting function of randomly generated peptide sequences correlates with predicted helical amphiphilicity, *J Biol Chem 264*, 20206-20215.

48. von Heijne, G. (1990) Protein targeting signals, *Curr Opin Cell Biol 2*, 604-608.

49. Kim, H. Y., and Gladyshev, V. N. (2004) Characterization of mouse endoplasmic reticulum methionine-R-sulfoxide reductase, *Biochem Biophys Res Commun 320*, 1277-1283.

50. Jung, S., Hansel, A., Kasperczyk, H., Hoshi, T., and Heinemann, S. H. (2002) Activity, tissue distribution and site-directed mutagenesis of a human peptide methionine sulfoxide reductase of type B: hCBS1, *FEBS Lett 527*, 91-94.

51. Muttenthaler, M., and Alewood, P. F. (2008) Selenopeptide chemistry, *J Pept Sci 14*, 1223-1239.

52. Hondal, R. J., Nilsson, B. L., and Raines, R. T. (2001) Selenocysteine in native chemical ligation and expressed protein ligation, *J Am Chem Soc 123*, 5140-5141.

53. Huber, R. E., and Criddle, R. S. (1967) Comparison of the chemical properties of selenocysteine and selenocystine with their sulfur analogs, *Arch Biochem Biophys 122*, 164-173.

54. Pearson, R. G., Sobel, H., and Songstad, J. (1968) Nucleophilic Reactivity Constants toward Methyl Iodide and Trans-[Pt(Py)2cl2], *Journal of the American Chemical Society 90*, 319-&.

55. Kim, H. Y., and Gladyshev, V. N. (2005) Different catalytic mechanisms in mammalian selenocysteine- and cysteine-containing methionine-R-sulfoxide reductases, *PLoS Biol 3*, e375.

56. Metanis, N., Keinan, E., and Dawson, P. E. (2006) Synthetic seleno-glutaredoxin 3 analogues are highly reducing oxidoreductases with enhanced catalytic efficiency, *J Am Chem Soc 128*, 16684-16691.

57. Bock, A., Forchhammer, K., Heider, J., and Baron, C. (1991) Selenoprotein synthesis: an expansion of the genetic code, *Trends Biochem Sci 16*, 463-467.

58. Bell, I. M., Fisher, M. L., Wu, Z. P., and Hilvert, D. (1993) Kinetic studies on the peroxidase activity of selenosubtilisin, *Biochemistry 32*, 3754-3762.

59. Stadtman, T. C. (1996) Selenocysteine, *Annu Rev Biochem 65*, 83-100.

60. Hatfield, D. L., and Gladyshev, V. N. (2002) How selenium has altered our understanding of the genetic code, *Mol Cell Biol 22*, 3565-3576.

61. Gromer, S., Johansson, L., Bauer, H., Arscott, L. D., Rauch, S., Ballou, D. P., Williams, C. H., Jr., Schirmer, R. H., and Arner, E. S. (2003) Active sites of thioredoxin reductases: why selenoproteins?, *Proc Natl Acad Sci U S A 100*, 12618-12623.

62. Yu, H. J., Liu, J. Q., Bock, A., Li, J., Luo, G. M., and Shen, J. C. (2005) Engineering glutathione transferase to a novel glutathione peroxidase mimic with high catalytic efficiency. Incorporation of selenocysteine into a glutathione-binding scaffold using an auxotrophic expression system, *J Biol Chem 280*, 11930-11935.

63. Fomenko, D. E., Xing, W., Adair, B. M., Thomas, D. J., and Gladyshev, V. N. (2007) High-throughput identification of catalytic redox-active cysteine residues, *Science 315*, 387-389.

64. Novoselov, S. V., Rao, M., Onoshko, N. V., Zhi, H., Kryukov, G. V., Xiang, Y., Weeks, D. P., Hatfield, D. L., and Gladyshev, V. N. (2002) Selenoproteins and selenocysteine insertion system in the model plant cell system, Chlamydomonas reinhardtii, *EMBO J 21*, 3681-3693.

65. Kim, H. Y., Fomenko, D. E., Yoon, Y. E., and Gladyshev, V. N. (2006) Catalytic advantages provided by selenocysteine in methionine-S-sulfoxide reductases, *Biochemistry 45*, 13697-13704.

66. Kim, H. Y., and Gladyshev, V. N. (2007) Methionine sulfoxide reductases: selenoprotein forms and roles in antioxidant protein repair in mammals, *Biochem J 407*, 321-329.

67. Boschi-Muller, S., Azza, S., Sanglier-Cianferani, S., Talfournier, F., Van Dorsselear, A., and Branlant, G. (2000) A sulfenic acid enzyme intermediate is involved in the catalytic mechanism of peptide methionine sulfoxide reductase from Escherichia coli, *J Biol Chem 275*, 35908-35913.

68. Boschi-Muller, S., Olry, A., Antoine, M., and Branlant, G. (2005) The enzymology and biochemistry of methionine sulfoxide reductases, *Biochim Biophys Acta 1703*, 231-238.

69. Olry, A., Boschi-Muller, S., and Branlant, G. (2004) Kinetic characterization of the catalytic mechanism of methionine sulfoxide reductase B from Neisseria meningitidis, *Biochemistry 43*, 11616-11622.

70. Kim, H. Y., and Kim, J. R. (2008) Thioredoxin as a reducing agent for mammalian methionine sulfoxide reductases B lacking resolving cysteine, *Biochem Biophys Res Commun 371*, 490-494.

71. Russel, M., and Model, P. (1986) The role of thioredoxin in filamentous phage assembly. Construction, isolation, and characterization of mutant thioredoxins, *J Biol Chem 261*, 14997-15005.

72. Lin, T. Y. (1999) G33D mutant thioredoxin primarily affects the kinetics of reaction with thioredoxin reductase. Probing the structure of the mutant protein, *Biochemistry 38*, 15508-15513.

73. Dobrovolska, O., Rychkov, G., Shumilina, E., Nerinovski, K., Schmidt, A., Shabalin, K., Yakimov, A., and Dikiy, A. (2012) Structural insights into interaction between mammalian methionine sulfoxide reductase B1 and thioredoxin, *J Biomed Biotechnol 2012*, 586539.

74. Lee, T. H., and Kim, H. Y. (2008) An anaerobic bacterial MsrB model reveals catalytic mechanisms, advantages, and disadvantages provided by selenocysteine and cysteine in reduction of methionine-R-sulfoxide, *Arch Biochem Biophys 478*, 175-180.

75. Bertini, I., Sigel, A., and Sigel, H. (2001) *Handbook on metalloproteins*, Marcel Dekker, New York.

76. Andreini, C., Bertini, I., Cavallaro, G., Holliday, G. L., and Thornton, J. M. (2008) Metal ions in biological catalysis: from enzyme databases to general principles, *J Biol Inorg Chem 13*, 1205-1218.

77. Gladyshev, V. N., Jeang, K. T., and Stadtman, T. C. (1996) Selenocysteine, identified as the penultimate C-terminal residue in human T-cell thioredoxin reductase, corresponds to TGA in the human placental gene, *Proc Natl Acad Sci U S A 93*, 6146-6151.

78. Mulkidjanian, A. Y., and Galperin, M. Y. (2009) On the origin of life in the zinc world. 2. Validation of the hypothesis on the photosynthesizing zinc sulfide edifices as cradles of life on Earth, *Biol Direct 4*, 27.

79. Auld, D. S. (2001) *Zinc Sites in Metalloenzymes and Related Proteins.*, Handbook on metalloproteins ed., Marcel Dekker, New York.

80. Auld, D. S. (2001) Zinc coordination sphere in biochemical zinc sites, *Biometals 14*, 271-313.

81. Vallee, B. L., and Auld, D. S. (1990) Zinc coordination, function, and structure of zinc enzymes and other proteins, *Biochemistry 29*, 5647-5659.

82. Korndorfer, I. P., Fessner, W. D., and Matthews, B. W. (2000) The structure of rhamnose isomerase from Escherichia coli and its relation with xylose isomerase illustrates a change between inter and intra-subunit complementation during evolution, *J Mol Biol 300*, 917-933.

83. Whitlow, M., Howard, A. J., Finzel, B. C., Poulos, T. L., Winborne, E., and Gilliland, G. L. (1991) A metal-mediated hydride shift mechanism for xylose isomerase based on the 1.6 A Streptomyces rubiginosus structures with xylitol and D-xylose, *Proteins 9*, 153-173.

84. Berry, M. B., and Phillips, G. N., Jr. (1998) Crystal structures of Bacillus stearothermophilus adenylate kinase with bound Ap5A, Mg2+ Ap5A, and Mn2+ Ap5A reveal an intermediate lid position and six coordinate octahedral geometry for bound Mg2+ and Mn2+, *Proteins 32*, 276-288.

85. Perrier, V., Surewicz, W. K., Glaser, P., Martineau, L., Craescu, C. T., Fabian, H., Mantsch, H. H., Barzu, O., and Gilles, A. M. (1994) Zinc chelation and structural stability of adenylate kinase from Bacillus subtilis, *Biochemistry 33*, 9960-9967.

86. Carfi, A., Pares, S., Duee, E., Galleni, M., Duez, C., Frere, J. M., and Dideberg, O. (1995) The 3-D structure of a zinc metallo-beta-lactamase from Bacillus cereus reveals a new type of protein fold, *EMBO J 14*, 4914-4921.

87.   Fabiane, S. M., Sohi, M. K., Wan, T., Payne, D. J., Bateson, J. H., Mitchell, T., and Sutton, B. J. (1998) Crystal structure of the zinc-dependent beta-lactamase from Bacillus cereus at 1.9 A resolution: binuclear active site with features of a mononuclear enzyme, *Biochemistry 37*, 12404-12411.

88.   Zhang, Y., and Gladyshev, V. N. (2011) Comparative genomics of trace element dependence in biology, *J Biol Chem 286*, 23623-23629.

89.   Makarova, K. S., Ponomarev, V. A., and Koonin, E. V. (2001) Two C or not two C: recurrent disruption of Zn-ribbons, gene duplication, lineage-specific gene loss, and horizontal gene transfer in evolution of bacterial ribosomal proteins, *Genome Biology 2*.

90.   Panina, E. M., Mironov, A. A., and Gelfand, M. S. (2003) Comparative genomics of bacterial zinc regulons: enhanced ion transport, pathogenesis, and rearrangement of ribosomal proteins, *Proc Natl Acad Sci U S A 100*, 9912-9917.

91.   Olry, A., Boschi-Muller, S., Yu, H., Burnel, D., and Branlant, G. (2005) Insights into the role of the metal binding site in methionine-R-sulfoxide reductases B, *Protein Sci 14*, 2828-2837.

92.   Kim, Y. K., Shin, Y. J., Lee, W. H., Kim, H. Y., and Hwang, K. Y. (2009) Structural and kinetic analysis of an MsrA-MsrB fusion protein from Streptococcus pneumoniae, *Mol Microbiol 72*, 699-709.

93.   Ranaivoson, F. M., Neiers, F., Kauffmann, B., Boschi-Muller, S., Branlant, G., and Favier, F. (2009) Methionine sulfoxide reductase B displays a high level of flexibility, *J Mol Biol 394*, 83-93.

94.   Dobrovolska, O., Shumilina, E., Holen, H. W., and Dikiy, A. (to be published) Structural and biochemical insights into mammalian cobalt-substituted methionine sulfoxide reductase B1.

95.   Lippard, S. J., and Berg, J. M. (1994) *Principles of bioinorganic chemistry*, University Science Books, Mill Valley, Calif.

96.   Bertini, I., Sigel, A., and Sigel, H. (2001) *Handbook on Metalloproteins*, Marcel Dekker.

97.   Gerdts, C. J., Elliott, M., Lovell, S., Mixon, M. B., Napuli, A. J., Staker, B. L., Nollert, P., and Stewart, L. (2008) The plug-based nanovolume Microcapillary Protein Crystallization System (MPCS), *Acta Crystallographica Section D 64*, 1116-1122.

98.   Lange, O. F., Rossi, P., Sgourakis, N. G., Song, Y., Lee, H. W., Aramini, J. M., Ertekin, A., Xiao, R., Acton, T. B., Montelione, G. T., and Baker, D. (2012) Determination of solution structures of proteins up to 40 kDa using CS-Rosetta with sparse NMR data from deuterated samples, *Proc Natl Acad Sci U S A 109*, 10873-10878.

Paper III

PONE-D-12-31959R1
Structural analysis of glutaredoxin domain of Mus musculus thioredoxin glutathione reductase
PLOS ONE

Dear Prof. Dikiy,

I am pleased to inform you that your manuscript has been deemed suitable for publication in PLOS ONE.

Your manuscript will now be passed on to our Production staff, who will check your files for correct formatting and completeness. After this review, they may return your manuscript to you so that you can make necessary alterations and upload a final version.

Before uploading, you should check the PDF of your manuscript very closely. There is no author proofing. You should therefore consider the corrected files you upload now as equivalent to a production proof. The text you supply at this point will be faithfully represented in your published manuscript exactly as you supply it. This is your last opportunity to correct any errors that are present in your manuscript files.

With kind regards,
Michael Massiah
Academic Editor
PLOS ONE

Structural analysis of glutaredoxin domain of *Mus musculus* thioredoxin glutathione reductase

Olena Dobrovolska[1], Elena Shumilina[1], Vadim N. Gladyshev[2] and Alexander Dikiy[1,*]

[1]Department of Biotechnology, Norwegian University of Science and Technology, N-7491 Trondheim, Norway

[2]Genetics Division, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA 02115, USA

Keywords:

Thioredoxin reductase, Glutaredoxin, Protein structure, Bioinformatics, Alignment, Protein function

[*]Corresponding author: Tel.:+4773597863, Fax: +4773591283; E-mail: alex.dikiy@biotech.ntnu.no

Abstract

Thioredoxin glutathione reductase (TGR) is a member of the mammalian thioredoxin reductase family that has a monothiol glutaredoxin (Grx) domain attached to the thioredoxin reductase module. Here, we report a structure of the Grx domain of mouse TGR, determined through high resolution NMR spectroscopy to the final backbone RMSD value of 0.48±0.10 Å. The structure represents a sandwich-like molecule composed of a four stranded β-sheet flanked by five α–helixes, with the CxxS active motif located on the catalytic loop. We structurally characterized the glutathione-binding site in the protein and describe sequence and structural relationships of the domain with glutaredoxins. The structure illuminates a key functional center that evolved in mammalian TGRs to act in thiol-disulfide reactions. Our study allows us to hypothesize that Cys105 might be functionally relevant for TGR catalysis. In addition, the data suggest that the N-terminus of Grx acts as a possible regulatory signal also protecting the protein active site from unwanted interactions in cellular cytosol.

Introduction

Thioredoxin (Trx) and glutaredoxin (Grx) systems are two major thiol pathways that control cellular redox homeostasis [1]. The Trx system is composed of thioredoxin reductase (TR), thioredoxin (Trx) and Trx peroxidase, whereas the Grx system consists of glutathione reductase (GR), glutathione (γ-Glu-Cys-Gly tripeptide; GSH), glutaredoxin (Grx) and glutathione peroxidase (GPx). In these systems, electron flow is directed from NADPH through GR and TR towards their respective protein substrates. GR and TR belong to the pyridine nucleotide-disulfide oxidoreductase family. They are homodimers and contain a tightly bound FAD molecule in each subunit [1]. Mammalian TR and GR were found to be structurally and functionally similar, although TR has an additional C-terminal selenocysteine-containing active site, which serves as a substrate for the N-terminal active site [2,3,4,5,6,7].

Three TRs genes have been identified in humans, including TXNRD1 (cytosolic TR, TR1), TXNRD2 (mitochondrial TR, TR3) and TXNRD3 (thioredoxin glutathione reductase, TGR) [8,9,10,11,12,13]. TGR is unusual among TRs in that it has an additional N-terminal Grx domain, which is fused to a canonical TR module [13,14,15,16]. The amino acid sequence of the TR module of TGR is more closely related to TR1 than to TR3 [13] and its Grx domain has a monothiol CPHS catalytic motif [13,14,15,17]. The active site motif of the Grx domain of TGR can receive electrons from either the TR module or from GSH, and the protein was proposed to function predominantly in disulfide bond formation and isomerization in sperm proteins during spermatogenesis [14]. Mammalian TGR exhibits broad substrate specificity and can reduce various components of both Trx and Grx systems [16]. In particular, it was demonstrated that TGR can catalyze reactions associated with Grx (deglutathionylation), GR (NADPH-dependent reduction of GSSG) and TR (NADPH-dependent reduction of Trx) activities. It was argued that Grx and GR activities of TGR are mediated by its Grx domain [16].

Structural characterization of proteins is an essential step for establishment of their functional peculiarities. Structures of platyhelminth TGR (pdb code 2V6O) and the Grx domain of human TGR (pdb code 3H8Q) have been recently determined. We previously reported NMR resonance assignments of full-length and shortened (lacking 22 N-terminal amino acids) forms of the Grx domain of *Mus musculus* TGR [18]. In the present work, we report solution structure of this Grx domain using high-resolution NMR spectroscopy. This

Trx-fold structure validates the model of the Grx domain [19] and is consistent with the structures of other Grx. We further used the structure to carry out comparative sequence, structure and charge distribution analyses of Grx and Grx domains in order to explain structural and functional peculiarities of the TGR's Grx domain.

Materials and methods

Reduced glutathione (GSH) was purchased from Acros Organics, and oxidized glutathione (GSSG) from Sigma Aldrich. Both compounds were dissolved in a buffer containing 10 mM sodium phosphate, 10 mM NaCl, pH 7.5.

Sample preparation

Protein expression and purification of a uniformly isotope labeled (^{15}N/^{13}C) His-tagged version of the full-length and shortened forms of the Grx domain of mouse TGR (hereafter Grx and sGrx, respectively) was carried out as described previously [18]. NMR samples of reduced 1 mM Grx or sGrx in 10 mM sodium phosphate buffer, 10 mM NaCl, 10 mM β-mercaptoethanol, in 95% H_2O/5% D_2O and 100% D_2O, pH 7.5, were analyzed by NMR.

NMR spectroscopy

NMR spectra were recorded at 298 K on a Bruker Avance 600 MHz spectrometer equipped with a 5-mm z-gradient TXI (H/C/N) cryoprobe. Three-dimensional ^{13}C- and ^{15}N-edited ^{1}H Nuclear Overhauser Effect Spectroscopy (NOESY) spectra were recorded in D_2O and H_2O, respectively. NMR data were processed using Bruker XWinNMR, version 3.5. NMR spectral analysis was performed using CARA version 1.8.4.2.

Structure calculation

NOE cross-peaks were identified, assigned and integrated in the aforementioned NOESY spectra using the CARA program. The CALIBA subroutine in CYANA 2.1 was used to convert cross peak intensities to distance constraints. Dihedral angle constraints were derived from secondary chemical shifts using the TALOS program [20]. Based on the input, the structure was calculated using the torsion angle dynamics program CYANA2.1 [21]. Twenty conformers with the lowest final CYANA target function values were further energy

minimized in vacuum using AMBER force field with the aid of AMBER 9 program [22]. The mean structure was generated using MOLMOL 2k.2.0 [23] and further energy minimized in AMBER.

Structure analysis

Quality of structures was analyzed using MOLMOL and PROCHECK - NMR [24]. The relevant figures and electrostatic potentials were prepared using MOLMOL version 2k.2.0.

NMR experiments

In order to characterize the glutathione-binding site of Grx NMR ^{15}N-^{1}H HSQC titration experiments were performed. A 1 mM sample of ^{15}N-labeled Grx in buffer, containing 10 mM sodium phosphate, 10 mM NaCl, pH 7.5, was titrated with unlabeled GSH and GSSG at room temperature in the following proportions: 1:1/3; 1:1/2; 1:2/3; 1:1; 1:2; 1:10, either in the absence or presence of 10 mM β-mercaptoethanol. In the membrane environmental modeling experiment, 5% w/v SDS was added to the solution of ^{15}N-labeled Grx and the ^{15}N-^{1}H HSQC spectra of the mixture were recorded at 30°C and 42°C. For water exchange experiments, samples of ^{15}N-labeled Grx or sGrx in the NMR buffer were lyophilized and further dissolved in D_2O. A course of subsequent ^{15}N-^{1}H HSQC spectra for each protein was recorded every 30 min.

Bioinformatics analysis

Protein multiple sequence alignments were performed with ClustalW [25]. Sequence similarity analysis was performed by the SIAS server (http://imed.med.ucm.es/Tools/sias.html). Structural superimposition was carried out using SuperPose [26]. An analysis of the N-terminal region (residues 1-22) was performed using iPSORT [27] and MITOPROT [28].

Results and Discussion

Structure description. Solution structure of the Grx domain of mouse TGR was calculated based on the NOE-derived geometrical constraints and dihedral angles obtained

from TALOS. The geometrical constraints used in the calculations are summarized in Table 1. In total, 894 NOE-based upper distance limits and 182 ψ and φ torsion angle restraints were used to derive the Grx structure. The resulting Grx family was further energy-minimized. The geometrical constraints and coordinate files of the minimized Grx family were deposited in the PDB under the code 2lv3. Figure 1 shows a superimposition of the final 20 minimized conformers with the lowest target function, together with a ribbon representation of the minimized conformer closest to the mean structure showing the secondary structure elements, active site cysteine C48, and C-terminal C105 (further discussed). The calculated structure is of high quality and fully corresponds to the experimentally determined constraints.

 The N-terminal region of Grx (first 22 amino acids) was excluded from the structure calculation as most of the corresponding HSQC and NOE signals were not detected [18]; hence, the structure of Grx starts with Ala 23. Analysis of the structure shows that the Grx domain is a compact Trx-like spherical molecule with a central core of four-stranded β-sheets flanked on either side by five α-helices arranged in the order α_1-β_1-α_2-β_2-α_3-β_3-β_4-α_4-α_5 (Figure 1). The N-terminal region begins with an α_1 (residues Arg 24 - Glu 36), followed by β_1 consisting of residues Val 40 to Ser 44. The active site Cys 48 - Ser 51 (-CPHS- motif) is situated on the unstructured loop between β_1 and α_2 (residues Arg 53 – Ser 59). The strand β_2 comprises residues Asn 66 to Glu 69; following a loop, α_3 consists of residues Gly 76 to Ser 87, followed by β_3 (Asn 94 – Val 97) and β_4 (Val 100 – Gly 103). The C-terminal region includes α_4 (residues Arg 107 – Asn 114) and α_5 (residues Leu 116 – Leu 120), connected through a hinge section. Strands β_1 and β_2 are parallel, and strand β_3 is antiparallel with β_1 and β_4. Helices α_1 and α_3 pack on one side of the β-sheet, whereas α_2, α_4 and α_5 are on the other. Packing of the sandwich-like architecture is mainly maintained by hydrophobic interactions between the sheet and helices. The determined Grx structure shares significant structural similarity with the modelled Grx domain of mouse TGR [19].

N-terminal region. As mentioned above, the N-terminal region of Grx was not detectable in HSQC and NOE spectra. The absence of the corresponding cross-peaks could be attributed to the higher mobility of this protein region. A decrease in temperature may slow down protein mobility and thus allow detecting the missing cross-peaks. Nevertheless, the ^{15}N-^{1}H HSQC spectra of Grx recorded at a lower temperature (8°C) did not show additional signals in the spectra (data not shown). Further analysis of the first 22 amino acids of Grx using iPSORT suggested the presence of a candidate mitochondrial targeting peptide. Indeed, the N-terminal sequence is rich in positively charged and hydrophobic residues that may

constitute the targeting helix [29]. Analysis of the N-terminus performed by MITOPROT predicted the cleavage site after the first 19 residues (MSSPPGRRARLASPGTSRP). However, analysis of cellular distribution of TGR suggested that the enzyme occurs in the cytosol of spermatids at the time of mitochondrial sheath formation [14]. In these spermatids, TGR is accumulated near the site of mitochondrial sheath assembly. It was shown that mammalian sperm is stabilized by disulfide bond (S-S) bridges cross-linking thiol-rich proteins present in the membranes of sperm mitochondria [30]. Thus, TGR, which combines the elements of Grx and Trx systems, might be involved in disulfide bond formation during spermatogenesis. The N-terminus can also act as a regulatory sequence. Showing little structural organization in solution, mitochondrial targeting sequences are predicted to form amphipathic α-helixes in the membrane or membrane-like environment. The amphipathic nature of these structures is thought to be important for their specific recognition by the protein import machinery [31]. A membrane-like media are prepared by self-association of surfactants in aqueous solutions, which are divided into two large groups: detergents (form micelles) and lipids (form bilayers) [32]. For preparation of micelles, which are widely used in NMR structural studies, negatively charged SDS detergent is often used [33]. In NMR spectroscopy, the formation of the protein's secondary structure results in the appearance of a set of well-dispersed HSQC cross-peaks. To further examine the N-terminal part of Grx, we recorded ^{15}N-^{1}H HSQC spectra of the SDS-treated Grx domain at 30°C and 42°C [34]. The obtained ^{15}N-^{1}H HSQC pattern was shifted towards the lower field of both dimensions; however, it did not change dramatically, and, as expected, contained additional well-dispersed cross-peaks, as shown in Figure 2. The number of new cross-peaks corresponded to (but did not exceed that of) the number of amino acids constituting the N-terminal segment, therefore, confirming that the positively charged N-terminus becomes structured in the negatively charged environment.

Although the ^{1}H-^{15}N HSQC patterns of Grx and sGrx resemble each other, indicating overall structural correspondence of the two protein forms, several differences between them were observed. First, the full-length protein had a higher stability than the shortened form. Second, the signals in ^{15}N-^{1}H HSQC, corresponding to residues C105 and D106, were not found in Grx, while they were present and assigned in sGrx [18]. Weak HSQC signals and any NOESY patterns were also found for residues D74 and A77 in the corresponding spectra of Grx [18]. These findings correlating with the presence/absence of the N-terminus further highlight structural differences between the two protein forms (see the following paragraphs).

Electrostatic potential. Figure 3 shows electrostatic potential calculated for the obtained structure of Grx. It is apparent from the analysis of this figure, that the missing NOE signals of D74, A77, C105, D106 (see above) amino acids belong to the negatively charged region closest to the N-terminus. As the N-terminal region of Grx is composed of positively charged (R7, R8, R10, R18) and polar (S2, S3, S13, T16, S17, S20, S21) amino acids, their involvement in electrostatic interaction with negatively charged protein surface can be suggested. The broadening of the NMR signals corresponding to D74, A77, C105, D106 due to this interaction may be a reason that the mentioned amino acid signals in the NMR spectra of the full length form of Grx where not detected. Indeed, the resonances belonging to D74, A77, C105, D106, not observed in Grx protein, were firmly detected in sGrx lacking the mentioned electrostatic interaction. Interestingly, the active site motif of Grx (C48, P49, H50, S51), as monitored by our structural studies, resides on the neighbouring loop near residues C105 and D106. Thus, the N-terminal region, positioned in proximity to the active site, could shield it from the solvent and, therefore, protect from unwanted reactions stabilizing the full-length protein.

D_2O exchange experiments. To further comparatively characterize both Grx and sGrx proteins, we carried out water exchange experiments monitoring behaviour of ^{15}N-^{1}H HSQC spectra in D_2O for both full-length and shortened forms of the ^{15}N-labeled Grx domain of TGR. Interestingly, during the first 30 minutes in D_2O, Grx exchanged ten residues more with respect to sGrx (D33, G37, N38, V40, S59, V63, D71, Q72, E85, T108). However, after 3 hours of incubation in D_2O, both Grx and sGrx reveal an identical pattern of exchanged/not exchanged residues. Therefore, our experiments show that while the final rate of water exchange is the same both for Grx and sGrx, the short term dynamics of the water exchange is different for these proteins. The observed differences mostly regard residues belonging to the negative patch involved in the suggested interaction with the positively charged N-terminus (see above). The fact that these residues exchanged within the first 30 min in Grx, while in sGrx they exchanged only 3 hours later indicates that the N-terminus in some way promotes faster rates of water diffusion into Grx protein.

Comparative sequence analysis of Grx. Grx occur in the majority of organisms in the three domains of life. Structures of many of these proteins were determined either by NMR or X-ray crystallography. Two main groups of Grx can be distinguished based on phylogeny, active site motifs, and domain structure: (i) 'classical' dithiol Grx containing the active site consensus sequence Cys-X-X-Cys (i.e., two Cys separated by two other amino acids); and (ii)

monothiol Grx with a Cys-X-X-Ser active site consensus sequence. The latter Grx utilize only the N-terminal active site Cys in their catalytic mechanism, which is used together with two glutathione molecules, while dithiol Grx can use either one or both Cys in the active site. Both types of disulfides formed during Grx catalysis are reduced *in vitro* by GSH or TRs [35].

We analyzed an alignment of the Grx domain of mouse TGR with both Grx and Grx domains of TGR from various organisms, which contain mono- or dithiol active sites (Figure 4). The active site residues (highlighted with a red rectangle), and residues involved in the interaction with GSH (marked with black rectangles) are conserved in mono- and dithiol Grx, including the Grx of mouse TGR [36].

Since interaction of the Grx domain with glutathione is assisted by electrostatic interactions, we next analysed the distribution of charged amino acids. As shown in the figure 4, the active site of Grx is surrounded by positively charged amino acids (marked in blue). Interestingly, only in *E. coli* Grx3 and the Grx domain of *Xenopus laevis* TGR, negative residues are found in the vicinity of the active site (Fig.5, marked in red). It was suggested that these negative residues influence the redox potential of these proteins [37]. In addition, the C-terminal segment of the Grx domain of TGR from human, mouse, *Danio rerio*, *Xenopus laevis* and Grx3 from *E. coli* harbor an additional Cys residue. Our structural data indicate that the distance between the active site cysteine (Cys 48) and C-terminal Cys105 (Figure 1, right panel) is 10-15 Å. Although the distance might be too large for the formation of an intra-molecular disulfide bridge, according to our data there is no steric hindrance between these two cysteines located on unstructured loops. It can be hypothesised that the formation of an intra-molecular disulfide bridge between the two cysteine residues (e.g., as observed in mammalian MsrB1 protein [38,39]) may have functional and, perhaps, even catalytic relevance.

Structural comparison with other Grx. Our analysis, as well as other structural studies of a set of GSH-dependent Grx [26] revealed common elements present in their binding sites: i) a CXXC/S active site motif; ii) a Tyr or a Phe in close proximity to the catalytic Cys; iii) a TVP motif with Pro in the *cis* conformation; iv) a GG kink in proximity to the active site; and v) conservation of charged residues at both edges of the substrate binding groove (GSH binding pocket). The structure of the Grx domain of mouse TGR was generally similar to other Grx structures, but varied in secondary structure elements attached to a common Grx core. For example, an N-terminal α-helix (α$_1$ in our structure) was present in many Grx

including human TGR (PDB code 3H8Q), human Grx2 (PDB code 2FLS), Grx of *Sm*TGR (PDB code 2V6O), monothiol *E. coli* Grx4 (PDB code 1YKA), and poplar Grx C1 (PDB code 2E7P), but not in the dithiol *E. coli* Grx3 (PDB code 1FOV). A lack of C- and/or N-terminal α-helices together with the length of the loop preceding the active site motif, was suggested to play a key role in constraining the degree of conformational adaptability for substrate binding displayed by Grx [40]. Minor structural differences were also observed between the Grx domain of mouse TGR and the Grx domain of *Sm*TGR: the latter was characterized by a shorter C-terminal α-helix [41].

Recent studies [42] demonstrated that *Sm*TGR may function *via* two catalytic mechanisms: monothiol and dithiol. In the monothiol mechanism, when the GSH concentration is high, glutathionylated catalytic Cys (Cys28) of Grx gets resolved by GSH. At low GSH concentrations, a second, dithiol mechanism applies, wherein the C-terminal Cys (Cys31) in the CXXC motif acts as a resolving group, breaking the disulfide bond between Cys28 and GSH, forming an internal Cys28-Cys31 disulfide and releasing GSH. An oxidized Grx can be further reduced by the redox-active Cys/Sec–Cys pair of the TR domain [43]. The same study analyzed the deglutathionylation activities of a *Sm*TGR variant, in which Cys31 was replaced with Ser (making it analogous to mouse TGR), which exhibited 22% of wild type *Sm*TGR activity. This study suggested that the role of the second Cys in the monothiol mechanism is to stabilize the thiolate anion of the N-terminal Cys through a hydrogen bond, thus facilitating its nucleophilic attack on GSSG.

Grx titration with GSH/GSSG. Figure 5 illustrates fragments of the ^{15}N-^{1}H HSQC titration of the Grx domain with reduced (Figure 5A) and oxidized (Figure 5B) glutathione. Upon interaction with these two molecules the chemical environment of the nuclei involved in the interaction changes, which results in perturbation of the corresponding NMR signal: i.e., chemical shift change or signal broadening occurs. By performing NMR titration of ^{15}N-labeled Grx with its unlabelled partners, GSH and GSSG, we monitored their interaction and mapped the Grx residues involved in binding with the respective partners.

Upon titration of Grx with either GSH or GSSG, nearly 90% of the signals remained unaltered; however, some of the signals appear changed (mostly broadened). The titration experiments of the Grx domain titration with GSH revealed signal broadening that corresponded to the following amino acids: S44, K45, C48, P49, H50, S51, T52, R53, V54, E81, T91, V92, P93, N94, G103, G104, R107. Grx titration with GSSG showed changes for

the following residues: N38, K45, C48, P49, H50, S51, T52, R53, V54, E81, T91, V92, P93, N94, G103, G104, and R107 (Figure 5C). Interestingly, amino acids observed in the two experiments almost coincided. These experimental data not only point to the glutathione binding site in the Grx domain of TGR, but also suggest that the binding sites for reduced and oxidized glutathione largely overlap.

Conclusions

This study describes the NMR solution structure of the monothiol Grx domain of mouse TGR. As expected, the protein possesses a Trx fold and consists of a four-stranded β-sheet flanked by five α-helices. The active site motif containing the catalytic redox-active Cys is located on the protruding loop connecting strand β1 and α2.

Analysis of the N-terminal segment of Grx, which was not included in the structure determination, showed that it has features of a targeting sequence or a regulatory region. It was found, by analyzing ^{15}N-^{1}H HSQC spectra, that this segment becomes structured when protein is treated with a detergent, thereby mimicking membrane-like environment. Based on the analysis of surface charge distribution of the protein, we suggest that the N-terminus resides near the active site, shielding it from redox interactions.

Sequence alignment of the domains with other Grx and Grx domains revealed a characteristic GSH-binding site, which was further characterized with the help of NMR. The data suggest a significant overlap between the GSH and GSSG binding sites. Further analysis of mammalian TGR function would require structural information of the entire enzyme.

Acknowledgements

References

1. Holmgren A (1989) Thioredoxin and glutaredoxin systems. J Biol Chem 264: 13963-13966.

2. Luthman M, Holmgren A (1982) Rat liver thioredoxin and thioredoxin reductase: purification and characterization. Biochemistry 21: 6628-6633.

3. Arscott LD, Gromer S, Schirmer RH, Becker K, Williams CH, Jr. (1997) The mechanism of thioredoxin reductase from human placenta is similar to the mechanisms of lipoamide dehydrogenase and glutathione reductase and is distinct from the mechanism of thioredoxin reductase from Escherichia coli. Proc Natl Acad Sci U S A 94: 3621-3626.

4. Tamura T, Stadtman TC (1996) A new selenoprotein from human lung adenocarcinoma cells: purification, properties, and thioredoxin reductase activity. Proc Natl Acad Sci U S A 93: 1006-1011.

5. Gladyshev VN, Jeang KT, Stadtman TC (1996) Selenocysteine, identified as the penultimate C-terminal residue in human T-cell thioredoxin reductase, corresponds to TGA in the human placental gene. Proc Natl Acad Sci U S A 93: 6146-6151.

6. Gasdaska PY, Gasdaska JR, Cochran S, Powis G (1995) Cloning and sequencing of a human thioredoxin reductase. FEBS Lett 373: 5-9.

7. Williams CH, Arscott LD, Muller S, Lennon BW, Ludwig ML, et al. (2000) Thioredoxin reductase two modes of catalysis have evolved. Eur J Biochem 267: 6110-6117.

8. Gasdaska PY, Berggren MM, Berry MJ, Powis G (1999) Cloning, sequencing and functional expression of a novel human thioredoxin reductase. FEBS Lett 442: 105-111.

9. Lee SR, Kim JR, Kwon KS, Yoon HW, Levine RL, et al. (1999) Molecular cloning and characterization of a mitochondrial selenocysteine-containing thioredoxin reductase from rat liver. J Biol Chem 274: 4722-4734.

10. Miranda-Vizuete A, Damdimopoulos AE, Pedrajas JR, Gustafsson JA, Spyrou G (1999) Human mitochondrial thioredoxin reductase cDNA cloning, expression and genomic organization. Eur J Biochem 261: 405-412.

11. Watabe S, Makino Y, Ogawa K, Hiroi T, Yamamoto Y, et al. (1999) Mitochondrial thioredoxin reductase in bovine adrenal cortex its purification, properties, nucleotide/amino acid sequences, and identification of selenocysteine. Eur J Biochem 264: 74-84.

12. Miranda-Vizuete A, Damdimopoulos AE, Spyrou G (1999) cDNA cloning, expression and chromosomal localization of the mouse mitochondrial thioredoxin reductase gene(1). Biochim Biophys Acta 1447: 113-118.

13. Sun QA, Wu Y, Zappacosta F, Jeang KT, Lee BJ, et al. (1999) Redox regulation of cell signaling by selenocysteine in mammalian thioredoxin reductases. J Biol Chem 274: 24522-24530.

14. Su D, Novoselov SV, Sun QA, Moustafa ME, Zhou Y, et al. (2005) Mammalian selenoprotein thioredoxin-glutathione reductase. Roles in disulfide bond formation and sperm maturation. J Biol Chem 280: 26491-26498.

15. Sun QA, Kirnarsky L, Sherman S, Gladyshev VN (2001) Selenoprotein oxidoreductase with specificity for thioredoxin and glutathione systems. Proc Natl Acad Sci U S A 98: 3673-3678.

16. Sun QA, Su D, Novoselov SV, Carlson BA, Hatfield DL, et al. (2005) Reaction mechanism and regulation of mammalian thioredoxin/glutathione reductase. Biochemistry 44: 14528-14537.

17. Sun QA, Zappacosta F, Factor VM, Wirth PJ, Hatfield DL, et al. (2001) Heterogeneity within animal thioredoxin reductases - Evidence for alternative first exon splicing. Journal of Biological Chemistry 276: 3106-3114.

18. Shumilina E, Solda A, Gerashchenko M, Gladyshev VN, Dikiy A (2012) (1)H, (1)(3)C, and (1)(5)N NMR resonance assignments of reduced full length and shortened forms of the Grx domain of Mus musculus TGR. Biomol NMR Assign 6: 103-107.

19. Ge Y, Qi Z, Wang Y, Liu X, Li J, et al. (2009) Engineered selenium-containing glutaredoxin displays strong glutathione peroxidase activity rivaling natural enzyme. Int J Biochem Cell Biol 41: 900-906.
20. Cornilescu G, Delaglio F, Bax A (1999) Protein backbone angle restraints from searching a database for chemical shift and sequence homology. J Biomol NMR 13: 289-302.
21. Guntert P, Braun W, Wuthrich K (1991) Efficient computation of three-dimensional protein structures in solution from nuclear magnetic resonance data using the program DIANA and the supporting programs CALIBA, HABAS and GLOMSA. J Mol Biol 217: 517-530.
22. Case DA, Cheatham TE, 3rd, Darden T, Gohlke H, Luo R, et al. (2005) The Amber biomolecular simulation programs. J Comput Chem 26: 1668-1688.
23. Koradi R, Billeter M, Wuthrich K (1996) MOLMOL: a program for display and analysis of macromolecular structures. J Mol Graph 14: 51-55, 29-32.
24. Laskowski RA, Macarthur MW, Moss DS, Thornton JM (1993) Procheck - a Program to Check the Stereochemical Quality of Protein Structures. Journal of Applied Crystallography 26: 283-291.
25. Thompson JD, Higgins DG, Gibson TJ (1994) Clustal-W - Improving the Sensitivity of Progressive Multiple Sequence Alignment through Sequence Weighting, Position-Specific Gap Penalties and Weight Matrix Choice. Nucleic Acids Research 22: 4673-4680.
26. Maiti R, Van Domselaar GH, Zhang H, Wishart DS (2004) SuperPose: a simple server for sophisticated structural superposition. Nucleic Acids Research 32: W590-594.
27. Bannai H, Tamada Y, Maruyama O, Nakai K, Miyano S (2002) Extensive feature detection of N-terminal protein sorting signals. Bioinformatics 18: 298-305.
28. Claros MG, Vincens P (1996) Computational method to predict mitochondrially imported proteins and their targeting sequences. Eur J Biochem 241: 779-786.
29. Neupert W (1997) Protein import into mitochondria. Annu Rev Biochem 66: 863-917.
30. Sutovsky P, Tengowski MW, Navara CS, Zoran SS, Schatten G (1997) Mitochondrial sheath movement and detachment in mammalian, but not nonmammalian, sperm induced by disulfide bond reduction. Mol Reprod Dev 47: 79-86.
31. Vonheijne G (1986) Mitochondrial Targeting Sequences May Form Amphiphilic Helices. Embo Journal 5: 1335-1342.
32. Warschawski DE, Arnold AA, Beaugrand M, Gravel A, Chartrand E, et al. (2011) Choosing membrane mimetics for NMR structural studies of transmembrane proteins. Biochim Biophys Acta 1808: 1957-1974.
33. Goncharuk SA, Goncharuk MV, Mayzel ML, Lesovoy DM, Chupin VV, et al. (2011) Bacterial Synthesis and Purification of Normal and Mutant Forms of Human FGFR3 Transmembrane Segment. Acta Naturae 3: 77-84.
34. Krueger-Koplin RD, Sorgen PL, Krueger-Koplin ST, Rivera-Torres IO, Cahill SM, et al. (2004) An evaluation of detergents for NMR structural studies of membrane proteins. J Biomol NMR 28: 43-57.
35. Gallogly MM, Starke DW, Mieyal JJ (2009) Mechanistic and kinetic details of catalysis of thiol-disulfide exchange by glutaredoxins and potential mechanisms of regulation. Antioxid Redox Signal 11: 1059-1081.
36. Couturier J, Koh CS, Zaffagnini M, Winger AM, Gualberto JM, et al. (2009) Structure-Function Relationship of the Chloroplastic Glutaredoxin S12 with an Atypical WCSYS Active Site. Journal of Biological Chemistry 284: 9299-9310.
37. Aslund F, Nordstrand K, Berndt KD, Nikkola M, Bergman T, et al. (1996) Glutaredoxin-3 from Escherichia coli. Amino acid sequence, 1H AND 15N NMR assignments, and structural analysis. J Biol Chem 271: 6736-6745.
38. Kim HY, Gladyshev VN (2005) Different catalytic mechanisms in mammalian selenocysteine- and cysteine-containing methionine-R-sulfoxide reductases. PLoS Biol 3: e375.
39. Aachmann FL, Sal LS, Kim HY, Marino SM, Gladyshev VN, et al. (2010) Insights into function, catalytic mechanism, and fold evolution of selenoprotein methionine sulfoxide reductase B1 through structural analysis. J Biol Chem 285: 33315-33323.

40. Berardi MJ, Bushweller JH (1999) Binding specificity and mechanistic insight into glutaredoxin-catalyzed protein disulfide reduction. J Mol Biol 292: 151-161.
41. Angelucci F, Miele AE, Boumis G, Dimastrogiovanni D, Brunori M, et al. (2008) Glutathione reductase and thioredoxin reductase at the crossroad: the structure of Schistosoma mansoni thioredoxin glutathione reductase. Proteins-Structure Function and Bioinformatics 72: 936-945.
42. Huang HH, Day L, Cass CL, Ballou DP, Williams CH, Jr., et al. (2011) Investigations of the catalytic mechanism of thioredoxin glutathione reductase from Schistosoma mansoni. Biochemistry 50: 5870-5882.
43. Gallogly MM, Starke DW, Leonberg AK, Ospina SM, Mieyal JJ (2008) Kinetic and mechanistic characterization and versatile catalytic properties of mammalian glutaredoxin 2: implications for intracellular roles. Biochemistry 47: 11144-11157.

Figure Legends

Figure 1. Solution structure of the reduced Grx domain of mouse TGR. Left: overview of backbone superimposition of 20 conformers with the lowest target function. Right: overview of the ribbon representation of the minimized conformer closest to the mean structure. The figure also shows the active site cysteine (C48) and C-terminal C105.

Figure 2. A fragment of ^{15}N-^{1}H HSQC spectra of the reduced Grx domain of mouse TGR. Green shows Grx HSQC spectrum at 30°C, light blue shows Grx HSQC spectrum in the presence of SDS at 30°C, and black corresponds to Grx HSQC spectrum in the presence of SDS at 42°C. For more details see the text.

**Figure 3. Surface charge distribution of the reduced Grx domain of mouse TGR.**

Figure 4. Amino acid sequence alignment of the Grx domain of mouse TGR with mono- and dithiol Grx from various organisms. The active site residues of Grx are highlighted with a red rectangle, and the residues involved in the interaction with glutathione are within black rectangles. Positive and negative charged amino acids are marked with blue and red colour, respectively. Cysteines residues are marked in green. Sequences abbreviation: HS, *Homo sapiens*; SM, *Schistosoma mansoni*; XL, *Xenopus laevis*; MM, *Mus musculus*; DR, *Danio rerio*; EC, *Escherichia coli*.

Figure 5. Fragments of ^{15}N-^{1}H HSQC spectra of ^{15}N-labeled reduced Grx domain of mouse TGR titrated with unlabelled GSH and GSSG (panels A and B, respectively). Green corresponds to free Grx and magenta to Grx incubated with GSH/GSSG. Only the residues for which alteration of NMR parameters upon titration was observed are marked. Panel C: qualitative representation of the data. Solid and dashed horizontal lines below the Grx amino acid sequence highlight the residues interacting with GSH and GSSG, respectively.The N-term of Grx is marked in blue. For more details see the text.

Table 1. Structural statistics and geometrical constraints derived from NMR for the reduced form of the Grx domain of mouse TGR.

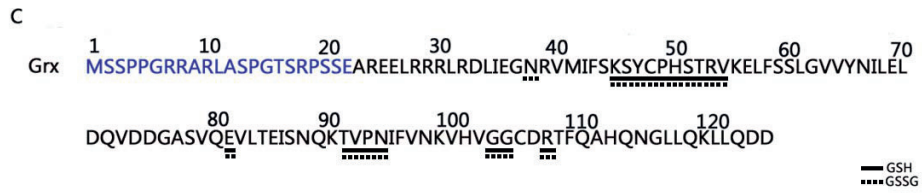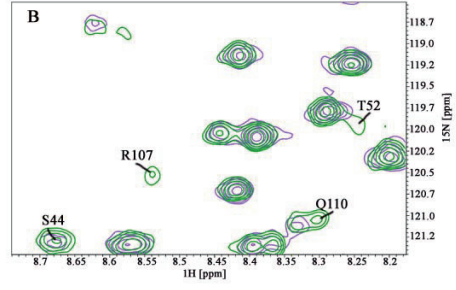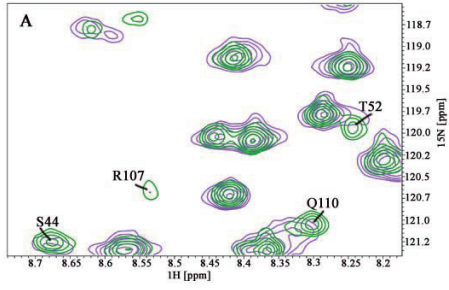| Restraints used in structure calculation	Number
Total number of NOE distance restraints	894
Intrarresidual	161
Medium range	473
Long range	260
Torsion angle constraints	182
Structure statistics, 20 conformers	
CYANA target function value ($\mathring{A}^2$)	6.48 ± 0.27
Maximal distance constraint violation ($\mathring{A}^2$)	0.44 ± 0.18
Maximal torsion angle constraint violation ($\mathring{A}^2$)	22.79 ± 0.9
AMBER energies in vacuum (kcal/mol)	$-2.95E +3$
PROCHECK – NMR Ramachandran statistics	
Residues in favourable regions (%)	87,8
Residues in additional allowed regions (%)	6,7
Residues in generously allowed regions (%)	3,3
Residues in disallowed regions (%)	2,2
Root mean square deviation to average coordinates ($\mathring{A}$)	
N, C^α, $C^{'}$ (23 -124)	$0,48 \pm 0,10$
Heavy atoms (23 -124)	$0.99 \pm 0,11$

Figures
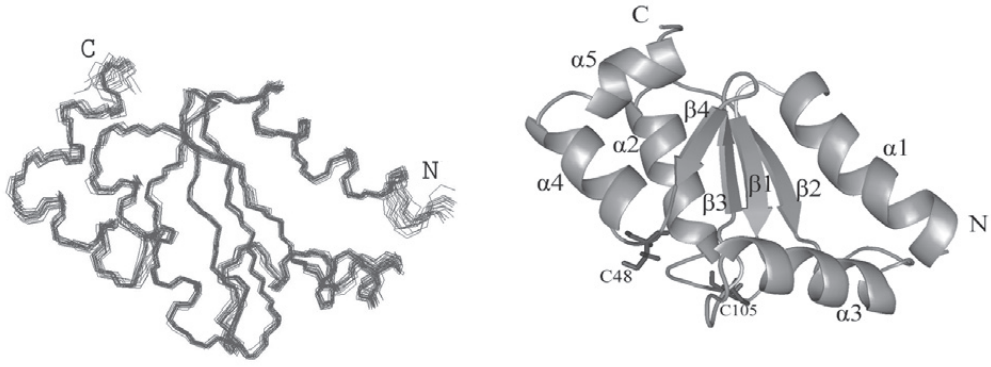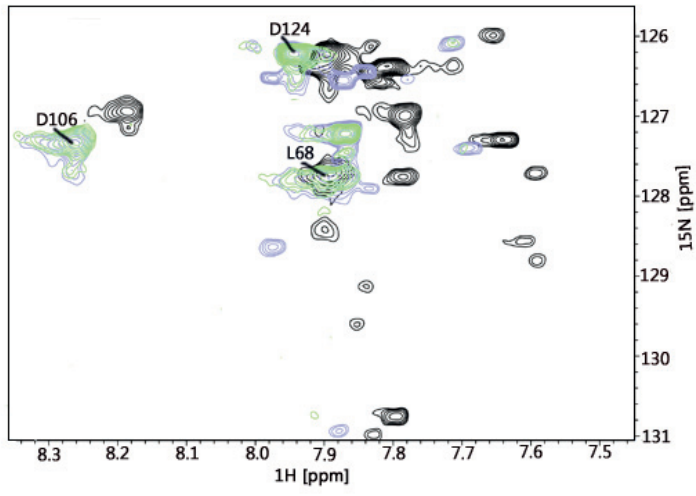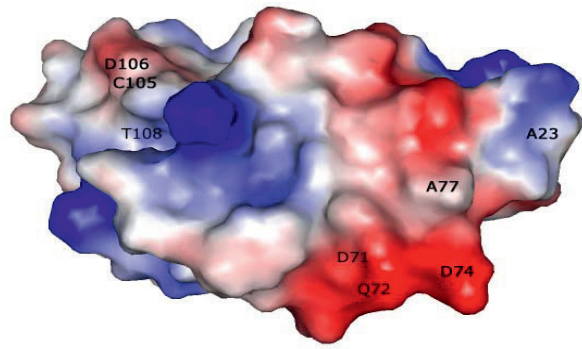
```
Grx-TGR MM  areelrrrlrdliegnrvmifsk sycphs trvkelfsslg---vvynileldqvddgasvq evlteisn-qk
Grx-TGR HS  areelrrhlvgliersrvvifsk sycphs trvkelfsslg---vecnvleldqvddgarvq evlseitn-qk
TGR DR      greqirskikelidssavvvfsk sfcpfc vkvkdlfkeln---vkyntieldlmedgtnyq dllhemtg-qk
TGR XL      grdllqarvkelidsnrvmvfsk sfcpyc drvkdlfsslg---aeyhsleldecddgsdiq ealqeltg-qk
Grx1 EC     --------------mqtvifgr pgcpyc vrakdlaeklsnerddfqyqyvdiraegitke dlqqkagkpve
Grx3 EC     -------------manveiytk etcpyc hrakallsskg-----vsfqelpidgnaakre emikrsgrt--
Grx2 HS     latapvnqiqetisdncvvifsk tscsyc tmakklfhdmn---vnykvveldlleygnqfq dalykmtg-er
TGR SM      padgtsqwlrktvdsaavilfsk ttcpyc kkvkdvlaeak---ikhatieldqlsngsaiq kclasfski-e


Grx-TGR MM  tvpnifvnkvhv ggcdr tfqahqngllqkllqdd
Grx-TGR HS  tvpnifvnkvhv ggcdq tfqayqsgllqkllqed
TGR DR      tvpnvfinkkhi ggcdn tmkahkdgvlqkllgeg
TGR XL      tvpnvfvnkthv ggcdk tlqahkdgslaklldn
Grx1 EC     tvpqifvdqqhi ggytd faawaken-lda
Grx3 EC     tvpqifidaqhi ggcdd lyaldarggldpllk
Grx2 HS     tvprifvngtfi ggatd thrlhkegkllplvhqc
TGR SM      tvpqmfvrgkfi gdsqt vlkyysndelagivnes
```

A

B

T52

R107

Q110

S44

15N [ppm]
118.7
119.0
119.2
119.5
119.7
120.0
120.2
120.5
120.7
121.0
121.2

1H [ppm]
8.7 8.65 8.6 8.55 8.5 8.45 8.4 8.35 8.3 8.25 8.2

C

```
       1          10         20         30         40         50         60         70
Grx   MSSPPGRRARLASPGTSRPSSEAREELRRRLRDLIEGNRVMIFSKSYCPHSTRVKELFSSLGVVYNILEL

              80         90        100        110        120
      DQVDDGASVQEVLTEISNQKTVPNIFVNKVHVGGCDRTFQAHQNGLLQKLLQDD
```

━━ GSH
▪▪▪ GSSG

Competitive Zinc for Cobalt Substitution in Mammalian MsrB1 Overexpressed in *E.coli*: Structural and Functional Insights.

Elena Shumilina[a,§], Olena Dobrovolska[a,§], Rebecca Del Conte[b], Henrik Waldal Holen[a] and Alexander Dikiy[a,*]

[a]Department of Biotechnology, Norwegian University of Science and Technology, N-7491 Trondheim, Norway; [c]CERM, University of Florence, Sesto Fiorentino,50019 Florence, Italy

[§] Authors have made an equal contribution to the work.

[*]Corresponding author: Tel.:+4773597863, Fax: +4773591283; E-mail: alex.dikiy@biotech.ntnu.no