THE SULFUR OXYGENASE REDUCTASE

FROM *ACIDIANUS AMBIVALENS*

FUNCTIONAL AND STRUCTURAL CHARACTERIZATION

OF A SULFUR-DISPROPORTIONATING ENZYME

vom Fachbereich Biologie der
Technischen Universität Darmstadt
zur Erlangung des akademischen Grades eines
Doctor rerum naturalium
genehmigte

DISSERTATION

vorgelegt von

Tim Urich
aus Bad König

1. Berichterstatter: Dr. habil. Arnulf Kletzin
2. Berichterstatterin: Prof. Dr. Felicitas Pfeifer

Eingereicht am: 09.05.2005
Tag der mündlichen Prüfung: 17.06.2005
D17

"...everything that living things do can be understood in terms of the jigglings and wigglings of atoms.“

Richard P. Feynman

Die vorliegende Arbeit wurde als Promotionsarbeit am Institut für Mikrobiologie und Genetik des Fachbereichs Biologie der Technischen Universität Darmstadt im Labor von Herrn Dr. Arnulf Kletzin in der Abteilung von Frau Prof. Felicitas Pfeifer im Zeitraum von Oktober 2001 bis Mai 2005 angefertigt.

Eidesstattliche Erklärung

Ich erkläre hiermit an Eides Statt, daß ich den materiellen Inhalt der vorliegenden Arbeit ohne Hilfe Dritter und ohne Benutzung anderer als der angegeben Hilfsmittel und Quellen angefertigt habe. Die aus fremden Quellen direkt oder indirekt übernommenen Gedanken und Fakten, sofern sie nicht wissenschaftliches Allgemeingut sind, wurden als solche kenntlich gemacht. Die Arbeit wurde bisher in gleicher oder ähnlicher Form keiner anderen Prüfungsbehörde vorgelegt und auch noch nicht veröffentlicht.

Darmstadt, den 13. September 2005,

Tim Urich

Lebenslauf

Persönliche Daten

Name	Tim Urich
Geburtsdatum	Geboren am 12. Januar 1974 in Erbach (Odenwald)
Familienstand	Ledig

Studium

1993-1996	**Grundstudium** der Biologie an der Technischen Universität Darmstadt (TUD)
1996-2001	**Hauptstudium** der Biologie an der TUD; Fächerkombination: Mikrobiologie, Biochemie, Immunologie, Forschungs- und Entwicklungsmanagement & Marketing (Gesamtnote: sehr gut) Thema der Diplomarbeit: Ein neuer Stoffwechselweg bei Archaea. Anaerobe Schwefeldisproportionierung bei *Acidianus ambivalens* und *Sulfurisphaera* sp.
seit Oktober 2001	**Promotion** im Fachbereich Biologie, Institut für Mikrobiologie und Genetik (TUD)
2002-2004	**Forschungsaufenthalte** am Instituto de Tecnologia Química e Biológica (ITQB), Universidade Nova de Lisboa, Oeiras, Portugal. AG Prof. Miguel Teixeira AG Prof. Maria Arménia Carrondo
Juni 2005	**Disputation**: summa cum laude

Schulbildung

1980-1984	Grundschule Bad König
1984-1986	Förderstufe Bad König
1986-1993	Gymnasium Michelstadt. Abschluss Abitur, Note: 2,1

ACKNOWLEDGEMENTS

I would like to thank my thesis supervisor **Dr. Arnulf Kletzin** for providing such an exciting research topic and for his friendship and encouragement. His extensive knowledge on thermophiles, microbial sulfur oxidation, biochemistry and beyond greatly supported this work. He always found time for helpful and inspiring discussions.

Special thanks go to **Prof. Dr. Felicitas Pfeifer** for setting up an excellent framework for "extreme" research and for taking the part of the co-supervisor. I very much appreciated the critical and constructive discussions and her continuous support.

I want to express my deepest gratitude to **Dr. Carlos Frazão** (Instituto de Tecnologia Química e Biológica, Universidade Nova de Lisboa, [ITQB], Oeiras, Portugal) for being my supervisor in the crystallographic part of this work and for never getting tired of teaching protein crystallography to a rookie. Spending many sleepless nights at synchrotrons, he solved the phase problem.

Prof. Dr. Maria Arménia Carrondo (ITQB) is acknowledged for making possible my stays in her crystallography laboratory and for the friendly uptake and her support.

Many thanks to Ricardo Coelho (ITQB) for the instructions on protein crystallization and for preparing crystals when I was not around. Prof. Dr. Claudio Soares is acknowledged for the help in electrostatic calculations. Peter Crowley (ITQB) is acknowledged for the preparation of the zink derivative. Thanks to all members of the xtal group (whom I can´t mention in detail because the list would be too long) for various kinds of support in crystallography and beyond.

Prof. Dr. Miguel Teixeira, Dr. Claudio Gomes, Dr. Tiago Bandeiras and Sónia Leal (ITQB) are very much acknowledged for a very productive and instructive collaboration and for providing the EPR and fluorescence spectroscopic data presented in this thesis. Many thanks for sharing your *ambivalens* treasures with our group, and for the encouragement in wild type SOR purification, despite the fact that "she" is colorless. Thanks for all the support during my time at ITQB.

Dr. Reinhard Rachel (Archaeenzentrum, Regensburg, Germany) is acknowledged for providing the electron microscopic data presented in this thesis. Anja Kroke (TUD) is acknowledged for the nice time during her diploma thesis and for providing mutant data to this thesis, as did Muriel Reuff and Kerstin Seyfarth during the time of their “Forschungspraktika”. Till Albrecht (TUD) provided some biochemical data while doing an enthusiastic "Forschungspraktikum".

This work was financially supported by a research grant from the Deutsche Forschungsgemeinschaft (contract Kl885/3-1, 3-2 and 3-3), and in part by a Marie Curie fellowship from EU (contract HPMT-CT-2000-00045) and by a travel grant from the Deutscher Akademischer Austauschdienst (contract 314-Al-p-dr).

Many thanks go to Christa Schleper (TUD) for a lot of inspiring and motivating discussions.
Special thanks go to Melanie Jonuscheit (TUD) for a great time and some stinky experiences. Sabrina, Alex, Achim and Torsten (TUD), my room mates in office 361, are acknowledged for making it a happy place to work in. Thanks to Fabi for exciting and funny times especially at ITQB.
I want to thank all the members of the Pfeifer group (TUD). I have spent some really great times with all of you during both work and leisure.
Tiago Bandeiras, João Vicente and Tiago Lourenço are deeply acknowledged for being outstanding examples for making friends through science. Thanks for making my stays in Portugal such wonderful experiences! The same goes to Daniele de Sanctis, and for sharing Italian pasta in all possible variations.

My parents **Traudel** and **Karlheinz Urich** gave me all kind of support during the last years, despite the fact that their son worked with something that produces the smell of rotten eggs.

....not to be forgotten the tiny, lovely SOR, for allowing fascinating insights into the beauty of nature.

Parts of the present work and of other projects during the time of this thesis have already been published or have been submitted for publication.

A. KLETZIN, **T. URICH**, F. MÜLLER, T.M. BANDEIRAS AND C.M. GOMES (2004). Dissimilatory oxidation and reduction of elemental sulfur in thermophilic archaea. J Bioenerg Biomembr 36 77-91. Review.

T. URICH, T.M. BANDEIRAS, S.S. LEAL, R. RACHEL, T. ALBRECHT, P. ZIMMERMANN, C. SCHOLZ, M. TEIXEIRA, C.M. GOMES AND A. KLETZIN (2004). The sulphur oxygenase reductase from *Acidianus ambivalens* is a multimeric protein containing a low-potential mononuclear non-haem iron centre. Biochem J 381 137-46.

F.H. MÜLLER, T.M. BANDEIRAS, **T. URICH**, M. TEIXEIRA, C.M. GOMES AND A. KLETZIN (2004). Coupling of the pathway of sulfur oxidation to dioxygen reduction: characterization of a novel membrane-bound thiosulfate:quinone oxidoreductase. Mol Microbiol 53(4) 1147-60.

T. URICH, R. COELHO, A. KLETZIN AND C. FRAZÃO (2005). The sulfur oxygenase reductase from *Acidianus ambivalens* is an icosatetramer as shown by crystallization and patterson analysis. Biochim Biophys Acta 1747(2) 267-270.

T. URICH, A. KROKE, C. BAUER, K. SEYFARTH, M. REUFF AND A. KLETZIN (2005). Identification of the core active site of the sulfur oxygenase reductase from *Acidianus ambivalens* by site-directed mutagenesis. FEMS Microbiol Lett 248: 171-176.

S.-V. ALBERS, M. JONUSCHEIT, S. BECKER, **T. URICH**, A. KLETZIN, R. TAMPÉ, A.J.M. DRIESSEN AND C. SCHLEPER (2005). Production of recombinant and tagged proteins in the hyperthermophilic archaeon *Sulfolobus solfataricus* with an inducible gene expression system. Appl Environment Microbiol. In review.

T. URICH, C.M. GOMES, A. KLETZIN AND C. FRAZÃO (2005). X-ray structure of a spherical self-compartmentalising sulfur-cycle metalloenzyme. Submitted for publication.

INDEX

SUMMARY

The microbial oxidation of reduced inorganic sulfur compounds and elemental sulfur to sulfate is one of the major reactions in the global sulfur cycle. Despite its importance, only limited information is available about molecular details of the enzymes involved. The present work was aimed to contribute to the understanding of the underlying molecular mechanisms by investigating the function and structure of the sulfur oxygenase reductase (SOR) from the thermoacidophilic crenarchaeote *Acidianus ambivalens*.

The expression of the *sor* gene in *Escherichia coli* resulted in active, soluble SOR and in inclusion bodies from which active SOR could be refolded as long as ferrous ions were present in the refolding solution. The wild type and recombinant SOR preparations possessed indistinguishable properties when analyzed for activity and by gel permeation chromatography, CD spectroscopy and electron microscopy. The analysis of the quaternary structure showed a multi-subunit shell-like assembly with a central hollow core. The subunits formed homodimers as the building blocks of the holoenzyme, as shown by denaturation experiments. Conformational stability studies showed that the apparent unfolding free energy in water was ~5 kcal mol^{-1}, at pH 7. Iron was found in the wild type enzyme at a stoichiometry of one iron per subunit. EPR spectroscopy of the colorless SOR resulted in a single isotropic signal at g = 4.3 characteristic of high-spin ferric iron. The signal disappeared upon reduction with dithionite or incubation with sulfur at elevated temperature. The iron center had a reduction potential of $E_0' = -268$ mV at pH 6.5. Protein database inspection identified five SOR protein homologues which allowed the prediction of amino acids putatively involved in catalysis.

The recombinant SOR was crystallized by the sitting drop vapor diffusion method. The crystal structure was determined at 1.7 Å resolution. The homo-icosatetrameric holoenzyme was a highly symmetrical hollow protein particle with 432 point group symmetry and a molecular mass of 871 kDa. The subunits were αβ-proteins and comprised a central β-barrel surrounded by α-helices. Each monomer contained one mononuclear non-heme iron site with the ligands H85, H89, E113 and two water molecules in an octahedral arrangement. The protein ligands formed a 2-His-1-carboxylate facial triad for iron binding. The cysteines C30, C100 and C103 were in the vicinity of the iron site and located along the same cavity within the interior of the subunit, therefore defining the enzyme´s active site. C30 was persulfurated. The 24 active sites

were spacially separated from each other, making an electronic interaction during catalysis unlikely. They were accessible solely via the inner compartment. Access of substrate to the inner compartment is most probably provided by six hydrophobic channels along the four-fold symmetry axes of the particle. Furthermore, the structure suggested that a linear polysufide species and not the cyclic α-S_8 is the substrate of the SOR. Crystal structures of the SOR in complex with the inhibitors p-hydroxy-mercury-benzoic acid and iodoacetamide identified the cysteines as the inhibitor binding sites. The iron-binding residues H85, H89 and E113 and the three cysteines C30, C100 and C103 were altered by site-directed mutagenesis and the mutant proteins were analyzed for activity and iron content. Mutations of the iron ligands and C30 resulted in inactive enzyme, whereas mutations of C100 and C103 resulted in a reduced activity. All mutations affected the oxygenase and reductase partial reactions to a similar degree.

These analyses allowed the first detailed insight into the mode of action of this self-compartmentalizing metalloenzyme. C30 is most probably the sulfur binding residue which aligns the substrate for the initial oxygenation catalyzed by the Fe site. The role of C100 and C103 is not clear, but they might act in the subsequent sulfur disproportionation reaction.

The comparison of the SOR with SoxAX, the only other sulfur compound oxidizing enzyme from prokaryotes for which a high resolution structure is available, showed no structural similarity. SoxAX and other sulfur oxidizing enzymes contain different cofactors, demonstrating the diversity of mechanistic approaches utilized for sulfur compound oxidation. In contrast, a basic functional principle seems to be the central role of cysteine residues, acting as covalent binding sites for the substrate.

1. GENERAL INTRODUCTION

Enzymes play a decisive role in almost all biological processes. They are potent biological catalysts, often enhancing the velocity of chemical reactions a million-fold or more. Diverse reactions, like the simple addition of water to carbon dioxide or the complex replication of whole genomes are catalyzed by specific enzymes. Enzymes are also the determinants of the amazing metabolic diversity of prokaryotic organisms. They enable prokaryotes to live under the most unlikely conditions, deriving energy for growth from the oxidation or reduction of such exotic substrates, as e.g. sulfur, metal ions, (aromatic) hydrocarbons or gaseous hydrogen. In particular, prokaryotes, respectively their enzymes catalyze essential steps of biogeochemical cycles like methanogenesis, sulfur oxidation, nitrogen fixation and denitrification.

The size and shape of enzymes are as diverse as their catalytic versatility, ranging from single-subunit enzymes with molecular masses of a few kilo dalton to huge multi-subunit enzyme complexes, which interact in a highly temporarly and spacially regulated way during catalysis. Like the tools and machines invented by mankind, enzymes are characterized by highly specialized structures enabling them to perform their biological function. Understanding their mode of action at the molecular and atomic level, respectively, is one of the most challenging tasks in life sciences and a prerequisite for their use in biotechnological processes or biomedical applications. Due to their complexity, an understanding requires detailed information about their catalytic function and about their three-dimensional structure. Many research disciplines contribute to this highly inter-disciplinary field, among others molecular genetics, biochemistry, biophysics and structural biology (Alberts, 1998).

The structure and function of the enzyme systems involved in the central metabolic pathways are understood to a certain degree (Sali & Chiu, 2005). In contrast, the energy-yielding metabolic pathways of many lithotrophic prokaryotes are often poorly understood at the enzymatic and molecular level, despite their biogeochemical and biotechnological significance. The present work is an attempt to shed light on the molecular mechanisms underlying microbial sulfur oxidation by investigating the function and structure of the sulfur oxygenase reductase (SOR), the initial enzyme in the dissimilatory sulfur oxidation pathway of the thermoacidophilic crearchaeote *Acidianus ambivalens*.

1.1. MICROBIAL SULFUR OXIDATION

The dissimilatory oxidation of reduced inorganic sulfur compounds and elemental sulfur to sulfate is one of the major reactions in the global sulfur cycle and is restricted to the prokaryotic domains of life (Kelly, 1979; Middelburg, 2000). Many lithotrophic micro-organisms derive energy for growth from these reactions, which are also essential in biotechnological processes such as the biomining of base and precious metal ores or the desulfurization of waste waters (Rohwerder *et al*. 2003; Olson *et al*., 2003).
Despite its significance and contrasting more than one century of research on microbial sulfur oxidation (Kelly *et al*., 1997), the process is still poorly understood at the molecular level. Especially for the oxidation of elemental sulfur few details are known about the enzymatic systems involved (Rohwerder & Sand, 2003).

1.1.1. BACTERIAL SULFUR OXIDATION PATHWAYS

The sulfur-oxidizing lithotrophic prokaryotes are phylogenetically diverse and belong to the archaeal and bacterial domains of life. Within the archaea, sulfur oxidation is restricted to the thermoacidophilic order *Sulfolobales*. In contrast, many different lithotrophic bacterial species are known including *Acidithiobacillus, Aquaspirillum, Aquifex, Bacillus, Beggiatoa, Methylobacterium, Paracoccus, Pseudomonas, Starkeya, Thermithiobacillus, Thiobacillus* and *Xanthobacter*. They are, besides *Aquifex,* mainly mesophilic (Friedrich *et al*., 2001). In addition, many phototrophic bacteria utilize reduced sulfur compounds as electron donors for anoxygenic photosynthesis. They include the genera *Allochromatium*, *Chlorobium*, *Rhodobacter*, *Rhodopseudomonas*, *Rhodovulum* and *Thiocapsa,* which are also mainly mesophilic (Friedrich *et al*., 2001).
Various sulfur compounds like sulfide, elemental sulfur, sulfite, thiosulfate and polythionates can serve as substrates (Brüser *et al*., 2000). The organisms differ in their ability to use different sulfur compounds and as a consequence, their sulfur oxidation pathways can be expected and are indeed found to be variable.
The oxidation of elemental sulfur (S^0) proceeds in at least two steps. S^0 is usually oxidized to sulfite (HSO_3^-) by a sulfur oxygenase (Suzuki, 1965a; Suzuki & Silver, 1966; Rohwerder & Sand, 2003) or a sulfur dehydrogenase (Rother *et al.*, 2001). In a second step sulfite is oxidized to sulfate (SO_4^{2-}) catalyzed by sulfite:acceptor oxidoreductases or dehydrogenases. Various intermediates may be formed like thiosulfate (SSO_3^{2-}) and tetrathionate ($O_3SSSSO_3^{2-}$; Kelly, 1982, 1985).
Most studies investigating S^0 and sulfur compound oxidation pathways focused on thiosulfate and sulfide oxidizing species because these soluble substrates render

laboratory investigations easier. There are at least two major pathways of sulfur compound oxidation postulated in bacteria, which have been reviewed repeatedly (e.g. Kelly, 1982, 1985; Pronk *et al*., 1990; Kelly & Wood, 1994; Kelly *et al*., 1997; Friedrich, 1998; Brüser *et al*., 2000; Friedrich *et al*., 2001). The best studied is the oxygen-independent SOX pathway (**s**ulfur **ox**idizing pathway) found in mesophilic bacteria growing at neutral pH, among others in *Paracoccus panthotrophus* and *Starkeya novella*. There, the oxidation of the sulfane (S^0) and sulfone (SO_3^{2-}) sulfur of thiosulfate to sulfate is performed by the periplasmic Sox complex without the formation of any free intermediate and is coupled to cytochrome C reduction (Friedrich *et al*., 2001; Kelly *et al*., 1997; Brüser *et al*., 2000). A model for the pathway has been recently developed (Fig. 1.1; Friedrich *et al*., 2001). SoxAX catalyzes the oxidative and covalent transfer of thiosulfate to the substrate carrier protein SoxYZ. The sulfone sulfur is cleaved off by the hydrolase SoxB, yielding sulfate, and subsequently the sulfane group is oxidized to sulfone by SoxCD and again hydrolyzed by SoxB. It was shown that the Sox complex is able to oxidize hydrogen sulfide (H_2S), sulfur (S^0) and sulfite besides thiosulfate in a cytochrome C dependent manner *in vitro* (Rother *et al*., 2001), indicating its central role in sulfur compound oxdiation in mesophilic bacteria growing at neutral pH.

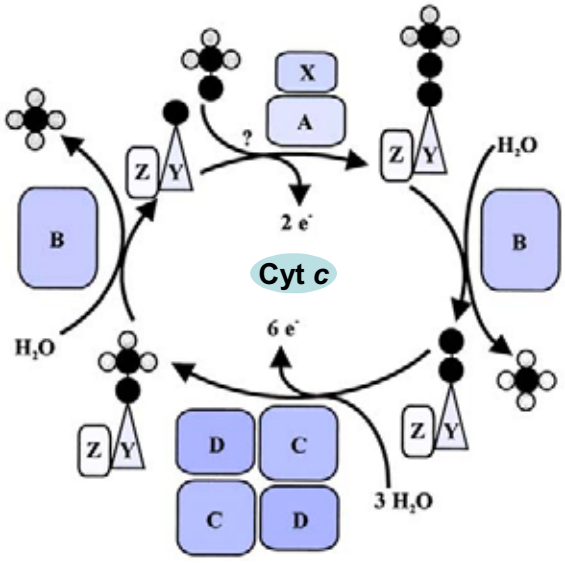


Figure 1.1: The Friedrich-Kelly model of thiosulfate oxidation by the periplasmic Sox complex found in neutrophilic bacteria (Friedrich *et al*., 2001). Refer to text for details.

A different type of sulfur compound oxidation pathway not involving the Sox complex seems to be present in mesophilic bacteria growing at acidic pH (pH 1-3) like *Acidithiobacillus* and *Acidiphilum*. Here, all thiosulfate is oxidized to the stable intermediate tetrathionate prior to cleavage by means of a tetrathionate hydrolase (Brüser

et al., 2000). Therefore, it has been termed S_4-intermediate pathway. Up to now, no conclusive model for this pathway has been developed and the central role of tetrathionate has come recently under dispute (Brüser *et al*., 2000). In addition, a different model not involving tetrathionate has been developed for the oxidation of S^0, where the initial step of sulfur oxidation is performed by periplasmic oxygen- and glutathione-dependent sulfur dioxygenases catalyzing the oxidation of elemental sulfur to sulfite (Fig. 1.2; Rohwerder & Sand 2003). Only limited information is available about these enzymes besides their oxygen- and glutathione-dependency (Suzuki, 1965b). A 46 kDa sulfur dioxygenase has been purified from *Acidithiobacillus thiooxidans* and *A. ferrooxidans* (Suzuki, 1965a; Silver & Lundgren, 1968) consisting of two subunits of 21 and 26 kDa or two 23 kDa proteins, respectively, but no additional molecular details are known. The resulting sulfite is postulated to be further oxidized to sulfate by a sulfite:acceptor oxidoreductase (Rohwerder & Sand, 2003).

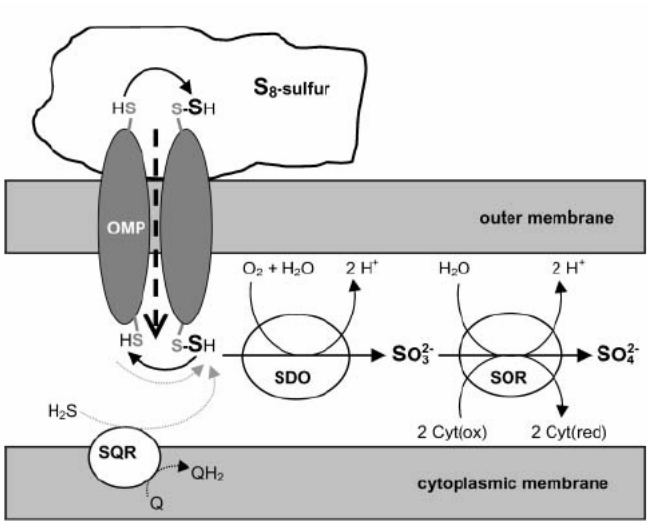


Figure 1.2: Hypothetical model for the oxidation of elemental sulfur in acidophilic bacteria (Rohwerder & Sand, 2003). Sulfur is transported to the periplasm via an outer membrane protein and oxidized to sulfite by a sulfur dioxygenase. OMP outer membrane protein, SDO sulfur dioxygenase, SOR sulfite:acceptor oxidoreductase, SQR sulfide:quinone oxidoreductase

In summary, although many bacterial enzymes have been characterized or postulated to be involved in the oxidation of inorganic sulfur compounds, information about their molecular details and mechanisms is limited.

1.1.2. SULFUR OXIDATION IN ARCHAEA

Although an entire order of archaeal species has been given its name due to their ability to oxidize sulfur (*Sulfolobales,* Brock *et al*., 1972), the knowledge about the enzymes involved has been for a long time even more limited than about enzymes from bacteria. Members of the *Sulfolobales* populate terrestrial hydrothermal vents, solfataras, hot springs and other habitats of volcanic origin and are well adapted to the often extreme and seemingly adverse growth conditions (Fig. 1.3; Barns *et al*., 1994; Stetter, 1996; Kletzin *et al*., 2004).

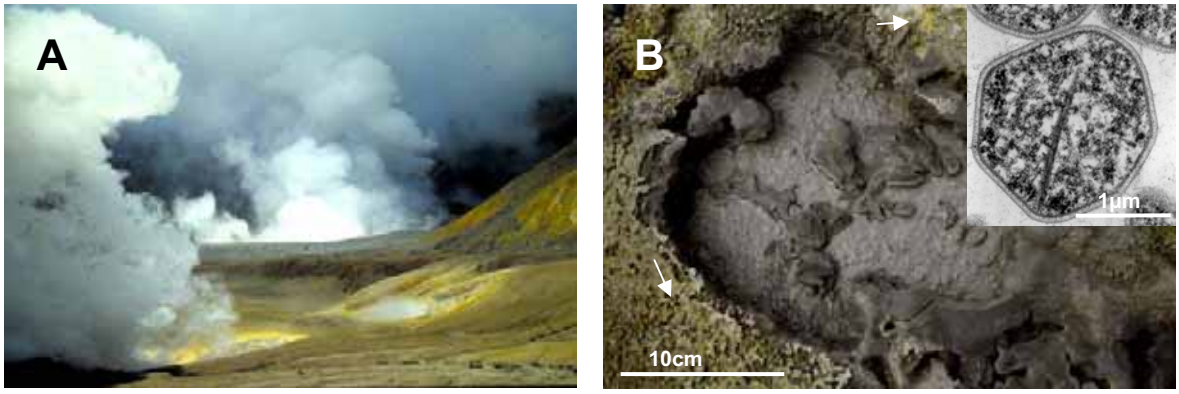


Figure 1.3: (A) Solfataric field with fumaroles and hot springs on White Island, New Zealand. Note the yellowish patches in the landscape, which are formed by S^0 deposites. **(B)** Small boiling pool from Furnas solfatara field (São Miguel, Açores, Portugal), heated by a stream of hot gas from the underground. The cell densities in these small holes with grayish or brown turbid liquid and S^0-containing yellow precipitants (arrows) reach up to 10^8 cells ml^{-1}, most of which belong to the *Sulfolobales*. Photos: A. Kletzin. Inset: Thin section electron micrograph of *Acidianus ambivalens.* Photo: W. Zillig.

Inorganic sulfur compounds and elemental sulfur are among the major energy sources in these habitats and are utilized by sulfur dependent chemolithoautotrophic organisms which constitute the basis of the food chains in these light-independent ecosystems. Only recently, *Acidianus ambivalens*, a thermoacidophilic and chemolithoautotrophic member of the *Sulfolobales* within the archaeal kingdom crenarchaeota (Fig. 1.3B, inset; Zillig *et al*., 1985, 1986; Fuchs *et al*., 1996) has developed into a model organism for these sulfur-dependent hyperthermophilic archaea (Laska *et al*., 2003; Kletzin *et al*., 2004; Urich *et al*., 2004; Müller *et al*., 2004; Gomes *et al*., 2001a, 2001b; Bandeiras *et al*., 2002).

1.1.3. THE SULFUR OXIDATION PATHWAY IN *ACIDIANUS AMBIVALENS*

A. ambivalens oxidizes elemental sulfur to sulfuric acid for energy conservation, when growing aerobically (Zillig *et al*., 1986). Both soluble and membrane-bond enzymes are

involved in the pathway (Fig. 1.4). Unlike in other sulfur oxidation pathways described, the initial step is catalyzed by a soluble cytoplasmic enzyme, the sulfur oxygenase reductase (SOR), which is the only sulfur oxidizing enzyme found in this organism. In consequence, sulfur has to be transported from the medium to the cytoplasm via the cytoplasmic membrane. It is not yet resolved how this takes place, but an excreted sulfur-binding protein has been recently identified (Lauber & Kletzin, unpublished). The SOR catalyzes the oxygen-dependent production of sulfite, thiosulfate and hydrogen sulfide from elemental sulfur. As no energy is conserved during this catalysis, the function of the SOR is most likely the generation of soluble sulfur species from the almost insoluble substrate elemental sulfur. The products of the SOR reaction are substrates for membrane bound enzymes like a thiosulfate:quinone oxidoreductase (TQO; Müller *et al*., 2004) and a sulfite:acceptor oxidoreductase (SAOR, Zimmermann *et al*., 1999), which are coupling the substrate oxidation to energy conservation. Furthermore, adenylylsulfate (APS) reductase and adenylylsulfate:phosphate adenylyltransferase (APAT), enzymes involved in the generation of ATP from sulfite by substrate level phosphorylation, have been identified (Zimmermann *et al*., 1999). For the third product of the SOR reaction, hydrogen sulfide, membrane-bound sulfide:quinone reductases (SQR) are known from bacteria and eukaryotes, which feed the electrons from sulfide oxidation into the quinone pool of the membrane. Although homologues to *sqr* genes can be found in the genome (Kletzin, personal communication), SQR activity has so far not been shown for *A. ambivalens*. The sulfite metabolizing enzymes have not yet been isolated, but the TQO has been extensively characterized (Müller *et al*., 2004). The enzyme catalyzes the generation of tetrathionate from two molecules of thiosulfate feeding the electrons into the quinone pool of the cytoplasmic membrane. In addition, a thiosulfate-driven consumption of oxygen has been shown, linking sulfur metabolism to the activity of the terminal quinole oxidase of *A. ambivalens*. No details are known about the fate of tetrathionate, but it has been postulated that a thiosulfate/tetrathionate cycle may exist, where the latter reacts non-enzymatically with hydrogen sulfide, one of the products of the SOR reaction, by which thiosulfate could be regenerated (Müller *et al*., 2004). Alternatively, a tetrathionate hydrolase may exist, as found in bacteria (Suzuki, 1999).

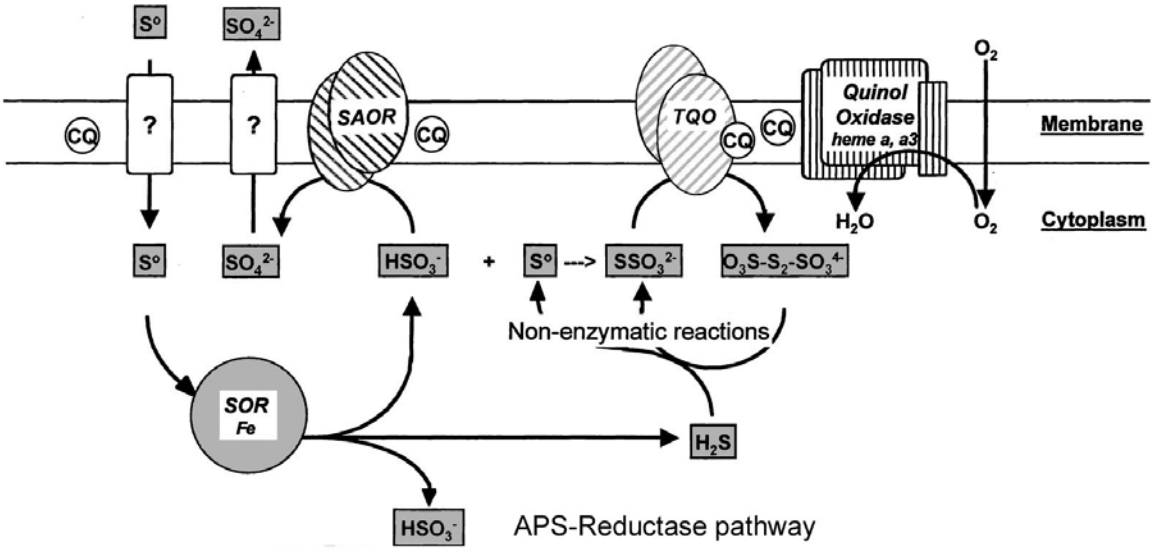


Figure 1.4: Hypothetical model of the sulfur oxidation pathway in *A. ambivalens* (Kletzin *et al*., 2004). Note the cytoplasmic position of SOR as the initial sulfur oxidizing enzyme. SOR: sulfur oxygenase reductase, SAOR sulfite:acceptor oxidoreductase, TQO thiosulfate:quinone oxidoreductase, CQ caldariella quinone.

1.2. THE SULFUR OXYGENASE REDUCTASE

The SOR is the only sulfur oxidizing enzyme known from archaea. It has been first described and purified from *A. ambivalens* (Kletzin, 1989) and the *sor* gene encoding the enzyme has been sequenced (Kletzin, 1992). Even before, an enzyme described as a sulfur oxygenase had been purified from a phylogenetically not classified isolate termed "*Sulfolobus brierleyi*" (Emmel *et al*., 1986). The SOR and the sulfur oxygenase were very similar with respect to the sizes of the holoenzymes and of the single subunits, the reaction products sulfite and thiosulfate and other properties (Table 1.1; Kletzin, 1994; Kletzin *et al*., 2004) but an S^0 reducing activity of the sulfur oxygenase had not been reported (Emmel *et al*., 1986). It was concluded from the similarities that both the SOR and the sulfur oxygenase actually catalyze the same reaction and that the sulfur reducing activity of the "*S. brierleyi*" enzyme had been overlooked (Kletzin, 1994). A moderate incorporation of ^{18}O from O_2 into sulfite has been demonstrated with the sulfur oxygenase (Emmel *et al*., 1986) as with a glutathione-dependent sulfur oxygenase from *Thiobacillus thiooxidans* (now *Acidithiobacillus thiooxidans*; Suzuki, 1965b). The *A. ambivalens* SOR is active only under air but not under H_2 or N_2 atmosphere (Kletzin, 1989). Both experiments show that the enzymes are indeed oxygenases.

Table 1.1: Biochemical properties of the SOR enzymes and the sulfur oxygenase.

	A. ambivalens	***"S. brierley"***	***A. tengchongensis***
	SOR	S-oxygenase	SOR
M_r Holoenzyme [Da][a]	550,000	560,000	n.r.
M_r Subunit [Da]	35,187 [b]	35,000	35,172[b]
pH range	4-8	n.r.	3.5-9
pH_{opt}	7-7.4	6.5-7.5	5
T_{opt}	85°C	65°C	70˚C
T_{max}	108°C	>80°C	> 90
Specific oxygenase activity	10.6 U/mg[c]	0.9 U/mg[c]	186.7 U/mg[d]
Specific reductase activity	2.6 U/mg[c]	n.r.	45.2 U/mg[d]
^{18}O-incorporation	n.r.	+	n.r.
Diameter	15.6 nm	n.r.	n.r.
Reference(s)	Kletzin, 1989, 1992	Emmel *et al*., 1986	He *et al*., 2000; Sun *et al*., 2003

[a]from gel permeation chromatography n.r. = not reported
[b]from sequence
[c]wild type enzyme
[d]recombinant enzyme

The same coupled oxygenase and reductase reaction has been found for a third enzyme, the SOR from *Acidianus tengchongensis* (He *et al*., 2000; Sun *et al*., 2003; He *et al*., 2004). The temperature optima of the *A. tengchongensis* and the "*S. brierleyi*" enzymes are 65-70°C, and of the *A. ambivalens* SOR 85°C (Kletzin, 1989), consistent with the growth temperatures of the host organisms (Tab. 1.1). The holoenzymes have high apparent molecular masses around 550 kDa and are each composed of a single 35-36 kDa subunit (Emmel *et al*., 1986; Kletzin, 1989; He *et al*., 2000). In electron micrographs of the *A. ambivalens* SOR hollow globular particles of ~15.5 nm in diameter appear (Fig. 1.5; Kletzin, 1989).

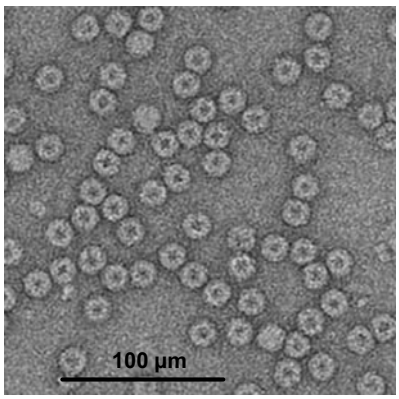


Figure 1.5: Electron micrograph of the SOR holoenzyme. Globular particles with a diameter of ~15.5 nm are visible (Kletzin, 1989).

The SOR catalyzes a unique reaction, since it simultaneously oxidizes and reduces S^0 (Kletzin, 1989; He *et al*., 2000). Only once a similar enzyme activity has been reported from the mesophilic bacterium *Thiobacillus thiooxidans* (Tano & Imai, 1968), where in cell-free extracts the production of thiosulfate and hydrogen sulfide from S^0 has been detected. Neither were the corresponding enzymes isolated, nor were these results confirmed independently.

The reaction products of the SOR are sulfite, thiosulfate and hydrogen sulfide. It is not clear whether thiosulfate is a product of catalysis or non-enzymically derived from the reaction of sulfite with excess sulfur under the assay conditions (Kletzin, 1989). No external electron donor is necessary *in vitro,* thus sulfur must be the electron donor as well as the electron acceptor of the reaction. It is difficult to determine a stoichiometry of the enzymic reaction because the product H_2S is highly volatile and quickly evaporates from the liquid phase. The use of H_2S-complexing zinc ions in the assay has resulted in a stoichiometry of 1 between reduced and oxidized products (Kletzin, 1989), which can be described by equation 1.1, assuming that thiosulfate is non-enzymically formed from sulfite and sulfur, and therefore not a primary product of catalysis.

Eq. 1.1 $$4\ S^0 + O_2 + 4\ H_2O \rightarrow 2\ H_2S + 2\ HSO_3^- + 2\ H^+$$

Formally, this reaction can be split into two partial reactions, a sulfur oxygenation (Eq. 1.2) and a sulfur disproportionation (Eq. 1.3).

Eq. 1.2 $$S^0 + O_2 + H_2O \rightarrow HSO_3^- + H^+$$

Eq. 1.3 $$3\ S^0 + 3\ H_2O \rightarrow HSO_3^- + 2\ H_2S + H^+$$

From these results, it has been concluded that the SOR catalyzes a S^0 disproportionation coupled to a sulfur oxygenation reaction (Kletzin, 1989).

P-chloromercurybenzoic acid, Zn^{2+} ions and iodoacetic acid are potent inhibitors of the SOR reaction, whereas N-ethylmaleimide shows a moderate inhibition (Kletzin, 1989). All these substances are known to be thiol-binding reagents. Therefore, it has been concluded that sulfur bound via the thiol groups of cysteines to the enzyme is the actual substrate of the SOR. No inhibitor specifically abolishes the oxygenase or reductase partial reaction, thus both seem to occur in a coupled fashion. The SOR is a colorless protein, in UV/Vis spectra no feature besides the 280 nm protein peak can be seen (Kletzin, 1989). This points to the absence of heme groups or iron-sulfur centers as redox-active components.

1.3. ELEMENTAL SULFUR

Elemental sulfur has some chemical and physical characteristics, which pose challenges to organisms utilizing S^0 as substrate but also to the researcher investigating the enzymatic processes. Chemically, sulfur is "one of the most interesting but also one of the most difficult elements" to investigate (Steudel, 2000). The complexity originates from the many oxidation states it can adopt (-2, -1, 0, +1, +4, +5, +6) and from the tendency of sulfur in the valence state 0 to self-associate to form rings and chains of a huge variety. This chemical reactivity makes the understanding of the metabolism of inorganic sulfur compounds difficult (Suzuki, 1999).

The most abundant allotrope of elemental sulfur is the orthorhombic ring-forming α-S_8 sulfur (Fig. 1.6) which is stable below 96°C (Steudel, 2000). At higher temperatures it is transformed into the monoclinic β-S_8 sulfur, which is stable up to 119°C. Elemental sulfur is practically insoluble in water at ambient temperature (5 µg/l H_2O; Boulegue, 1978). Therefore, microorganisms have to activate sulfur prior to oxidation to make it better available (Rohwerder & Sand, 2003). This activation could be an opening of the S_8 ring by nucleophilic reagents, resulting in the formation of linear organic or inorganic polysulfanes, or the reduction of elemental sulfur to water-soluble polysulfides (Fig. 1.6). Both reactions could be carried out by the thiol groups of cysteine residues within proteins.

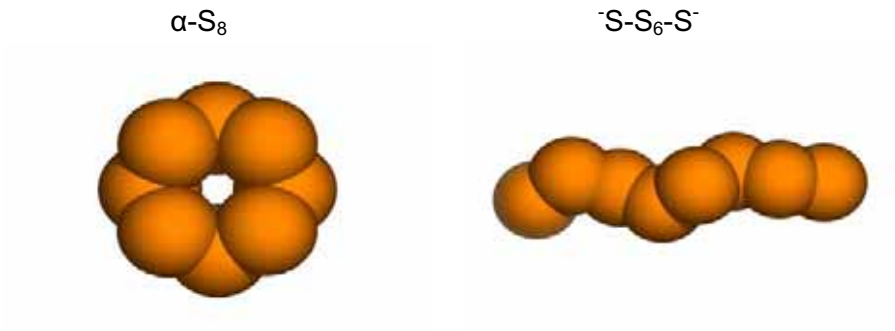


Figure 1.6: Elemental sulfur (left) and polysulfide (right). The chain length of the polysulfide is hypothetical, as it depends on the equilibrium of sulfur and hydrogen sulfide (Klimmek *et al*., 1991).

Another important aspect in the context of this work is the thermodynamical instability of elemental sulfur in water at pH values above 7. S^0 disproportionates to thiosulfate and sulfide in the presence of hydroxyl ions according to equation 1.4 (McCampbell Hamilton, 1991).

Eq. 1.4 $\frac{1}{2} S_8 + 4\ OH^- \rightarrow S_2O_3^{2-} + 2HS^- + H_2O$

Whereas at ambient temperature a pH above 11.5 is necessary, the reaction occurs at 80°C at a pH of 7.6 due to the increased chemical reactivity of sulfur. Secondary reactions are the formation and disproportionation of polysulfides (Eqs. 1.5, 1.6), however the latter one occurs only at temperatures above 160°C.

Eq. 1.5 $n/8\ S_8 + HS^- + OH^- \rightarrow S_{n-1}^{2-} + H_2O$

Eq. 1.6 $S_5^{2-} + 3\ OH^- \rightarrow S_2O_3^{2-} + 3\ HS^-$

In acidic and neutral solutions the disproportionation of sulfur proceeds as well, but temperatures exceeding 150°C are necessary (Steudel, 2000).

1.4. AIMS OF THIS STUDY

For none of the enzymes involved in the biological oxidation of reduced inorganic sulfur compounds a detailed study focusing on its structure-function relationship had been reported at the onset of this work in October 2001. The overall objective of the present study was to shed some light on basic structural and functional principles underlying the sulfur oxidation mechanism. The SOR from *A. ambivalens* was chosen as an appropriate candidate, because there was already, albeit limited, information available about its reaction mechanism and structure (Kletzin, 1989, 1992). Furthermore, the unique sulfur disproportionation catalyzed by the SOR was of special interest, and in particular how this huge homo-oligomeric protein complex performs this reaction. Different molecular genetic, biochemical, spectroscopic and structural methods were applied to address these questions.

At the beginning of this work no high resolution structure of any enzyme involved in the oxidation of reduced sulfur compounds was available. Since then, crystal structures of two of at least four proteins of the Sox complex from *Rhodovulum sulfidophilum*, of the triheme cytrochrome SoxAX (Bamford *et al*., 2002) and of the thiosulfate-binding protein SoxYZ have been solved (Sauvé *et al*., 2004) however, information on details of the Sox reaction mechanism is still limited.

A prerequisite for all detailed biochemical, spectroscopic and structural analyses is the availability of sufficient amounts of pure protein. This was achieved by heterologously expressing the *sor* gene in *E. coli* and establishing appropriate purification procedures, which is described in chapter 2. The first goal was then the identification of redox-active cofactors of the SOR. Hints had already existed for the presence of iron as a cofactor

(Scholz, 1999). These results were confirmed and the iron site was characterized as a mononuclear non-heme iron center present in each subunit. In addition, the enzyme´s quaternary and secondary structure were analyzed by various biochemical methods and a model for the assembly of the holoenzyme was developed.

In chapter 3, the crystallization and the X-ray structure determination of the SOR are reported. The results showed that the SOR is a highly symmetrical icosatetramer enclosing a hollow inner compartment. Furthermore the active site(s) could be identified and first insights in the mode of action of the SOR were gained.

Chapter 4 focuses on the functional investigation of proposed active site residues by site-directed mutagenesis. The minimal catalytic unit composed of the mononuclear iron site and one cysteine was identified in the course of this work.

Finally, in a general discussion section (chapter 5) all data collected during this work are summarized and discussed with respect to the structure-function relationship of the SOR and the enzyme´s catalytic mechanism. The characteristics of the SOR are compared with other enzymes involved in microbial sulfur compound oxidation pathways to derive general mechanistic principles.

2. BIOCHEMICAL, SPECTROSCOPIC AND STRUCTURAL CHARACTERIZATION OF WILD TYPE AND RECOMBINANT SOR

2.1. INTRODUCTION

Although the SOR from *Acidianus ambivalens* is the best-studied member of the family of sulfur oxygenases reductases, comparatively little information is available about the reaction mechanism and the cofactors involved in catalysis (Emmel *et al.*, 1986; Kletzin 1989, 1992; He *et al.*, 2000).

The availability of sufficient amounts of pure protein is essential for a detailed biochemical analysis. *A. ambivalens* grows only to cell densities of ~ 5 x 10^8 cells ml^{-1}, thus the cultivation of large amounts of cells for biochemical purposes is rather time-consuming. Furthermore, no genetic transformation system that would be useful for functional studies is available for *A. ambivalens* to date. In consequence, a first goal of the present study was the set up of a heterologous expression system of the *sor* gene in *E. coli* and of an appropriate protein purification protocol. Furthermore, the wild type SOR was purified from *A. ambivalens*. The properties of the recombinant and the wild type SOR were compared by several biochemical and spectroscopic methods and by electron microscopy.

The SOR was then analyzed for the presence of cofactors essential for catalysis. The secondary and quaternary structure was investigated by circular dichroism (CD) spectroscopy and electron microscopy. In addition, the conformational stability of the SOR upon denaturant-induced unfolding was monitored by fluorescence spectroscopy, gel permeation chromatography and blue native poly acryl amide gel electrophoresis (BN PAGE). Furthermore, the known SOR amino acid sequences were analyzed for amino acid residues putatively involved in catalysis. The results of this chapter are published in Urich *et al.* (2004).

2.2. RESULTS

2.2.1. HETEROLOGOUS GENE EXPRESSION IN *E. COLI*

A *sor* gene expression plasmid (pASK-sor.05) has been constructed previously by the ligation of the *sor* gene with the pASK75 vector (Skerra, 1994), resulting in a fusion of the *sor* gene with a sequence at the 3' end, encoding a 10 amino acid Strep-tag that could be used for protein purification by affinity chromatography (Zimmermann, 1998; see chapter 6.2.1 for details). The plasmid was transformed into *E. coli* strain BL21 Codon plus (DE3-RIL). The “Codon plus” plasmid encodes tRNA genes for the arginine codons AGA and AGG, isoleucine AUA and leucine CUA, which are underrepresented in *E. coli* but often used in the AT-rich organism *A. ambivalens*.

The expression of the *sor* gene was induced by the addition of anhydrotetracycline (200 µg/l) to a culture growing in LB medium at 37 ˚C at OD_{600} = 0.8. Two hours after induction of *sor* gene expression the occurrence of the first inclusion bodies within the *E. coli* cells was observed microscopically, indicating a high level of *sor* gene expression.

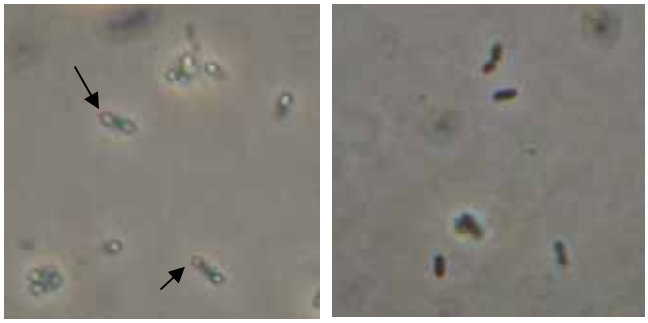


Figure 2.1: Left, phase-contrast micrograph of *E. coli* BL21 Codon plus cells harboring the pASK sor.05 plasmid five hours after induction of *sor* gene expression. Arrows indicate inclusion bodies. Right, phase-contrast micrograph of identically treated *E. coli* BL21 Codon plus cells harboring the original pASK75 plasmid. No inclusion body formation was observed.

2.2.2. PURIFICATION OF RECOMBINANT SOR

Soluble *E. coli* extracts harvested from a 15 l culture 16 h after induction of *sor* gene expression showed an enzyme activity of 464 mU (mg protein)$^{-1}$ (oxygenase) and 145 mU (mg protein)$^{-1}$ (reductase). From the extract (3.5 g of total protein) 120 mg of SOR were isolated by heat treatment, anion exchange chromatography and ultra-centrifugation, which corresponds to a yield of 8 mg per liter of culture (Fig. 2.2; see for details chapter 6.3.1).

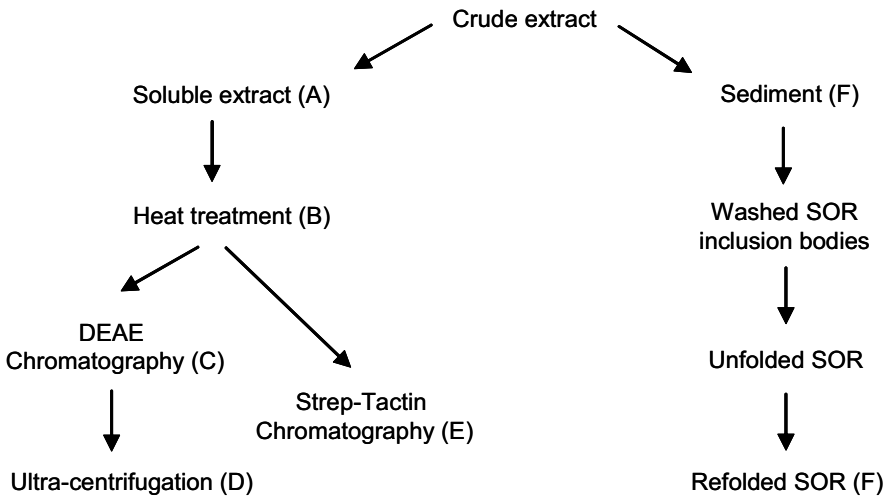


Figure 2.2: Purification scheme of the recombinant SOR from *E. coli*. The labelling refers to the lanes in Fig. 2.3.

The protein was electrophoretically homogeneous (Fig. 2.3D). Apart from the major 36 kDa band, a band corresponding to a dimer and two faint low molecular mass bands were observable. The latter were also found in wild type SOR preparations, and N-terminal sequencing revealed that they were proteolysis products of the 36 kDa subunit (not shown). The specific activity of the purified recombinant enzyme was 2.82 U (mg protein)$^{-1}$ (oxygenase) and 0.66 U (mg protein)$^{-1}$ (reductase). The lowered reductase activity, when compared to the oxygenase activity, was caused by the quick evaporation of hydrogen sulfide from the assay buffer (Kletzin, 1989), thus preventing the exact determination of the reductase activity (see chapter 6.2.5 for details). This effect was seen in all SOR preparations.

In later stages of this work, the time-consuming anion exchange chromatography and ultra-centrifugation steps were replaced by a rapid Strep-tactin affinity chromatography which like-wise resulted in pure protein (Fig. 2.3E; see for details chapter 6.2.2.1). These protein preparations exhibited a higher specific activity of ~11.5 U (mg protein)$^{-1}$ (oxygenase) and ~4.2 U (mg protein)$^{-1}$ (reductase).

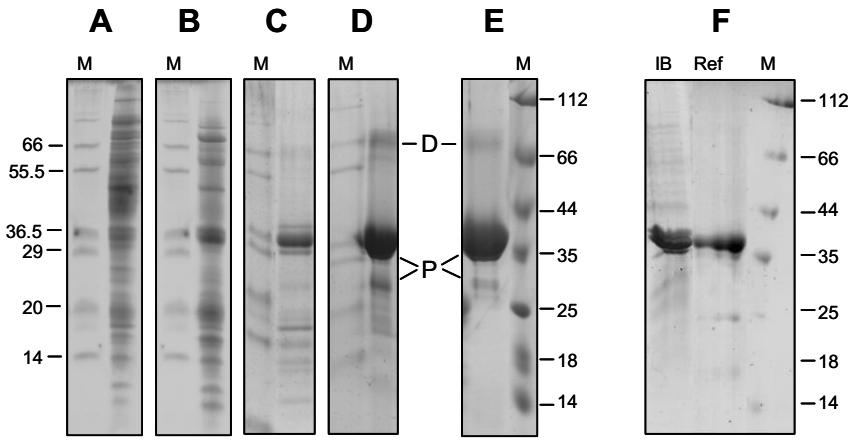


Figure 2.3: (A) - **(E)** 10 % SDS polyacrylamide gels stained with colloidal coomassie blue from the different purification stages of recombinant soluble SOR. A, soluble crude extract; B, supernatant after heat treatment; C, pool after anion-exchange chromatography; D, fraction after ultra-centrifugation; E, Strep-Tactin eluate. Labels: M, molecular mass marker; D, dimer band; P, proteolysis products of SOR subunit. **(F)** 10 % SDS gel with SOR inclusion body fraction (IB) and a fraction after refolding (Ref).

Approximately 80% of the protein in the pellet fraction after cell lysis consisted of SOR (Fig. 2.3F). After extensive washing the yield was 2 g inclusion body fraction (wet mass) per liter of *E. coli* culture. The unfolding and refolding of the SOR inclusion body fraction (see below, section 2.2.3) resulted in > 95 % pure protein (Fig. 2.3F).

2.2.3. UNFOLDING AND REFOLDING OF THE SOR FROM INCLUSION BODIES

Washed inclusion bodies (~ 5 g) were dissolved in 50 ml 8 M urea containing 30 mM dithiotreitol (DTT) as reductant. After denaturation at 60˚C for 1 h the solution was centrifuged to remove particular material. The supernatant containing unfolded SOR was subsequently used for refolding (see chapter 6.3.2 for details).

An initial screening for appropriate refolding conditions was done by refolding denatured SOR aerobically or anaerobically by dialysis against a buffer (50 mM Tris/HCl pH 7.2, 2 mM β-mercaptoethanol) without urea and in the presence or absence of 10 µM ferric citrate, as iron was a possible cofactor of the SOR (Scholz, 1999). Additionally, the influence of an intermediate dialysis step at 2 M urea was investigated. Active SOR was obtained in all tested conditions (Fig. 2.4). High specific activities (≥ 1 U oxygenase / mg protein) were detected in all samples where ferric citrate was present, indicating the importance of ferric ions. Aerobic and anaerobic refolding conditions gave similar results. The one-step refolding without ferric citrate resulted in lower specific activities (~ 0.1 U oxygenase / mg protein), whereas two-step refolding resulted in samples with specific activities comparable to the activity of samples refolded with ferric citrate. After refolding,

the samples were kept at 4°C for up to 8 d and again analyzed for activity. The specific activities of all samples refolded with ferric citrate were similar or even increased, whereas the samples without ferric citrate showed reduced specific activities. Taken the results together, this showed the positive effect of the presence of ferric ions on the refolding efficiency.

Subsequently, refolding was performed aerobically with an increased concentration of ferric citrate (100 µM) and a one-step dialysis. In addition, ferric citrate (100 µM) was added to the unfolding buffer. The activities varied from sample to sample and were up to 5.98 U (mg protein)$^{-1}$ (oxygenase) and 1.35 U (mg protein)$^{-1}$ (reductase).

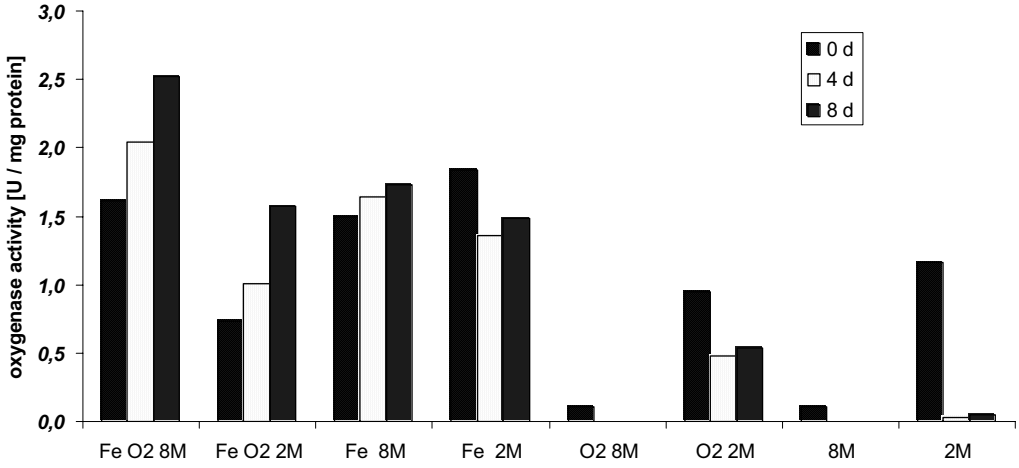


Figure 2.4: Effect of different refolding conditions on the sulfur oxygenase activity. The sulfur reductase activity assays gave similar results (not shown). Labels: Fe, refolded with 10 µM ferric citrate; O2, aerobically dialyzed; 8M, one-step dialysis; 2M, two-step dialysis with 2M urea intermediate step.

In a different experiment, the unfolded protein solution was extensively dialysed against 8 M urea prior to refolding to remove residual iron and no ferric citrate was added to the refolding buffer. The specific activities were 0.006 U (mg protein)$^{-1}$ (oxygenase) and 0 U (mg protein)$^{-1}$ (reductase). An identically treated sample, but refolded in the presence of 10 µM ferric citrate had specific activities of 0.29 U (mg protein)$^{-1}$ (oxygenase) and 0.06 U (mg protein)$^{-1}$ (reductase), showing that the presence of iron is essential for catalysis.

To identify the iron species actually incorporated into SOR, unfolding and refolding was performed with 100 µM ferric citrate in the presence or absence of a reductant. The sample containing ferric (FeIII) ions had no activity, whereas the parallel one containing DTT to reduce ferric to ferrous iron (FeII) showed activity (1.1 U (mg protein)$^{-1}$ (oxygenase) and 0.21 U (mg protein)$^{-1}$ (reductase)). Thus, exclusively ferrous iron is incorporated into the SOR.

2.2.4. PURIFICATION OF WILD TYPE SOR

Wild type SOR was purified from *A. ambivalens* soluble extracts applying two different protocols. Protocol I involved sucrose density gradient centrifugation followed by hydrophobic interaction chromatography and anion exchange chromatography (Scholz, 1999). Protocol II consisted of two anion exchange chromatographic steps followed by gel permeation chromatography (see chapter 6.3.3 for details).

Using protocol I, 10.5 mg of SOR were purified from 175 mg protein of soluble *A. ambivalens* cell extract with specific activities of 2.96 U (mg protein)$^{-1}$ (oxygenase) and 0.60 U (mg protein)$^{-1}$ (reductase). With protocol II 20 mg of SOR were purified from 16.85 g of soluble protein with specific activities of 0.79 U (mg protein)$^{-1}$ (oxygenase) and 0.22 U (mg protein)$^{-1}$ (reductase). SDS-PAGE showed that the purification procedures yielded pure enzymes (Fig. 2.5). Apart from the major 36 kDa band, two very faint low molecular mass bands were observable, but N-terminal sequencing revealed that they were proteolysis products of the 36 kDa subunit (not shown).

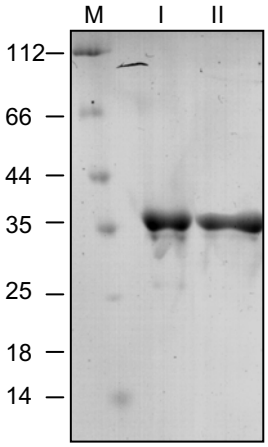


Figure 2.5: 10 % SDS polyacrylamide gel with purified wild type SOR from protocols I and II, stained with colloidal coomassie blue. Amounts of protein: 6.3 µg. M: Molecular mass markers.

2.2.5. MOLECULAR MASS DETERMINATION

The apparent molecular masses of wild type, recombinant soluble and refolded holoenzymes determined by gel permeation chromatography were approximately 550 kDa showing that the enzyme preparations from *E. coli* adopted a near-native state (Fig. 2.6; see chapter 6.3.7 for details). A minor peak corresponding to ~75 kDa equivalent to a dimer of subunits as shown by SDS-Page was also observed in varying intensities in all samples.

The determination of the molecular mass of the holoenzyme by non-denaturing Blue Native (BN) PAGE using the Ferguson plot (Ferguson, 1964) gave a different value of 732 kDa (Fig. 2.6; see chapter 6.2.2.9 for details).

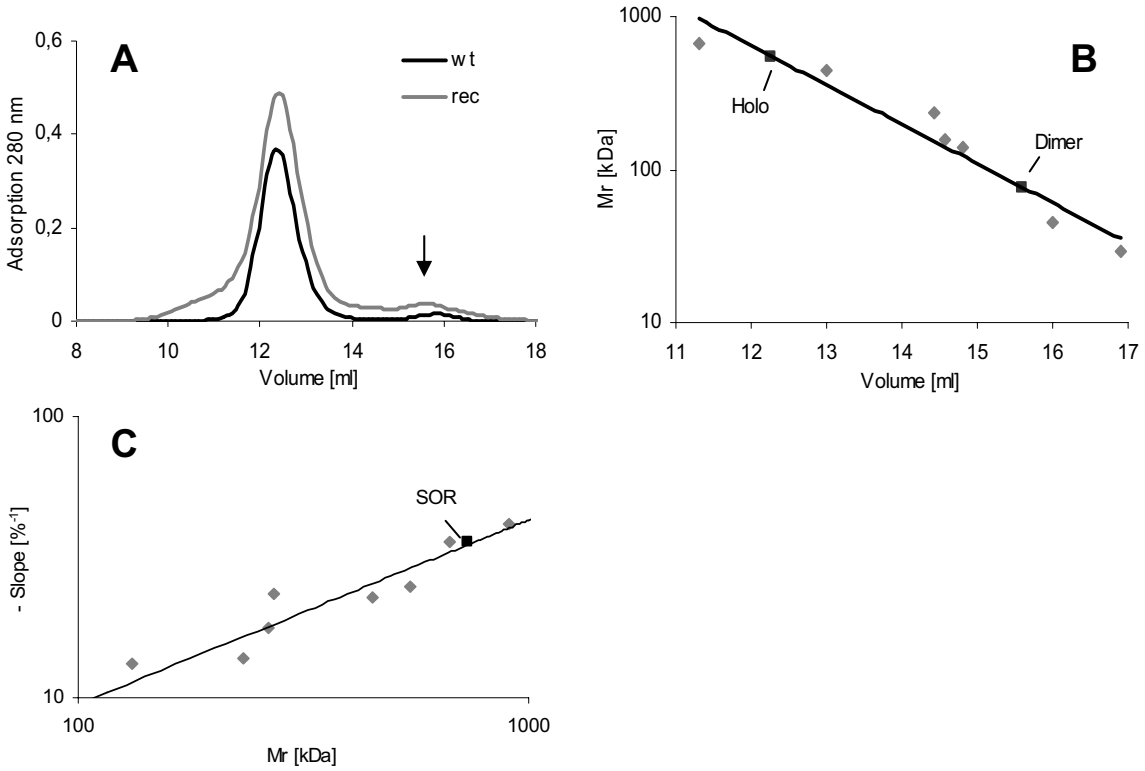


Figure 2.6: (A) Gel permeation chromatograms of a wild type and a recombinant soluble SOR preparation run on a Superose 6 column. Arrow, SOR dimer peaks. **(B)** Calibration plot of the Superose 6 column. Diamonds, Mr of proteins used for calibration plotted against their elution volumes. Rectangles, SOR holoenzyme and dimer. **(C)** Ferguson calibration plot used for molecular mass determination of SOR by BN PAGE. Diamonds, negative slopes of the migration behaviour of calibration proteins on gels with different acrylamide concentrations plotted against their molecular masses (see chapter 6.3.9 for details). Rectangle, SOR.

2.2.6. IRON QUANTIFICATION

The iron content of wild type and recombinant SOR preparations was investigated to determine the amount of iron incorporated into each monomer. A colorimetric method based on the reaction of ferrous iron with the dye 2,4,6-tripyridyl-1,3,5-triazine (TPTZ) was applied (Fischer & Price, 1964; see chapter 6.3.6 for details).

The determination of the iron content of the two differently purified wild type SOR preparations resulted in a ratio of 1.0 Fe atom per SOR monomer for both preparations. For the recombinant refolded SOR preparation a ratio of 1.4 Fe per monomer was calculated. The recombinant soluble SOR preparations had varying iron contents, ranging from 0.5 to 2.2 Fe per monomer (see for comparison chapter 4.2.3.2).

2.2.7. EPR SPECTROSCOPY

The UV/Visible spectra of SOR did not show any feature besides the 280 nm protein absorption in the reduced or oxidized state (not shown). Electron paramagnetic resonance (EPR) spectra of wild type SOR in the "as prepared" state performed at liquid helium temperature showed an isotropic resonance at g = 4.3, consistent with the presence of a mononuclear non-heme iron center in the ferric high-spin state and a rhombicity of 0.33 (E/D; Fig. 2.7). Both recombinant soluble and refolded proteins showed the same type of signal, although with a higher intensity. Adequate controls showed that the observed signal was from the enzyme and not from the buffers used.

The intensity of the g = 4.3 signal decreased upon incubation with elemental sulfur at 80°C up to 20 minutes. The signal disappearance was not reversible after the SOR had been washed with oxidised buffer following the reaction. This observation suggests that an intermediate may remain bound to the metal reaction site upon catalysis. The midpoint reduction potential of the mononuclear Fe center was determined in an EPR-monitored redox titration experiment to be E'_o = –268 mV (Fig. 2.7, inset).

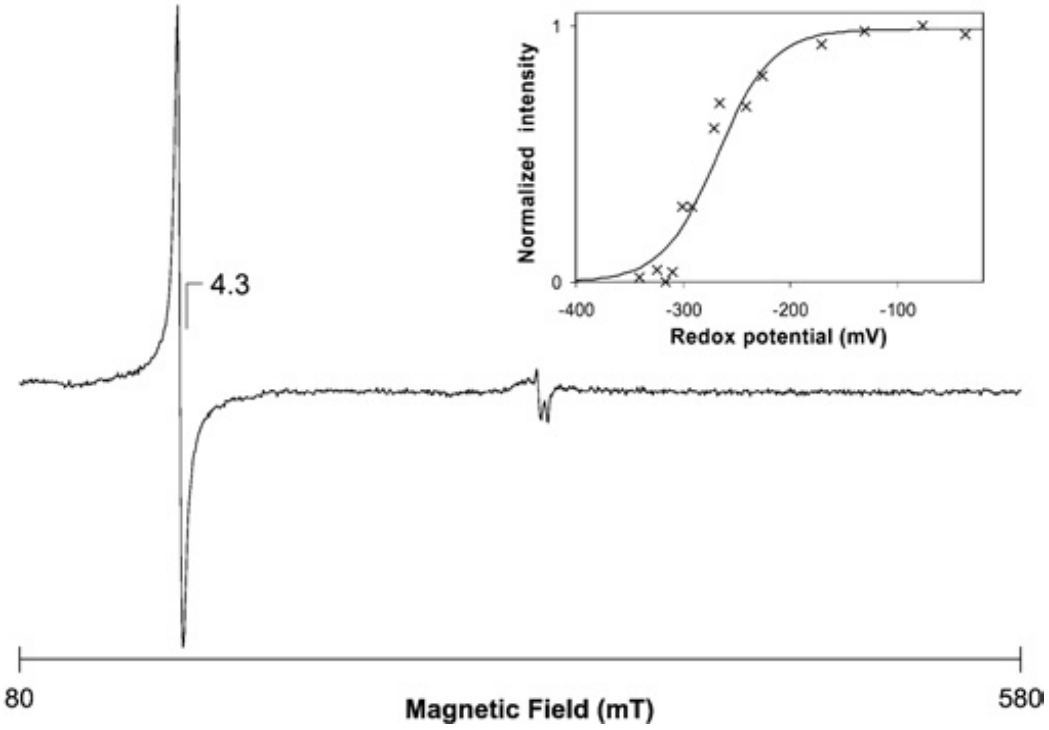


Figure 2.7: EPR spectrum of wild type *A. ambivalens* SOR. Temperature, 4.6 K; protein concentration, 226 µM; microwave power – 2.4 mW; modulation amplitude – 1 mT. **Inset:** Redox titration of the SOR mononuclear iron center. The normalised intensities of the g = 4.3 resonance are plotted against the reduction potential. The line represents a Nernst equation for a one electron step transition, with an E_0' = -268 mV.

2.2.8. CATALYTIC PROPERTIES OF SOR

The apparent K_M values for sulfur, determined with the oxygenase and the reductase reaction separately at 65 ˚C were 23 mM and 13 mM, respectively, calculated on the basis of S^0, not S_8 (Fig.2.8). The k_{cat} values were 2.2 s^{-1} and 0.1 s^{-1} (oxygenase and reductase, respectively).

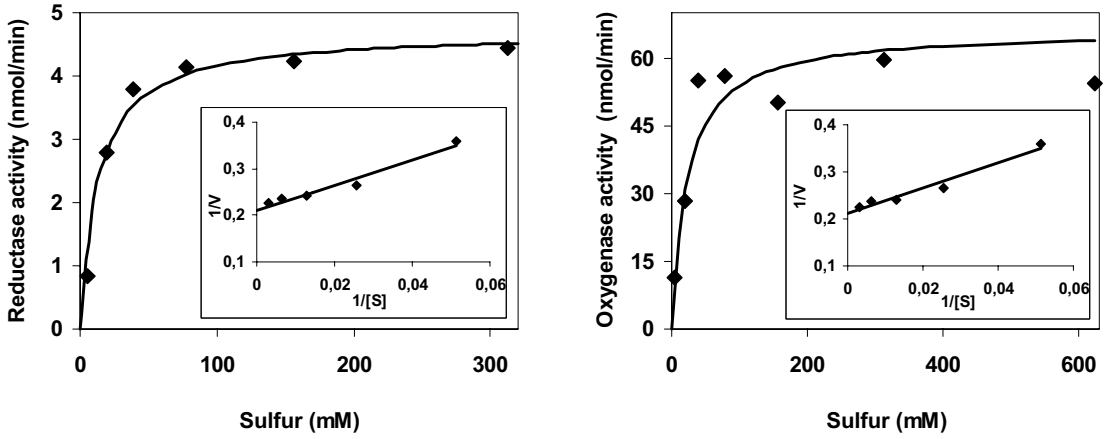


Figure 2.8: K_M determination of SOR. **(A)** Michaelis-Menten plot of reductase partial reaction. The line represents a Michaelis-Menten equation for a K_M of 13 mM. Inset: lineweaver-Burk plot used for V_{max} determination. **(B)** Michaelis-Menten plot of oxygenase partial reaction. The line represents a Michaelis-Menten equation for a K_M of 23 mM. Inset: Lineweaver-Burk plot used for V_{max} determination.

Several substrates and potential inhibitors were tested for their effects on SOR activity (Tab. 2.1). The SOR was not active with tetrathionate. Metal chelators like EDTA and sodium citrate were not inhibitory even in concentrations as high as 100 mM in the case of EDTA. In contrast, EDTA showed a slight stimulatory effect (Tab. 2.1), which was also seen in enzymatic assays with partially denatured SOR (not shown).

It was tried to determine whether the positive effect of EDTA was due to the binding of potentially inhibitory divalent cations. Ca^{2+} and Mg^{2+} had no effect on the enzyme activity (Tab. 2.1) whereas Zn^{2+} showed a strong inhibition as observed earlier (Kletzin, 1989). The effect of Zn^{2+} was abolished when EDTA was also added to the assay mixture. To determine whether inhibitory traces of transition metals in the substrate sulfur caused the stimulatory effect of EDTA, sulfur was washed with EDTA-containing buffer. The incubation of washed sulfur with additional EDTA in the assay buffer resulted in the same increase in activity as determined with the unwashed sulfur (Tab. 2.1).

The addition of low amounts of Fe^{2+} or Fe^{3+} did neither inhibit, nor increase the activity in low concentrations, but interfered with the colorimetric assays, thus preventing the use of higher concentrations.

Table 2.1: Effects of various substrates and reagents on SOR activity.

Substrate	Reagent	Reagent conc. (mM)	% activity oxygenase	% activity reductase
Sulfur (S_8)	-	78	100	100
	EDTA	1	127	143
		10	150	141
		50	160	108*
		100	134	13*
Sulfur (S_8), EDTA-washed	-	78	104	94
	EDTA	10	147	124
	Zn^{2+}	0.5	5	16
	EDTA/Zn^{2+}	1/0.5	114	86
	Fe^{III} Citrate	0.1	102	19*
		0.01	96	1*
	$[(NH_4)_2Fe^{II}(SO_4)_2]$	0.1	105	150
		0.01	110	90
Sulfur (S_8)	Ca^{2+}	1	82	101
	Mg^{2+}	1	110	92
	Na_3citrate	1	103	85
K_2-Tetrathionate	-	10	0	0

* interference with the colorimetric assays

2.2.9. CONFORMATIONAL STABILITY OF SOR

The conformational stability of wild type *A. ambivalens* SOR was investigated by induced unfolding at neutral pH using the denaturants guanidinium hydrochloride (GuHCl) and urea. The experiment was done by Claudio Gomes and Sonia Leal (Instituto de Tecnologia Química e Biológica (ITQB), Oeiras, Portugal). Protein unfolding was monitored following the emission fluorescence spectrum of tryptophane residues, which were used as intrinsic probes for protein stability. SOR was resistant to denaturation until concentrations of 2 M GuHCl and 4 M urea (Fig. 2.9). The reversible denaturation curves were found to have midpoints at 3.4 M GuHCl and 5.9 M urea when analyzed using a two-state model for the folding/unfolding reaction. The protein apparent folding free energy in water was determined to be ~5 kcal mol^{-1}, at pH 7 (Fig. 2.9, see legend for details). This low value is explained by the fact that although the native conformation is essential for activity, its stability with respect to the unfolded state is very low, typically ranging from 5-15 kcal/mol (Shirley, 1995). Dialysis of the unfolded protein restored the initial emission spectra, indicating that the observed chemical denaturants induced unfolding transition is

reversible. This was further corroborated by the fact that active SOR could be obtained from inclusion bodies dissolved in 8 M urea (see chapter 2.2.3).

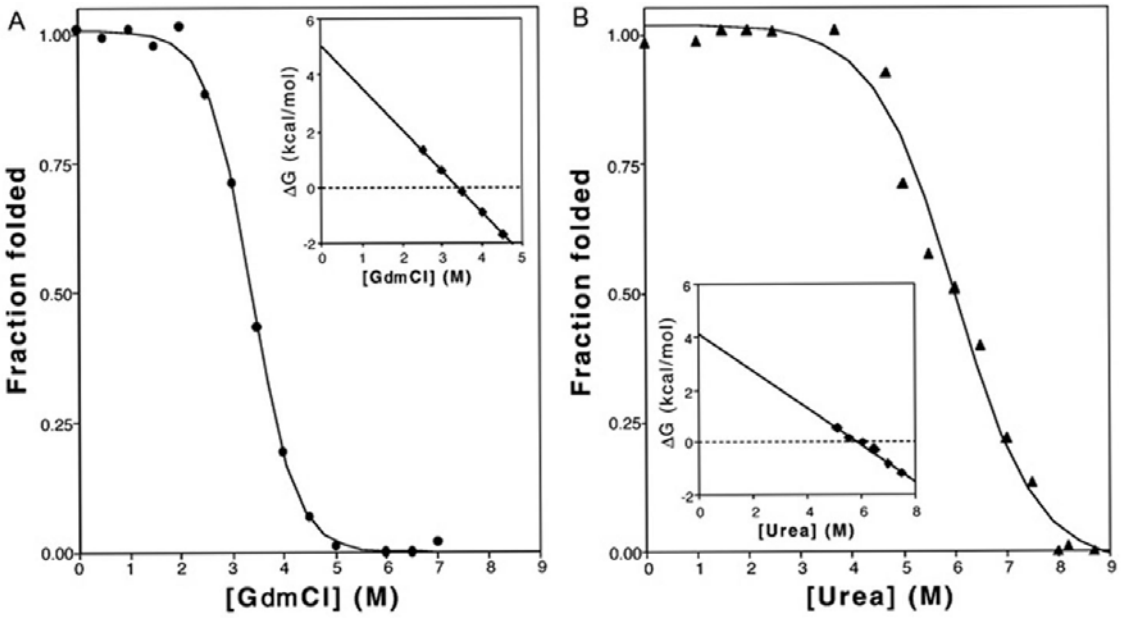


Figure 2.9: Guanidine hydrochloride (left) and urea (right) denaturation curves. The solid lines are a least squares analysis fit of the experimental data to a two-state Folded↔Unfolded model yielding the following parameters: $\Delta G(H_2O)$ = 5.04 kcal/mol, *m*= 1.49 kcal/mol/M and $[GuHCl]_{½}$=3.38 M (for GuHCl, panel A); $\Delta G(H_2O)$ = 4.13 kcal/mol, *m*= 0.7 kcal/mol/M and $[urea]_{½}$=5.89 M (for urea, panel B). It should be noted that *m* depends on the amount of polypeptide chain that is exposed to solvent upon unfolding, and thus varies with denaturants (Shirley, 1995). The insets are plots of the variation of ΔG_U in the transition region as a function of denaturant concentration, according to the $\Delta G_U = \Delta G(H_2O)$ - *m*.[denaturation] equation. The figures are a courtesy of C.M. Gomes and S.S. Leal (ITQB, Oeiras, Portugal).

In a different experiment recombinant SOR preparations were extensively dialysed against different urea concentrations and analysed by gel permeation chromatography. An additional peak corresponding to a SOR dimer appeared in preparations containing 2 M urea or above, whereas the holoenzyme peak decreased with increasing urea concentration (Fig. 2.10A). A peak corresponding to 35 kDa or the SOR monomer appeared in preparations dialysed against 6 M urea or above whereas the dimer and holoenzyme peaks diminished. Comparative analysis of the peak fractions by SDS-PAGE under reducing and non-reducing conditions showed only the 36 kDa band; the absence of a dimer band indicated that inter-subunit disulfide bridges were not present (not shown).

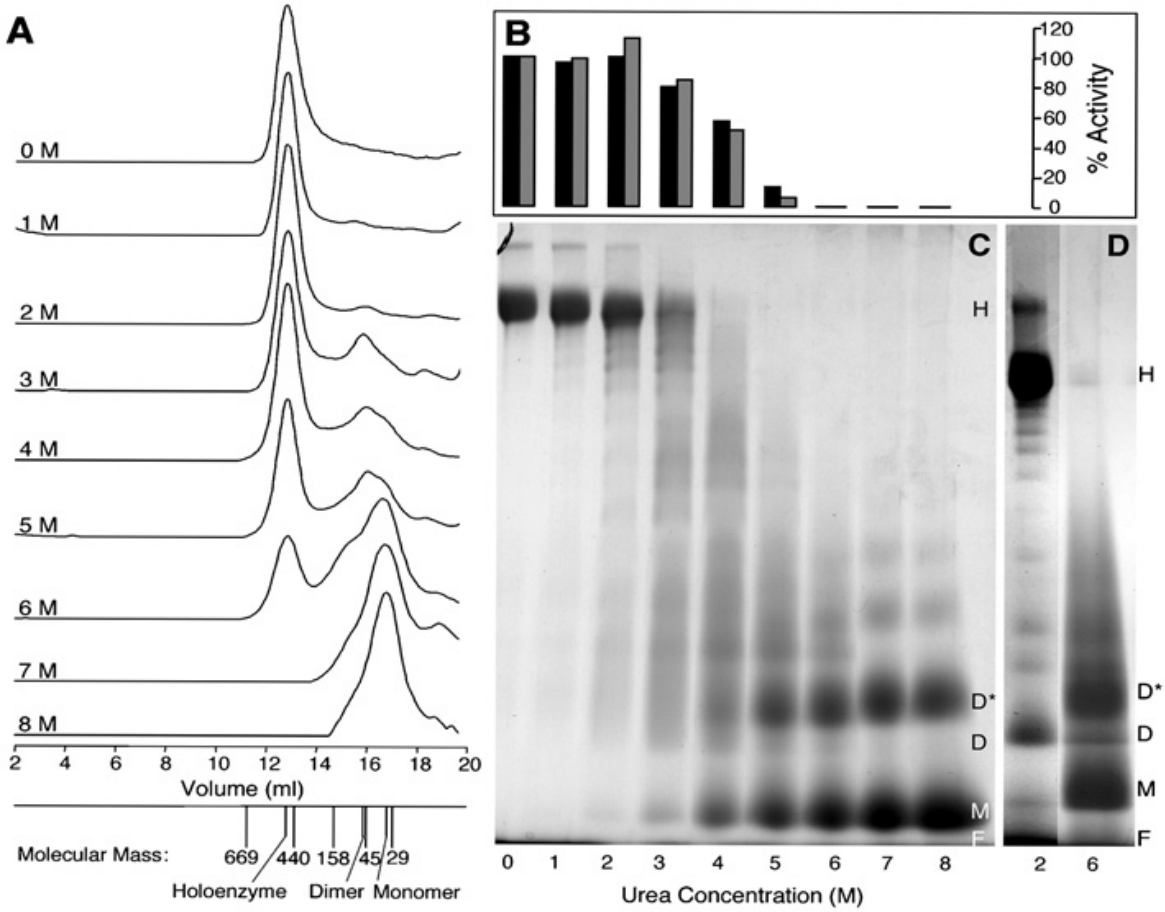


Figure 2.10: (A): Merged gel permeation chromatograms of recombinant soluble SOR dialysed against different urea concentrations. The protein peaks were monitored by the absorbance at 280 nm. The elution volumes of thyroglobulin, ferritin, aldolase, chicken albumin and carbonic anhydrase are given with their molecular masses together with the holoenzyme, dimer and monomer peaks. **(B)** Relative specific activities of the SOR fractions after dialysis against different urea concentrations as indicated below the gel (C). The oxygenase (black) and reductase activities (gray) of the 0 M urea control were each set to 100 %. **(C)** BN gradient gel (4-15 %) of SOR preparations dialysed against different urea concentrations restained with colloidal Coomassie blue after electrophoresis. **(D)** Ditto showing a ladder of oligomers of the partially denatured enzyme; H, holoenzyme; D* dimer denatured; D, dimer; M, monomer; F, migration front.

Identically treated samples separated with non-denaturing BN PAGE showed that the holoenzyme band dominated in samples dialysed against 0-3 M urea (Fig. 2.10C and 2.10D). Intensely stained bands corresponding to monomer and dimer sizes and, in less intensity to higher oligomers were seen in samples dialysed against 6 M urea and above. In the lanes containing the samples incubated with 2 - 4 M urea a faint ladder of partially denatured protein was observed (Fig. 2.10C, D). A dominant dimer band appeared in these lanes, which exhibited a different migration behaviour compared to those of the fully denatured samples (Fig. 2.10D). In parallel, the samples were analyzed for activity. SOR activity was observed up to 5 M urea but not above (Fig. 2.10B).

2.2.10. ELECTRON MICROSCOPY

In transmission electron micrographs of negatively stained wild type SOR preparations globular particles of 15.5 nm in diameter were visible (Fig. 2.11A and Kletzin, 1989). The inner core appeared densely stained. Electron micrographs of both recombinant protein preparations gave essentially the same results (Fig. 2.11B). Single particle averaging of individual wild type particles revealed two distinct subpopulations occurring in similar amounts. One subpopulation showed a 4-fold (Fig. 2.11C), the other a 2- or 3-fold symmetry axis (Fig. 2.11D) indicating that the holoenzyme is a highly symmetrical particle.

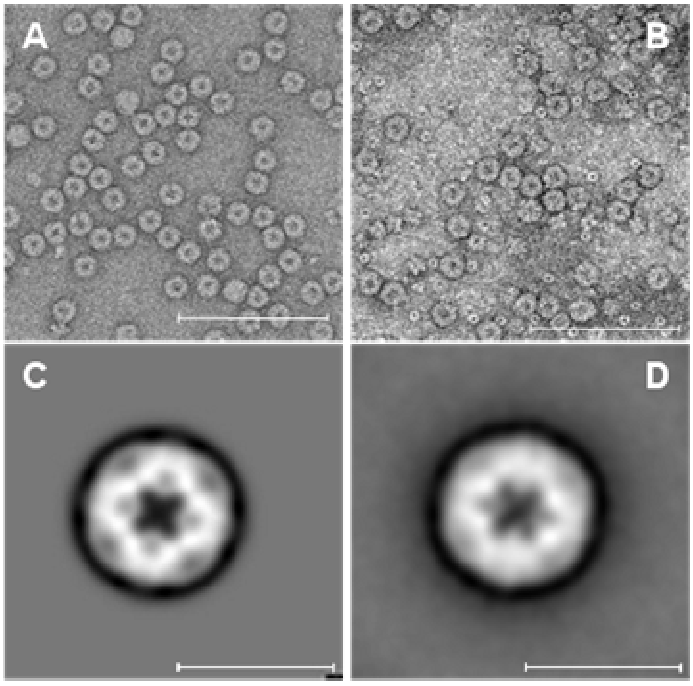


Figure 2.11: Electron micrographs of negatively stained SOR, and image reconstruction. **(A), (B),** Micrographs of wild type and recombinant SOR, respectively; the bar represents 100 nm; **(C)**, **(D)**, results of single particle averaging of SOR particles with 4-fold and 2- or 3-fold symmetry axis, respectively; the bar represents 15 nm. The figures are a courtesy of R. Rachel (Archaeenzentrum, Regensburg, Germany).

2.2.11. SECONDARY STRUCTURE ANALYSIS

The secondary structure of wild type and recombinant SOR was analyzed by circular dichroism (CD) spectroscopy. The CD spectra of the different preparations were rather similar, with minima around 220 and 207 nm and a maximum around 195 nm and minor differences in the molar ellipticity at 187-197 nm and at 220 nm (Fig. 2.12). Secondary structure quantification of the wild type SOR spectrum resulted in 23.8 (+/- 1.5) % α -helix and 23.4 (+/- 0.8) % β-sheet content. The recombinant and refolded preparations showed an α-helical content of 25.2 (+/- 0.6) % and 25.7 (+/- 0.9) % and a β -sheet content of

26.2 (+/- 0.5) % and 24.2 (+/- 0.2) %, respectively. The average length of the helices predicted was 9 amino acids. Secondary structure prediction programs gave 29 % α-helical and 27 % β-sheet content (Fig. 2.13), which is in good agreement with those calculated from the CD spectra.

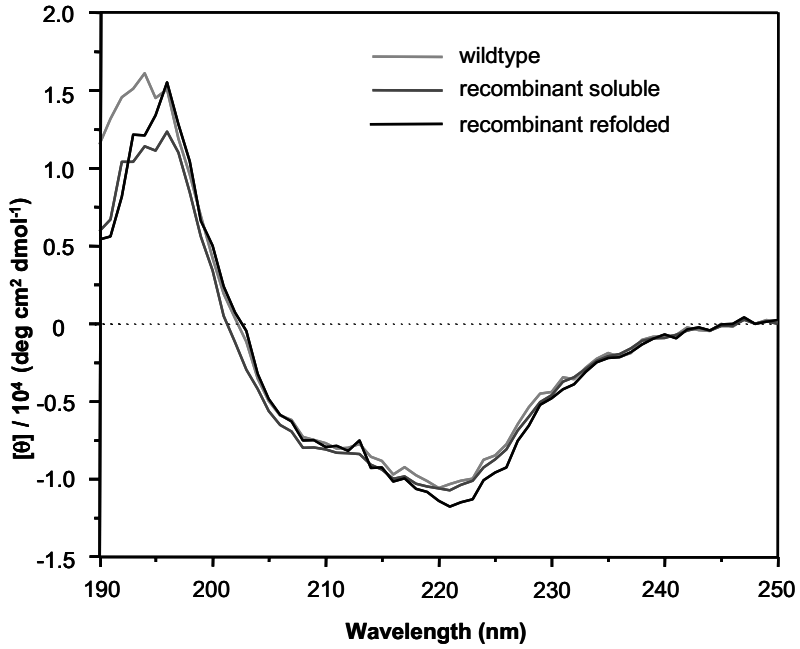


Figure 2.12: Circular dichroism spectra of wild type, recombinant soluble and recombinant refolded SOR. All the CD spectra were displayed relative to the buffers used. The y-axis represents the mean residue molar ellipticity obtained using the protein concentration, the path length and the number of protein residues for calculation (wild type: 308; recombinant: 318).

2.2.12. COMPARISON OF SOR AMINO ACID SEQUENCES

Homologous *sor* genes were found in BLAST searches in the genomes of the archaea *Sulfolobus tokodaii, Ferroplasma acidarmanus, Picrophilus torridus* and the hyperthermophilic bacterium *Aquifex aeolicus*, in addition, the *sor* gene from *Acidianus tengchongensis* had already been described (He *et al.*, 2000). The pairwise similarities of the deduced amino acid sequences compared with the *A. ambivalens* enzyme ranged from 39 % (*Aquifex aeolicus*) to 88 % (*Acidianus* S5; Fig. 2.13). Extensive database searches using PSIBLAST and other search tools did not provide any further motifs or significant similarities to other proteins. No similar gene was found when searching the genome sequence of the sulfur and iron oxidizing mesophilic bacterium *Acidithiobacillus ferrooxidans*. Three cysteines were conserved in all sequences, in addition to a putative metal binding sequence motif H^{86}-E-E/D-M-H^{90} (Fig. 2.13).

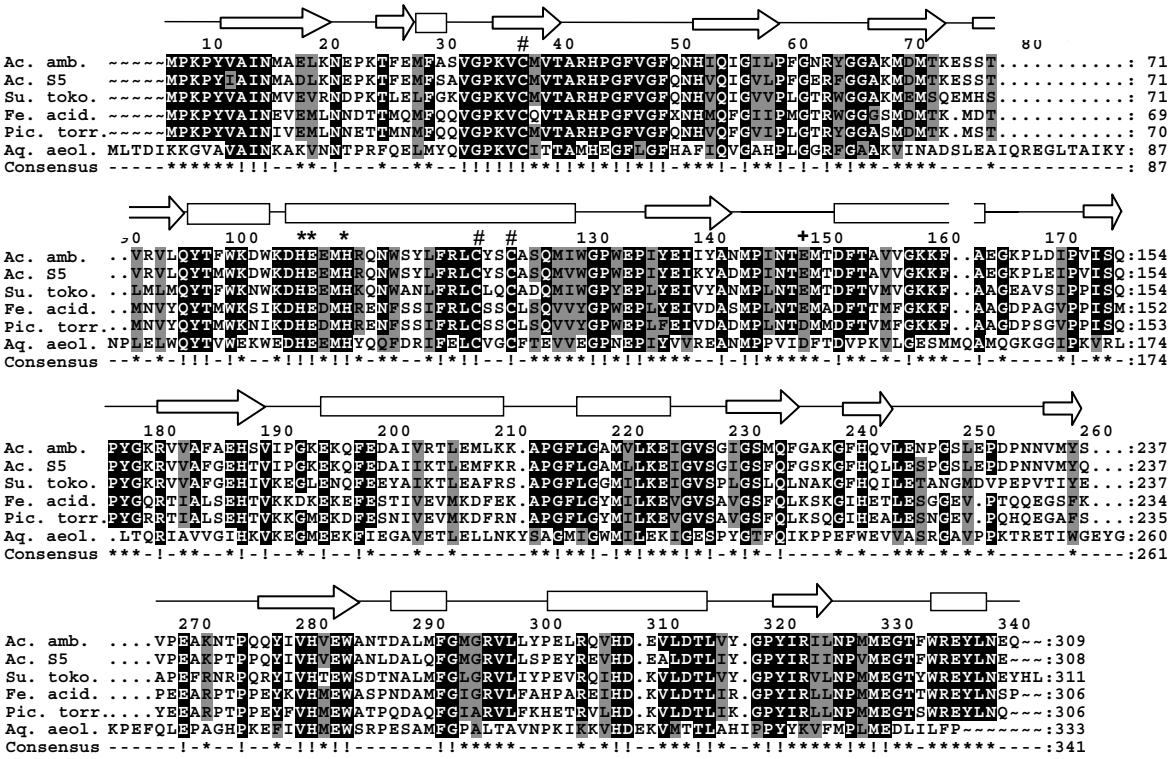


Figure 2.13: Multiple alignment of SOR sequences. #, conserved cysteine residues, * H-E-x-x-H motif known from proteins with non-heme iron centers; rectangular boxes, predicted alpha-helical regions; open arrows, beta sheet regions for the *A. ambivalens* SOR. Species names: *Ac.amb*., *Acidianus ambivalens* (Accession number X56616; *Ac.* S5., *Acidianus tengchongensis* (AF267286); *Su.toko*., *Sulfolobus tokodaii* (NC_003106); *Pic. torr., Picrophilus torridus* (NC_005877.1); *Fe.acid*., *Ferroplasma acidarmanus* (Faci1674 encoded by NZ_AABC02000001); *Aq.aeol*., *Aquifex aeolicus* (NC_000918). Immediately upstream of the *Ferroplasma acidarmanus* Faci1674 ORF an in-frame stop codon was found (X, Pos. 49). It was treated as an unknown residue, because of the similarity of the deduced amino acid sequences of the upstream region to the other SOR sequences. It is not known whether this represents a pseudo gene or whether *Ferroplasma acidarmanus* produces active SOR under suitable conditions.

2.3. DISCUSSION

The biochemical and functional properties of the recombinant and the wild type *A. ambivalens* SOR have been characterized here focusing on the enzymatic activity, on the first description of a redox-active center, and on the secondary and quaternary structure of the enzyme.

2.3.1. RECOMBINANT AND WILD TYPE SOR HAVE SIMILAR PROPERTIES

The specific activities of several preparations of the recombinant soluble and refolded SOR and of the wild type SOR were comparable, ranging in the same order of magnitude (see below, Tab. 2.2). The application of two different purification protocols for the recombinant soluble SOR showed that the specific activity of the purified enzyme depends largely on a rapid purification procedure, which also explains the lower specific activities of the wild type enzyme purified by time-consuming protocols.

Wild type and recombinant SOR were analyzed by electron microscopy, gel permeation chromatography, CD and EPR spectroscopy, and for their iron content (Tab. 2.2). The results showed that the recombinant enzyme possessed similar properties to the wild type, indicating that the C-terminal 10 amino acid Strep-tag did not impair the function of the enzyme. Furthermore, it showed that the refolding procedure resulted in an enzyme with a near-native state.

Table 2.2: Comparison of recombinant and wild type SOR properties.

	Recombinant soluble	**Recombinant refolded**	**Wild type**
Specific activity [U/mg]			
Sulfur oxygenase	2.8 - 11.5	6.0	0.8 - 3.0
Sulfur reductase	0.7 - 4.2	1.4	0.2 - 0.6
M_r [Da] (Gel permeation)	550,000	550,000	550,000
Particle size (EM)	15.5 nm	15.5 nm	15.5 nm
Secondary structure			
α-helix	25 %	26 %	24 %
β-sheet	26 %	24 %	23 %
Fe / monomer	0.5 -2.2	1.4	1.0

The specific activity of the wild type sulfur oxygenase from "*Sulfolobus brierleyi*" was lower (0.66 U/mg; Emmel *et al.*, 1986), whereas for the recombinant SOR from *Acidianus tengchongensis* different specific activities were reported (187 U/mg; Sun *et al.*, 2003; and 4.9 U/mg; Shen *et al.*, 2005). The reason for the differences in activity is not known.

The apparent K_M values for sulfur determined with the oxygenase and the reductase reaction separately at 65 ˚C were 23 mM and 13 mM, respectively (calculated with S^0, not

S_8). The lower than optimal temperatures were necessary for this experiment to avoid sulfur precipitation in the course of the incubation. The k_{cat} values were 2.2 s^{-1} and 0.1 s^{-1}, respectively. The values suggest a low substrate affinity and enzyme velocity. A similar K_M of 50 mM had been reported for the oxygenase reaction of the "*Sulfolobus brierleyi*" enzyme (Emmel *et al.*, 1986). A possible explanation for the unfavourable values lies in the poor solubility of sulfur (Boulegue, 1978) and therefore, low substrate availability despite the large excess in the assay thus preventing the enzyme from reaching maximal velocity.

2.3.2. A MONONUCLEAR NON-HEME IRON CENTER IS AN ESSENTIAL COFACTOR

The results of the refolding experiments with different iron concentrations showed that its presence was essential for the catalytic activity of the refolded SOR. Slightly higher iron/subunit ratios were observed in the recombinant enzyme preparations as compared to the wild type however, the strong EPR signal of the oxidized, as prepared SOR showed that iron is most likely present in a mononuclear center in each subunit (Fig. 2.7).

The lack of any absorbance of the enzyme in the UV/Visible spectrum, besides the 280 nm absorption (Kletzin, 1989), points to a histidine and/or carboxylate but not cysteine ligation of the iron as observed in many oxygenases with mononuclear non-heme iron centers (e.g. Lange & Que, 1998). These conclusions were later confirmed by the crystal structure (chapter 3).

The reduction potential of the iron is the lowest observed for mononuclear iron centers so far (E_m = -268 mV). Other proteins known to harbor mononuclear non-heme iron centers show clearly positive midpoint potentials, like lipoxygenase (E_m = +600 mV; Nelson, 1988), phenylalanine hydroxylase (E_m = +207 mV, Hagedoorn et al, 2001), neelaredoxins (E_m = +200-300 mV, Abreu *et al.*, 2002) and Fe-superoxide dismutase (E_m = +220 mV, Vance & Miller, 2001). This low reduction potential most probably originates from a different protein environment around the iron center. Intriguingly, the reduction potential is low enough to explain the observed sulfur reductase activity of the enzyme (E'_o H_2S/S^0 = −270 mV; Thauer *et al.*, 1977).

A tight binding of the iron is suggested by the observation that EDTA was not inhibitory at all; in contrast, it stimulated the enzyme activity even in assays with partially denatured enzyme (not shown). The removal of potentially inhibitory traces of transition metals bound to the SOR in the “as isolated” state could be a reason for this effect. This explanation is supported by the observations that the inhibitory effect of Zn^{2+} could be reversed by addition of EDTA to the reaction mixture (Table 2.1).

2.3.3. SOR SEQUENCE ANALYSIS AND ACTIVE SITE PREDICTIONS

Four of the six SOR sequences had been grouped in a PFAM domain (PF_07682; http://www.sanger.ac.uk/Software/Pfam). No overlap with other domains was detected. Similarly, our analyses did not provide evidence that the SOR proteins are phylogenetically related to any other enzyme family.

A multiple alignment of the six SOR sequences showed a H-X_3-H site in the sequence H^{86}-E-E/D-M-H^{90} as a putative metal binding sequence motif (Fig. 2.13), which is similar to the one found in aromatic acid hydroxylases and Rieske dioxygenases containing a mononuclear non-heme iron center (Costas *et al.*, 2004). In these enzymes, the histidines are part of the structural motif of the 2-His-1-carboxylate facial triad, which anchors the iron in the active site. Secondary structure prediction identified a putative extensive α-helix including the H-X_3-H box in the SOR sequences (Fig. 2.13). A helix projection suggested that both histidines are facing the same side as observed in aromatic acid hydroxylases and Rieske dioxygenases (Fig. 2.14) and that the SOR may adopt a similar structural motif for the iron binding. The crystal structure confirmed that both histidines indeed are ligands of the iron center (chapter 3).

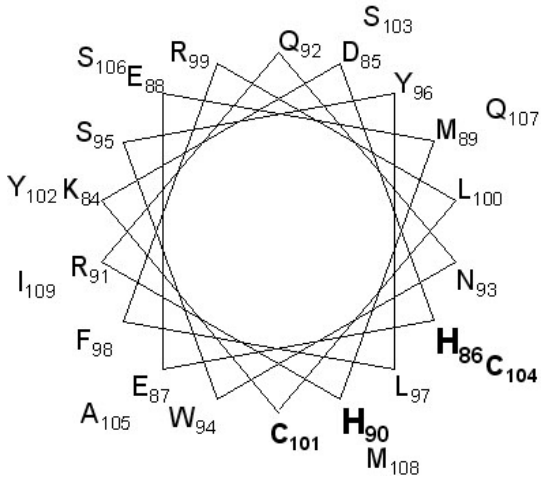


Figure 2.14: Helical wheel representation of the putative 26 amino acids long α-helix predicted in the *A. ambivalens* SOR. His86 and His90 are located on the same side of the helix, as are Cys101 and Cys104.

Each SOR has three conserved cysteine residues in common, two of which are arranged as a C-X-X-C-motif (Fig. 2.13) residing on the same α-helix as the putative iron-binding histidines (Fig. 2.14). The activity of purified SOR was inhibited by thiol-binding reagents such as N-ethylmaleimide and Zn^{2+} ions (Kletzin, 1989 and this work). In consequence, the cysteines or some of them must be essential, although it is improbable that they might coordinate the metal site as explained above. It is more likely that they play a role in sulfur binding and activation either as enzyme-bound polysulfide in a similar way as it has been

described for reduced glutathione (Suzuki *et al.*, 1965; Silver *et al.*, 1968) and for the periplasmatic bacterial sulfur oxygenase (Rohwerder & Sand, 2003) or with a disulfide redox chemistry as postulated for the sulfide:quinone oxidoreductase (Griesbeck *et al.*, 2002).

2.3.4. A SHELL-LIKE QUATERNARY STRUCTURE WITH A CENTRAL HOLLOW CAVITY

The well-contrasted dot in the middle of the spherical SOR particles observed in EM images (Fig. 2.11) points to a hollow interior filled with uranyl acetate used for negative staining. Other classes of proteins are known to have this structural feature, most notably the ferritins, bacterioferritins and the ferritin-like DNA-binding Dps proteins (Ratnayake *et al.*, 2000; Kim *et al.*, 2001; Macedo *et al.*, 2003; Tonello *et al.*, 1999). Electron micrographs of these particles resemble SOR particles, thus strengthening the hypothesis that the SOR is a shell-like multimeric protein complex with a hollow core. The *Desulfovibrio desulfuricans* bacterioferritin is slightly smaller in diameter than the SOR (13.0 nm vs. 15.5 nm, internal hollow core: 8.5 nm). It consists of 24 subunits of a 179 amino acid protein arranged in dimers and possess a 432 symmetry (Mr 500 kDa; Macedo *et al.*, 2003), whereas Dps proteins are homo-dodecamers with external and internal diameters of 9.0 and 4.5 nm, respectively, and a 23 symmetry (Roy *et al.*, 2003).
It was so far unknown of how many subunits the SOR holoenzyme is actually composed. The molecular mass of the SOR determined by gel permeation chromatography would point to a 15-mer (550 kDa, 308 amino acid residues without the N-terminal methionine; Kletzin, 1989, 1992), whereas the results of the BN PAGE point to a higher molecular mass (732 kDa), equivalent to 20 subunits. A structural model explaining either number of subunits would require the presence of a 5-fold symmetry axis, which was not seen in electron micrographs. In contrast, the existence of a 4-fold and a 3- or 2-fold symmetry axis (Fig. 2.11) as well as the diameter of the SOR particles point to a quaternary structure being an icosatetramer with the same overall 432 symmetry as the ferritins, although this conclusion is in contrast to the determined apparent molecular masses. The icosatetrameric quaternary structure was later on confirmed by the results from x-ray crystallography (chapter 3).
The results from the CD spectra and from secondary structure prediction programs both suggested that the SOR has an α-helical and β-sheet content of around 25 % each. Thus, the SOR exhibits a different pattern of secondary structure elements as the ferritins, which are almost completely made of a four-helix bundle (Macedo *et al.*, 2003).

2.3.5. A DIMER AS BUILDING BLOCK OF THE HOLOENZYME

The tryptophane fluorescence-followed experiments showed that the SOR is relatively resistant to unfolding up to urea concentrations of 4 M whereas at urea concentrations above 7 M the fluorescence reached its minimum indicating a complete denaturation of the enzyme (Fig. 2.9). Monitoring by gel permeation chromatography showed that a predominant dimer peak was observed until 5 M urea besides the holoenzyme peak (Fig. 2.10A), whereas the monomer peak appeared at higher urea concentrations. The tryptophane fluorescence-followed denaturation curve and the signal intensities of the holoenzyme, dimer and monomer peaks in the gel permeation experiment could be superpositioned without significant differences, when assuming that the dimer is non-denatured and the monomer is denatured. This finding points to a dimer being a stable intermediate during the assembly of the holoenzyme, which is supported by the results of the BN PAGE showing the dimer band being most prominent in the 2 M urea sample (Fig. 2.10D). The partially denatured enzyme formed a ladder of bands presumably showing the holoenzyme in various states of disintegration, whereas the fully denatured protein migrated and stained differently in the BN PAGE suggesting that the denatured monomers expose more of the hydrophobic residues and aggregate non-specifically (Fig. 2.10C, D). Interestingly, SOR activity was only observed as long as higher oligomeric states were present, indicating that a cooperation between subunits may be required for catalysis (Fig. 2.10B, D).

Taken the experiments together, the dimeric state as observed in the partially denatured samples in gel permeation and BN gels is properly folded, whereas the complete unfolding of SOR is in good correlation with the appearance of the monomeric state. Furthermore, the data indicate that the assembly of the holoenzyme proceeds via oligomerization of dimers as the building blocks (see also chapter 3.3.2).

The existence of a stable dimeric intermediate has also some implications on observations from the refolding screening, where SOR activity increased in a time-dependent manner when the protein solution was kept at 4°C after refolding (Fig. 2.3). The increase in activity could be explained by the assembly of dimers to form the catalytically active sphere, although no additional evidence supports this conclusion.

3. THE CRYSTAL STRUCTURE OF THE SOR

3.1. INTRODUCTION

High resolution information of the structure of proteins is essential for a deeper understanding of their modes of action and reaction mechanism. X-ray crystallography is an excellent technique for detailed structural studies, both for small or large protein structures, especially when high resolution data is available, although only limited information about their dynamic behavior is usually obtained. The method can be divided roughly into three distinct stages.

(1) Protein crystallization to produce a highly ordered array of many protein molecules in single crystals, which act as signal amplifiers for X-ray diffraction data collection.

(2) Structure solution, meaning the calculation of an interpretable electron density map from the diffraction data.

(3) Construction of a protein model within the electron density and a subsequent refinement and analysis of the model.

The crystallization and structure solution of the SOR by X-ray crystallography, including its analysis with respect to implications on its function, is reported in this chapter. As a wealth of data was obtained by this method, several questions could be addressed, focusing on important aspects of the catalytic function of the enzyme. Besides the tertiary and quaternary structure of the enzyme, the identification and analysis of its active site(s) was of interest. One focus was on the protein ligands and the geometry of the previously detected mononuclear non-heme iron center and the positions and orientations of the cysteine residues. By analyzing SOR structures complexed with inhibitors, substrates and products of the enzymatic reaction, functionally important residues were identified. The data allowed a first detailed insight into the structure-function relationship of the SOR.

3.2. RESULTS

The results from the different stages of X-ray crystallography are presented in this section. The three-letter code for the amino acids is used within the text. Furthermore, their numbering refers to the amino acids in the mature protein and not to the codons of the gene. They differ in numbering, because the N-terminal methionine is post-translationally cleaved off (Kletzin, 1992). The atomic coordinates of the SOR crystal structures are not yet deposited in the Protein Data Bank (PDB, Berman *et al.*, 2000), therefore they are provided on a supplementary CD-Rom (see section 3.4 for details).

3.2.1. CRYSTALLIZATION

Proteins crystallize by a slow, controlled precipitation from aqueous solutions under conditions that do not denature the protein. Ionic compounds or organic solvents are used as precipitants. Recombinant SOR, containing a C-terminal 10 amino acid Strep-tag, was purified by heat-treatment, anion exchange chromatography and ultra-centrifugation (chapter 2) and concentrated to 13.1 mg ml^{-1} in 50 mM Tris/HCl pH 7.5. Initial crystallization attempts were performed at 22°C using the Jena Bioscience Screens I, II, IV and V (144 conditions), but they did not deliver recognizable crystals (see for details chapter 6.7.1). Additional trials were performed at 22°C using the Hampton Research MembFac Screen (64 conditions). Conditions number 1, 12, 16 and 19 delivered mostly multiple, small crystals within 3-5 d. In order to control the number of crystal nucleations and their growing size, the same conditions were then tried at 4°C, but no crystal formation was observed. However, sizable crystals were obtained at 32°C within 2-5 d. Two optimized crystallization conditions were obtained by a refinement of the above mentioned conditions, one of which was 0.1 M sodium acetate pH 4.5 with 8% MPD (2-Methyl-2,4-pentanediol) as precipitant and 50 mM NaCl, and the other one was 0.1 M sodium citrate pH 5.5 and 0.3 M $MgSO_4$ as precipitant (Urich *et al.*, 2005a). Both conditions delivered colorless and polyhedral single crystals with maximal dimensions 0.7, 0.4 and 0.3 mm (Fig. 3.1). Cryoprotecting conditions appropriate to allow flash freezing in liquid nitrogen were obtained by soaking the crystals for 1 min in their crystallization solution complemented with glycerol 25% (v/v).

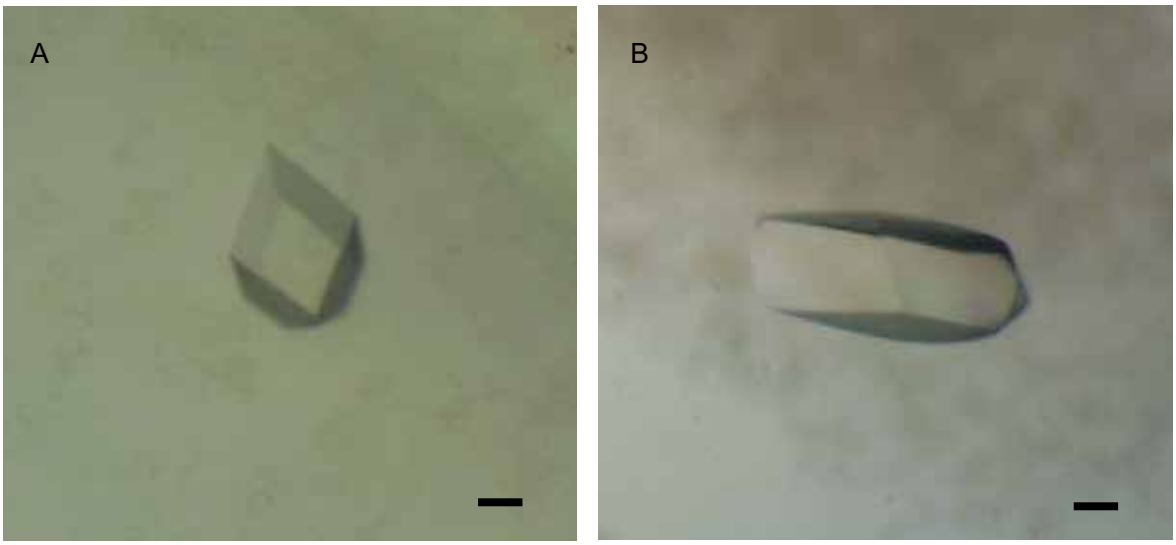


Figure 3.1: SOR crystals. (**A)** Grown in 0.1 M sodium acetate pH 4.6, 8% MPD and 50 mM NaCl; (**B)** grown in 0.1 M sodium citrate pH 5.5 and 0.3 M $MgSO_4$. The bar represents 0.1 mm.

3.2.2. CRYSTAL CHARACTERIZATION

The crystals from both conditions diffracted to 1.7 Å resolution (Urich *et al.*, 2005a). Their crystal lattice belonged to the tetragonal body-centered space group I4 (International Tables for Crystallography number: 79; Hahn, 2002) with unit cell dimensions a = b ≈162 Å and c ≈154 Å (Tab. 3.1). I4 is a doubly-primitive space group, containing two lattice points, one in the center of the unit cell and the other distributed over the eight cell vertices, each contributing with a 1/8th (Fig. 3.2). The two lattice points correspond to 2x4 asymmetric units (AU, the largest volume in the unit cell that possesses no symmetry elements) within the I4 unit cell, which can be juxtaposed over one-another using the space group symmetry operations.

Electron microscopic studies had shown that SOR homo-oligomers are highly symmetrical spherical particles with a diameter of ~15.5 nm, which is in the order of magnitude of the SOR´s unit cell dimensions (see chapter 2 and Urich *et al.*, 2004). It is therefore reasonable that each unit cell of SOR I4 crystals, with its similar dimensions of a = b and c, could contain two such spherical particles packed together in a distorted body centered cubic fashion.

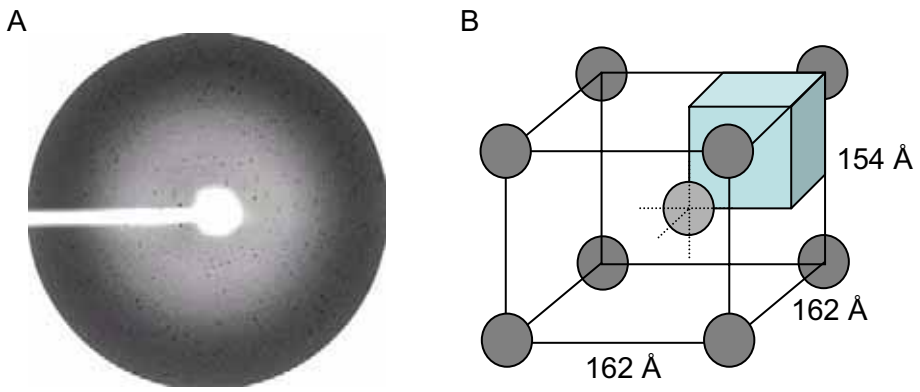


Figure 3.2: (A) Exemplified X-ray diffraction image derived from a SOR crystal. **(B)** The SOR crystal unit cell and its dimensions. One asymmetric unit is indicated as a box within and contains 1/8 of the unit cell.

Estimating the number of molecules in the AU is one of the first steps in a macromolecular structure determination. The calculation of the possible number of SOR monomers in the AU (Matthews, 1968) resulted in 5 to 8 monomers, which corresponds to a solvent fraction in the range of 59 % to 35 % (Kantardjieff & Rupp, 2003).

Crystals of oligomeric proteins often possess, besides the crystallographic symmetries, additional symmetry relations between the monomers (non-crystallographic symmetry, NCS), which can be detected without phase information by rotating the diffraction data set in distinct angles (self rotation Patterson function; Drenth, 1999). Although the asymmetric unit obviously contains multiple SOR monomers, its native Patterson space did not reveal any off-origin peak. The self-rotation Patterson function, on the other hand, showed significant peaks at 90°, 120° and 180° Kappa sections (Fig. 3.3) giving evidence for the existence of 4-, 3-, and 2-fold non-crystallographic rotation axes in a relative orientation characteristic for the 432 point group symmetry. The orders of the non-crystallographic rotation axes suggest 6 as the most probable number of SOR monomers in the asymmetric unit, which corresponds to an estimated V_M = 2.5 $Å^3$ Da^{-1} (V_M: Matthews coefficient; crystal volume per unit of protein molecular weight; Matthews, 1968) and a solvent content of 51 %. Taking into consideration the 4-fold crystallographic rotation axis, each SOR particle should then be composed of 24 monomers arranged in a highly symmetrical order with 432 point group symmetry.

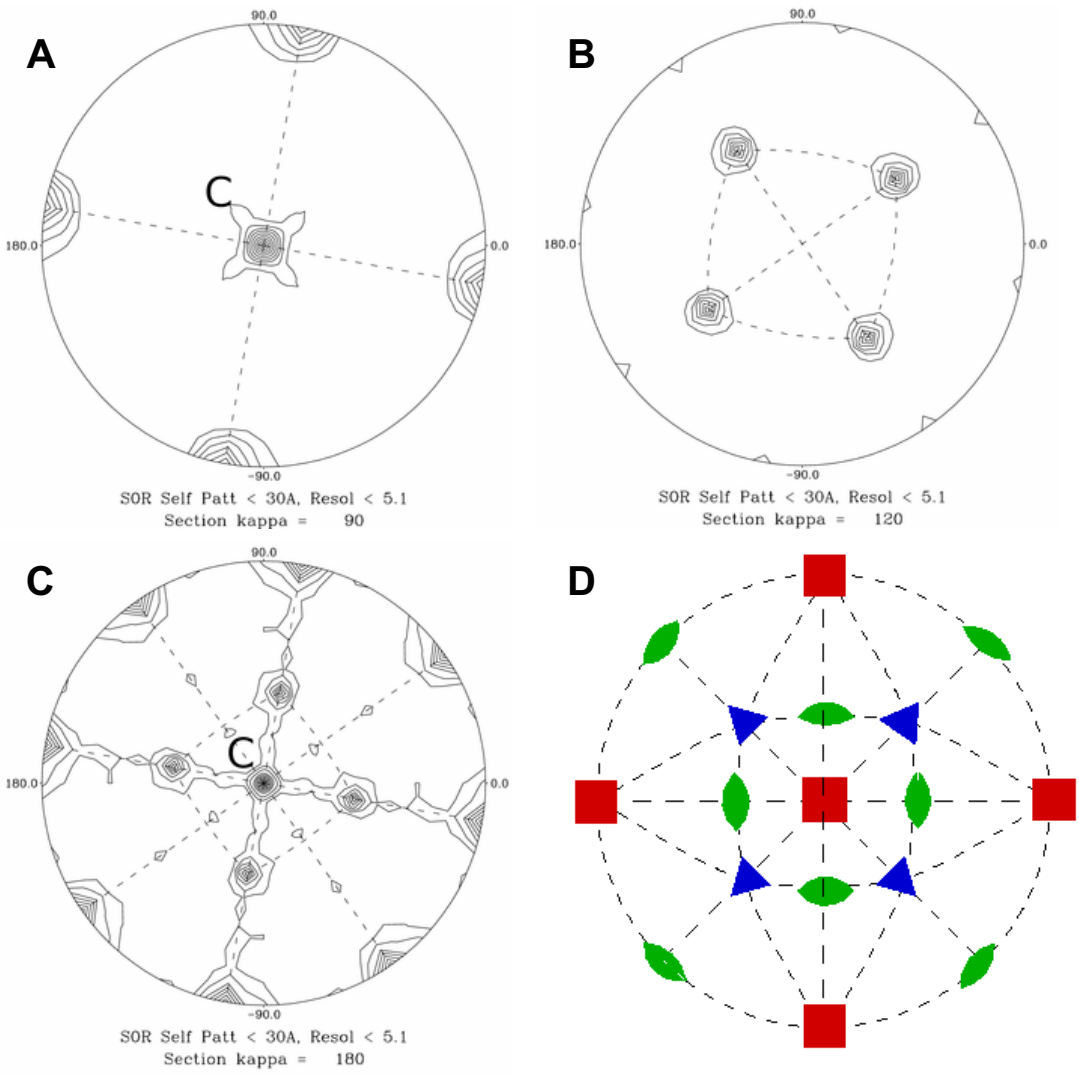


Figure 3.3: Self-rotation Patterson function plots of SOR, using Patterson vectors until 30 Å (resolution range 49.2 < d < 5.1 Å). The sections at **(A)** K = 90°, **(B)** K = 120° and **(C)** K = 180° are represented with maximum value normalized to 100 (origin peak) and contour levels drawn at 10-units intervals. The crystallographic rotation axes of space-group I4 are labeled with C, all other peaks represent non-crystallographic rotation axes within the asymmetric unit. **(D)** relative orientation of symmetry relations in the 432 point group, with 4-, 3- and 2-fold rotation axes shown as red squares, blue triangle and green ellipsoids, respectively (Figure taken from http://ruby.chemie.uni-freiburg.de/Vorlesung/ symmetrie_oktaeder.html).

Table 3.1: Diffraction data processing and model refinement statistics.

Data	SOR-Native	SOR-Hg^{2+}	SOR-Au^{3+}	Fe -edge
Source	DESY PX11	ESRF ID14-3	ESRF ID14-3	DESY BW7-A
Space group	I4 (79)	I4 (79)	I4 (79)	I4 (79)
Unit cell parameters (Å)	a = b =162.15 c = 154.92	a = b =161.88 c = 154.37	a = b =161.84 c = 153.90	a =b =161.58 c = 154.34
Wavelength (Å)	0.811	0.934	0.934	1.736
Nr. unique intensities	218,799	68,559	35,731	106,905
Redundancy	7.3	15.0	7.2	12.0
Resolution (outer shell) (Å)	49.2-1.70 (1.76-1.70)	48.8-2.5 (2.54-2.5)	38.1-3.1 (3.16-3.1)	37.93-2.15 (2.21-2.15)
Completeness (outer shell) (%)	99.2 (98.0)	99.9(99.2)	99.9(100.0)	99.9
R_{merge} (outer shell) (%) [a]	6.6 (60.7)	5.8	10.9	3.1
I/σ(I) (outer shell)	21.3 (1.4)	33.0 (4.31)	17.4 (2.6)	31.33 (3.76)
Refinement (49.2-1.70 Å)				
Refined structure	1842 residues (6 monomers, residues 1-307) + 6 Fe, + 477 water molecules			
R_{crys} (%) [b]	18.3			
R_{free} (%) [c]	21.0			
Average B-factors ($Å^2$)	31.3			
R.m.s. deviations from standard geometry				
Bonds (Å)	0.022			
Angles (°)	1.718			

[a] $R_{merge} = \sum | I_o - \langle I \rangle | / \sum I_o$, where $\langle I \rangle$ is the average of symmetry equivalent reflections and the summation extends over all observations I_o for all unique reflections.
[b] $R_{cryst} = \sum \| F_o | - | F_c \| / \sum | F_o |$, where F_o and F_c are the observed and calculated structure factor amplitudes.
[c] R_{free} = R_{cryst} calculated for reflections not included in the refinement. Because of the high NCS within the asymmetric unit and to minimized bias, reflections for R_{free} were chosen randomly in thin resolution shells. The test set size was 0.6 % of the total reflections.

3.2.3. PHASING, STRUCTURE DETERMINATION AND MODEL BUILDING

Obtaining phases from the experimental diffraction data is a crucial step in structure determination. The phasing was done by Carlos Frazão (ITQB, Oeiras, Portugal). Initially, phasing was tried by MAD (multiple wavelength anomalous dispersion; Drenth, 1999) using diffraction data collected at the iron -edge, but without success, as no iron atom sites could be detected. Heavy atom sites and initial phase estimates were then determined by MIRAS (multiple isomorphous replacement with anomalous scattering; Drenth, 1999; see chapter 6.7.4 for details) with the native diffraction dataset and derivative datasets from a mercury (Hg^{2+}) and a gold (Au^{3+})-soaked crystal (19 Au^{3+} and 13 Hg^{2+} sites; see statistics in Tab. 3.1). After phase improvement, Fourier maps calculated from these phases showed a clear and contigous electron density (Fig. 3.4).

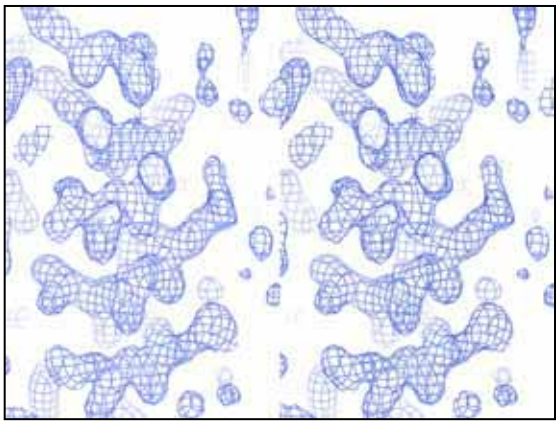


Figure 3.4: Stereo view of a sigmaA-weighted electron density map generated in an early stage of the refinement contoured at 1σ level. An α-helical region within the protein is shown.

An initial model of one monomer was built which contained the residues 2 - 308 out of 319 possible ones (see chapter 6.7.5 for details). No density for the N-terminal Met was present. Probably it is post-translationally cleaved off, as it is the case for the wild type enzyme (Kletzin, 1992). Therefore, Pro2 (codon numbering) was set as the first residue of the model. The 11 C-terminal residues could not be modeled, although their presence was confirmed by MALDI-TOF analysis (not shown), indicating a high structural flexibility of these residues presumably due to a disordering in the surrounding solvent. As 10 of these belong to the affinity tag, the model contained nearly the whole wild type sequence and solely the C-terminal glutamine could not be included.

The six monomers in the AU were generated by applying the NCS operators derived from 12 heavy atom sites of the Hg derivative. Subsequently one Fe ion per monomer was added. Model refinement resulted in a final model with 1,842 residues, 477 water molecules and 6 Fe ions with an R_{cryst} of 18.3 % and R_{free} of 21.0 % (Fig. 3.5, Tab. 3.1).

A global analysis of the 6 crystallographically independent SOR monomers in the AU indicated that their stereochemistry is similar or better than the reference set of structures in PROCHECK (Laskowski *et al.*, 1993) at similar resolution, with 94% and 6% of SOR amino-acids located in the most favored and in the additionally allowed regions of the Ramachandran diagram, respectively. The six crystallographically independent monomers had a very similar three-dimensional structure (Fig. 3.5), as judged from the average r.m.s. displacement of the α carbon main chain atoms of 0.12 Å. A model of the 24 subunit holoenzyme was generated from the model within the AU by applying the crystallographic symmetry operators of the space group I4 (Fig. 3.5).

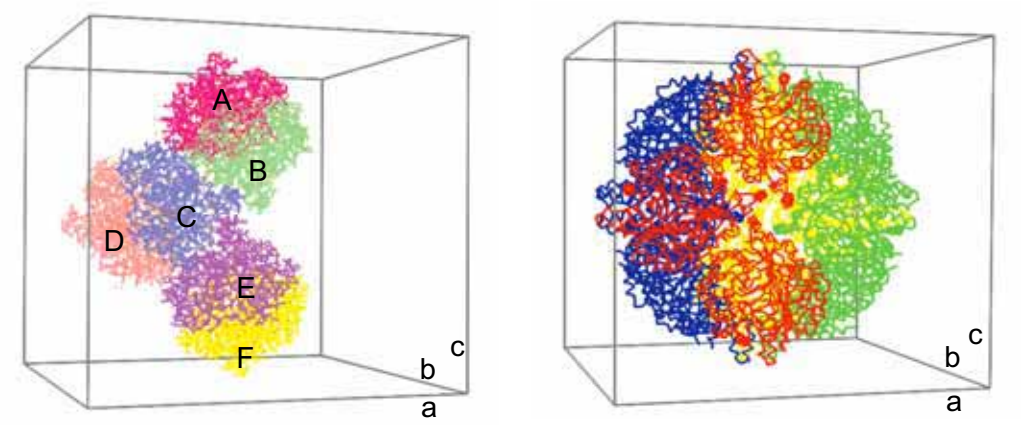


Figure 3.5: Left: The model in the asymmetric unit with the six crystallographically independent monomers colored differently. The unit cell is shown as grey box. Right: The model of the 24 subunit holoenzyme. It was generated by applying the rotational symmetry operations at the crystallographic four-fold axis along (c) to the model in the AU. The resulting models in the four AU are colored red, blue, yellow and green.

3.2.4. THE ARCHITECTURE OF THE HOLOENZYME

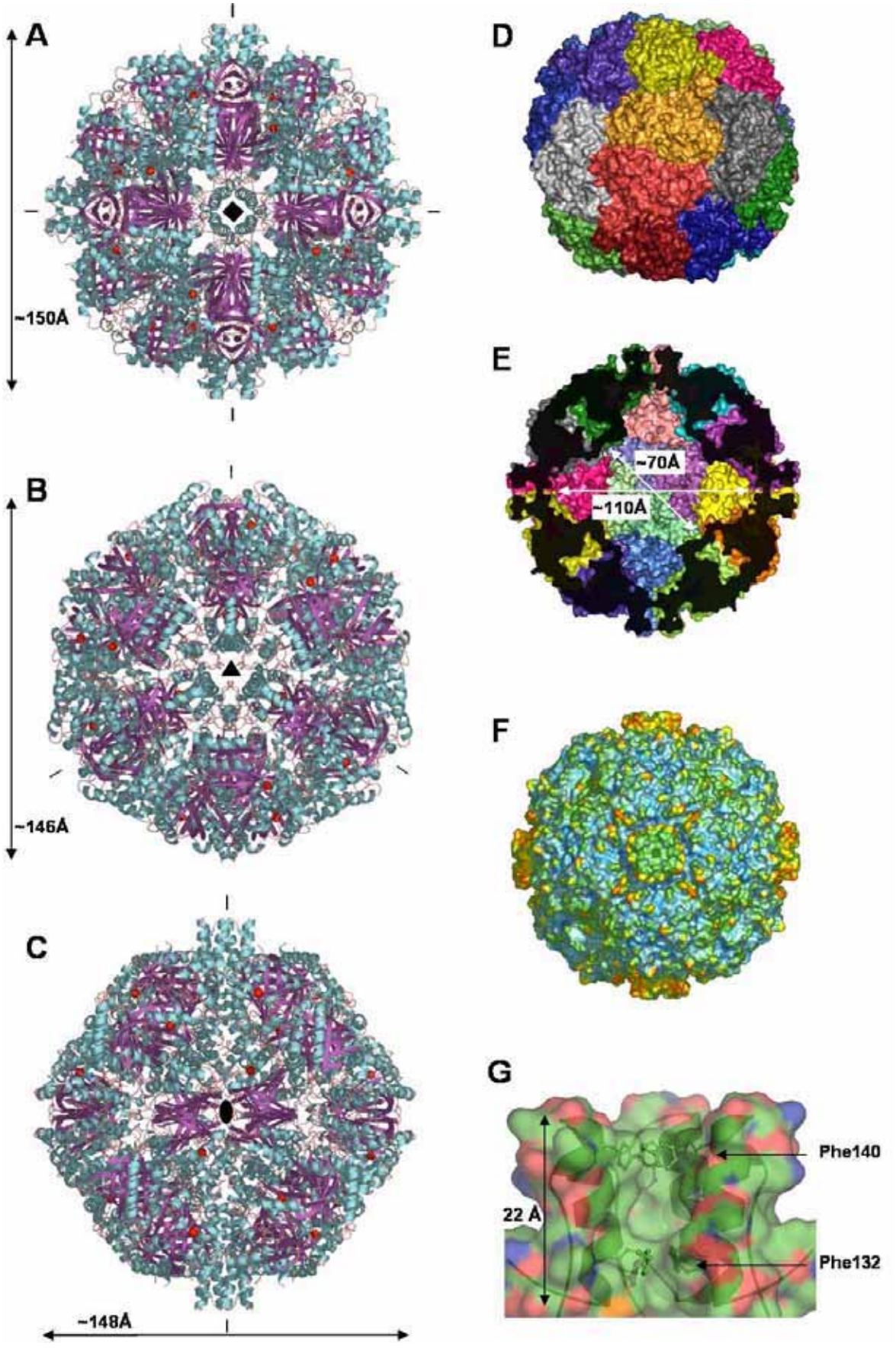
The recombinant holoenzyme is a homo-oligomer composed of 24 subunits which forms a highly symmetrical hollow sphere with 432 point group symmetry and a molecular mass of 871 kDa (Fig. 3.6A-C). The particle is almost perfectly symmetric, although minor distortions exist along the 3-fold symmetry axes (Fig. 3.6B). The root cause for this may be attributable to effects of crystal lattice packing contacts. The sphere has an external and internal diameter of ~150 Å and ~90 Å, respectively, and occupies a volume of 1,333 nm^3 (inner cavity included). The subunits have a distorted triangular shape and are equally distributed within the sphere, oriented tangential to the particle surface (Fig. 3.6D). The most striking structural features of the outer surface are six chimney-like protrusions at the 4-fold symmetry axes of the particle surrounded by ring-shaped grooves (Fig. 3.6D−F). No continuous channel from the exterior to the interior is observed in the structure, as these “chimneys” are occluded by a small barrier (Fig. 3.6E). The “chimneys” are ~22 Å in length and are each composed by the residues 129 to 153 of four neighboring monomers. Solely hydrophobic residues contribute to the inner chimney surface. The side chains of four Phe residues (Phe132) create the barrier to the inner compartment, above which a pocket of 9 Å in diameter is located. The outer entry is composed of four Phe residues (Phe140) connecting the pocket and the outer surrounding by a narrow pore (Fig. 3.6G). In contrast to the outer surface, the inner one is less compact in shape, giving rise to the existence of several separated compartments (Figs. 3.6E, 3.7A). The central cavity has an approximately star-like shape with dimensions of ~71 Å to ~107 Å and encloses a volume of 347 nm^3 (26 % of the total particle volume). Smaller cavities exist between the monomers, furthermore two cavities

are present within each monomer. All of them are connected via channels to the central inner compartment (not shown).

The analysis of the electrostatic properties of the particles´ outer surface revealed a patchwork-like distribution of positive and negative charges, besides positive areas around neutral patches at the “chimneys” at the 4-fold axes (Fig. 3.7A). The calculation of the electrostatic field showed a strong positive field surrounding these “chimneys”, originating from the previously described positive surface area (Fig. 3.7B). The inner surface of the sphere was dominated by positive areas which was mainly caused by the charges of the Fe(III) ions (Fig. 3.7A). Especially the surfaces of the cavities at the iron sites had a strong positive charge (not shown). Solely the “chimneys” at the 4fold axis had a neutral surface charge and the inter-subunit cavities had a partially negative surface (Fig. 3.7A).

Following page:

Figure 3.6: The Holoenzyme. **(A) – (C):** Cartoon representations of views along the crystallographic 4-fold, and the non-crystallographic 3-, and 2-fold axes. Symmetries are indicated by bars and their respective symmetry symbols. α-helices are colored cyan, β-sheets purple; red spheres: Fe ions. Note the protrusions at the 4-fold symmetry axes. **(D)** Molecular surface representation with monomers colored differently (solvent probe radius 1.4 Å). Each monomer has five neighbors. **(E)** Half of the SOR sphere after a cut along the 4-fold crystallographic symmetry axis. Note the inter-subunit cavities and the small protein barrier within the chimney-like protrusions separating the internal compartment from the exterior. **(F)** Molecular Surface representation to highlight the outer surface topology; the coloring is according to the atomic temperature factors, ranging from lower (blue) to higher (red) atomic vibrating sites. **(G)** The chimney-like structure as surface representation showing three out of four contributing monomers. The α-helices A4 contribute to the inner surface. The side chains of Phe132 and Phe140 are shown as ball-and-stick representation. Atom color code: carbon, green; sulfur, yellow; oxygen, red; nitrogen, blue.


A
~150Å
B
~146Å
C
~148Å
D
E
~70Å
~110Å
F
G
Phe140
22 Å
Phe132

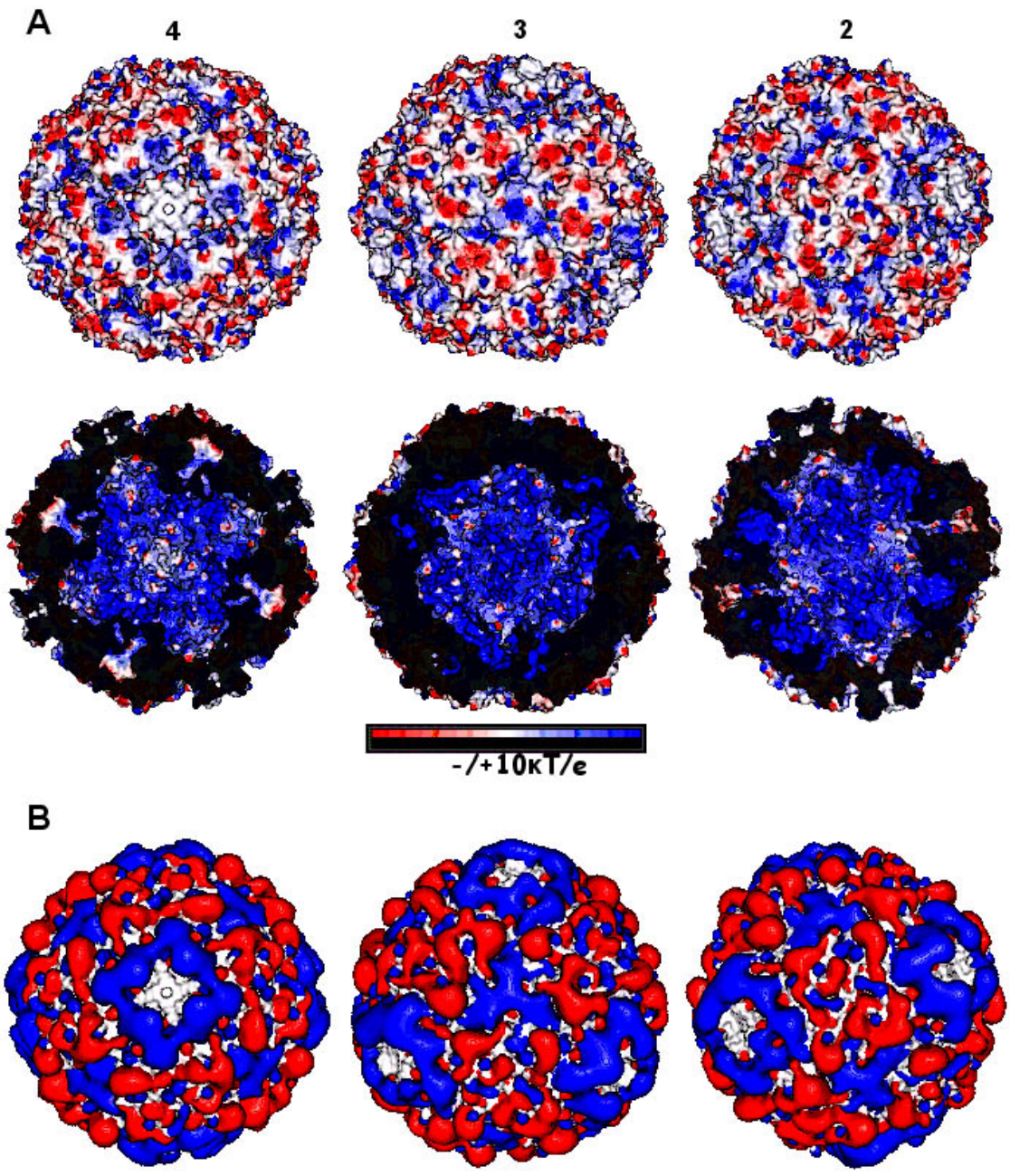


Figure 3.7: Electrostatic properties of the SOR. **(A)** Electrostatic potential surface views along the crystallographic 4-fold, and non-crystallographic 3-, and 2-fold symmetry axes. External and internal surfaces are shown. The molecular surface is color coded according the electrostatic potentials. Potentials less than -10 T/e are red, those greater than +10 T/e are blue and neutral potentials are white. Note the strong positive charge on the inner sphere surface. **(B)** Electrostatic fields at ±1kT/e viewed along the symmetry axes as indicated in (A). The electrostatic potentials and fields were calculated with MEAD using a protein and external dielectric constant of 4 and 80, respectively, a temperature of 300 K and an ionic strength of 0 M.

3.2.5. THE MONOMER

Each SOR monomer is an αβ protein with approximately triangular shape and is composed of an internal β-barrel partially surrounded by α-helices (Figs. 3.9A-B). It is oriented within the holoenzyme such that the barrels longitudinal axis runs tangential by the spherical surface of the particle (Fig. 3.9B). The barrel is formed by 8 anti-parallel β−strands distributed on two β-sheets with 4 antiparallel β-strands each. It has a distorted shape, with a squeezed “top” and an opened “base”. The α-helices are distributed in two sets at opposite sides of the barrel, in addition to two helices near its squeezed “top”. The α-helix A4 on top of the barrel is involved in the formation of the chimney-like protrusions of the holoenzyme. The N- and C-terminus are in vicinity to each other on top of the β−barrel and interact via a hydrogen bond formed by the imino nitrogen of Pro1 and the carbonyl group of Tyr304. The mononuclear iron site is located close to α-helix A2 adjacent to the N-terminal β-sheets (Fig. 3.9A).

The monomer contains 30.4 % β-strand 35.9 % helical regions (α-helices and 3.10 helices), which are higher values than previously predicted from the CD-spectra and secondary structure prediction programs (24-27 % β-strand and 24-29 % helical regions; chapter 2.2.10). Furthermore, the distribution of the secondary structural elements differs from the prediction (Fig. 3.8; see for comparison chapter 2, Fig. 2.13).

The monomer has a remarkable internal pseudo-symmetry (Figs. 3.9B-C), a 2-fold rotation through the barrel axis, that splits the structure into N- and C-terminal halves of similar sizes (residues 1-153 and 154-307, respectively). Both halves have the same major secondary structures in a similar spatial arrangement and with the same topological connections, besides one additional helix in the C-terminal half (A6, Fig. 3.9C). Both halves have a low sequence identity of 15.7 %. Each half belongs to the α-β sandwich group of proteins and comprises a ferredoxin-like split $(\beta\alpha\beta)_2$-fold (Orengo & Thornton, 1993), with two βαβ motifs forming an antiparallel sheet in which strand B1 lays in between strands B3 and B4 (Figs. 3.9D-E). The β-sheets form the β-barrel by extensive interaction of strand B2 with the antiparallel strand B4 of the opposite β sheet (B2 / B8, residues 42-51 / 290-298 and B4 / B6, residues 114-122 / 195-203, Figs. 3.9F-H). In addition, the regions between the two βαβ motifs of both half barrels contribute to the interaction of the N- and C-terminal halves by forming loop regions which interact with the α-helices of the second βαβ motifs of the opposite half.

The overall fold of the SOR monomer can therefore be interpreted as a duplication of the prototype ferredoxin fold.

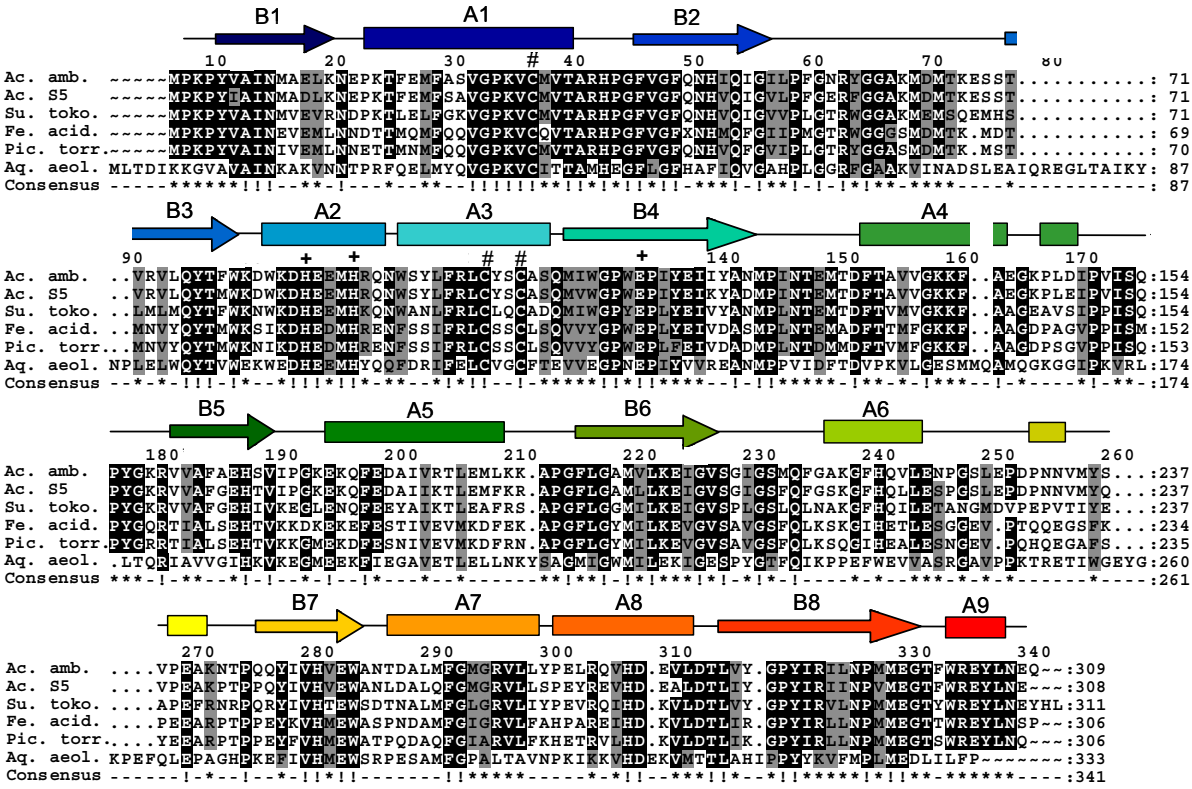
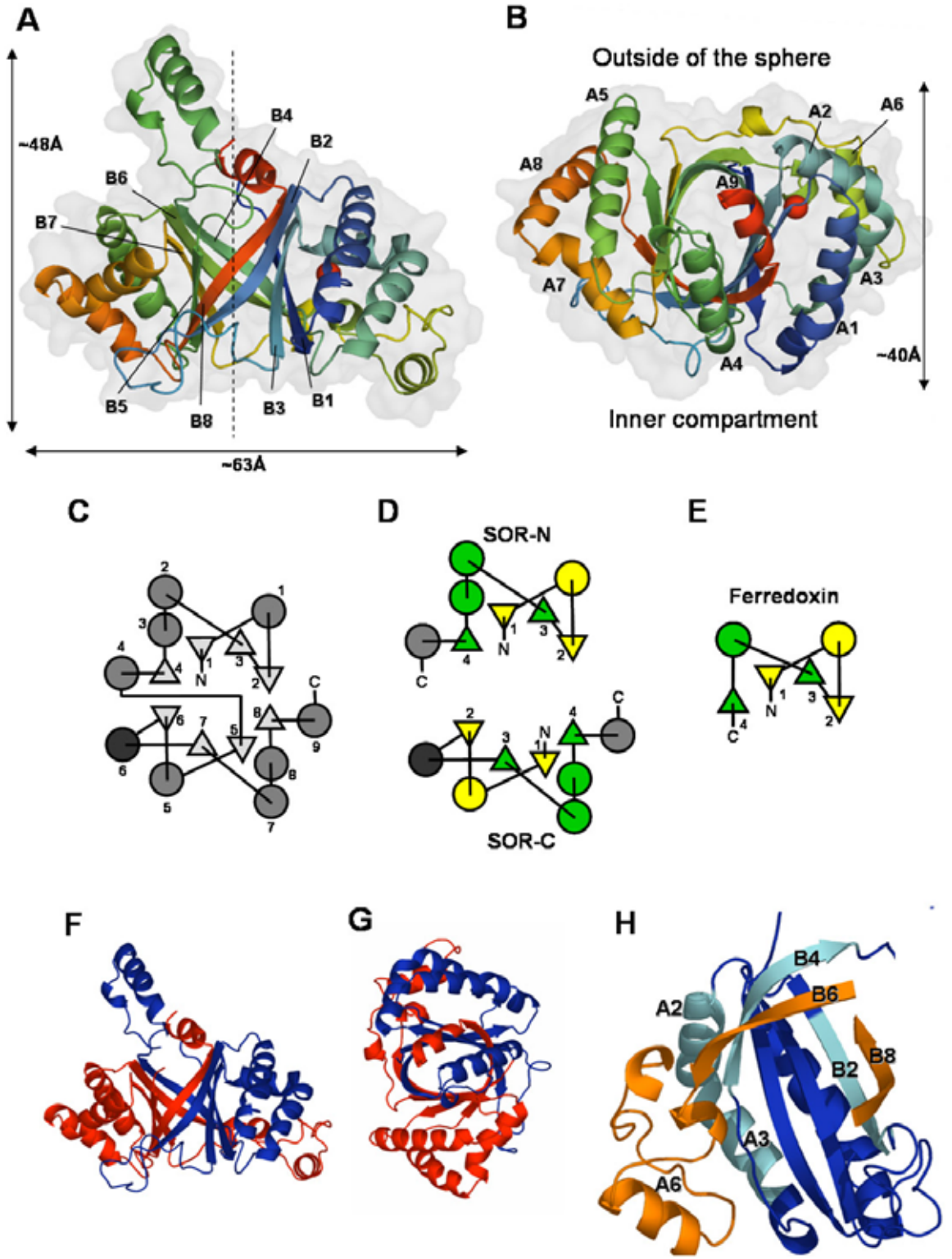


Figure 3.8: Multiple alignment of SOR protein sequences drawn according to the topology of *A. ambivalens* SOR with the upper two and lower two rows representing the N-terminal and C-terminal half barrels of the *A. ambivalens* SOR. Secondary structures are shown above the alignment (arrows, β-strands; rectangles, α-helices) with coloring and numbering according to Fig 3.8A and B. #, cysteine residues, + iron ligands. See Fig. 2.13 for further details of the alignment.

Following page:

Figure 3.9: The architecture of the SOR monomer. **(A)** side view **(B)** top view of a cartoon representation. The secondary structural elements (A, α-helices; B, β-sheets) are shown with coloring from blue (N-terminal) to red (C-terminal). Red sphere: iron ion. The monomer solvent-accessible surface is indicated in grey. Dashed line in A: pseudo-2fold symmetry axis. **(C)** Topology scheme. Triangles represent β-strands, circles α-helices; connections are shown as lines; short helical regions (≤ 4 residues) are omitted. The figure was drawn with the TOPS program (Flores *et al.*, 1994). The pseudo-2-fold symmetry axis along the β barrel is perpendicular to the plane of the paper and parallel to the barrel axis. Both halves have the same topological connections. The additional helix in the C-terminal half barrel of SOR is highlighted in dark gray. **(D)** Topology scheme highlighting the split $(\beta\alpha\beta)_2$ ferredoxin fold in each monomer half. The two βαβ motifs are colored yellow and green, respectively. **(E)** Topology scheme of Ferredoxin II from *Desulfovibrio gigas* (pdb code 1FXD; Kissinger *et al.*, 1991) as an example of the prototype ferredoxin fold. **(F)**, **(G)** Views on the monomer with the N-terminal and C-terminal halves colored in blue and red, respectively. Note their similar three-dimensional arrangement. **(H)** The regions mainly involved in half barrel interaction. Blue: N-terminal half with its interacting areas shown in light blue. Orange: regions of C-terminal half contributing to the interaction. Refer to text for details.


A
~48Å
B4
B2
B6
B7
B5
B8
B3
B1
~63Å
B
Outside of the sphere
A5
A2
A6
A8
A9
A7
A3
A1
A4
~40Å
Inner compartment
C
D
SOR-N
SOR-C
E
Ferredoxin
F
G
H
B4
B6
A2
B8
B2
A3
A6

A search for structural homologues was performed using the programs DALI (Holm & Sander, 1993), DEJAVU (Kleywegt & Jones, 1997) and CE (Shindyalov & Bourne, 1998) and revealed no homologue having the duplicated ferredoxin fold identified in the SOR. The highest structural similarity was found to the ferredoxin-like protein ActVA-orf6 from *Streptomyces coelicor* (Sciara, 2003; pdb code 1lq9; DALI: z-score 6.0; positional r.m.s. deviation 2.8 Å within 105 aligned residues), which is one of the rare examples for homo-dimerization within this fold category (see for details, the Structural Classification of Proteins [SCOP] database; http://scop.berkeley.edu/). Interestingly, both chains of ActVA-orf6 could be superimposed with the SOR monomer and showed the same strand interactions to form the β-barrel structure as the SOR (not shown).
In summary, none of the ferredoxin proteins known to date contains the duplicated $(\beta\alpha\beta)_2$ motif found in SOR, which makes the SOR the first example of this variation of the ferredoxin fold.

3.2.5.1 Distribution of conserved amino acids

The six known SOR proteins are characterized by a high sequence similarity. The five other SOR sequences have 39 % to 88 % identical positions, compared to the *A. ambivalens* SOR (Fig. 3.8), which implies that they most probably have a similar three-dimensional fold. The *A. ambivalens* SOR monomer was analyzed for the structural distribution of conserved amino acids. 77 amino acids are identical in all SOR proteins (25 % of the *A. ambivalens* SOR). These residues are not equally distributed in the three-dimensional structure of the monomer and clusters of conserved residues can be identified (Fig. 3.10A). Notably, all 11 residues in a radius of 6 Å around the iron ion are identical in all SOR sequences, and even in a radius of 8 A, 19 out of 23 (82.6 %) are conserved, indicating the importance of the iron site and its primary and secondary ligand sphere for function. In addition, 21 out of 38 residues (55 %) lining the active site cavity are also conserved (see below, section 3.2.6).

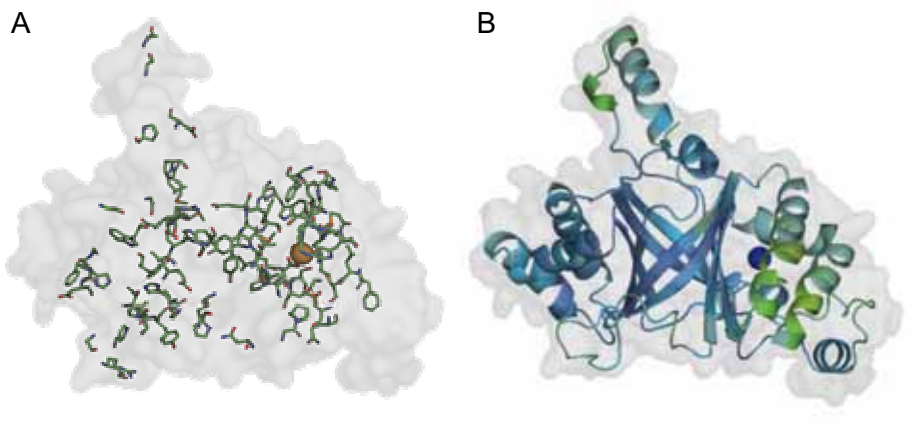


Figure 3.10: (A) Distribution of amino acids identical in all SOR sequences. Conserved residues are shown as stick representations; the iron atom is indicated as sphere. The shape of the monomer is indicated by its solvent-accessible surface area shown in grey. **(B)** Cartoon representation of one monomer colored according to the B-factors of α-carbon main chain atoms (blue low, yellow high B-factors). Refer to text for further details.

3.2.5.2 Conformational stability

Although the model obtained by X-ray crystallography is largely static, some insights about the dynamics and the rigidity of the protein is obtained from the atomic temperature factors (B-factors), which are measures of how much an atom oscillates around the position specified in the model (Drenth, 1999).

Analysis of the monomer structure using the B-factors of the α-carbon main chain atoms showed that the central β-barrel is the most rigid part, having B-factors ranging mainly from 25 to 30 $Å^2$ (Fig. 3.10B). The helices on the C-terminal side of the barrel have B-factors up to 35 $Å^2$, whereas the helices on the N-terminal side are less well defined. Especially the helices A1 and A3 have adjacent parts with B-factors ranging from 41 to 53 $Å^2$ (Glu15 - Gly26, Cys100 - Ser105), indicating an increased structural flexibility of this region, which lines the active site cavity and involves two of the cysteines (see below, section 3.2.6). Another flexible part of the structure with B-factors above 40 $Å^2$ is located close to the helix on top of the barrel (Glu142 - Asp147), the region of the monomer being involved in the formation of the chimney-like protrusions at the outer surface of the sphere.

Ion pairs (salt bridges) are considered to be involved in the increased thermal stability of proteins from hyperthermophilic organisms (Karshikoff & Ladenstein, 2001). The monomer A was analyzed for the presence of ion pairs using a cutoff distance of 4.0 Å between oppositely charged residues (Wang *et al.*, 2003). 15 salt bridges were identified within the SOR monomer, in addition to 6 inter-subunit salt bridges (Tab. 3.2). This yields a ratio of 6.9 ion pairs / 100 residues (4.9 intra- and 2.0 inter-subunit ion pairs / 100

residues). Remarkably, 19 out of 21 ion pairs are located on the surface of the monomer. The C-terminus of the monomer contributes to the formation of the two remaining ion pairs.

Table 3.2: Ion pairs within the monomer A and inter-subunit ion pairs with distances. Residues involved in ion pair networks are highlighted in bold.

ion pairs within the monomer			Intersubunit ion pairs		
residues involved		distance (Å)	residues involved (chain)		distance (Å)
Glu 184	Lys 187	3.07	Arg 35 (a)	Glu 221 (h)	2.79
Glu 11	Arg 72	3.18	Glu 221 (a)	Arg 35 (j)	3.01
Glu 15	Lys 13	3.20	Glu 221 (a)	Arg 273 (b)	2.85
Glu 67	Lys 61	3.56	Arg 265 (a)	Glu 303(m)	3.03
Glu 86	Arg 90	3.02	Arg 273 (a)	Glu 221 (b)	2.79
Glu 128	Lys 157	2.90	Glu 303 (a)	Arg 265 (g)	3.11
Glu 142	Lys 139	3.60			
Glu 164	Arg 290	2.72			
Glu 176	Lys 199	2.77			
Asp 177	Arg 181	2.43			
Glu 227	Arg 98	2.91			
Asp 229	Lys 214	2.92			
Glu 252	Arg 302	2.96			
Glu 278	Arg 181	2.72			
Glu 297	Arg 158	2.81			

3.2.5.3 The nature of the subunit interfaces

The SOR has a highly symmetrical quaternary structure, which implies that also the subunits are distributed in a highly regular pattern across the sphere. Analysis of the inter-subunit contacts in the crystal structure showed that each monomer has contacts to five neighboring monomers (Fig. 3.11).

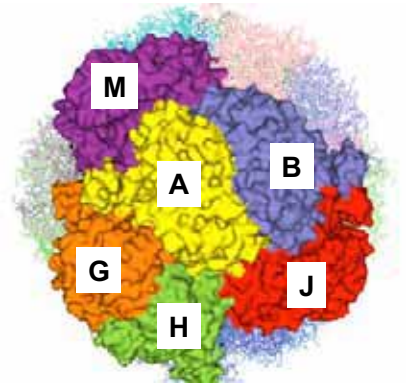


Figure 3.11: Subunit contacts in the holoenzyme, exemplified with monomer A and the surrounding monomers shown as surface representations.

The nature of the subunit interfaces was analyzed with monomer A and its surrounding partners using the Protein-Protein interaction server (Jones & Thornton, Tab. 3.3). A surface area of 5625 $Å^2$ of monomer A is buried in the holoenzyme, which corresponds to 39.5 % of the total monomer surface area (14240 $Å^2$). The neighboring subunits participate in different intensities in the interaction. Monomer A has the biggest interface accessible surface area with monomer B (2285 $Å^2$), which is 16.0 % of the whole monomer surface area and nearly the double interface area than with monomers G and M (8.5 % and 8.8 %, respectively). Minor interfaces exist with the monomers H and J (2.8 % and 3.1 %). A similar trend is seen, when analyzing the packing of the interfaces, indicated by the gap volume indices (GVI), which are measures for the complementary of interfaces (Jones & Thornton, 1996). The interface between A and B is well packed (GVI 1.39) followed by less complementary interfaces with G and M (GVI 2.29 and 2.28) and H and J (GVI 5.36 and 3.84). All subunit interfaces are predominantly hydrophobic in nature, composed of 59 % to 76 % apolar atoms (with J and G, respectively), but also hydrogen bonds and salt bridges are involved in the interactions.

Table 3.3: Analysis of the nature of the interface of monomer A with neighboring subunits.

Monomer	B	G	M	H	J
Interface accessible surface area ($Å^2$)	2285	1223	1276	402	439
% of total surface area	16.0	8.5	8.8	2.8	3.1
% Polar atoms in interface	40	24	29	35	41
% Apolar atoms in interface	60	76	71	65	59
Salt bridges	2	1	1	1	1
Gap Volume ($Å^3$)	6176	5680	5669	3744	3353
Gap Volume Index[a]	1.39	2.29	2.28	5.36	3.84

[a]Gap Volume Index = Gap Volume / (2 x Interface accessible surface area)

3.2.6. THE ACTIVE SITE(S) OF THE SOR

3.2.6.1 The iron site

One mononuclear iron site was identified in each monomer, which was located on the N-terminal half, between β-strand B4 and α-helix A2 (Fig. 3.9A). The ligand iron environment had a hexa-coordinatated octahedral geometry with a bidentate Glu (Glu113) and two His (His85 and His89) residues as protein ligands in all monomers (Fig. 3.12). The remaining sites were occupied by two water molecules, as judged from their distance to the iron ion (2.23 and 2.37 Å). Both water molecules were oriented towards a solvent-filled cavity. The protein ligands are conserved in all SOR sequences known to date (Fig. 3.8). The type of protein ligands and their spacial arrangement is typical for mononuclear non-heme iron(II) centers, which have a common structural motif, where the iron(II) center is coordinated by

three protein ligands, two His and a Glu or Asp, constituting one face of an octahedron, a motif referred to as the 2-His-1-carboxylate facial triad (Fig. 3.12C; Costas *et al.*, 2004). The sites were buried in the interior of each subunit with minimal distances of 8 Å to the outer surface and located at far end of a cavity which was exclusively accessible from the interior of the sphere. The minimal distance between the single iron sites in the holoenzyme was 37 Å.

As shown by EPR spectroscopy, the SOR in the "as isolated" state contained an Fe(III) ion and not Fe(II), (chapter 2; Urich *et al.*, 2004). The Fe(III) signal of the iron site disappeared in EPR spectra upon incubation with sulfur or addition of dithionite due to its reduction to Fe(II) (chapter 2.2.7, Urich *et al.*, 2004). Therefore, a diffraction dataset of a dithionite-reduced crystal was collected to investigate a possible ligand exchange at the iron site in the reduced state (see Tab. 3.5 for statistics). The reduced structure was essentially isomorphous to the "as isolated" one and in particular, no changes in the iron ligand sphere were detected (not shown), thus the site kept its hexa-coordinated octahedral geometry in the reduced state.

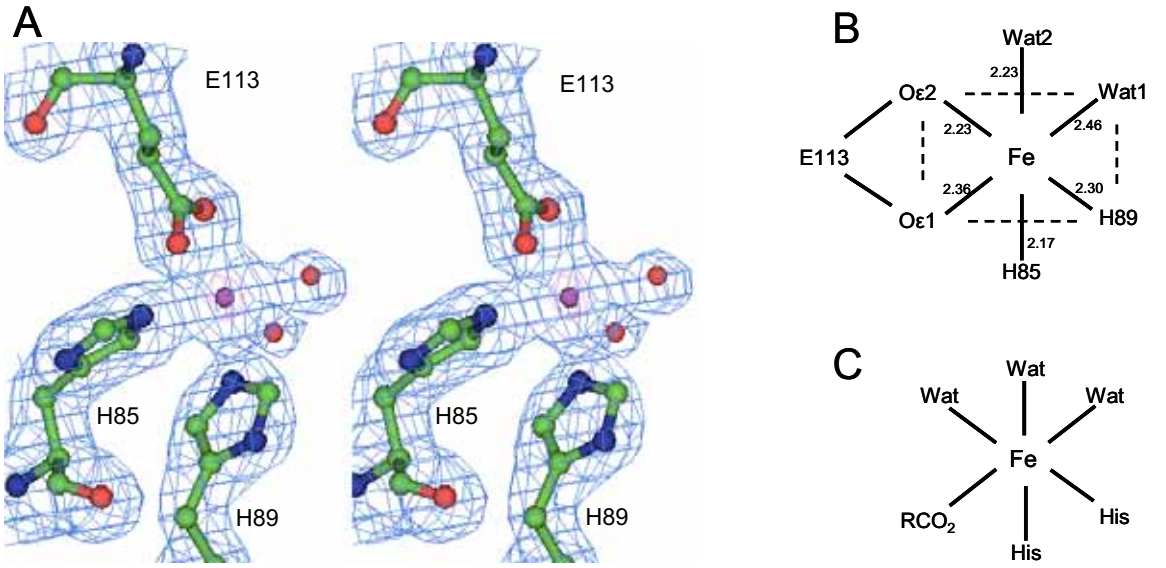


Figure 3.12: The mononuclear non-heme iron center. **(A)** Stereo view with Fe ligands and electron density in a sigmaA-weighted 2 mIFoI – DIFcI map contoured at 1σ, where σ represents the r.m.s. electron density for the unit cell. The Fe ion is indicated as purple sphere. **(B)** Schematic representation of the Fe center to highlight the hexa-coordinated octahedral geometry. Continuous lines represent Van der Waals ligand distances (average values for the six crystallographically independent monomers). Dashed lines indicate the plane of the octahedron. **(C)** Schematic representation of the 2-His-1-carboxylate facial triad structural motif. O_2CR, glutamate or aspartate ligand.

3.2.6.2 Posttranslational modification of a cysteine

All three cysteine residues of the monomer were located in close vicinity to the iron site, at the surface of the same cavity (Fig. 3.13A). Cys30, which is located close to the monomer surface on helix A1, is in a distance of ~ 10 Å to the Fe site whereas Cys100 is 7.8 Å away, located on helix A3. Cys103 is located 6.0 Å from Cys100 in the same helix and with 12.7 Å in even greater distance to the Fe site. The distance between Cys100 and Cys103 is too wide to allow the formation of a disulfide bridge. Cys100 and Cys103 are located in a region of high structural flexibility (Fig. 3.10B).

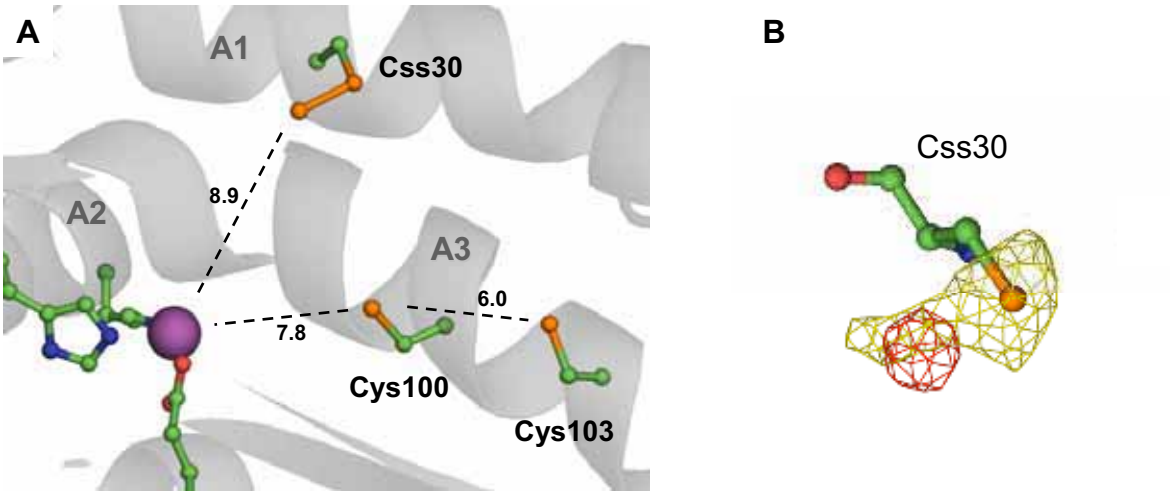


Figure 3.13: The active site cysteines. **(A)** Cartoon and stick representation of the positions of the three cysteines with Van der Waals distances between their terminal sulfur atoms and the Fe atom (purple sphere) in Å. **(B)** Identification of an additional sulfur atom at Cys 30. The position of the additional atom is indicated by an mIFoI–DIFcI electron density map averaged about the six crystallographically independent monomers and contoured at 5 σ (red mesh). Yellow mesh: averaged anomalous difference Fourier map calculated from a dataset collected at the iron -edge. The map is contoured at 3 σ and shows an anomalous scattering signal at the position of the additional atom.

Fourier maps using a fully refined structural model revealed unexplained density in a distance of ~ 2.0 Å to the Ṣ atom of Cys30 in each subunit, of ~ 6.7 σ in an m|Fo|-D|Fc| difference electron density map, indicating the presence of one additional atom (Fig. 3.13B). Taking into account the geometry of the modified amino acid as indicated by the electron density, cysteine persulfide (Cys-S-S) and cysteine sulfenic acid (Cys-S-O) were the most probable modifications (Claiborne, 1999; Bamford *et al.*, 2002). A refinement with dummy atoms in REFMAC resulted in an average distance of 2.2 Å to Ṣ of Cys30, which was indicative for the additional atom being sulfur (S-S distance 2.08 Å, compared to S-O distance of 1.68 Å). Analysis of averaged anomalous difference Fourier maps calculated using an iron -edge wavelength data set (see Tab. 3.1) showed an anomalous scattering signal at the site of unexplained density, which identified the

additional atom as sulfur, because oxygen has no such signal. In consequence, the Cys30 was refined as Cysteine-persulfide (Css). The terminal Sδ atom was oriented towards the iron site in a distance of 8.9 Å. (Fig. 3.13A).

3.2.6.3 Topology of the active site cavity

The active site cavity had a flattened and approximately triangular shape with dimensions of about 18 Å x 15 Å x 6 Å and is connected via a narrow channel to the central cavity (Fig. 3.14A, B). The side chains of 29 residues contribute to the topology of the cavity, whereas seven others contribute with their main chain atoms (Tab. 3.4). The side chains of two hydrophobic residues, Met296 and Met297, located at the entrance of the cavity, appear to be flexible, as judged from their B-factors (not shown). The walls of the cavity are mainly hydrophobic, although predominantly hydrophilic and hydrophobic surface areas on opposite sides of the cavity can be distinguished (Fig. 3.14C, D). The polar area generates a patch connecting the cavity entry with the iron site and is mainly formed by the side chains of Asn44, Gln75 and Asn8 and the main chain atoms of Tyr76, whereas Cys100, Cys103, Val25, Val29, Leu96 and to a minor extend Leu12, Phe22, Trp93 and Phe97 contribute to the apolar area.

Table 3.4: Residues lining the active site cavity. MC: main chain atoms contribute to cavity surface; NM: residue belongs to neighboring monomer.

Asn 8	Met 9 (MC)	Ala 10	Leu 12	Met 21
Phe 22	Val 25	Gly 26 (MC)	Val 29	Css 30
Thr 33	Asn 44	His 45 (MC)	Ile 46	Val 73
Gln 75	Tyr 76 (MC)	Thr 77	His 85	Met 88
His 89	Trp 93	Leu 96	Phe 97	Leu 99 (MC)
Cys 100	Tyr 101	Cys 103	Ala 104	Met 107
Pro 111	Glu 113	Ile 206 (MC)	Gly 207 (MC)	Met 209
Met295	Met 296	Arg 56 (NM)		

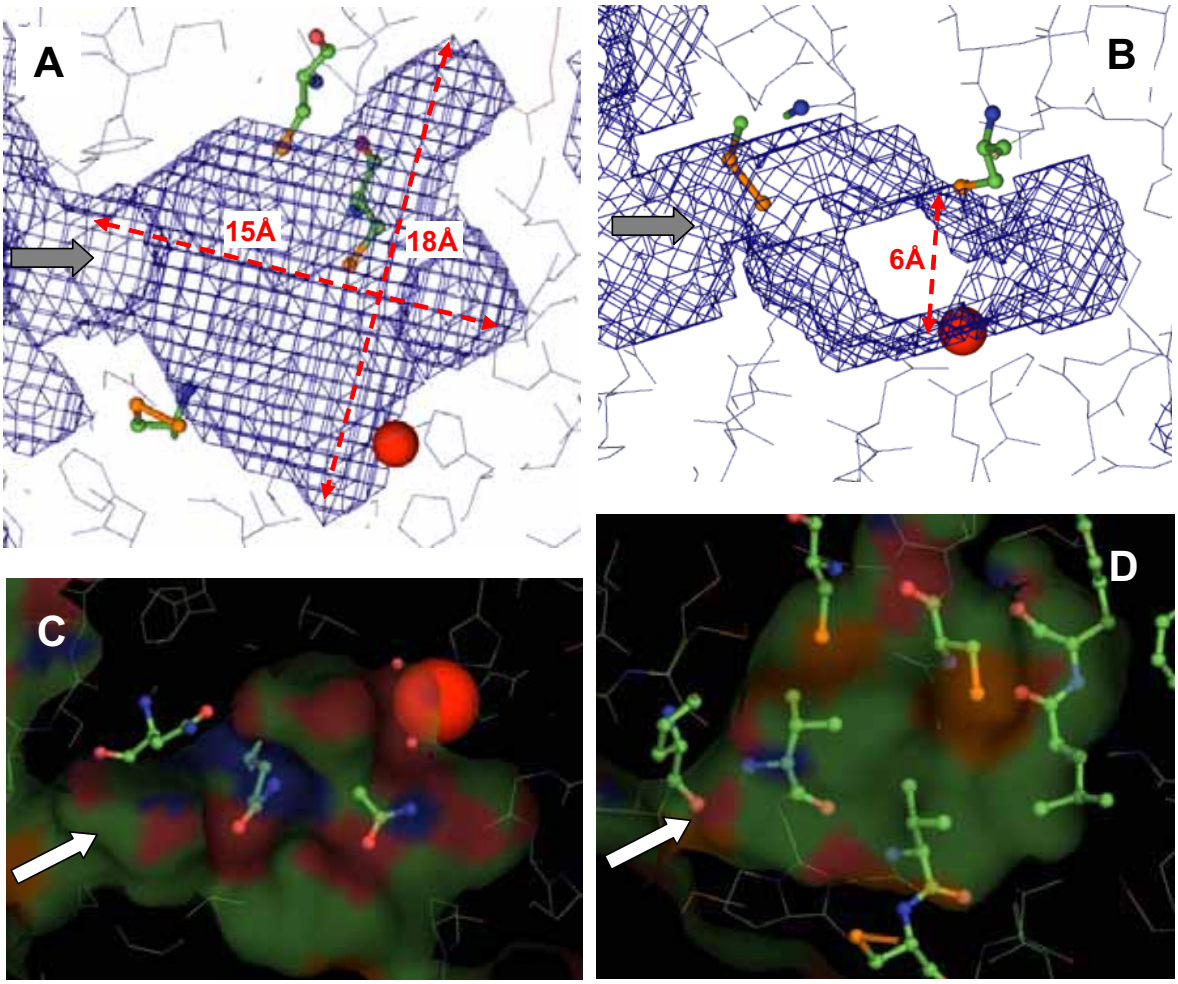


Figure 3.14: The active site cavity. **(A)** Mesh representation of the cavity generated with VOIDOO using a rolling probe of 1.4 Å radius. The iron atom (red sphere) and the cysteines are shown for orientation. The cavity entrance is indicated by an arrow. **(B)** Ditto, view perpendicular to A **(C)** predominantly hydrophilic, **(D)** predominantly hydrophobic cavity surface areas. Residues mainly contributing to the areas are shown as sticks, others are indicated as lines. The iron atom is shown as red sphere with its water ligands indicated as small red spheres. Color code: C, green, S, yellow, O, red, N, blue.

3.2.7. STRUCTURES COMPLEXED WITH SUBSTRATES, PRODUCTS AND INHIBITORS

Crystals of the SOR were soaked with sulfur, polysulfide and the enzymatic products sulfide, sulfite and thiosulfate to identify substrate binding sites and intermediate steps of catalysis (see chapter 6.7.2 for details). In addition, the SOR was incubated with sulfur at 70°C until a catalytic reaction occurred and subsequently cocrystallized with sulfur. After data collection and processing, the native structure was used as template for the refinement (see Tab. 3.5 for statistics). The final electron density maps were subsequently screened for additional areas of electron density not explained by the models, where the respective chemicals could fit in.

All models were essentially isomorphous to the native structure (not shown). Unfortunately, in none of the structures treated with substrate or products significant additional electron density was found (not shown), thus no reaction intermediate or bound substrate could be identified.

Following page:

Table 3.5: Diffraction data processing and model refinement statistics of SOR crystal derivatives. All data sets were collected at ESRF beam line ID14-3. All crystals had space group I4 (79). The numbers in brackets refer to the outer shell. The upper line of the unit cell parameter section corresponds to a=b dimensions, the second to the c dimension.

[a] $R_{merge} = \sum | I_o - \langle I \rangle | / \sum I_o$, where $\langle I \rangle$ is the average of symmetry equivalent reflections and the summation extends over all observations I_o for all unique reflections.

[b] $R_{cryst} = \sum \| F_o | - | F_c \| / \sum | F_o |$, where F_o and F_c are the observed and calculated structure factor amplitudes.

[c] $R_{free} = R_{cryst}$ calculated for reflections not included in the refinement.

Data	SOR-ioac	SOR-ditred	SOR-scoc	SOR-sinc	SOR-sfd	SOR-sft	SOR-ts	SOR-zn	SOR-ps	SOR-hg
Derivative	iodo-acetamide	dithionite reduced	sulfur co-crystallized	sulfur incubated	sulfide soaked	sulfite soaked	thiosulfate soaked	zinc actetate soaked	polysulfide soaked	mercury soaked
Unit cell parameters (Å)	161.90 154.27	161.95 154.25	161.82 154.21	161.43 153.99	161.68 154.31	161.70 154.17	161.68 154.33	162.07 154.24	161.67 154.35	161.88 154.37
Wavelength (Å)	0.934	0.934	0.934	0.934	0.934	0.934	0.934	0.934	0.934	0.934
Nr. unique intensities	172,862	130,286	150,469	69,907	154,316	153,306	154,119	154,918	154,143	68,537
Redundancy	2.8	2.6	3.0	3.2	2.9	2.9	2.8	2.8	2.9	15.0
Resolution (Å)	38.05 - 1.80 (1.85-1.80)	63.97 - 2.00 (2.05-2.0)	48.80-1.90 (1.95-1.9)	55.90-2.40 (2.46-2.40)	55.81-1.90 (1.95-1.9)	55.90-1.90 (1.95-1.90)	64.55-1.90 (1.95-1.90)	43.03-1.90 (1.95-1.90)	65.47- 1.90 (1.95-1.90)	49.04-2.50 (2.54-2.5)
Completeness (%)	94.4 (92.5)	97.5 (95.1)	96.7 (97.9)	90.7 (91.5)	99.3 (98.8)	98.7 (97.9)	99.2 (99.1)	99.3 (94.1)	99.1 (98.8)	99.9 (98.9)
R_{merge} (%) [a]	5.5 (26.2)	7.3 (35.3)	6.8 (30.2)	9.2 (31.5)	5.8 (33.3)	7.4 (33.5)	8.2 (29.1)	6.3 (45.9)	6.1 (21.5)	2.9 (47.0)
I/σ (I)	16.70 (1.79)	11.46 (1.67)	12.73 (2.5)	10.36 (1.93)	16.94 (2.38)	13.23 (2.03)	11.93 (2.36)	15.57 (3.33)	11.46 (1.67)	32.97 (4.31)
Refinement										
Refined structure	1842 aa 477 waters	1842 aa 471 waters	1842 aa 471 waters	1842 aa 471 waters	1842 aa 471 waters	1842 aa 471 waters	1842 aa 471 waters	1842 aa 477 waters	1842 aa 471 waters	1842 aa 477 waters
R_{crys} (%) [b]	17.0	17.3	17.2	18.2	17.3	17.0	16.6	16.0	17.0	15.9
R_{free} (%) [c]	19.0							18.9		21.5
Average B-factors ($Å^2$)	29.70	29.11	28.75	33.54	29.37	32.78	29.64	29.55	28.91	30.99
r.m.s.d.										
Bonds (Å)	0.022	0.027	0.024	0.024	0.024	0.024	0.025	0.124	0.023	0.036
Angles (°)	1.710	1.982	1.832	1.830	1.843	1.837	1.860	2.098	1.800	2.298

To determine the binding sites of inhibitors within the SOR (Kletzin, 1989) crystals were soaked with Hg^{2+}, Zn^{2+} and iodoacetamide. The structure of mercury- (Hg^{2+}) complexed SOR, solved at 2.5 Å resolution showed the presence of two Hg^{2+} ions per monomer (Fig. 3.15A). One was bound in a distance of 2.12 Å (averaged about the six crystallographically independent monomers) to the Ṣ atom of Css30, whereas the other was located between Cys100 and Cys103, although it appeared in a distance of ~5 Å to both Ṣ atoms not to be covalently bound. Thus, the inhibition of the enzyme by Hg compounds (Kletzin, 1989) is caused by modification of the active site cysteines.

The cocrystallization with the thiol-modifying reagent iodoacetamide resulted in additional electron density resembling acetamide in covalent distance to Cys100 (not shown). The density was present with low occupancy only in three out of six monomers in the asymmetric unit, which is presumably due to a too short incubation time during crystallization (2 d). No additional density was observed at Css30 or Cys103. In consequence, it appears that Cys 100 is the primary target for alkylation with iodoacetamide, although this conclusion is tentative due to the low reactivity between enzyme and inhibitor in this experiment.

In the Zn^{2+} complexed structure of the SOR, solved at 1.8 Å resolution, one Zn^{2+} ion per monomer was detected, although not in close vicinity to the active site, but bound in a distance of 2.1 Å (averaged about the six crystallographically independent monomers) to His276, together with His165 part of a 2-His motif on the opposite side of the beta barrel core of the monomer (Figs. 3.15B-D). This motif is located at the end of a narrow pocket guided by mainly hydrophobic residues, which is connected to the active site residue Css30 of the neighboring monomer.

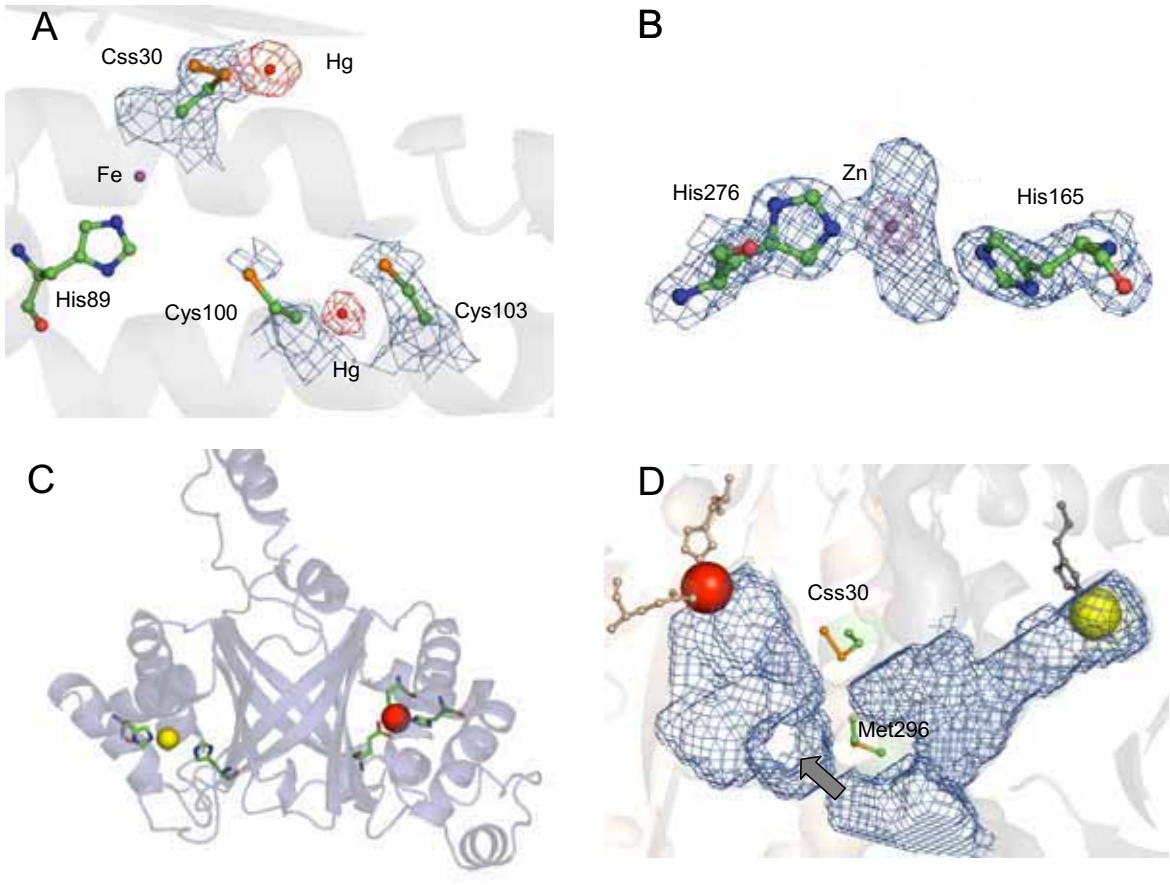


Figure 3.15: Binding sites of inhibitors. **(A)** Mercury sites. Hg^{2+} ions are indicated as red spheres. Red mesh: 2mIFoI – DIFcI electron density map around Hg^{2+} ions contoured at 5 σ. Blue mesh: 2mIFoI – DIFcI electron density map around the cysteines contoured at 1 σ. **(B)** $Zink^{2+}$ (purple sphere) binding site at His276. Blue mesh: 2mIFoI – DIFcI electron density map contoured at 1 σ indicating the presence of two additional water ligands of Zn^{2+}. Purple mesh: 2mIFoI – DIFcI electron density map contoured at 5 σ. **(C)** Position of the Zn^{2+} (yellow sphere) binding site within the monomer. Note the distance to the active site (red sphere: iron ion). **(D)** The Zn^{2+} binding site of one monomer (gray) in context with the active site of the neighboring monomer (pink). Zn^{2+} is located at the end of a tunnel leading to the active site Css 30 of the neighboring monomer. Blue mesh: representation of cavities detected with VOIDOO using a rolling probe of 1.4 Å. The residues separating both cavities (Css30 and Met296) are shown as stick representation, as are the metal binding ligands. The arrow indicates the active site cavity entrance.

3.3. DISCUSSION

The results from crystallography are discussed mainly in respect to functional implications. Although the discussion focuses on the SOR from *A. ambivalens*, it is reasonable that, due to the high sequence similarity, most of the conclusions hold also true for the other members of the SOR protein family.

3.3.1. THE SOR IS A SELF-COMPARTMENTALIZING ICOSATETRAMER

The molecular mass and subunit composition of SOR have been under debate for a long time. Biochemical analyses did not gain conclusive answers to this topic, with relative molecular masses (M_r) and stoichiometries ranging from 550 kDa to 730 kDa and 15 to 20 subunits, respectively (chapter 2.3.4; Urich *et al.*, 2004). The present data show that the recombinant SOR crystallizes as an icosatetramer, having an M_r of 871 kDa. As the recombinant monomer is slightly bigger than the wild type monomer (318 amino acids compared to 308 amino acids), the wild type holoenzyme should have an M_r of 845 kDa. Thus, the SOR has approximately 1/3 of the mass of the 70S ribosome (M_r: ~2.5 MDa) and is one of the biggest protein complexes for which X-ray structural information is available (Dutta & Berman, 2005). The dimension of the SOR oligomer (~150 Å in diameter) is in accordance with the size of particles previously described by electron microscopy (see chapter 2.2.10), thus showing that the crystal structure represents the biologically active molecule.

The highly symmetric overall architecture of the complex with its 432 point group symmetry closely resembles the ones known from ferritins or bacterioferritins, which are spherical icosatetramers with M_r of ~450 kDa, a diameter of ~110 Å and, like the SOR, encapsulate an inner compartment (Fig. 3.16). Thus the SOR represents the second example of a protein with this architecture, although with nearly the double M_r and particle volume than ferritins (SOR 1,333 nm^3, bacterioferritin from *E. coli* 720 nm^3). In contrast to their similar quaternary structure, ferritins and SOR differ in their tertiary structure (Fig. 3.16). The ferritin monomers have a four-helix bundle fold with no β-sheet part, whereas the SOR is an αβ protein with a duplicated ferredoxin fold.

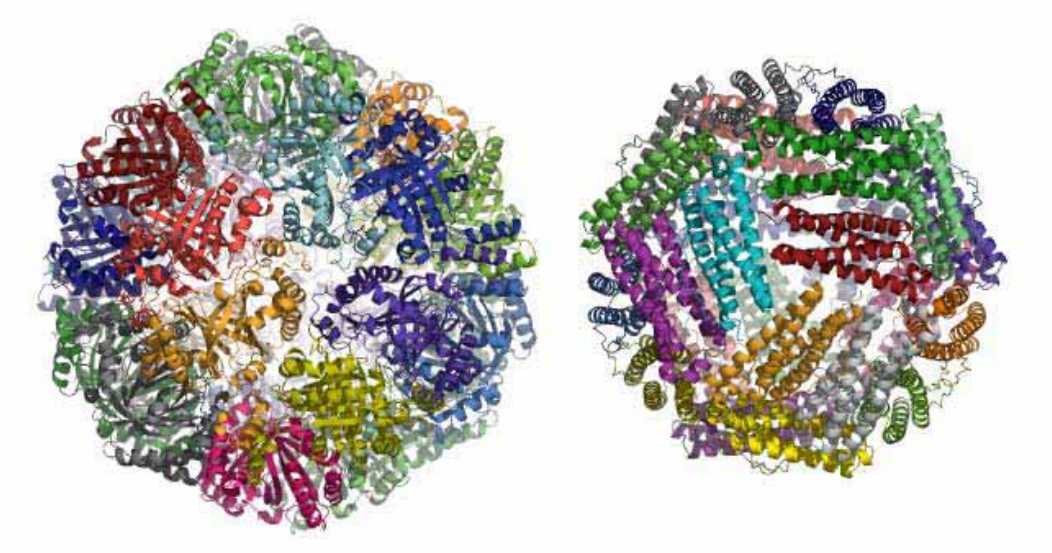


Figure 3.16: Comparison of size and architecture of the SOR and ferritins, exemplified by the bacterioferritin from *E. coli* (1BFR.pdb, Frolow *et al.*, 1994). The molecules are set to the same scale and viewed along a 3-fold symmetry axis. Cartoon representation with monomers colored differently.

Ferritins are involved in the iron metablism of many organisms and their inner compartment serves as iron reservoir (Theil, 2001), whereas the SOR is involved in sulfur oxidation, which implies that the formation of its inner compartment is due to different functional requirements.

Self-compartmentalization can be advantageous for several reasons. As judged from the structure, the catalytic sites in the monomers are not directly accessible from the exterior of the sphere but solely via pores from the inner compartment. In consequence, the sulfur substrate has to enter the sphere prior to catalysis, which allows for the pre-selection of the "correct" substrate by a gate-keeping structure and thus for a highly controlled access to the catalytic sites. The most probable entry routes are six chimney-like protrusions equally distributed within the particle, whose inner surfaces are composed of hydrophobic residues. As sulfur is likewise hydrophobic, a transfer towards the inner compartment seems possible to occur via these routes (Fig. 3.6G; and see below, section 3.3.4).

A second advantage of self-compartmentalization might be to segregate the sites of catalysis from the exterior, which might provide a spacial protection of the cytoplasmic environment from reactive oxygen species generated during catalysis.

In addition, the dimensions of the inner compartment (~347 nm^3) could also allow the storage of considerable quantities of substrate sulfur compounds. Within this microenvironment distinct conditions are present, as judged from the strongly positive charged surface (Fig. 3.7). From there, the substrate could efficiently be delivered to the

24 active sites surrounding the compartment. A spacial clustering of many catalytic sites within one particle might be another positive effect of the oligomerization.
The biological principle of providing a functional control by self-compartmentalization has also been identified in a number of proteases like the proteasome, the tricorn peptidase and ClpAP (Lupas *et al.*, 1997). These proteolytic macromolecular complexes form compartments through self-association of multiple subunits and segregate their proteolytic active sites to the interior of these compartments to provide a controlled access of substrate proteins (Baumeister *et al.*, 1998).

3.3.2. A MODEL FOR THE HOLOENZYME ASSEMBLY

The formation of a hollow spherical protein particle composed of 24 subunits appears to be a delicate task and implies the presence of distinct intermediate stages in its assembly pathway. Denaturant-induced unfolding experiments had identified a stable, properly folded dimeric intermediate during the disassembly of the holoenzyme, which was interpreted such that the holoenzyme formation might proceed via the assembly of homodimeric building blocks (see chapter 2.3.5 and Urich *et al.*, 2004). The analysis of the inter-subunit contacts in the crystal structure provided further evidence for this model, as two monomers share 16.0 % of their total surface area, which is nearly the double than with other monomers (Tab. 3.3). The highest value of interfaces area together with a highly complementary shape of the participating interfaces strongly suggests that these monomers make up the previously described homodimeric building block (Fig. 3.17A). The assembly of twelve dimers to form the icosatetramer then may proceed via the less complementary subunit interfaces (Fig. 3.17B). These are of predominantly hydrophobic nature (Tab. 3.3), which is considered to be a main driving force for oligomer formation by the exclusion of solvent molecules upon the burying of hydrophobic surface patches (Bahadur *et al.*, 2003).

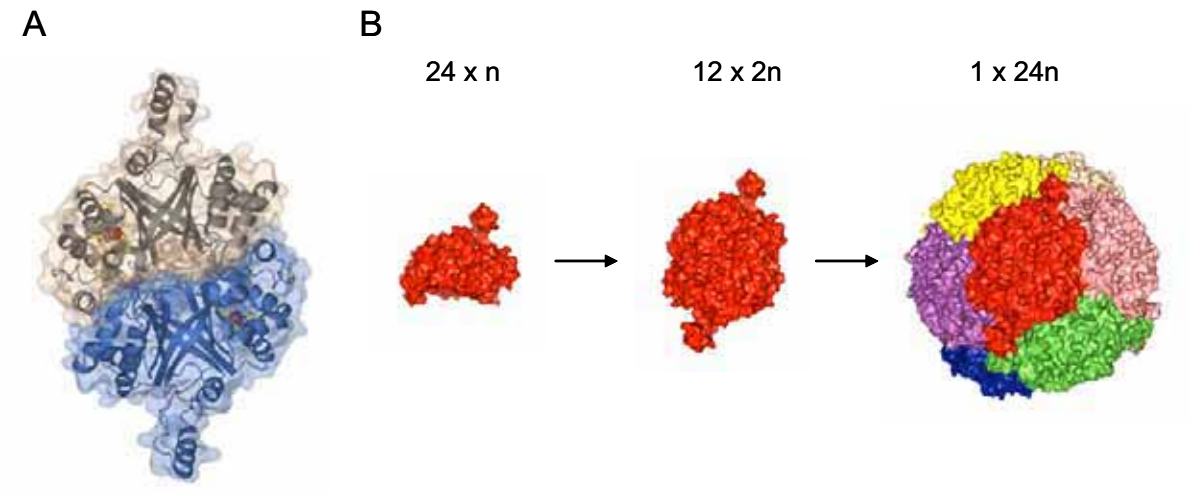


Figure 3.17: Assembly of the holoenzyme. **(A)** The homodimeric building block. **(B)** Oligomerization scheme. Refer to text for detailed explanation.

3.3.3. THE ACTIVE SITE(S) OF THE SOR

It had been postulated previously that a mononuclear non-heme iron site together with three cysteines participate in catalysis (chapter 2.3.3 and Urich *et al.*, 2004) and the structure indeed supports this hypothesis, as all components are located in close vicinity to each other along a cavity. In consequence, this region was considered to be the active site within the SOR monomer. The distribution of the active sites within the SOR holoenzyme particle, with minimal Fe-Fe distances of ~37 Å, makes electronic interaction between the single centers during catalysis almost impossible. In consequence, 24 separate active sites are present within the SOR holoenzyme and most probably these centers do not interact in catalysis, although cooperative effects cannot be excluded.
From the crystal structure, one can derive possible functions for the single components involved in catalysis.

3.3.3.1 The iron site

The Fe(III) ion is coordinated by two histidines (His85 and His89) and a glutamate (Glu113), which serves as a bidentate ligand. The type of residues and their spacial distribution identifies them as a variation of the 2-His-1-carboxylate facial triad motif (Hegg & Que, 1997; Costas *et al.*, 2004). There, the Fe(II) ion is coordinated by two His and one acidic Asp or Glu residue, constituting one face of the octahedron, which leaves the opposite face of the coordination sphere free for the addition of exogenous ligands (Costas *et al.*, 2004). A diverse collection of metalloenzymes with non-related enzyme activities utilizes this structural motif for iron binding (Tab. 3.6; Costas *et al.*, 2004).

Although the carboxylate ligand is monodentate in most examples, there are examples for a bidentate carboxylate ligand in the literature. For example, a bidentate aspartate ligand was found In the crystal structure of naphtalene 1,2 dioxygenase from *Pseudomonas putida* (Karlsson *et al.*, 2003), as it was the case for a glutamate in the enzyme-substrate complex of human phenylalanine hydroxylase (Andersen *et al.*, 2002).

Sequence comparisons showed that the 2-His-1-carboxylate facial triad is conserved within each class of enzymes but not among different classes, making an identification by sequence comparison difficult (Tab. 3.6). Whereas His85 and His89 had been previously suggested as iron ligands by sequence analysis, Glu113 was not predicted (chapter 2.3.3, Urich *et al.*, 2004). The utilization of different sequence motifs for the creation of the 2-His-1-carboxylate facial triad was interpreted as a convergent evolution towards a particularly favored metal-binding motif that is versatile in promoting a variety of different reactions (Costas *et al.*, 2004). The SOR protein family is another example for this finding, as within all known SOR sequences the three ligand residues are conserved (Fig. 3.8), as it is the case for the respective sequence motif, H-X_3-H-X_{23}-E, which has no similarity to any of the previously described ones (Tab. 3.6).

Table 3.6: Iron binding sequence motifs of enzyme classes with a crystallographically established 2-His-1-carboxylate facial triad.

Enzyme class	dominant sequence motif	Reference
Extradiol Catechol Dioxygenases	H-$X_{55(8)}$-H-$X_{50(2)}$-E	Costas *et al.*, 2004
α-Keto Acid-dependent Enzymes	H-X-(D/E)-X_n-H	Costas *et al.*, 2004
Pterin-dependent Enzymes	H-X_4-H-$X_{40(1)}$-E	Costas *et al.*, 2004
Rieske Dioxygenases	H-X_4-H-X_{148}-D	Costas *et al.*, 2004
Sulfur Oxygenases Reductases	H-X_3-H-X_{23}-E	this study

All enzymes with the 2-His-1-carboxylate facial triad have in common that the iron center is the site of oxygen activation and that the site is in the Fe(II) state during catalysis (Costas *et al.*, 2004). This distinguishes them from the mononuclear non-heme Fe(III) enzymes, where the iron site is in the Fe(III) state throughout catalysis and acts in substrate activation without directly interacting with O_2 (i.e. lipoxygenases, intradiol cleaving catechol dioxygenases; Costas *et al.*, 2004). The latter do not have a common, general iron binding motif. In consequence, the identification of the 2-His-1-carboxylate facial triad in SOR points to the Fe site being responsible for oxygen activation. Furthermore it suggests that the Fe(III) state of the “as isolated” enzyme gets reduced to Fe(II) prior to O_2 binding.

Therefore, the geometry of the iron center in the reduced Fe(II) state was investigated by the determination of the structure of the dithionite-reduced SOR. The site kept its hexa-

coordinated octahedral geometry and no iron ligand changes were observed. However, the relevance of this observation gained with the chemically reduced enzyme in the absence of its physiological substrate remains unclear.

One striking feature of the SOR iron site is the low reduction potential (E'_0 = -268mV, chapter 2), which is unusual for mononuclear non-heme iron centers. For the soy bean lipoxygenase with a strongly positive reduction potential of E'_0 = +600 mV (Nelson, 1988, Costas *et al.*, 2004) it was speculated that the high potential was a consequence of the iron-coordination by four basic histidine ligands. In contrary, the SOR should then have a predominance of acidic residues in the vicinity of the iron site. This was not found in the structure, where a balanced ratio of acidic and basic residues was present, even in a radius of 8 Å around the iron site. Thus, different, not yet identified factors appear to influence the reduction potential of the Fe site.

3.3.3.2 The cysteines

All three cysteine residues of the monomer are located in the vicinity of the iron site and their importance for catalysis had been shown by the inhibitory effect of thiol-binding reagents (Kletzin, 1989 and chapter 2.2.8). One probable function is their involvement in binding of the sulfur substrate. Furthermore, it has been discussed that Cys100 and Cys103 could act as a second redox-active center in addition to the iron site, forming a reversible disulfide bridge (chapter 2.3.3), as it had been postulated for the sulfide:quinone oxidoreductase from *Rhodobacter capsulatus* (Griesbeck *et al.*, 2002). However, with a distance of ~6 Å between their Sγ atoms in the native structure, disulfide-bridge formation is impossible, and therefore, no second redox-active center is present in the SOR monomer. Nevertheless, the inhibitor Hg^{2+} was bound almost exactly between both cysteines, indicating (1) that they are important for catalysis and (2) that they may function in a coupled manner. Interestingly, the inhibitor iodoacetamide was found to bind, although with a low occupancy, solely at Cys 100, contrasting the hypothesis of a coupled action of both cysteines in catalysis.

One intriguing finding was the presence of an additional sulfane sulfur atom at Cys30, generating a Cys-persulfide (Css). It should be noted that the modification was found in the enzyme that was heterologously produced in *E. coli* and that the modification has not yet been confirmed for the wild type SOR.

From the chemical point of view, sulfane sulfur compounds contain a labile, highly reactive sulfur atom in a reduced oxidation state with a valence of 0 or -1 covalently bound to another sulfur atom. Sulfur with such features easily leaves the compound structure and

can be transferred onto acceptor groups (Beinert, 2000). Furthermore, Cys-persulfides are more easily deprotonated at physiological pH, compared to the corresponding ionization of cysteines (Iciek & Wlodek 2001), which results in reactive perthiolate anions.

Css are generated in the catalytic cycles of various enzymes (Tab. 3.7). Structurally, most of these proteins can be interpreted as variations of the prototype rhodanese. These enzymes act as sulfur transferases in the biosynthesis of sulfur-containing biomolecules like Fe/S clusters, biotin, thiamine, molybdopterin and sulfur-containing bases in RNA (Beinert, 2000).

Interestingly, Css were also identified in enzymes involved in the dissimilatory sulfur metabolism. The rhodanese-type protein Sud acts as a polysulfide:sulfur transferase in the dissimilatory sulfur reduction pathway of *Wolinella succinogenes* (Lin *et al.*, 2004). Furthermore, the thiosulfate oxidoreductase SoxAX, a c-type cytochrome from *Rhodovulum sulfidophilum* contained a Css in its active site (Bamford *et al.*, 2002).

Table 3.7: Proteins carrying one or several sulfane sulfur atoms.

Protein	Organism	Catalytic Function	Physiological Pathway
Rhodanese (Thiosulfate:cyanide sulfurtransferase)	*Bos taurus* (Ploegman, 1978*[+]), *Azotobacter vinelandii* (Bordo, 2000*[+])	Sulfur transferase	Fe/S cluster assembly (Cereda, 2003)
SoxAX	*Rhodovulum sulfidophilum* (Bamford, 2002*[+])	Thiosulfate oxidase	Dissimilatory thiosulfate oxidation
Sud	*Wolinella succinogenes* (Lin, 2004*[+])	Polysulfide:sulfur transferase	Dissimilatory sulfur reduction
NifS	*A. vinelandii* (Yuvaniyama, 2000), *Thermotoga maritima* (Kaiser, 2000*)	Cysteine desulfurylase / Sulfur transferase	Fe/S cluster assembly
IscS	*E. coli* (Kambampati, 1999)	Cysteine desulfurylase / Sulfur transferase	Fe/S cluster assembly 4-Thiouridine synthesis
ThiI	*E. coli* (Palenchar, 1999)	Sulfur transferase	4-Thiouridine synthesis Thiamine synthesis
SseA	*E. coli* (Spallarosa, 2004)	3-Mercaptopyruvate sulfurtransferase	?

* 3D structures available
[+] Cys-persulfide confirmed by X-ray crystallography / NMR spectroscopy
? unknown function

A common feature in all above mentioned proteins is the central role of the persulfurylated cysteine in catalysis, although it is under dispute whether the Css is an intermediate or is formed prior to catalysis (see, for example, Bamford *et al.*, 2002; Bordo *et al.*, 2000). The substrate sulfur is bound at this cysteine and subsequently transferred from there onto

acceptor molecules. As a consequence, the identification of the persulfurylated Cys30 in the SOR indicates a central role for this residue, most probably acting in substrate sulfur binding.

Additional residues are probably involved in the sulfur disproportionation reaction catalyzed by the SOR. Possible candidates are the polar residues Asn8, Asn44 and Gln75 located at the active site cavity (Fig. 3.14) conserved in all SOR sequences known to date (Fig. 3.8). Further contributing residues are indicated by the binding site of the inhibitor Zn^{2+}, which did not bind within the active site cavity but to His276, part of a 2-His motif located at the end of a narrow pocket guided by mainly hydrophobic residues connected to the active site residue Css30 of the neighboring monomer (Fig. 3.15D). At first sight, the inhibitory effect of Zn^{2+} binding to His276 is enigmatic and several aspects have to be addressed. First, the crystals were soaked with Zn^{2+} at 22°C and not at elevated temperatures, in consequence it is not clear whether the binding of Zn^{2+} to His 276 reflects the situation at elevated temperatures and whether Zn^{2+} then could bind elsewhere in the structure, i.e. at the active site cysteines or could exchange the iron ion at the metal center. On the other hand, the site at His276 was clearly the only binding site of Zn^{2+}, as no other significant additional electron density was found in the structure, which indicates a highly specific binding. Both histidines are conserved in all SOR sequences (Fig. 3.8) and are not involved in a hydrogen bonding network important for structural integrity. A possible explanation for the inhibition might be their involvement in the binding of the sulfur substrate. A long-chain polysulfide, i.e. $HS\text{-}S_6\text{-}S^-$ could fit into the hydrophobic pocket and could bind with its negatively charged terminal sulfur atom to one of the histidines, therefore being aligned for a nucleophilic attack of Css30 on sulfur. Only a small barrier exists between this pocket and the active site cavity being composed of Css30 and Met296 (Fig. 3.15). One could imagine a structural rearrangement of these residues upon substrate binding allowing for the entry of the latter to the active site cavity. On the other hand, a different entry site to the active site cavity does already exist. Additional data are necessary to pinpoint the function of these histidines and to elucidate the inhibitory effect of Zn^{2+}.

3.3.4. THE SUBSTRATE OF THE SOR

I also tried to identify the actual substrate of the SOR. As none of the soaking experiments with cyclic $\alpha\text{-}S_8$ sulfur or polysulfide were successful, these substrates were modeled *in silico* into the active site cavity and the chimney guiding the entry to the inner compartment. Besides cyclic $\alpha\text{-}S_8$ sulfur a linear $^-S\text{-}S_6\text{-}S^-$ species was modeled, although

the chain length of this polysulfide species is hypothetical and depends largely on the equilibrium between α-S_8 sulfur and sulfide (Klimmek *et al.*, 1991). Usually, $^{-}$S-S_2-S^{-} and $^{-}$S-S_3-S^{-} are the predominant polysulfide species, but $^{-}$S-S_6-S^{-} is the first product of the reductive ring opening of α-S_8 sulfur. The structures were modeled manually according to spacial restraints based on Van der Waals distances between protein and substrate molecules.

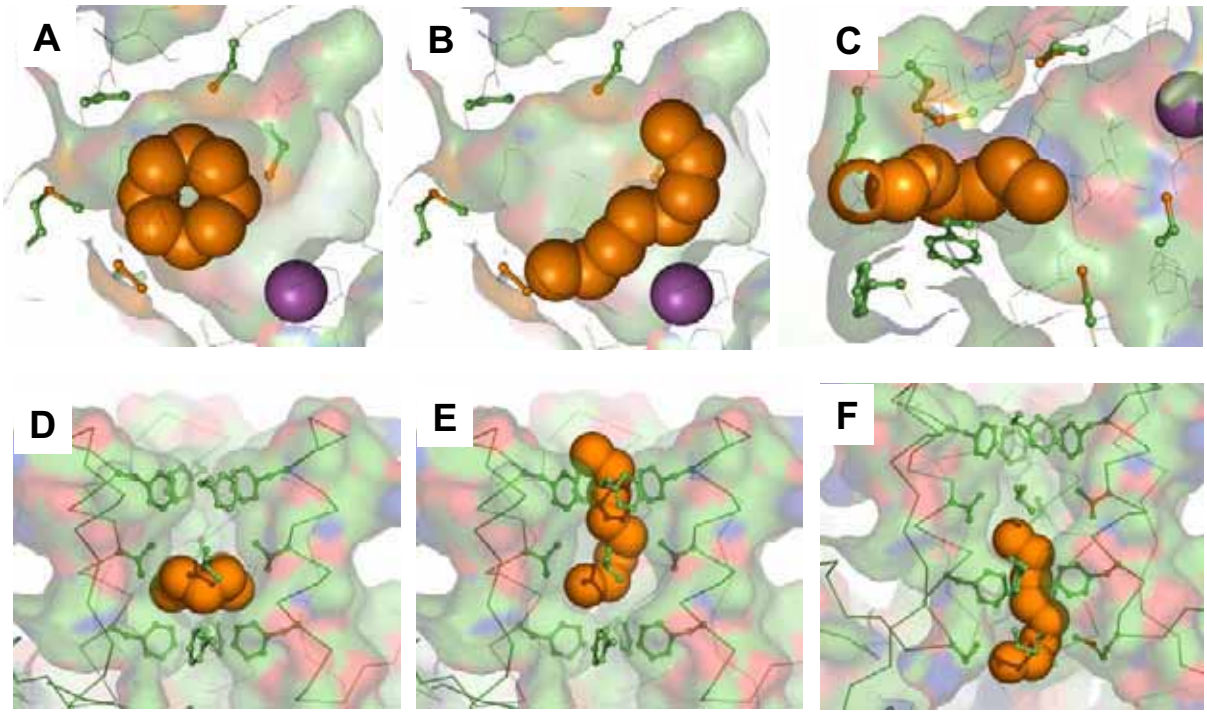


Figure 3.18: Modeling of putative sulfur substrate molecules into SOR. Substrates are represented by their Van der Waals radii. **(A)** - **(C)**: The active site cavity as surface representation. Purple sphere: Fe ion; balls-and-sticks: side chains of gate-keeping residues at the entry pore and active site cysteines. Note that α-S_8 sulfur can be modeled at several positions within the cavity, as it can be done with the linear polysulfide molecule, in addition to several conformations. **(B)** Polysulfide molecule modeled with its S_1 atom in covalent distance to the terminal Sδ of Css 30. **(D)** - **(F)**: The entry to the inner compartment shown as surface representation (solvent probe radius 1.4 Å) showing three out of four contributing monomers. Balls-and-sticks: side chains of residues mainly contributing to the hydrophobic inner chimney surface. Atom color code: carbon, green; sulfur, yellow; oxygen, red; nitrogen, blue.

The dimensions of the active site cavity are sufficient to allow binding of α-S_8 sulfur and the linear polysulfide species (Fig 3.18A, B). In contrast, only the linear polysulfide could pass the narrow entry pore of the cavity (Fig 3.18C), while the α-S_8 molecule would be too voluminous to enter without considerable structural rearrangements of the residues guiding the entrance. The same trend is observed for the chimney-like structure, which is the most probable pathway through the protein shell to the inner compartment. α-S_8 sulfur would fit into the central pocket of the chimney but could not pass the narrow pore above

and the barrier underneath (Fig 3.18D). The linear polysulfide could pass the outer pore and it appears much easier to pass the barrier below, as only minor rearrangements of the gate-keeping side chains of Phe132 could allow for the entry to the inner compartment (Fig 3.18E, F).

These findings point to a linear polysufide species and not to α-S_8 sulfur being the actual substrate of the SOR, as polysulfides are less sterically hindered in entering the inner compartment and the active site cavity. It should be noted that these conclusions are somewhat constrained due to the fact that the crystal structure provides only a snapshot of the enzymes´ structural state and does not reflect the flexibility and dynamics of the protein. Furthermore, SOR behaves differently at 85°C (the temperature optimum of catalysis) than at -163°C (110 K), the temperature of the crystals diffraction measurements.

Additional support for the polysulfide hypothesis comes from the strongly positive charged inner particle surface (Fig. 3.7), whose biological function could be the compensation of the negative charges of the polysulfide molecules. A similar charge complementary had been described for the ferritins, whose inner surfaces are negatively charged to compensate for the positive charges of the stored iron ions (Macedo *et al.*, 2003). In addition, the positive surface charges in the active sites cavity could result in the attraction of the negatively charged polysulfide substrates to the sites of catalysis.

3.3.5. THERMOSTABILITY

One of the striking properties of the SOR is the extreme thermal stability of the protein. The optimal temperature for catalysis was 85°C and the protein was still active at 108°C (Kletzin, 1989). The thermostability of the SOR is even more remarkable, when compared with the proteins´ marginal stability against chemical denaturants, having a free energy difference between the native and the denatured state of ~5 Kcal mol^{-1} at pH 7 and 25°C (see chapter 2, Fig 2.7).

Many factors have been suggested to contribute to the elevated stability of hyperthermophilic proteins. These factors include e.g. an increased number of salt bridges, hydrogen and disulfide bonds, the shortening of surface loops, and a decrease of the accessible surface area (see, for review, e.g. Jaenicke & Böhm, 1998). Determinants of thermostability were mainly identified by comparing the structures of thermophilic and mesophilic protein counterparts. These analyses revealed that thermal stability was achieved by subtle but relevant exchanges at different locations in the structures (Karshikoff & Ladenstein, 2001) rather than by a universal principle. Furthermore, different

solutions for the problem of thermostability were identified in different protein families (Jaenicke & Böhm, 1998).

Since no structure of a mesophilic SOR is available for comparison, the identification of factors contributing to the thermostability of the *A. ambivalens* SOR is somewhat hampered and has to be restricted to rather general factors previously deduced by the comparison of thermophilic and mesophilic proteins.

A global comparison of the primary structure revealed that thermophilic proteins are characterized by a decrease of polar and an increase of charged residues, when compared to mesopilic proteins (Jaenicke & Böhm, 1998). Analysis of the primary structure of the SOR revealed an amino acid composition typical for mesophilic and not for thermophilic proteins (Tab. 3.8).

Table 3.8: The relative amino acid composition of the SOR and of mesophilic and thermophilic proteins. The values for mesophilic and thermophilic proteins were compiled from a global comparison of mesophilic and thermophilic prokaryotic genomes (Singer & Hickey, 2003).

Amino acid	SOR	Mesophiles	Thermophiles
Charged residues (DEKRH)	23.5 %	23.8 %	27.5 %
Polar/uncharged residues (GSTNQYC)	30.1 %	31.2 %	27.6 %
Hydrophobic residues (LMIVWPAF)	46.4 %	45.0 %	44.9 %

The tertiary structure of the SOR is not characterized by a tight packing or a compact surface, as it might be expected from previous studies of hyperthermostable proteins (Pushkaev *et al.*, 2002). In contrast, two cavities are present in each monomer, in addition to an extended loop region including α-helix A4 on top of the β-barrel. This is reflected by the calculation of the solvent-accessible atoms showing that 1305 out of 2518 atoms in the monomer (51.8 %) are solvent-exposed, which is a value typical for mesophilic proteins (Chan *et al.*, 1995). Furthermore, no stabilizing disulfide bridges exist in the SOR.

A factor possibly contributing to the thermostability of SOR might be the unusual duplicated $(\beta\alpha\beta)_2$ ferredoxin fold, as this results in a rigid β-barrel core within the monomer. This hypothesis is in accordance with a study of Hoecker *et al.* (2004), who investigated the stability of the $(\beta\alpha)_8$- or TIM barrel protein HisF from *Thermotoga maritima*. They compared a monomeric HisF having a complete $(\beta\alpha)_8$-barrel fold, with the respective $(\beta\alpha)_4$-barrel protein of identical sequence, which formed the TIM barrel upon dimerization. They found that the monomeric form of HisF had an increased stability against chemical denaturing compared to the homodimeric protein.

Another factor contributing to the thermostability of the SOR might be that the N- and C−termini do not have extended loop regions. Especially the C-terminus is stabilized by three ion pairs.

A total of 15 ion pairs were found within the monomer, in addition to six inter-subunit ion pairs, which gives a value of 6.9 ion pairs / 100 residues (4.9 intra- and 2.0 inter-subunit ion pairs / 100 residues). This value is higher than the average found in proteins (~4, Barlow & Thornton, 1983; Karshikoff & Ladenstein, 2001) and in agreement with the one reported for hyperthermophilic proteins (7 - 10 ion pairs / 100 residues; Karshikoff & Ladenstein, 2001). In consequence, this suggests that ion pairs may contribute to the thermostability of the SOR, although they might not represent the major determinant of thermostability.

A low surface area to volume ratio is considered to be another determinant of extreme thermostability due to the minimization of unfavorable surface energy (Chan *et al.*, 1995). Hyperthermostable proteins often realize this by the association of subunits at the level of the quaternary structure (Jaenicke & Böhm, 1998). This appears to be also the case for the SOR, as within the shell-like icosatetrameric complex approximately 40 % of the monomer surface area is buried from the solvent (Tab. 3.3). Intriguingly, the fraction of solvent-exposed monomer atoms in the complex is reduced from 51.8 % to 27.5 %, which is a remarkably low value (see for comparison Chan *et al.*, 1995).

In summary, it appears that several factors contribute to the thermostability of the *A. ambivalens* SOR, as it has been found in other proteins (Jaenicke & Böhm, 1998). Besides ion pairs, especially the shell-like quaternary structure of the holoenzyme might be a major determinant of the extreme thermostability.

3.3.6. STRUCTURAL INDICATIONS FOR A GENE DUPLICATION / FUSION EVENT DURING SOR EVOLUTION

The tertiary structure of the SOR monomer can be interpreted as a duplication of the $(\beta\alpha\beta)_2$ ferredoxin fold. It shows a remarkable internal 2-fold pseudo-symmetry along the β−barrel axis (Fig. 3.9A-C), that splits the structure into N- and C-terminal half barrels of similar sizes, which are characterized by a very similar fold and topology (Fig. 3.19A). Both halves were superimposed onto each other for a further analysis of their structural homology using the program MODELLER (Sali & Blundell, 1993). It revealed that large parts of both halves could indeed be superimposed (Fig. 3.19B). 107 of the α carbon main chain atoms (70 %) had identical spacial positions (r.m.s. deviation: 1.6 Å), which involved mainly the two βαβ super-secondary structural elements of the ferredoxin fold. Positional differences were restricted to the extended loop regions between the single βαβ elements

and to the C-terminal α-helical/loop regions of both halves. In summary, this analysis highlighted the very similar three-dimensional fold of both halves.

The presence of two structurally very similar domains in one protein had first been described for bovine liver rhodanese (Ploegman *et al.*, 1978) and since then it was found in a vast number of different protein structures (e.g. Lang *et al.*, 2000; Bamford *et al.*, 2002; Abramson *et al.*, 2003), most of which having in common a very low sequence identity between the two domains. This phenomenon is widely accepted to be an evidence for a gene duplication followed by a gene fusion event, which resulted in a duplicated ancestor protein. Subsequently, the sequence similarity between the two domains was lost due to mutational adaption during evolution, whereas the tertiary structure was conserved, because nature selects for structural and functional integrity instead of amino acid conservation (Dietmann & Holm, 2001). Gene duplication is considered to play an important role in enzyme evolution (Doolittle, 1995).

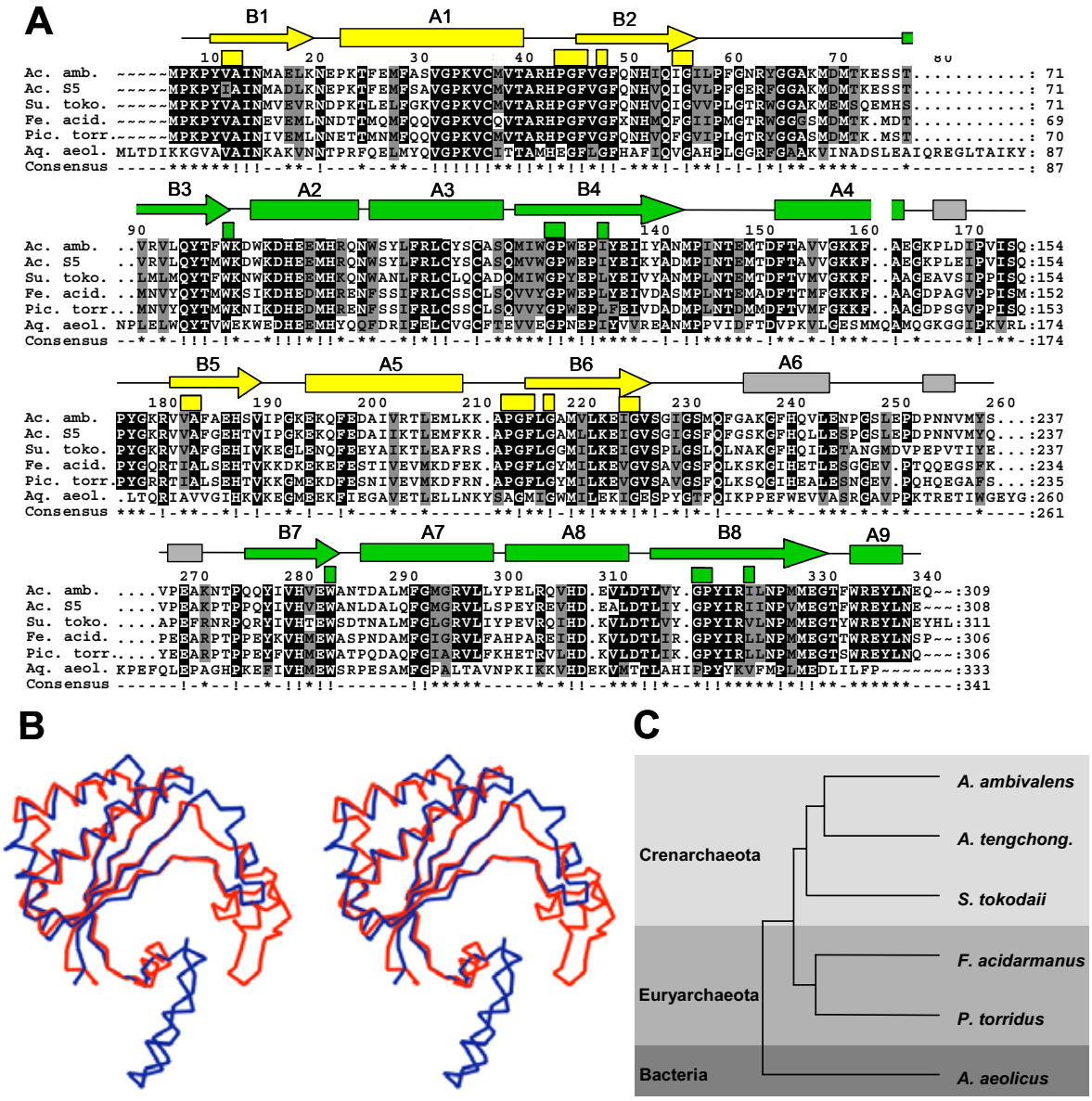


Figure 3.19: Indications for a gene duplication/fusion event. **(A)** Multiple alignment of SOR protein sequences drawn according to the topology of *A. ambivalens* SOR with the upper two and lower two rows representing the N-terminal and C-terminal half barrels, respectively. Topologically similar secondary structural elements in both halves are shown in yellow and green above the alignment, as are identical residues at structurally identical positions in both halves (colored boxes). See Fig. 2.13 for further details of the alignment. **(B)** Stereo view of the superposition of the N-terminal (blue) and C-terminal (red) monomer halves obtained by MODELLER shown as ribbon representation. **(C)** Tree representation of the phyletic relationship between SOR proteins. The phylogenetic position of the respective host organisms is indicated by grey boxes

Evolutionary recent combinatorial events can be detected by sequence resemblances, but for ancient domain rearrangements these resemblances often have been eroded over time (Doolittle, 1995). The sequence identity between the SOR halves is 11.3 % within the structurally equivalent residues (12 out of 107, 6 Pro or Gly) and thus negligible, being indicative for an ancient gene duplication / fusion event (Fig. 3.19A).

The high sequence similarity between the six published SOR sequences points to the presence of the duplicated fold in all of them (Fig. 3.19A). A phylogenetic tree of the sequences is in accordance to the phylogenetic positions of their host organisms based on 16S rDNA analysis (Fig. 3.19C), which indicates that no horizontal transfer of the *sor* gene has occurred between these organisms and that a duplicated ancestral gene might already have been present in a common ancestor organism. These findings date back the gene duplication / fusion event prior to the splitting of the archaeal and bacterial lineages, probably more than two billion years ago (Woese, 2000).

From these data, a model for the evolution of the SOR protein family can be proposed, where a gene duplication of an ancestral half barrel sequence is followed by a gene fusion into the complete barrel gene and a mutational adaption of the initially identical half barrel sequences (Fig. 3.20).

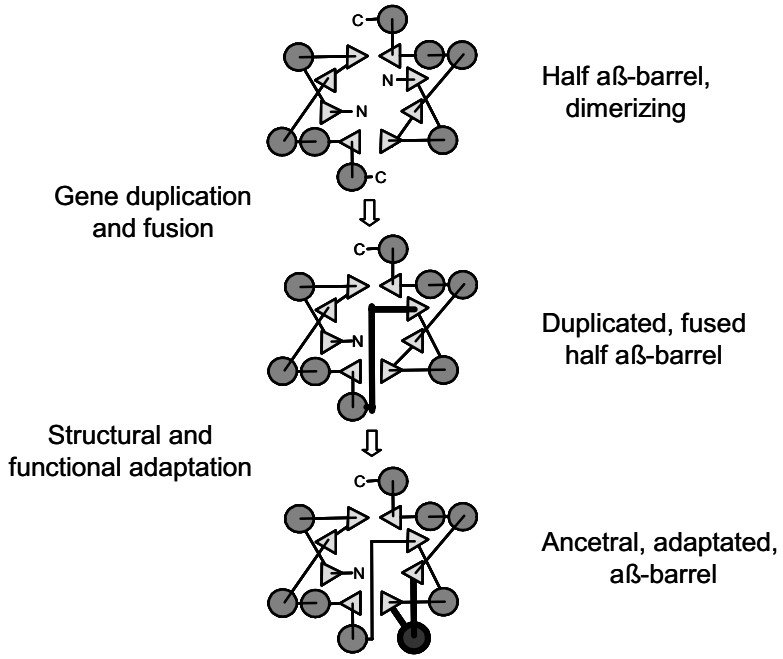


Figure 3.20: Topology diagram showing a hypothetical model for the evolution of SOR αβ-barrel by gene duplication and fusion (adopted from Lang *et al.*, 2000). Initially, an ancestral SOR half barrel exists that may dimerize into the full barrel. A gene duplication/fusion event generates two fused half barrels, which are initially identical and subsequently adapted to form the ancestral SOR monomer. The topology scheme of SOR shows β strands as triangles and α helices as circles; evolutionary events are highlighted in bold.

In conclusion, the SOR proteins can be seen as another example for the evolution of new protein types by duplication and fusion of a precursor protein gene followed by divergent evolution.

3.4. APPENDIX

The atomic coordinates of the SOR crystal structures are not yet publicly available via the Protein Data Bank (PDB, Berman *et al.*, 2000), therefore they are provided in the supplementary archive "urich_suppl.zip". The archive contains the coordinate files (in PDB format) of the model in the asymmetric unit and of the holoenzyme model. Furthermore, it contains the models of structures determined after incubation of the SOR with several inhibitors, substrates and products of the enzymatic reaction.

Table 3.9: The coordinate files in the supplementary archive "urich_suppl.zip".

Coordinate file	Description
sor_au.pdb	SOR model in the asymmetric unit
sor_holo.pdb	SOR holoenzyme
sor_ditred.pdb	SOR dithionite-reduced
sor_hg.pdb	SOR complexed with p-hydroxy mercury bezoic acid
sor_zn.pdb	SOR complexed with zinc actetate
sor_ioac.pdb	SOR complexed with iodoacetamide
sor_sinc.pdb	SOR incubated and cocrystallized with elemental sulfur
sor_scoc.pdb	SOR cocrystallized with elemental sulfur
sor_ps.pdb	SOR soaked with polysulfide solution
sor_sft.pdb	SOR soaked with potassium sulfite solution
sor_sfd.pdb	SOR soaked with potassium sulfide solution
sor_ts.pdb	SOR soaked with potassium thiosulfate solution

The coordinate files can be visualized with molecular graphics programs like RASMOL or PYMOL (see http://www.pdb.mdc-berlin.de/pdb/software-list.html#Graphics for links to various molecular graphics programs).

4. SITE-DIRECTED MUTAGENESIS OF ACTIVE SITE RESIDUES

4.1. INTRODUCTION

Despite a wealth of biochemical and structural information obtained (see chapters 2 and 3), details on the catalytic reaction of the SOR were still limited. In the previous chapters it was shown that a mononuclear non-heme iron site is involved in catalysis, which is coordinated by two histidines (His85 and His89) and a glutamate residue (E113). Furthermore, the cysteine residues Cys30, Cys100 and Cys103 are considered to be involved in catalysis. All elements are located in close vicinity to each other along a cavity within each SOR monomer, thus defining the enzyme´s active site.

Especially for the three cysteines, no clear function could be assigned yet, although they most probably are involved in sulfur binding and/or catalysis. Therefore, a mutational study was initiated to investigate their catalytic function. After the construction of the mutated *sor* genes the proteins were purified and analyzed for their activity and iron content.

A more detailed picture of the SOR´s sulfur disproportionation reaction could be drawn from the results.

4.2. RESULTS

The residues postulated to be involved in catalysis were specifically mutated. The cysteines C30, C100 and C103 were independently exchanged with the analogous polar amino acid serine and with the apolar alanine. Besides single mutants, double mutants of C100 and C103 to alanine and serine were also produced. Furthermore, the iron ligands H85, H89 and E113 were exchanged with alanine to investigate their role in iron binding. E113 was additionally exchanged to the analogous acidic amino acid aspartate (Tab. 4.1). In this chapter, the term wild type SOR refers to the non-mutated SOR purified from *E. coli* and not to the SOR purified from *A. ambivalens*.

Table 4.1: The twelve mutants constructed and analyzed.

C30A	C100A	C103A	C100/103A	H85A	E113A
C30S	C100S	C103S	C100/103S	H89A	E113D

4.2.1. CONSTRUCTION OF SOR MUTANTS BY SITE-DIRECTED MUTAGENESIS

The mutated *sor* genes were constructed according to the overlap-extension method (Ho *et al.*, 1989). Forward and reverse mutagenesis oligonucleotides carrying the desired mutations and flanking oligonucleotides were combined in two separate PCR with the pASK-SOR.05 plasmid as template. The amplified DNA fragments showed in all cases the expected sizes (Fig. 4.1). The forward and reverse PCR products were used as template in a third PCR with the flanking oligonucleotides to yield the complete mutated *sor* gene of 960 bp in length (Fig. 4.2)

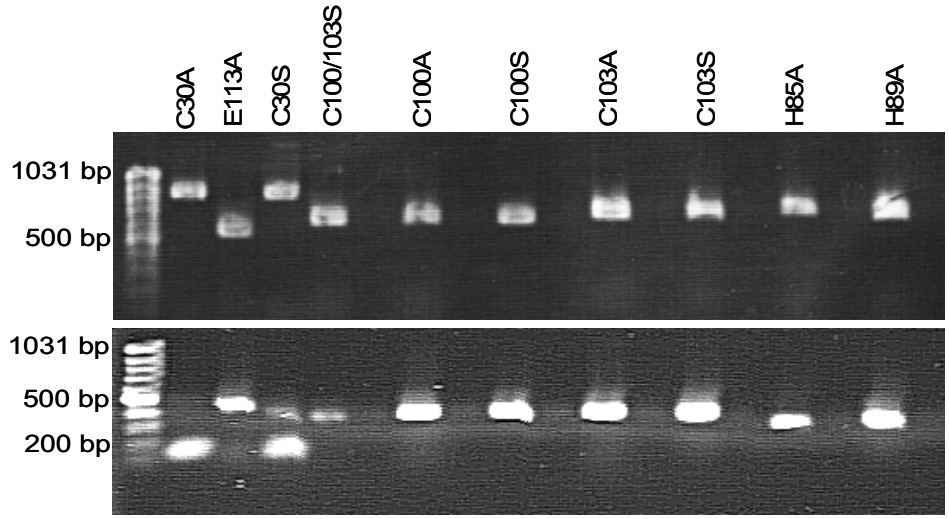


Figure 4.1: Ethidiumbromide-stained 1% TAE agarose gels with amplified DNA from the mutagenesis PCR. **Upper panel**: PCR products of forward mutagenesis oligonucleotides as indicated above; **lower panel**: PCR products of reverse mutagenesis oligonucleotides. All products exhibited the expected sizes. E113D and C110/103A gave similar results (not shown).

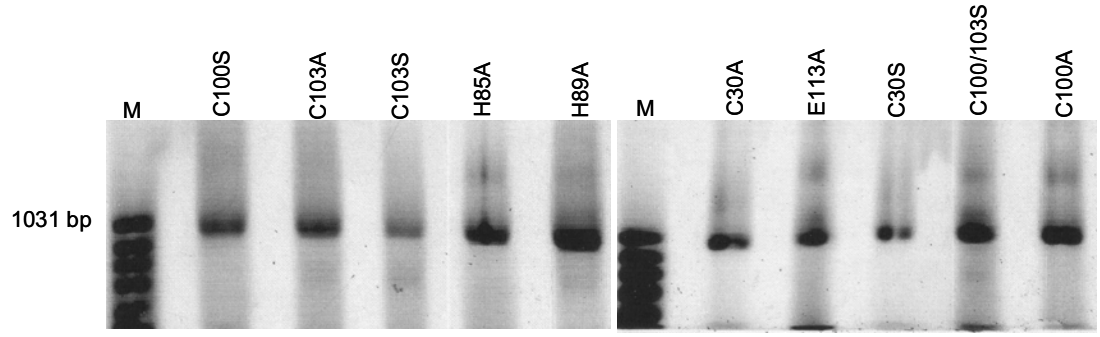


Figure 4.2: 1% TAE agarose gels, cyber green-stained with amplified DNA from the fusion PCR reactions. All products exhibited the expected size of 960 bp. E113D and C110/103A gave similar results (not shown).

The desired DNA fragments were isolated and digested with *Eco*47III and *Xba*I and after purification used for the ligation with the dephosphorylated pASK75 vector digested with the same enzymes. After transformation of competent *E. coli* cells, transformants carrying the desired inserts were detected by plasmid mini prep followed by restriction with *Xba*I and *Eco*47III or by colony PCR. Positive constructs were transformed into *E. coli* BL 21 Codon plus cells (DE3-RIL) and subsequently sequenced for the verification of the introduced mutations. All constructs carried the desired mutations (see appendix, section 4.4). The construct C30S carried an additional mutation where the codon for Leu51 (TTG) was exchanged to a ser codon (TCG). Since the Leu51 residue is far away from the active site, no influence of the additional mutation was expected.

4.2.2. HETEROLOGOUS GENE EXPRESSION AND PURIFICATION

Cultures were inoculated from single colonies of the clones and the expression of the sor genes was induced at OD ~0.6 -1 by the addition of anhydrotetracycline to the medium. After cell harvesting and breaking, the desired soluble proteins were purified by heat treatment and Strep-Tactin affinity chromatography (see chapter 6.2.2.1 for details) using an FPLC-connected Superflow column (V_c 10 ml) or gravity flow columns (V_c 1 ml). After affinity chromatography, the proteins were apparently homogenous (>95 %), as judged by SDS gels (Fig. 4.3). Minor smaller bands were attributable to proteolytic digestion of the major subunit, as the same bands appeared also in wild type SOR preparations (see chapter 2.2.4). The yield ranged from 0.5 mg to 10 mg pure protein per liter *E. coli* culture.

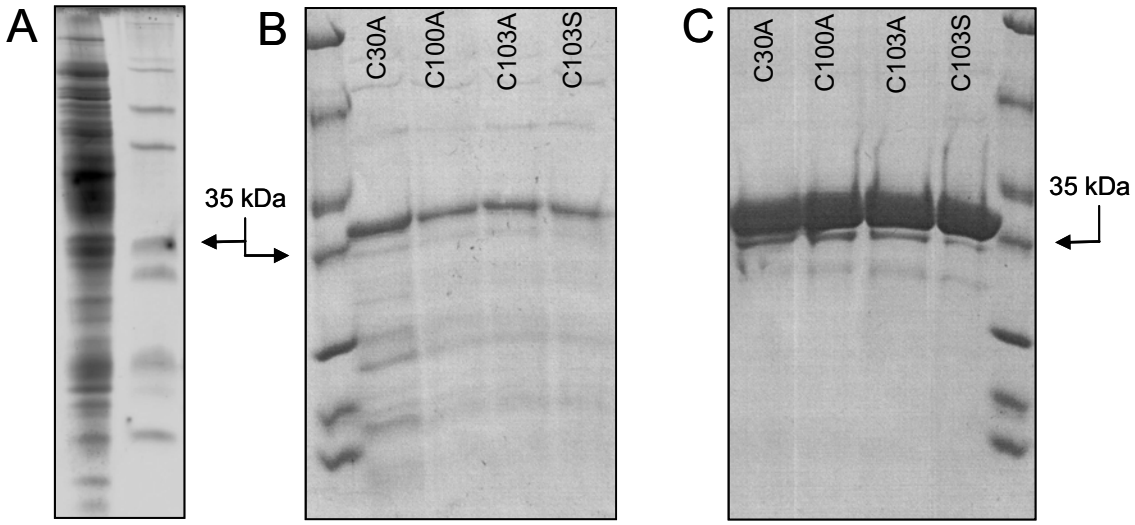


Figure 4.3: Exemplified 10% SDS polyacrylamide gels stained with colloidal coomassie blue. **(A)** soluble crude extract, **(B)** supernatants after heat-treatment, **(C)** Strep-Tactin eluates. Amounts of protein: A 20 µg; B 5 µg, C 10 µg.

4.2.3. CHARACTERIZATION OF THE MUTANTS

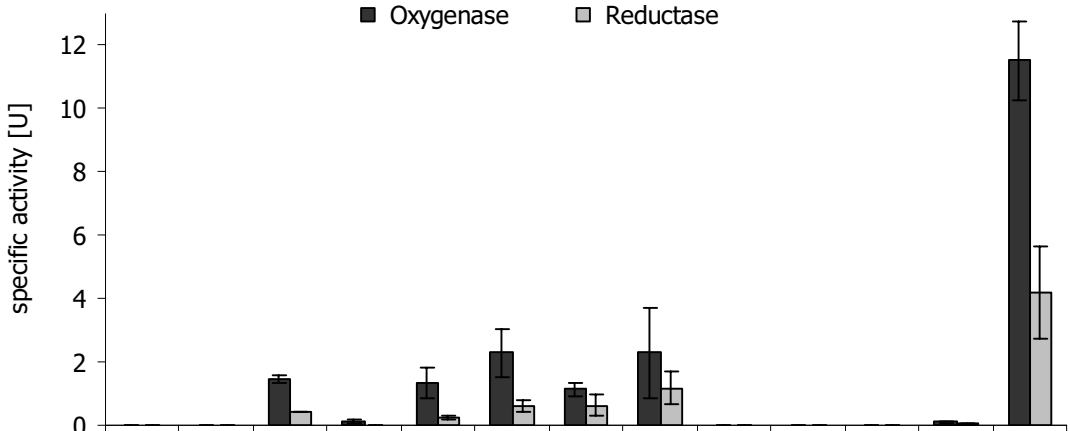
The purified mutant proteins were analysed in respect to their specific activity and iron content to investigate the effects of the point mutations on the catalytic activity and on the iron incorporation.

4.2.3.1 Activity measurements

All mutants were prepared independently at least twice from *E. coli* cells and analyzed for activity. The purified proteins were incubated in a sulfur containing buffer and the amounts of thiosulfate (oxygenase) and sulfide (reductase) were determined (see chapter 6.2.2.4 for details). In most cases 5 µg protein / ml buffer was incubated for 0, 4, 8, 12, and

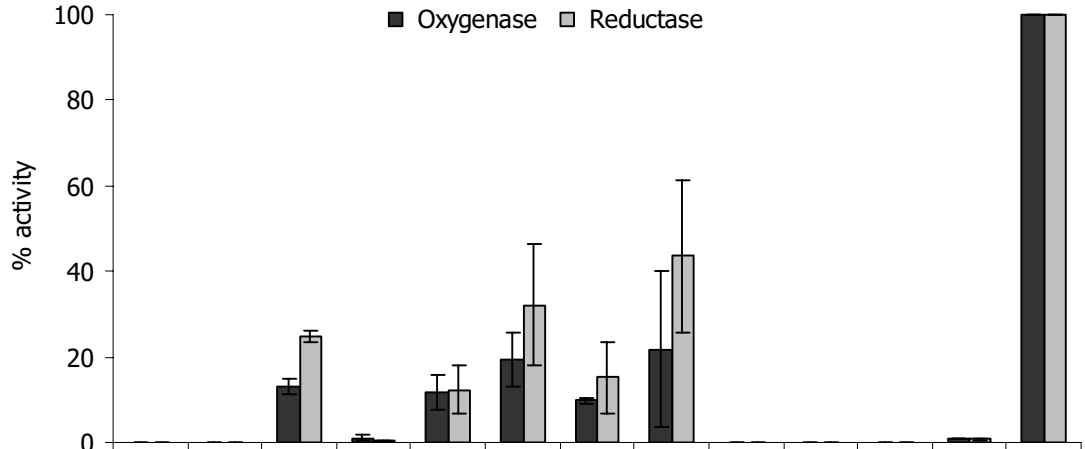
16 min and the specific activity was calculated from the linear increase of product concentrations. The protein concentration was increased up to 50 µg protein / ml in cases of mutants with low activity. Appropriate positive and negative controls were determined for each series. In all samples, a lowered reductase activity was observed, when compared to the oxygenase activity. This was caused by the quick evaporation of hydrogen sulfide from the assay buffer (Kletzin, 1989), thus preventing the exact determination of the reductase activity (see chapter 6.2.5 for details).

The results of the activity assays are summarized in Fig. 4.4 and Fig. 4.5. In no case the oxygenase or reductase activity were affected separately. The mutants C30A, C30S, H85A, H89A and E113A showed a complete loss of oxygenase and reductase activity. E113D yielded little residual activity of 0.11 U (oxygenase) and 0.03 U (reductase), as had C100S (0.12 U oxygenase, 0.02 U reductase). The mutants C100A, C103A, C103S, C100/103A and C100/103S yielded higher residual activities, ranging between 1.12 U and 2.29 U for the oxygenase and 0.23 U and 1.17 U for the reductase activity. When comparing relative activities, E113D and C100S had 1.1 % of the wild type oxygenase activity and even lower reductase activities. C100A had a more than tenfold increased activity compared to C100S (13% oxygenase, 25 % reductase activity). The mutants C103S and C100/103S had a approximately twice as high activity as C103A and C100/103A. The oxygenase activities ranged from 10 to 20 % and the reductase activities from 10 to 40 % of the wild type activity.



	C30A	C30S	C100A	C100S	C103A	C103S	C100/ 103A	C100/ 103S	H85A	H89A	E113A	E113D	WT
Oxygenase	0,00	0,00	1,47	0,12	1,35	2,28	1,12	2,29	0,00	0,00	0,00	0,11	11,518
Reductase	0,00	0,00	0,43	0,02	0,23	0,60	0,64	1,17	0,00	0,00	0,00	0,03	4,19

Figure 4.4: Specific activities of SOR mutants. Dark grey: oxygenase activity, light grey: reductase activity. The activities are mean values of 2 - 5 measurements using independently purified protein preparations; the respective error bars are shown. Exact values are given below the chart. WT: non-mutated SOR purified from *E. coli*.



	C30A	C30S	C100A	C100S	C103A	C103S	C100/ 103A	C100/ 103S	H85A	H89A	E113A	E113D	WT
Oxygenase	0,0	0,0	13,0	1,1	11,6	19,5	9,8	21,8	0,0	0,0	0,0	1,1	100,0
Reductase	0,0	0,0	24,9	0,4	12,3	32,1	15,2	43,6	0,0	0,0	0,0	0,8	100,0

Figure 4.5: Relative activities as % wild type SOR activity. Dark grey: oxygenase activity, light grey: reductase activity. Exact values are given below the chart. WT: non-mutated SOR purified from *E. coli*.

4.2.3.2 Iron quantification

The iron content of the mutants was determined chemically by the 2,4,6-tripyridyl-1,3,5-triazine (TPTZ) method (Fischer & Price, 1964) to investigate the effect of the introduced mutations on iron incorporation (see chapter 6.2.2.6 for details). Duplicates of samples were quantified. Adsorption changes below 0.01 at 593 nm (corresponding to iron

amounts below 0.5 nmol) were set to zero due to the detection limit of the TPTZ assay and the spectrophotometer used.

Table 4.2: Iron / protein ratios for SOR wild type and mutants calculated as mol iron per mol monomer. ± error. WT: non-mutated SOR purified from *E. coli*.

Sample	WT	C30A	C30S	C100A	C100S	C103A	C103S
Fe / Protein ratio	0.45 ± 0.05	0.47 ± 0.05	0.42 ± 0.03	0.22 ± 0.01	0 *	0.19 ± 0.01	0.30 ± 0.06

Sample	C100/103A	C100/103S	H85A	H89A	E113A	E113D
Fe / Protein ratio	0.56 ± 0.11	0.40 ± 0.03	0 *	0 *	0 *	0.02 ± 0.003

* Fe content below detection limit

The results of the iron quantifications are summarized in Tab. 4.2 and Fig. 4.6. The wild type SOR had an iron content of 0.45 mol/mol, which is lower than one mol Fe per mol SOR monomer, as it would be expected for full iron incorporation into each monomer (see for comparison, chapter 2.2.6 and Urich *et al.*, 2004). A possible explanation is the high level expression in *E. coli*, which due to iron limitations, eventually prevented the complete occupation of the iron site in the wild type SOR purified from *E.coli*.

The C30A, C30S, C100/103A and C100/103S mutants had Fe contents comparable to the wild type SOR, whereas in the cases of C100A, C103A and C103S the Fe content was significantly reduced to ratios between 0.19 and 0.30 mol/mol. No iron was detected in the mutants C100S, H85A, H89A and E113A. E113D had a low iron content of 0.02 mol/mol. EPR spectroscopy of this mutant gave a weak g=4.3 signal as found in the wild type SOR (Fig. 4.7, see chapter 2.2.7 for comparison).

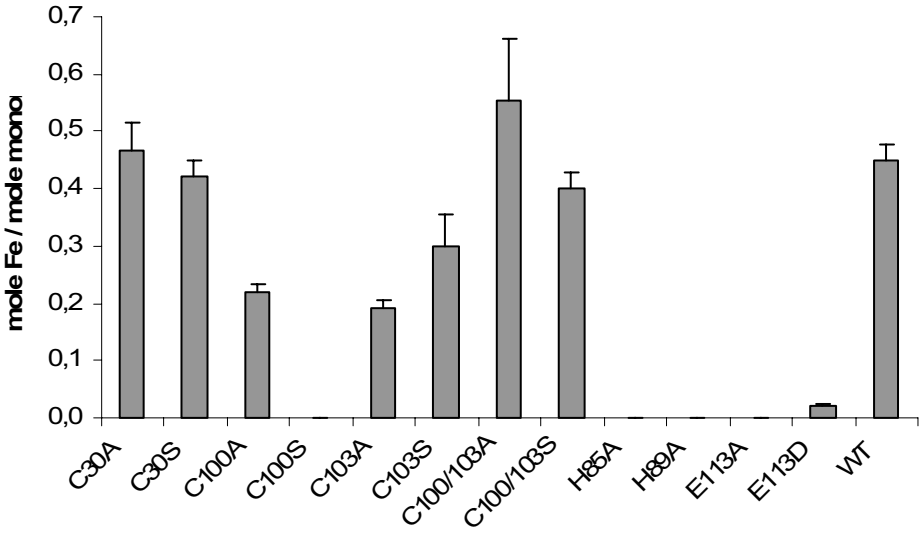


Figure 4.6: Graphical representation of Fe / protein ratios of mutants. WT: non-mutated SOR produced in *E. coli*.

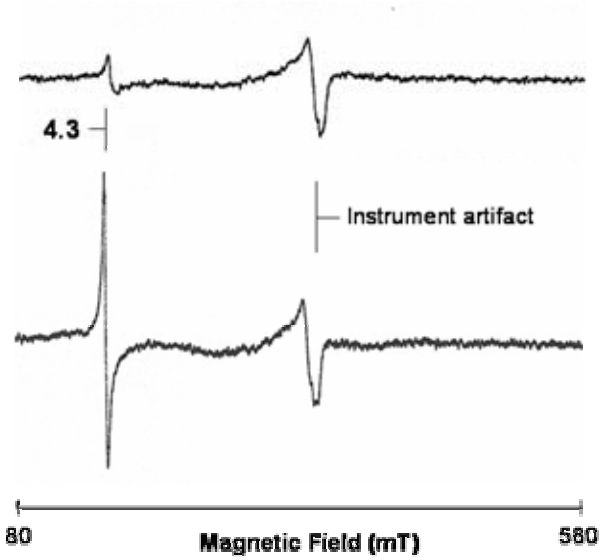


Figure 4.7: EPR spectra. Upper spectrum: SOR E113D mutant. Lower spectrum: SOR WT. Temperature, 5.5 K; protein concentration, 200 μM; microwave power – 2.4 mW; modulation amplitude – 1 mT. Baseline subtraction not performed.

4.3. DISCUSSION

The function of active site residues of SOR was investigated by site-directed mutagenesis. The purified mutant proteins were characterized with respect to activity and iron content. Fig. 4.8 summarizes the results of both analyses.

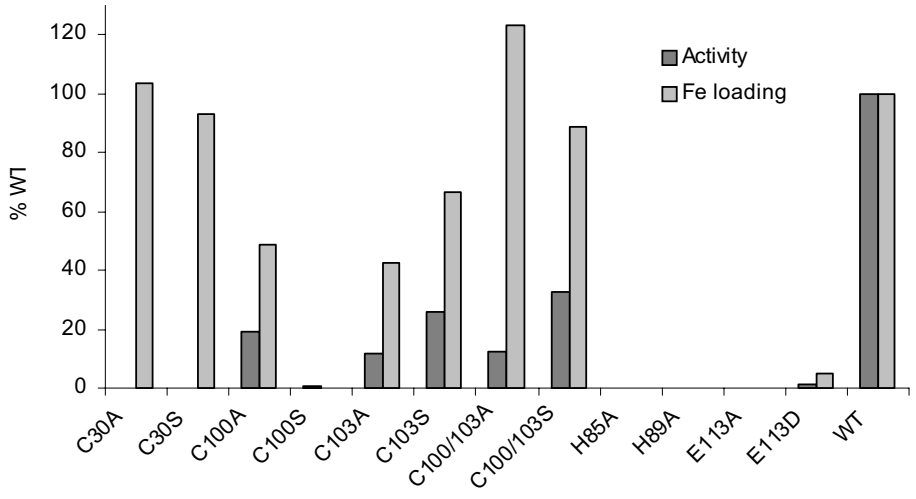


Figure 4.8: Comparison of SOR activity and iron loading of mutants. Values are shown as % of the wild type control. The activities are mean values of oxygenase and reductase partial reaction. WT: non-mutated SOR purified from *E. coli*.

4.3.1. MUTANTS OF IRON LIGANDS

The mutation of each of the iron ligands H85, H89 and E113 to alanine lead to inactive enzyme with no iron detectable (Fig. 4.8). In contrast, an E113D mutant had a residual activity of approximately 1 % of the recombinant wild type enzyme and showed both sulfur-oxidizing and sulfur-reducing activities. The iron content of the mutant was likewise drastically reduced to 4 % of the wild type.

These results showed that the exchange of one of the members of the iron chelating motif to a non-functional alanine residue abolished its ability to bind the metal and that the loss of activity was caused by the loss of the mutants´ iron-binding capability. The replacement of the glutamate by an analogous aspartate did not abolish enzyme activity and showed that the lowered activity of the E114D mutant was the result of a low occupancy of iron at the iron site. In summary, it appears that the H-X_3-H-X_{23}-E motif conserved in all SOR proteins is highly optimized for iron binding and that any exchange results in a drastically loss of affinity to the metal.

4.3.2. MUTANTS OF CYSTEINES

The mutants C30A and C30S both did not have any activity despite an iron loading comparable to the wild type control (Fig. 4.8), showing that the lack of iron incorporation was not the cause for the loss of activity. In consequence, C30 must be essential for the SOR´s catalytic mechanism.

The mutation of the C100 and C103 residues to alanine and serine revealed a more complex picture both for activity and iron content. Five of the six mutants had residual activities ranging between 9.8 and 21.8 % (oxygenase) and 12.3 and 43.6 % (reductase) of the wild type, respectively (Fig. 4.8). The sole exception was C101S with an almost complete loss of activity (1.1 % oxygenase and 0.4 % reductase). The iron content of this mutant was below the detection limit, showing that the amino acid exchange strongly affected the iron incorporation and thus lead to the observed low activities (Fig 4.8). In contrast, an exchange of C100 to alanine resulted in higher residual activity (19 %) and iron content (44 %, Fig. 4.8), indicating that the presence of the polar serine residue interfered stronger than alanine with the iron incorporation. Mutation of C103 to alanine or serine resulted in reduced activities and iron contents comparable to C100A. Therefore, from these results no definite statement can be put forward towards a role in catalysis. In contrast, the alanine and serine double mutants had reduced activities despite iron contents comparable to the wild type control (Fig. 4.8), clearly showing that the effect of the mutations on activity is not caused by a reduced iron incorporation. In consequence, C100 and C103 participate in catalysis, although they are, unlike C30, not essential.

Recently, it has been reported for the SOR from *A. tengchongensis* that the mutation of C30 to alanine or serine resulted in a complete loss of activity, whereas the mutation of C100 and C103 to serine retained 1.6 % and 0.8 % of activity (Chen *et al.*, 2005). The iron contents of the mutants were not reported. These results are in good agreement with the ones presented here, although higher residual activities for most of the C100 and C103 mutants were found in this work.

One important point to mention is that the oxygenase and reductase activities were parallely affected in all mutants. A separation of both activities was never observed, implying that both activities are inseparably coupled.

In summary, it can be concluded that the minimal catalytic unit of the SOR consists of the iron site and Cys30. No evidence was found for a spacial separation of the oxygenase and reductase partial reactions, which suggests that the SOR is a true sulfur-disproportionating enzyme.

4.4. APPENDIX

Pair-wise alignments of the mutated *sor* gene sequences against the non-mutated sequence (pASK-sor.05). Mutated codons are shaded in grey with the original and resulting amino acid indicated below and above, respectively. The translational start is underlined for orientation.

pASK-sorC30A

```
                .         .         .         .         .
C30A:  1 .....................................AAATGCCGaAACCA 8
                                              |||||||||||||
WT:  101 GAATAGTTCGACAAAAATCTAGATAACGAGGGCAAAAAATGCCGAAACCA 150
                .         .         .         .         .
       9 TACGTTGCTATAAACATGGCAGAATTAAAGAATGAACCTAAAACTTTTGA 55
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     151 TACGTTGCTATAAACATGGCAGAATTAAAGAATGAACCTAAAACTTTTGA 200
                .         .         A         .         .
      56 AATGtTTGCCTCAGTAGGACCGAAGGTCGCCATGGTAACAGCAAGGCATC 105
         ||||||||||||||||||||||||||||  ||||||||||||||||||||
     201 AATGTTTGCCTCAGTAGGACCGAAGGTCTGCATGGTAACAGCAAGGCATC 250
                .         .         C         .         .
     106 CGGGCTTTGTTGGTTTTCAAAACCATATACAAATAGGAATTTTGCCATTC 155
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     251 CGGGCTTTGTTGGTTTTCAAAACCATATACAAATAGGAATTTTGCCATTC 300
                .         .         .         .         .
     156 GGAAACAGATACGGCGGAGcTAAAATGGACATGACTAAGGAAAGTAGTAC 205
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     301 GGAAACAGATACGGCGGAGCTAAAATGGACATGACTAAGGAAAGTAGTAC 350
                .         .         .         .         .
     206 TGTCAGAGTTTTACAGTACACCTTCTGGAAAGATTGGAAAGACCATGAAG 255
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     351 TGTCAGAGTTTTACAGTACACCTTCTGGAAAGATTGGAAAGACCATGAAG 400
                .         .         .         .         .
     256 AAATGCACAGGCAAAACTGGAGTTACTTATTCAGGCTATGCTATTCATGC 305
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     401 AAATGCACAGGCAAAACTGGAGTTACTTATTCAGGCTATGCTATTCATGC 450
                .         .         .         .         .
     306 GCTTCACAAATGATATGGGGACCCTGGGAGCCAATTTATGAAATAATCTA 355
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     451 GCTTCACAAATGATATGGGGACCCTGGGAGCCAATTTATGAAATAATCTA 500
                .         .         .         .         .
     356 CGCAAACATGCCTATAAACACTGAAATGACCGaCTTCACTGCAGTTGTAG 405
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     501 CGCAAACATGCCTATAAACACTGAAATGACCGACTTCACTGCAGTTGTAG 550
                .         .         .         .         .
     406 GAAAGAAGTTCGCAGAAGGAAAGCCTTTAGATATTCCAGTTATTTCACAA 455
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     551 GAAAGAAGTTCGCAGAAGGAAAGCCTTTAGATATTCCAGTTATTTCACAA 600
                .         .         .         .         .
     456 CCATATGGAAAGAGAGTTGTTGCCTTTGCAGAGCACTCAGTAATTCCAGG 505
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     601 CCATATGGAAAGAGAGTTGTTGCCTTTGCAGAGCACTCAGTAATTCCAGG 650
                .         .         .         .         .
     506 CAAAGAGAAGCAATTTGAGGACGCAATAGTTAGGACTTTAGAAATGTTAA 555
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     651 CAAAGAGAAGCAATTTGAGGACGCAATAGTTAGGACTTTAGAAATGTTAA 700
```

```
                .         .         .         .         .
556 AGAAAGCTCCTGGCTTCTTAGGTGCAATGGTATTAAAGGAAATAGGAGTT 605
    |||||||||||||||||||||||||||||||||||||||||||||||||||
701 AGAAAGCTCCTGGCTTCTTAGGTGCAATGGTATTAAAGGAAATAGGAGTT 750
                .         .         .         .         .
606 TCCGGAATTGGAAGCATGCAATTCGGTGCCAAGGGATTCCATCAAGTCTT 655
    |||||||||||||||||||||||||||||||||||||||||||||||||||
751 TCCGGAATTGGAAGCATGCAATTCGGTGCCAAGGGATTCCATCAAGTCTT 800
                .         .         .         .         .
656 AGAGAACCCTGGATCACTTGAGCCAGATCCAAATAATGTAATGTATTCAG 705
    |||||||||||||||||||||||||||||||||||||||||||||||||||
801 AGAGAACCCTGGATCACTTGAGCCAGATCCAAATAATGTAATGTATTCAG 850
                .         .         .         .         .
706 TCCCAGAAGCAAAGAATACTCCACAACAATACATAGTTCATGTAGAATGG 755
    |||||||||||||||||||||||||||||||||||||||||||||||||||
851 TCCCAGAAGCAAAGAATACTCCACAACAATACATAGTTCATGTAGAATGG 900
                .         .         .         .         .
756 GCAAATACTGATGCTTTAATgTTTGGAATGGGTAGAGTACTaTTATATCC 805
    |||||||||||||||||||||||||||||||||||||||||||||||||||
901 GCAAATACTGATGCTTTAATGTTTGGAATGGGTAGAGTACTATTATATCC 950
```

pASK-sorC30S

```
                .         .         .         .         .
C30S : 1 .................TCTAGATAACGAGGGCAAAAAATGCCGAAACCA 33
                          |||||||||||||||||||||||||||||||||
WT:  101 GAATAGTTCGACAAAAATCTAGATAACGAGGGCAAAAAATGCCGAAACCA 150
                .         .         .         .         .
      34 TACGTTGCTATAAACATGGCAGAATTAAAGAATGAACCTAAAACTTTTGA 83
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     151 TACGTTGCTATAAACATGGCAGAATTAAAGAATGAACCTAAAACTTTTGA 200
                .         .       S         .         .
      84 AATGTTTGCCTCAGTAGGACCGAAGGTCTCCATGGTAACAGCAAGGCATC 133
         |||||||||||||||||||||||||||| ||||||||||||||||||||
     201 AATGTTTGCCTCAGTAGGACCGAAGGTCTGCATGGTAACAGCAAGGCATC 250
                .         .         C         .  S      .
     134 CGGGCTTTGTTGGTTTTCAAAACCATATACAAATAGGAATTTCGCCATTC 183
         |||||||||||||||||||||||||||||||||||||||||| |||||||
     251 CGGGCTTTGTTGGTTTTCAAAACCATATACAAATAGGAATTTTGCCATTC 300
                .         .         .         .  L      .
     184 GGAAACAGATACGGCGGAGCTAAAATGGACATGACTAAGGAAAGTAGTAC 233
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     301 GGAAACAGATACGGCGGAGCTAAAATGGACATGACTAAGGAAAGTAGTAC 350
                .         .         .         .         .
     234 TGTCAGAGTTTTACAGTACACCTTCTGGAAAGATTGGAAAGACCATGAAG 283
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     351 TGTCAGAGTTTTACAGTACACCTTCTGGAAAGATTGGAAAGACCATGAAG 400
                .         .         .         .         .
     284 AAATGCACAGGCAAAACTGGAGTTACTTATTCAGGCTATGCTATTCATGC 333
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     401 AAATGCACAGGCAAAACTGGAGTTACTTATTCAGGCTATGCTATTCATGC 450
                .         .         .         .         .
     334 GCTTCACAAATGATATGGGGACCCTGGGAGCCAATTTATGAAATAATCTA 383
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     451 GCTTCACAAATGATATGGGGACCCTGGGAGCCAATTTATGAAATAATCTA 500
                .         .         .         .         .
     384 CGCAAACATGCCTATAAACACTGAAATGACCGACTTCACTGCAGTTGTAG 433
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     501 CGCAAACATGCCTATAAACACTGAAATGACCGACTTCACTGCAGTTGTAG 550
```

```
                .         .         .         .         .
     434 GAAAGAAGTTCGCAGAAGGAAAGCCTTTAGATATTCCAGTTATTTCACAA 483
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     551 GAAAGAAGTTCGCAGAAGGAAAGCCTTTAGATATTCCAGTTATTTCACAA 600
                .         .         .         .         .
     484 CCATATGGAAAGAGAGTTGTTGCCTTTGCAGAGCACTCAGTAATTCCAGG 533
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     601 CCATATGGAAAGAGAGTTGTTGCCTTTGCAGAGCACTCAGTAATTCCAGG 650
```

pASK-sorH85A

```
                .         .         .         .         .
H85A:  1 ...........CAAAAATCtAGATAACGAGGGCAAAAAATGCCGAAACCA 39
                    |||||||||||||||||||||||||||||||||||||
WT:  101 GAATAGTTCGACAAAAATCTAGATAACGAGGGCAAAAAATGCCGAAACCA 150
                .         .         .         .         .
      40 TACGTTGCTATAAACATGGCAGAATTAAAGAATgAACCTAAAACTTTTGA 89
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     151 TACGTTGCTATAAACATGGCAGAATTAAAGAATGAACCTAAAACTTTTGA 200
                .         .         .         .         .
      90 AATGTTTGCCTCAGTAGGACCgAAGGTCTGCATGGTAACAgCAAGGCATC 139
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     201 AATGTTTGCCTCAGTAGGACCGAAGGTCTGCATGGTAACAGCAAGGCATC 250
                .         .         .         .         .
     140 CGGGCTTTGTTGGTTTTCAAAACCATATACAAATAGGAATTTTGCCATTC 189
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     251 CGGGCTTTGTTGGTTTTCAAAACCATATACAAATAGGAATTTTGCCATTC 300
                .         .         .         .         .
     190 GGAAACAGATACGGCGGAGcTAAAATGGACATGACTAAGGAAAGTAGTAC 239
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     301 GGAAACAGATACGGCGGAGCTAAAATGGACATGACTAAGGAAAGTAGTAC 350
                .         .         .         .    A    .
     240 TGTCAGAGTTTTACAGTACACCTTCTGGAAAGATTGGAAAGACGCTGAAG 289
         |||||||||||||||||||||||||||||||||||||||||||  |||||
     351 TGTCAGAGTTTTACAGTACACCTTCTGGAAAGATTGGAAAGACCATGAAG 400
                .         .         .         .    H    .
     290 AAATGCACAGGCAAAACTGGAGTTACTTATTCAGGCTATGCTATTCATGC 339
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     401 AAATGCACAGGCAAAACTGGAGTTACTTATTCAGGCTATGCTATTCATGC 450
                .         .         .         .         .
     340 GCTTCACAAATGATATGGGGACCCTGGGAGCCAATTTATGAAATAATCTA 389
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     451 GCTTCACAAATGATATGGGGACCCTGGGAGCCAATTTATGAAATAATCTA 500
                .         .         .         .         .
     390 CGCAAACATGCCTATAAACACTGAAATGACCGACTTCACTGCAGTTGTAG 439
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     501 CGCAAACATGCCTATAAACACTGAAATGACCGACTTCACTGCAGTTGTAG 550
                .         .         .         .         .
     440 GAAAGAAGTTCGCAGAAGGAAAGCCTTTAGATATTCCAGTTATTTCACAA 489
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     551 GAAAGAAGTTCGCAGAAGGAAAGCCTTTAGATATTCCAGTTATTTCACAA 600
                .         .         .         .         .
     490 CCATATGGAAAGAGAGTTGTTGCCTTTGCAGAGCACTCAGTAATTCCAGG 539
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     601 CCATATGGAAAGAGAGTTGTTGCCTTTGCAGAGCACTCAGTAATTCCAGG 650
                .         .         .         .         .
     540 CAAAGAGAAGCAATTTgAGGACGCAATAGTTAGGACTTTAGAAATGTTAA 589
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     651 CAAAGAGAAGCAATTTGAGGACGCAATAGTTAGGACTTTAGAAATGTTAA 700
```

```
                .         .         .         .         .
     590 AGAAAGCTCCTGGCTTCTTAGGTGCAATGGTATTAAAGGAAATAGGaGTT 639
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     701 AGAAAGCTCCTGGCTTCTTAGGTGCAATGGTATTAAAGGAAATAGGAGTT 750
                .         .         .         .         .
     640 TCCgGAATTgGAAGCATGCAATTCGGTGCCAAGGGATTCCATCAAGTCTT 689
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     751 TCCGGAATTGGAAGCATGCAATTCGGTGCCAAGGGATTCCATCAAGTCTT 800
                .         .         .         .         .
     690 AGAGAACCCTGGATCACTTgAGCCAgATCCAAATAATgTAATGTATTCAG 739
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     801 AGAGAACCCTGGATCACTTGAGCCAGATCCAAATAATGTAATGTATTCAG 850
```

pASK-sorH89A

```
                .         .         .         .         .
H89A:  1 ........CGACAAAAAtCTAGATAACGAGGGCAAAAAATGCCGAAACCA 42
                 |||||||||||||||||||||||||||||||||||||||||
WT:  101 GAATAGTTCGACAAAAATCTAGATAACGAGGGCAAAAAATGCCGAAACCA 150
                .         .         .         .         .
      43 TACGTTGCTATAAACATGGCAGAATTAAAGAATgAACCTAAAACTTTTGA 92
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     151 TACGTTGCTATAAACATGGCAGAATTAAAGAATGAACCTAAAACTTTTGA 200
                .         .         .         .         .
      93 AATGTTTGCCTCAGTAGGACCGAAGGTCTGCATGGtAACAGCAAGGCATC 142
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     201 AATGTTTGCCTCAGTAGGACCGAAGGTCTGCATGGTAACAGCAAGGCATC 250
                .         .         .         .         .
     143 CGGGCTTTGTTGGTTTTCAAAACCAtAtACAAAtAGGAATTTTGCCATTC 192
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     251 CGGGCTTTGTTGGTTTTCAAAACCATATACAAATAGGAATTTTGCCATTC 300
                .         .         .         .         .
     193 GGAAACAGAtACGGCGGAGcTAAAATGGACATGACTAAGGAAAGTAGTAC 242
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     301 GGAAACAGATACGGCGGAGCTAAAATGGACATGACTAAGGAAAGTAGTAC 350
                .         .         .         .         .
     243 TGTCAGAGTTTTACAGTACACCTTCTGGAAAGATTGGAAAGACCATGAAG 292
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     351 TGTCAGAGTTTTACAGTACACCTTCTGGAAAGATTGGAAAGACCATGAAG 400
               A   .         .         .         .         .
     293 AAATGGCCAGGCAAAACTGGAGTTACTTATTCAGGCTATGCTATTCATGC 342
         |||||  |||||||||||||||||||||||||||||||||||||||||||||
     401 AAATGCACAGGCAAAACTGGAGTTACTTATTCAGGCTATGCTATTCATGC 450
               H   .         .         .         .         .
     343 GCTTCACAAATGATATGGGGaCCCTGGGAGCCAATTTATGAAATAATCTA 392
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     451 GCTTCACAAATGATATGGGGACCCTGGGAGCCAATTTATGAAATAATCTA 500
                .         .         .         .         .
     393 CGCAAACATGCCTATAAACACTGAAATGACCGaCTTCACTGCAGTTGTAG 442
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     501 CGCAAACATGCCTATAAACACTGAAATGACCGACTTCACTGCAGTTGTAG 550
                .         .         .         .         .
     443 GAAAGAAGTTCGCAGAAGGAAAGCCTTTAGATATTCCAGTTATTTCACAA 492
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     551 GAAAGAAGTTCGCAGAAGGAAAGCCTTTAGATATTCCAGTTATTTCACAA 600
                .         .         .         .         .
     493 CCATATGGAAAGAGAGTTGTTGCCTTTgCAGAGCACTCAGTAATTcCAGG 542
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     601 CCATATGGAAAGAGAGTTGTTGCCTTTGCAGAGCACTCAGTAATTCCAGG 650
```

```
                .         .         .         .         .
      543 cAAAGAGaAGCAATTTgAGGaCGCAATAGTTAGGACTTTAGAAATGTTAA 592
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      651 CAAAGAGAAGCAATTTGAGGACGCAATAGTTAGGACTTTAGAAATGTTAA 700
                .         .         .         .         .
      593 AGAAAGCTCCTGGCTTCTTAGGTGCAATGGTATTAAAGGAAATAGGAGTT 642
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      701 AGAAAGCTCCTGGCTTCTTAGGTGCAATGGTATTAAAGGAAATAGGAGTT 750
                .         .         .         .         .
      643 TCCGGAATTgGAAGCATGCAATTCGGTGCCAAGGGATTCCATCAAGTCTT 692
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      751 TCCGGAATTGGAAGCATGCAATTCGGTGCCAAGGGATTCCATCAAGTCTT 800
                .         .         .         .         .
      693 AGRGRACCCTGGRTCACTTgAGCCAGATCCAAATAATGTAATGTATTCAG 742
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      801 AGAGAACCCTGGATCACTTGAGCCAGATCCAAATAATGTAATGTATTCAG 850
```

pASK-sorC100A

```
                .         .         .         .         .
C100A:  1 ..................AAGATAACGAGGGCAAAAAATGcCGAAACCA 31
                           ||||||||||||||||||||||||||||||
WT:   101 GAATAGTTCGACAAAAATCTAGATAACGAGGGCAAAAAATGCCGAAACCA 150
                .         .         .         .         .
       32 TACGTTGCTATAAACATGGCAGAATTAAAGAATGAACCTAAAACTTTTGA 81
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      151 TACGTTGCTATAAACATGGCAGAATTAAAGAATGAACCTAAAACTTTTGA 200
                .         .         .         .         .
       82 AATGTTTGCCTCAGTAGGACCGAAGGTCTGCATGGTAACAGCAAGGCATC 129
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      201 AATGTTTGCCTCAGTAGGACCGAAGGTCTGCATGGTAACAGCAAGGCATC 250
                .         .         .         .         .
      130 CGGGCTTTGTTGGTTTTCAAAACCATATACAAATAGGAATTTTGCCATTC 179
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      251 CGGGCTTTGTTGGTTTTCAAAACCATATACAAATAGGAATTTTGCCATTC 300
                .         .         .         .         .
      180 GGAAACAGATACGGCGGAGcTAAAATGGACATGACTAAGGAAAGTAGTAC 229
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      301 GGAAACAGATACGGCGGAGCTAAAATGGACATGACTAAGGAAAGTAGTAC 350
                .         .         .         .         .
      230 TGTCAGAGTTTTACAGTACACCTTCTGGAAAGATTGGAAAGACCATGAAG 279
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      351 TGTCAGAGTTTTACAGTACACCTTCTGGAAAGATTGGAAAGACCATGAAG 400
                                              A
                .         .         .         .         .
      280 AAATGCACAGGCAAAACTGGAGTTACTTATTCAGGCTAGCCTATTCATGC 329
          |||||||||||||||||||||||||||||||||||||  ||||||||||
      401 AAATGCACAGGCAAAACTGGAGTTACTTATTCAGGCTATGCTATTCATGC 450
                                              C
                .         .         .         .         .
      330 GCTTCACAAATGATATGGGGACCCTGGGAGCCAATTTATGAAATAATCTA 379
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      451 GCTTCACAAATGATATGGGGACCCTGGGAGCCAATTTATGAAATAATCTA 500
                .         .         .         .         .
      380 CGCAAACATGCCTATAAACACTGAAATGACCGaCTTCACTGCAGTTGTAG 429
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      501 CGCAAACATGCCTATAAACACTGAAATGACCGACTTCACTGCAGTTGTAG 550
                .         .         .         .         .
      430 GAAAGAAGTTCGCAGAAGGAAAGCCTTTAGATATTCCAGTTATTTCACAA 479
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      551 GAAAGAAGTTCGCAGAAGGAAAGCCTTTAGATATTCCAGTTATTTCACAA 600
```

```
                .         .         .         .         .
     480 CCATATGGAAAGAGAGTTGTTGCCTTTGCAGAGCACTCAGTAATTCCAGG 529
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     601 CCATATGGAAAGAGAGTTGTTGCCTTTGCAGAGCACTCAGTAATTCCAGG 650
                .         .         .         .         .
     530 CAAAGAGAAGCAATTTGAGGACGCAATAGTTAGGACTTTAGAAATGTTAA 579
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     651 CAAAGAGAAGCAATTTGAGGACGCAATAGTTAGGACTTTAGAAATGTTAA 700
                .         .         .         .         .
     580 AGAAAGCTCCTGGCTTCTTAGGTGCAATGGTATTAAAGGAAATAGGAGTT 629
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     701 AGAAAGCTCCTGGCTTCTTAGGTGCAATGGTATTAAAGGAAATAGGAGTT 750
                .         .         .         .         .
     630 TCCGGAATTGGaAGCATGCAATTCGGTGCCAAGGGATTCCATCAAGTCTT 679
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     751 TCCGGAATTGGAAGCATGCAATTCGGTGCCAAGGGATTCCATCAAGTCTT 800
                .         .         .         .         .
     680 AGAGAACCCTGGATCACTTGAGCCAGATCCAAATAATGTAATGTATTCAG 729
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     801 AGAGAACCCTGGATCACTTGAGCCAGATCCAAATAATGTAATGTATTCAG 850
                .         .         .         .         .
     730 TCCCAGAAGCAAAGAATACTCCACAACAATACATAGTTCATGTAGAATGG 779
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     851 TCCCAGAAGCAAAGAATACTCCACAACAATACATAGTTCATGTAGAATGG 900
                .         .         .         .         .
     780 GCAAATACTGATGCTTTAATG.........................  801
         |||||||||||||||||||
     901 GCAAATACTGATGCTTTAATGTTTGGAATGGGTAGAGTACTATTATATC  950
```

pASK-sorC100S

```
                .         .         .         .         .
C100S: 1 ..........ACAAAAATCTAGATAACGAGGGCAAAAAATGCCGAAACCA 40
                    |||||||||||||||||||||||||||||||||||||
WT:  101 GAATAGTTCGACAAAAATCTAGATAACGAGGGCAAAAAATGCCGAAACCA 150
                .         .         .         .         .
      41 TACGTTGCTATAAACATGGCAGAATTAAAGAATGAACCTAAAACTTTTGA 90
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     151 TACGTTGCTATAAACATGGCAGAATTAAAGAATGAACCTAAAACTTTTGA 200
                .         .         .         .         .
      91 AATGTTTGCCTCAGTAGGaCCGaAGGTCTGCATGGtAACAGCAAGGCATC 138
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     201 AATGTTTGCCTCAGTAGGACCGAAGGTCTGCATGGTAACAGCAAGGCATC 250
                .         .         .         .         .
     139 CGGGCTTTGTTGGTTTTCAAAACCATATACAAATAGGAATTTTGCCATTC 188
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     251 CGGGCTTTGTTGGTTTTCAAAACCATATACAAATAGGAATTTTGCCATTC 300
                .         .         .         .         .
     189 GGAAACAGATACGGCGGAGcTAAAATGGACATGACTAAGGAAAGTAGTAC 238
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     301 GGAAACAGATACGGCGGAGCTAAAATGGACATGACTAAGGAAAGTAGTAC 350
                .         .         .         .         .
     239 TGTCAGAGTTTTACAGTACACCTTCTGGAAAGATTGGAAAGACCATGAAG 288
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     351 TGTCAGAGTTTTACAGTACACCTTCTGGAAAGATTGGAAAGACCATGAAG 400
                .         .         .         S         .
     289 AAATGCACAGGCAAAACTGGaGTTACTTATTCAGGCTATCCTATTCATGC 338
         |||||||||||||||||||||||||||||||||||| ||||||||||
     401 AAATGCACAGGCAAAACTGGAGTTACTTATTCAGGCTATGCTATTCATGC 450
                                             C
```

```
     339 GCTTCACAAATGATATGGGGaCCCtGGGaGCCAATTTATGAAATAATCTA 388
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     451 GCTTCACAAATGATATGGGGACCCTGGGAGCCAATTTATGAAATAATCTA 500
                 .         .         .         .         .
     389 CGCAAACATGCCTATAAACACTGaAATGaCCGACTTCACTGCAGTTGTAG 438
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     501 CGCAAACATGCCTATAAACACTGAAATGACCGACTTCACTGCAGTTGTAG 550
                 .         .         .         .         .
     439 GAAAGAAGTTCGCAGAAGGAAAGCCTTTAGATATTCCAGTTATTTCACAA 488
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     551 GAAAGAAGTTCGCAGAAGGAAAGCCTTTAGATATTCCAGTTATTTCACAA 600
                 .         .         .         .         .
     489 CCATATGGAAAGAGAGTTGTTGCCTTTGCAGAGCACTCAGTAATTCCAGG 538
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     601 CCATATGGAAAGAGAGTTGTTGCCTTTGCAGAGCACTCAGTAATTCCAGG 650
                 .         .         .         .         .
     539 CAaAGAGAAGCAATtTGAGGACGCAATAGTTAGGACTTTAGAAATGtTAA 588
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     651 CAAAGAGAAGCAATTTGAGGACGCAATAGTTAGGACTTTAGAAATGTTAA 700
                 .         .         .         .         .
     589 AGAAAGCTCCTGGCT..................................... 603
         |||||||||||||||
     701 AGAAAGCTCCTGGCTTCTTAGGTGCAATGGTATTAAAGGAAATAGGAGTT 750
```

pASK-sorC103A

```
                 .         .         .         .         .
C103A:   1 .....................................ATGCCGAAACCA  11
                                                ||||||||||||
WT:    101 GAATAGTTCGACAAAAATCTAGATAACGAGGGCAAAAAATGCCGAAACCA 150
                 .         .         .         .         .
      12 TACGTTGCTATAAACATGGCAGAATTAAAGAATGAACCTAAAACTTTTGA  61
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     151 TACGTTGCTATAAACATGGCAGAATTAAAGAATGAACCTAAAACTTTTGA 200
                 .         .         .         .         .
      62 AATGTTTGCCTCAGTAGGACCGAaGGTCTGCATGGTAACAGCAAGGCATC 109
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     201 AATGTTTGCCTCAGTAGGACCGAAGGTCTGCATGGTAACAGCAAGGCATC 250
                 .         .         .         .         .
     110 CGGGCTTTGTTGGTTTTCAAAACCATATACAAATAGGAATtttGCCATTC 159
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     251 CGGGCTTTGTTGGTTTTCAAAACCATATACAAATAGGAATTTTGCCATTC 300
                 .         .         .         .         .
     160 GGAAACAGATACGGCGGAGcTAAAATGGACATGACTAAGGAAAGTAGTAC 209
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     301 GGAAACAGATACGGCGGAGCTAAAATGGACATGACTAAGGAAAGTAGTAC 350
                 .         .         .         .         .
     210 TGTCAGAGTTTTACAGTACACCTTCTGGAAAGATTGGAAAGACCATGAAG 259
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     351 TGTCAGAGTTTTACAGTACACCTTCTGGAAAGATTGGAAAGACCATGAAG 400
                 .         .         .         .        A.
     260 AAATGCACAGGCAAAACTGGAGTTACTTATTCAGGCTATGCTATTCAGCC 309
         |||||||||||||||||||||||||||||||||||||||||||||||| |
     401 AAATGCACAGGCAAAACTGGAGTTACTTATTCAGGCTATGCTATTCATGC 450
                 .         .         .         .        C.
     310 GCTTCACAAATGATATGGGGACCCTGGGAGCCAATTTATGAAATAATCTA 359
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     451 GCTTCACAAATGATATGGGGACCCTGGGAGCCAATTTATGAAATAATCTA 500
```

```
                .         .         .         .         .
     360 CGCAAACATGCCTATAAACACTGaAATGaCCGaCTTCACTGCAGTTGTAG 409
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     501 CGCAAACATGCCTATAAACACTGAAATGACCGACTTCACTGCAGTTGTAG 550
                .         .         .         .         .
     410 GAAAGAAGTTCGCAGAAGGAAAGCCTTTAGATATTCCAGTTATTTCACAA 459
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     551 GAAAGAAGTTCGCAGAAGGAAAGCCTTTAGATATTCCAGTTATTTCACAA 600
                .         .         .         .         .
     460 CCATATGGAAAGAGAGTTGTTGCCTTTGCAGAGCACTCAGTAATTCCAGG 509
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     601 CCATATGGAAAGAGAGTTGTTGCCTTTGCAGAGCACTCAGTAATTCCAGG 650
                .         .         .         .         .
     510 CAAAGAGAAGCAATTTGAGGACGCAATAGTTAGGACTTTAGAAATGTTAA 559
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     651 CAAAGAGAAGCAATTTGAGGACGCAATAGTTAGGACTTTAGAAATGTTAA 700
                .         .         .         .         .
     560 AGAAAGCTCCTGGCTTCTTAGGTGCAATGGTATTAAAGGAAATAGGAGTT 609
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     701 AGAAAGCTCCTGGCTTCTTAGGTGCAATGGTATTAAAGGAAATAGGAGTT 750
                .         .         .         .         .
     610 TCCGGAATTGGaAGCATGCAATTCGGTGCCAAGGGATTCCATCAAGTCTT 659
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     751 TCCGGAATTGGAAGCATGCAATTCGGTGCCAAGGGATTCCATCAAGTCTT 800
                .         .         .         .         .
     660 AGAGAACCCTGGATCACTTGAGCCAGATCCAAATAATGTAATGTATTCAG 709
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     801 AGAGAACCCTGGATCACTTGAGCCAGATCCAAATAATGTAATGTATTCAG 850
                .         .         .         .         .
     710 TCCCAGAAGCAAAGAATACTcCACAACAATACATAGTTCATGTAGAATgG 759
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     851 TCCCAGAAGCAAAGAATACTCCACAACAATACATAGTTCATGTAGAATGG 900
                .         .         .         .         .
     760 GCAAATACTGATGCTTTAATGTTTGGAaTGGGTAGGGTACTa........ 801
         |||||||||||||||||||||||||||||||||||||||||
     901 GCAAATACTGATGCTTTAATGTTTGGAATGGGTAGAGTACTATTATATCC 950
```

pASK-sorC103S

```
                .         .         .         .         .
C103S:   1 .....................GATAACGAGGGCAAAAaATGCCGAAACCA 29
                                |||||||||||||||||||||||||||||
WT:    101 GAATAGTTCGACAAAAATCTAGATAACGAGGGCAAAAAATGCCGAAACCA 150
                .         .         .         .         .
      30 TACGTTGCTATAAACATGGCAGAATTAAAGAATGAACCTAAAACTTTTGA 79
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     151 TACGTTGCTATAAACATGGCAGAATTAAAGAATGAACCTAAAACTTTTGA 200
                .         .         .         .         .
      80 AATGTTTGCCTCAGTAGGACCGAAGGTCTGCATGGTAACAGCAAGGCATC 127
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     201 AATGTTTGCCTCAGTAGGACCGAAGGTCTGCATGGTAACAGCAAGGCATC 250
                .         .         .         .         .
     128 CGGGCTTTGTTGGTTTTCAAAACCATATACAAATAGGAATTTTGCCATTC 177
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     251 CGGGCTTTGTTGGTTTTCAAAACCATATACAAATAGGAATTTTGCCATTC 300
                .         .         .         .         .
     178 GGAAACAGATACGGCGGAGcTAAAATGGACATGACTAAGGAAAGTAGTAC 227
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     301 GGAAACAGATACGGCGGAGCTAAAATGGACATGACTAAGGAAAGTAGTAC 350
```

```
                .         .         .         .         .
     228 TGTCAGAGTTTTACAGTACACCTTCTGGAAAGATTGGAAAGACCATGAAG 277
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     351 TGTCAGAGTTTTACAGTACACCTTCTGGAAAGATTGGAAAGACCATGAAG 400
                .         .         .         .        S.
     278 AAATGCACAGGCAAAACTGGAGTTACTTATTCAGGCTATGCTATTCATCC 327
         ||||||||||||||||||||||||||||||||||||||||||||||||| |
     401 AAATGCACAGGCAAAACTGGAGTTACTTATTCAGGCTATGCTATTCATGC 450
                .         .         .         .        C.
     328 GCTTCACAAATGATATGGGGACCCTGGGAGCCAATTTATGAAATAATCTA 377
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     451 GCTTCACAAATGATATGGGGACCCTGGGAGCCAATTTATGAAATAATCTA 500
                .         .         .         .         .
     378 CGCAAACATGCCTATAAACACTGAAATGACCGaCTTCACTGCAGTTGTAG 427
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     501 CGCAAACATGCCTATAAACACTGAAATGACCGACTTCACTGCAGTTGTAG 550
                .         .         .         .         .
     428 GAAAGAAGTTCGCAGAAGGAAAGCCTTTAGATATTCCAGTTATTTCACAA 477
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     551 GAAAGAAGTTCGCAGAAGGAAAGCCTTTAGATATTCCAGTTATTTCACAA 600
                .         .         .         .         .
     478 CCATATGGAAAGAGAGTTGTTGCCTTTGCAGAGCACTCAGTAATTCCAGG 527
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     601 CCATATGGAAAGAGAGTTGTTGCCTTTGCAGAGCACTCAGTAATTCCAGG 650
                .         .         .         .         .
     528 CAAAGAGAAGCAATTTGAGGACGCAATAGTTAGGACTTTAGAAATGTTAA 577
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     651 CAAAGAGAAGCAATTTGAGGACGCAATAGTTAGGACTTTAGAAATGTTAA 700
                .         .         .         .         .
     578 AGAAAGCTCCTGGCTTCTTAGGTGCAATGGTATTAAAGGAAATAGGAGTT 627
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     701 AGAAAGCTCCTGGCTTCTTAGGTGCAATGGTATTAAAGGAAATAGGAGTT 750
                .         .         .         .         .
     628 TCCGGAATTGGAAGCATGCAATTCGGTGCCAAGGGATTCCATCAAGTCTT 677
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     751 TCCGGAATTGGAAGCATGCAATTCGGTGCCAAGGGATTCCATCAAGTCTT 800
                .         .         .         .         .
     678 AGAGAACCCTGGATCACTTGAGCCAGATCCAAATAATGTAATGTATTCAG 727
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     801 AGAGAACCCTGGATCACTTGAGCCAGATCCAAATAATGTAATGTATTCAG 850
                .         .         .         .         .
     728 TCCCAGAAGCAAAGAATACTcCACAACAATACATAGTTCATGTAGAATgG 777
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     851 TCCCAGAAGCAAAGAATACTCCACAACAATACATAGTTCATGTAGAATGG 900
```

pASK-sorC100/103A

```
                .         .         .         .         .
C100/103A:        CAAaaaTCTaGATAaCGAGGGCAAAAAATGCCGAAACCA 45
                  |||||||||||||||||||||||||||||||||||||||
WT : 101 GAATAGTTCGACAAAAATCTAGATAACGAGGGCAAAAAATGCCGAAACCA 150
                .         .         .         .         .
      46 TACGTTGcTATAAACATGGCAGaATTAAAGAATGAACCTAAAACTTTTGA 95
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     151 TACGTTGCTATAAACATGGCAGAATTAAAGAATGAACCTAAAACTTTTGA 200
                .         .         .         .         .
      96 AATGTTTGCCTCAGTAGGaCCGaAGGTCTgCATGGTAACAGCAAGGcATC 143
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     201 AATGTTTGCCTCAGTAGGACCGAAGGTCTGCATGGTAACAGCAAGGCATC 250
```

```
                .         .         .         .         .
     144 CGGGCTTTGTTGGTTTTCAAAACCATATACAAATAGGAATTTTGCCATTC 193
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     251 CGGGCTTTGTTGGTTTTCAAAACCATATACAAATAGGAATTTTGCCATTC 300
                .         .         .         .         .
     194 GGaAACAGATACGGCGGaGcTAAAATGGaCATGaCTAAGGAAAGTAGTAC 243
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     301 GGAAACAGATACGGCGGAGCTAAAATGGACATGACTAAGGAAAGTAGTAC 350
                .         .         .         .         .
     244 TGTCAGAGTTTTACAGTACACCTTCTGGAAAGATTGGAAAGACCATGaAG 293
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     351 TGTCAGAGTTTTACAGTACACCTTCTGGAAAGATTGGAAAGACCATGAAG 400
                .         .         .         A         A.
     294 AAATGCACAGGCAAAACTGGaGTTACTTATTCAGGCTAGCCTATTCAGCC 343
         |||||||||||||||||||||||||||||||||||||  |||||||  |
     401 AAATGCACAGGCAAAACTGGAGTTACTTATTCAGGCTATGCTATTCATGC 450
                .         .         .         C         C
     344 GCTTCACAAATGATATGGGGACCCTGGGAGCCAATTTATGAAATAATCTA 393
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     451 GCTTCACAAATGATATGGGGACCCTGGGAGCCAATTTATGAAATAATCTA 500
                .         .         .         .         .
     394 CGCAAACATGCCTATAAACACTGAAATGACCGACTTCACTGCAGTTGTAG 443
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     501 CGCAAACATGCCTATAAACACTGAAATGACCGACTTCACTGCAGTTGTAG 550
                .         .         .         .         .
     444 GAAAGAAGTTCGCAGAAGGAAAGCCTTTAGATATTCCaGTTATTTCACAA 493
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     551 GAAAGAAGTTCGCAGAAGGAAAGCCTTTAGATATTCCAGTTATTTCACAA 600
                .         .         .         .         .
     494 CCaTA..................................................... 498
         |||||
     601 CCATATGGAAAGAGAGTTGTTGCCTTTGCAGAGCACTCAGTAATTCCAGG 650
```

pASK-sorC100/103S

```
                .         .         .         .         .
C100/103S:.....................AGATAACGAGGGCAAAAaATGCCGAAACCA 30
                               ||||||||||||||||||||||||||||||
WT:  101 GAATAGTTCGACAAAAATCTAGATAACGAGGGCAAAAAATGCCGAAACCA 150
                .         .         .         .         .
      31 TACGTTGCTATAAACATGGCAGAATTAAAGAATGAACCTAAAACTTTTGA 80
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     151 TACGTTGCTATAAACATGGCAGAATTAAAGAATGAACCTAAAACTTTTGA 200
                .         .         .         .         .
      81 AATGTTTGCCTCAGTAGGACCGAAGGTCTGCATGGTAACAGCAAGGCATC 128
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     201 AATGTTTGCCTCAGTAGGACCGAAGGTCTGCATGGTAACAGCAAGGCATC 250
                .         .         .         .         .
     129 CGGGCTTTGTTGGTTTTCAAAACCATATACAAATAGGAATTTTGCCATTC 178
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     251 CGGGCTTTGTTGGTTTTCAAAACCATATACAAATAGGAATTTTGCCATTC 300
                .         .         .         .         .
     179 GGAAACAGATACGGCGGAGcTAAAATGGACATGACTAAGGAAAGTAGTAC 228
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     301 GGAAACAGATACGGCGGAGCTAAAATGGACATGACTAAGGAAAGTAGTAC 350
                .         .         .         .         .
     229 TGTCAGAGTTTTACAGTACACCTTCTGGAAAGATTGGAAAGACCATGAAG 278
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     351 TGTCAGAGTTTTACAGTACACCTTCTGGAAAGATTGGAAAGACCATGAAG 400
```

```
                .         .         .         S         S.
    279 AAATGCACAGGCAAAACTGGAGTTACTTATTCAGGCTATCCTATTCATCC 328
        |||||||||||||||||||||||||||||||||||||| |||||||| |
    401 AAATGCACAGGCAAAACTGGAGTTACTTATTCAGGCTATGCTATTCATGC 450
                .         .         .         C         C.
    329 GCTTCACAAATGATATGGGGACCCTGGGAGCCAATTTATGAAATAATCTA 378
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    451 GCTTCACAAATGATATGGGGACCCTGGGAGCCAATTTATGAAATAATCTA 500
                .         .         .         .         .
    379 CGCAAACATGCCTATAAACACTGAAATGACCGACTTCACTGCAGTTGTAG 428
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    501 CGCAAACATGCCTATAAACACTGAAATGACCGACTTCACTGCAGTTGTAG 550
                .         .         .         .         .
    429 GAAAGAAGTTCGCAGAAGGAAAGCCTTTAGATATTCCAGTTATTTCACAA 478
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    551 GAAAGAAGTTCGCAGAAGGAAAGCCTTTAGATATTCCAGTTATTTCACAA 600
                .         .         .         .         .
    479 CCATATGGaAAGAGAGTTGTTGCCTTTGCAGAGCACTCAGTAATTCCAGG 528
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    601 CCATATGGAAAGAGAGTTGTTGCCTTTGCAGAGCACTCAGTAATTCCAGG 650
                .         .         .         .         .
    529 CAAAGAGAAGCAATTTGAGGACGCAATAGTTAGGACTTTAGAAATGTTAA 578
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    651 CAAAGAGAAGCAATTTGAGGACGCAATAGTTAGGACTTTAGAAATGTTAA 700
                .         .         .         .         .
    579 AGAAAGCTCCTGGCTTCTTAGGTGCAATGGTATTAAAGGAAATAGGAGTT 628
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    701 AGAAAGCTCCTGGCTTCTTAGGTGCAATGGTATTAAAGGAAATAGGAGTT 750
                .         .         .         .         .
    629 TCCGGAATTGGAAGCATGCAATTCGGTGCCAAGGGATTCCATCAAGTCTT 678
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    751 TCCGGAATTGGAAGCATGCAATTCGGTGCCAAGGGATTCCATCAAGTCTT 800
                .         .         .         .         .
    679 AGAGAACCCTGGATCACTTGAGCCAGATCCAAATAATGTAATGTATTCAG 728
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    801 AGAGAACCCTGGATCACTTGAGCCAGATCCAAATAATGTAATGTATTCAG 850
                .         .         .         .         .
    729 TCCCAGAAGCAAAGAATACTCCACAACAATACATAGTTCATGTAGAATGG 778
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    851 TCCCAGAAGCAAAGAATACTCCACAACAATACATAGTTCATGTAGAATGG 900
                .         .         .         .         .
    779 GCAAATACTGATGCTTTAATGTT..........................  801
        |||||||||||||||||||||||
    901 GCAAATACTGATGCTTTAATGTTTGGAATGGGTAGAGTACTATTATATCC 950
```

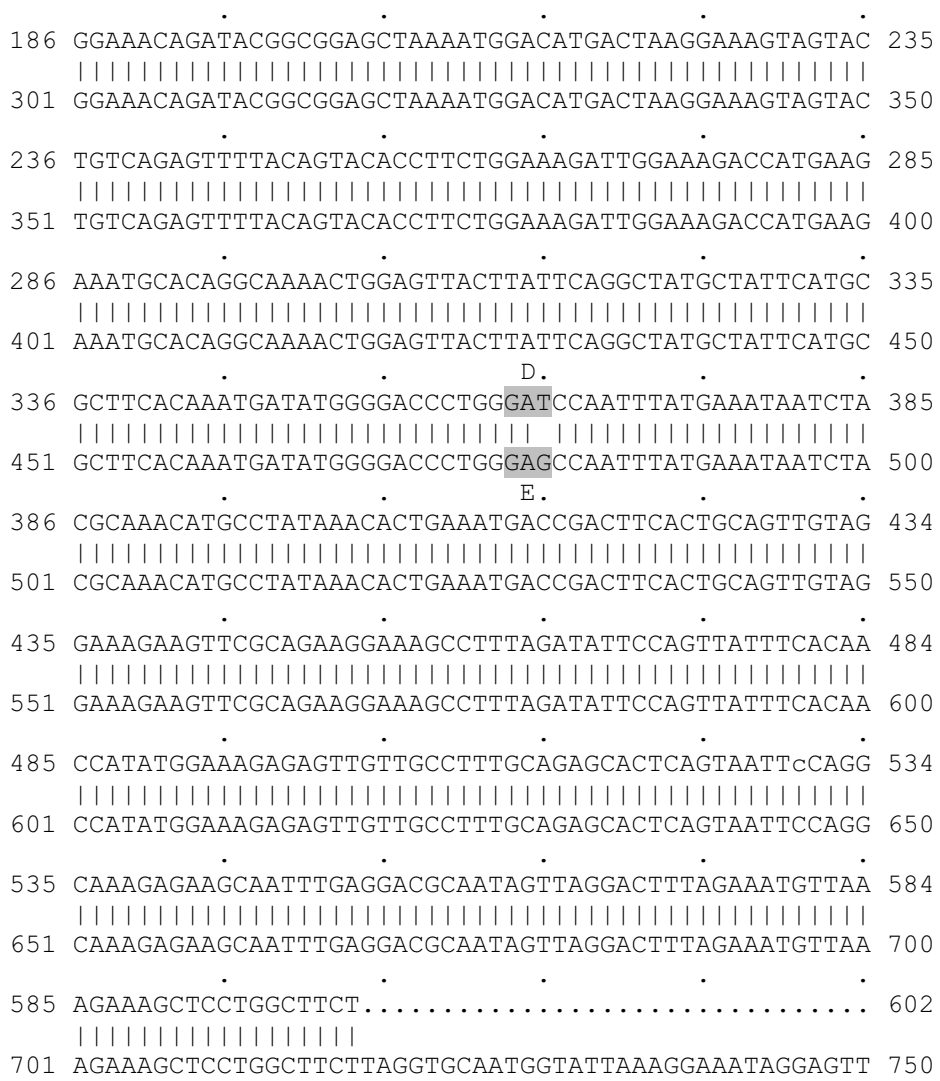
pASK-sorE113A

```
                .         .         .         .         .
E113A:  1 ......................ATAACGAGGGCAAAAAATGCCGAAACCA  28
                                 ||||||||||||||||||||||||||||
WT :  101 GAATAGTTCGACAAAAATCTAGATAACGAGGGCAAAAAATGCCGAAACCA 150
                .         .         .         .         .
       29 TACGTTGCTATAAACATGGCAGAATTAAAGAATgAACCtAAAACTTTTGA  78
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      151 TACGTTGCTATAAACATGGCAGAATTAAAGAATGAACCTAAAACTTTTGA 200
                .         .         .         .         .
       79 AAtGTTTGCCTCAGtAGGACCGAAGGTCTGCATGGTAACAGCAAGGCATC 127
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      201 AATGTTTGCCTCAGTAGGACCGAAGGTCTGCATGGTAACAGCAAGGCATC 250
```

```
                .         .         .         .         .
     128 CGGGCTTTGTTGGTTTTCAAAACCATATACAAATAGGAATTTTGCCATTC 177
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     251 CGGGCTTTGTTGGTTTTCAAAACCATATACAAATAGGAATTTTGCCATTC 300
                .         .         .         .         .
     178 GGAAACAGATACGGCGGAGcTAAAATGGACATGACTAAGGAAAGTAGTAC 227
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     301 GGAAACAGATACGGCGGAGCTAAAATGGACATGACTAAGGAAAGTAGTAC 350
                .         .         .         .         .
     228 TGTCAGAGTTTTACAGTACACCTTCTGGAAAGATTGGAAAGACCATGAAG 277
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     351 TGTCAGAGTTTTACAGTACACCTTCTGGAAAGATTGGAAAGACCATGAAG 400
                .         .         .         .         .
     278 AAATGCACAGGCAAAACTGGAGTTACTTATTCAGGCTATGCTATTCATGC 327
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     401 AAATGCACAGGCAAAACTGGAGTTACTTATTCAGGCTATGCTATTCATGC 450
                .         .        A.         .         .
     328 GCTTCACAAAtGaTATGGGGaCCCTGGGCGCCAATTTATGAAATAATCTA 377
         |||||||||||||||||||||||||| ||||||||||||||||||||||
     451 GCTTCACAAATGATATGGGGACCCTGGGAGCCAATTTATGAAATAATCTA 500
                .         .        E.         .         .
     378 CGCAAACATGCCTATAAACACTGaAATGaCCGACTTCACTGCAGTTGTAG 427
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     501 CGCAAACATGCCTATAAACACTGAAATGACCGACTTCACTGCAGTTGTAG 550
                .         .         .         .         .
     428 GAAAGAAGTTCGCAGAAGGAAAGCCTTTAGATATTCCAGTTATTTCACAA 477
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     551 GAAAGAAGTTCGCAGAAGGAAAGCCTTTAGATATTCCAGTTATTTCACAA 600
                .         .         .         .         .
     478 CCATATGGAAAGAGAGTTGTTGCCTTTGCAGAGCACTCAGTAATTcCAGG 527
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     601 CCATATGGAAAGAGAGTTGTTGCCTTTGCAGAGCACTCAGTAATTCCAGG 650
                .         .         .         .         .
     528 CAAAGAGAAGCAATTTGAGGACGCAATAGTTAGGACTTTAGAaATGTTAA 577
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     651 CAAAGAGAAGCAATTTGAGGACGCAATAGTTAGGACTTTAGAAATGTTAA 700
                .         .         .         .         .
     578 AGAaAGCTCCTGGCTTCTtAGGtGC..........................  602
         |||||||||||||||||||||||||
     701 AGAAAGCTCCTGGCTTCTTAGGTGCAATGGTATTAAAGGAAATAGGAGTT 750
```

pASK-sorE113D

```
                .         .         .         .         .
E113D:   1 ..............AAATCTAGATAACGAGGGCAAAAAATGCCGAAACCA  36
                         |||||||||||||||||||||||||||||||||||
WT :   101 GAATAGTTCGACAAAAATCTAGATAACGAGGGCAAAAAATGCCGAAACCA 150
                .         .         .         .         .
        37 TACGTTGCTATAAACATGGCAGAATTAAAGAATGAACCTAAAACTTTTGA  86
           ||||||||||||||||||||||||||||||||||||||||||||||||||
     151 TACGTTGCTATAAACATGGCAGAATTAAAGAATGAACCTAAAACTTTTGA 200
                .         .         .         .         .
        87 AATGTTTGCCTCAGTAGGACCGAAGGTCTGCATGGTAACAGCAAGGCATC 135
           ||||||||||||||||||||||||||||||||||||||||||||||||||
     201 AATGTTTGCCTCAGTAGGACCGAAGGTCTGCATGGTAACAGCAAGGCATC 250
                .         .         .         .         .
     136 CGGGCTTTGTTGGTTTTCAAAACCATATACAAATAGGAATTTTGCCATTC 185
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     251 CGGGCTTTGTTGGTTTTCAAAACCATATACAAATAGGAATTTTGCCATTC 300
```


186 GGAAACAGATACGGCGGAGCTAAAATGGACATGACTAAGGAAAGTAGTAC 235
301 GGAAACAGATACGGCGGAGCTAAAATGGACATGACTAAGGAAAGTAGTAC 350
236 TGTCAGAGTTTTACAGTACACCTTCTGGAAAGATTGGAAAGACCATGAAG 285
351 TGTCAGAGTTTTACAGTACACCTTCTGGAAAGATTGGAAAGACCATGAAG 400
286 AAATGCACAGGCAAAACTGGAGTTACTTATTCAGGCTATGCTATTCATGC 335
401 AAATGCACAGGCAAAACTGGAGTTACTTATTCAGGCTATGCTATTCATGC 450
D.
336 GCTTCACAAATGATATGGGGACCCTGGGATCCAATTTATGAAATAATCTA 385
451 GCTTCACAAATGATATGGGGACCCTGGGAGCCAATTTATGAAATAATCTA 500
E.
386 CGCAAACATGCCTATAAACACTGAAATGACCGACTTCACTGCAGTTGTAG 434
501 CGCAAACATGCCTATAAACACTGAAATGACCGACTTCACTGCAGTTGTAG 550
435 GAAAGAAGTTCGCAGAAGGAAAGCCTTTAGATATTCCAGTTATTTCACAA 484
551 GAAAGAAGTTCGCAGAAGGAAAGCCTTTAGATATTCCAGTTATTTCACAA 600
485 CCATATGGAAAGAGAGTTGTTGCCTTTGCAGAGCACTCAGTAATTcCAGG 534
601 CCATATGGAAAGAGAGTTGTTGCCTTTGCAGAGCACTCAGTAATTCCAGG 650
535 CAAAGAGAAGCAATTTGAGGACGCAATAGTTAGGACTTTAGAAATGTTAA 584
651 CAAAGAGAAGCAATTTGAGGACGCAATAGTTAGGACTTTAGAAATGTTAA 700
585 AGAAAGCTCCTGGCTTCT..... 602
701 AGAAAGCTCCTGGCTTCTTAGGTGCAATGGTATTAAAGGAAATAGGAGTT 750

5. GENERAL DISCUSSION

Comparatively little information is available on molecular details of the biological oxidation of inorganic sulfur compounds in spite of its biogeochemical and biotechnological significance. In recent years, some progress has been made in understanding the enzyme systems present in mesophilic organisms (Friedrich *et al*., 2001; Bamford *et al*., 2002). The work presented here focused on the enzymatic mechanisms of elemental sulfur oxidation in thermoacidophilic organisms like *Acidianus ambivalens*. The sulfur oxygenase reductase, the initial enzyme in *A. ambivalens´* sulfur oxidation pathway was investigated with respect to its structure and function.

The data collected during this work are summarized in this chapter and discussed in the light of the structure-function correlation of the SOR and the enzyme´s catalytic mechanism. In addition, the characteristics of the SOR will be compared to other enzymes involved in the biological oxidation of inorganic sulfur compounds.

5.1. THE PRESENCE OF A SULFUR-DISPROPORTIONATING ENZYME IS AN ADAPTATION TO THE HIGH GROWTH TEMPERATURE OF *A. AMBIVALENS*

The SOR catalyzes a unique reaction where sulfite, thiosulfate and hydrogen sulfide are produced from elemental sulfur in an oxygen-dependent manner. No external electron donors or cofactors are necessary (Kletzin, 1989).

Taking into consideration that thiosulfate is probably derived from a non-enzymatic reaction of sulfite with excess elemental sulfur, an overall reaction has been formulated (Kletzin, 1989).

Eq. 5.1 $4\ S^0 + O_2 + 4\ H_2O \rightarrow 2\ H_2S + 2\ HSO_3^- + 2\ H^+$

The stoichiometry of Eq. 5.1 was explained by the presence of two coupled partial reactions, a sulfur oxygenation (Eq. 5.2) and a sulfur disproportionation (Eq.5.3), the latter one explaining the formation of the reduced product hydrogen sulfide (Kletzin, 1989).

Eq. 5.2 $S^0 + O_2 + H_2O \rightarrow HSO_3^- + H^+$

Eq. 5.3 $3\ S^0 + 3\ H_2O \rightarrow HSO_3^- + 2\ H_2S + H^+$

This type of reaction is unique, as none of the enzymes isolated from mesophilic sulfur oxidizing organisms is capable of a simultaneous oxidation and reduction of sulfur (Rohwerder & Sand, 2003, Friedrich *et al*., 2001), although some bacteria are known that disproportionate elemental sulfur for energy conservation (Finster *et al*., 1998). The present work supports the original postulate of a coupled oxygenation and disproportionation. The mutagenesis study suggested a strong coupling of the oxygenase and reductase activity, as none of the mutations specifically abolished one partial reaction (chapter 4). A non-enzymatic sulfur disproportionation is known to occur at elevated temperatures under moderate alkaline conditions, whereas pH values above 11.5 are necessary at 20°C (McCampell Hamilton, 1991). Therefore, it appears that the utilization of a sulfur disproportionating enzyme is an adaptation of *A. ambivalens* to the increased reactivity of sulfur at its growth temperature (80°C).

The presence of a sulfur disproportionating enzyme might also explain one fundamental difference between the sulfur oxidation pathway of *A. ambivalens* and the ones found in mesophilic sulfur oxidizing organisms. In *A. ambivalens*, the SOR as the initial enzyme is localized in the cytoplasm, whereas the sulfur oxidizing pathway of the mesophilic sulfur oxidizers resides in the periplasm (Rohwerder & Sand, 2003, Friedrich *et al*., 2001). The basic cause for this difference might be the pH-dependency of the sulfur disproportionation reaction catalyzed by the SOR, which requires moderate alkaline conditions, whereas a strongly acidic pH is present in the surrounding medium (pH 2.5 to <1) and, in less extend, in its pseudo-periplasm.

5.2. THE SOR IS A SELF-COMPARTMENTALIZING ICOSATETRAMER

Especially the X-ray structure provided a deeper insight into the structure-function correlation of the SOR (chapter 3). Briefly summarized, the crystal structure showed that the oligomeric enzyme is a huge and highly symmetrical hollow protein particle of ~150 Å in diameter composed of 24 identical subunits and a molecular mass of 845 kDa (calculated for the wild type holoenzyme). Each monomer contains one sulfur oxygenation/disproportionation apparatus, composed of a mononuclear non-heme iron site and three cysteines, located along a cavity within the interior of the subunit. The single active sites are spacially separated from each other, making their interaction during catalysis unlikely. The 24 active sites are not directly accessible from the exterior of the sphere but solely via the inner compartment. Substrate access to the inner compartment is most probably provided by the six hydrophobic channels located at the four-fold axes of the particle.

The particular architecture of the holoenzyme has several functional consequences and provides a first sketch of its mode of action (Fig. 5.1).

(1) Six chimney-like structures mediate the transfer of the probably linear polysulfide substrate (section 5.5) via a hydrophobic channel to the inner compartment (Fig. 5.1A).

(2) After entering the inner compartment, the polysulfide is delivered to one of the 24 active sites of the oligomeric complex, which are accessible via narrow pores from the inner compartment (Fig. 5.1B). Four active site entrances are in the vicinity of each "chimney". An efficient supply of substrate is mediated by the strongly positive surface charge of the active site cavity, resulting in the electrostatic attraction of the negatively charged polysulfides. Additional polysulfide may be stored within the cavity.

(3) The active sites catalyze the oxygen-dependent disproportionation resulting in the formation of sulfite and hydrogen sulfide from polysulfide (Fig. 5.1C). Thiosulfate, as the third product of the SOR reaction could results from a non-enzymatic reaction of sulfite with additional polysulfide in the inner compartment. The latter could be interpreted as an enclosed microenvironment to allow this non-enzymatic reaction.

(4) Subsequently, the products leave the sphere via not yet identified routes to serve as substrates for the enzymes downstream in the sulfur oxidation pathway of *A. ambivalens*.

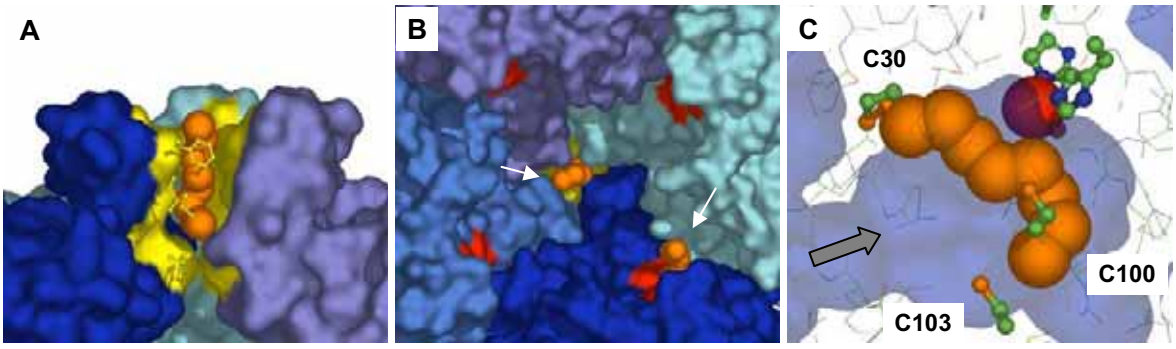


Figure 5.1: A hypothetical sketch of the SOR´s mode of action. **(A)** The sulfur substrate, modeled as a linear polysulfide (orange) is transferred to the inner compartment via a chimney-like protrusion. Three out of four subunits contributing to the formation of the "chimney" are shown. Yellow, hydrophobic inner "chimney" surface. The residues of the fourth subunit contributing are shown as balls-and-sticks representation. **(B)** Upon entering the inner sphere, the substrate is transferred to one of the four active site entrances in the proximity of the "chimney". View from the interior of the sphere with monomers colored differently. The yellow area represents the "chimney"; red areas indicate the active site entrances. Arrows indicate substrate molecules entering the sphere and the active site, respectively. **(C)** The active site in the interior of the subunits catalyzes the oxygen dependent sulfur disproportionation reaction. Important residues are shown as balls-and-sticks representation, the iron ion as red sphere. The cavity surface is indicated in blue. The arrow indicates the cavity entrance.

The architecture of the holoenzyme has additional functional advantages, as it shields the cytoplasmic environment from harmful oxygen species generated during catalysis. In

addition, it provides an effective spacial clustering of 24 active sites within one enzyme particle.

5.3. THE SULFUR-DISPROPORTIONATION APPARATUS

Two catalytically essential components of the SOR were identified in the course of this work (Fig. 5.2). It was shown by EPR spectroscopy that a mononuclear non-heme iron center was present as a redox-active cofactor (chapter 2). The midpoint potential of the center was $E_0´$ = -268 mV, which is unusually low for this type of metal center. However it is in the range of the redox couple H_2S / S^0 ($E_0´$ = -270mV), suggesting that the Fe center is the site of sulfur reduction. The results from the Fe quantitation and X-ray crystallography showed that each subunit contained one center with the side chains of His85, His89 and Glu113 as protein ligands. The refolding of denatured SOR in the absence of iron resulted in inactive enzyme. Furthermore, mutation of the iron ligands abolished both iron incorporation and activity, demonstrating the essential role of the iron for catalysis.

The protein ligands and their spacial arrangement identified the Fe center as a 2-His-1-carboxylate facial triad (chapter 3; Hegg & Que, 1997). A common feature of all enzymes with this type of center is that they are catalytically active in the Fe(II) but not in the Fe(III) state (Costas *et al*., 2004). In contrast, the SOR´s Fe site is in the Fe(III) state in the "as isolated" enzyme, as shown by EPR. EPR revealed that upon incubation with sulfur the site gets reduced to Fe(II), and stays reduced even after extensive washing with oxidized buffer. This observation might point to Fe(II) as catalytically active state, and the Fe(III) state being a result of the protein purification under oxic conditions, as it had been found in aromatic amino acid hydroxylases (Costas *et al*., 2004).

The Fe site is most probably responsible for the oxygen activation, as it is the case for all other enzymes with the 2-His-1-carboxylate facial triad. This type of center has been found in various oxygenases and hydroxylases (Que & Ho, 1996; Costas *et al*., 2004), but never before in the context of the oxidation of inorganic sulfur compounds, thus its identification in the SOR adds to the remarkable versatility of the 2-His-1-carboxylate facial triad to catalyze a variety of chemical conversions.

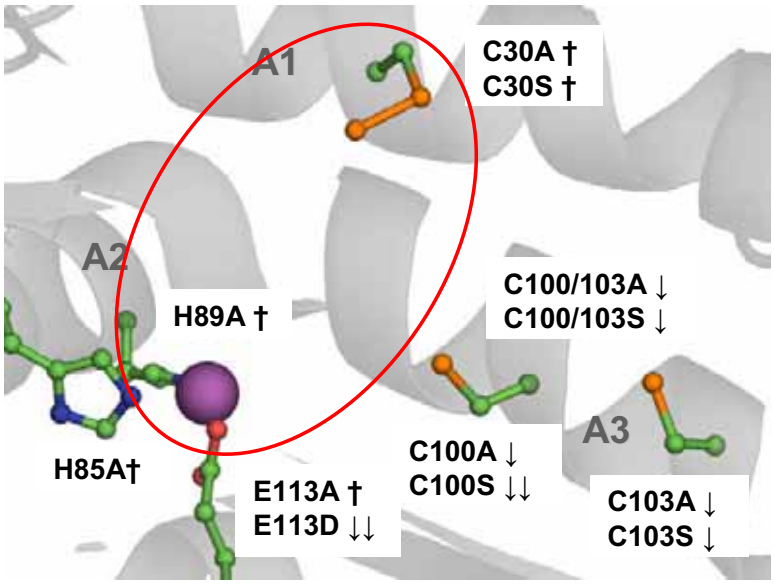


Figure 5.2: The sulfur-disproportionation apparatus of the SOR and the effects of amino acid exchanges. Symbols behind the mutants: † abolished activity; ↓, reduced activity; ↓↓, strongly reduced activity. Red circle: the core active site composed of the Fe site and Cys30.

Three cysteine residues are conserved in all SORs known to date. In the *A. ambivalens* SOR they create the active site together with the mononuclear non-heme iron center (Fig. 5.2). Cys30 was found to be persulfurated in the recombinant SOR used for crystallization. It is not known whether this is also the case within the wild type enzyme, but this finding points to Cys30 as the sulfur binding site. Cysteine persulfides are generated during the catalytic cycles of several enzymes involved in sulfur metabolism (Beinert, 2000). It is under dispute, whether the cysteine persulfide is generated as an intermediate during catalysis or formed prior to it, however it was shown that these cysteines are the sulfur-binding residues (Bordo *et al*., 2000; Bamford *et al*., 2002; Lin *et al*., 2004). Additional evidence for the role of Cys30 came from mutational studies, as its mutation resulted in inactive enzyme, whereas the exchange of the other two cysteines resulted in lowered but not in abolished activity. Although a direct proof for the covalent binding of sulfur to Cys30 was not obtained by crystallography, the above mentioned observations strongly suggest that Cys30 is the binding site of the sulfur substrate.

It was speculated that Cys100 and Cys103 might form a second redox-active site by transient disulfide formation (chapter 2.3.3, Urich *et al*., 2004). The X-ray structure showed that this is not possible due to the distance of 6 Å between their thiol groups and their location in an α-helix (chapter 3.2.6). A role in sulfur binding is also unlikely, as single and double mutants showed residual activity (Fig. 5.2). Eventually, those cysteines are necessary for an optimal sterical alignment of the sulfur substrate in the active site pocket,

or they might act in catalysis steps subsequent to the proposed initial reaction of sulfur with oxygen at the iron site.

Although it is probable that additional residues are involved in the catalytic reaction (e.g. Asn8, Asn44 and Gln75), it appears that the major players have been identified. The results from the mutagenesis study point to a core catalytic apparatus composed of the iron center and Cys30 (Fig. 5.2). Furthermore, all mutations affected simultaneously the sulfur oxygenase and reductase activity, indicating that both activities are inseparably coupled. These results point to the SOR being a true sulfur disproportionating enzyme with no spacially separated sulfur oxidation and reduction sites.

5.4. INSIGHTS INTO THE CATALYTIC MECHANISM OF THE SOR

Although no intermediate steps of catalysis were identified in the course of this work, e.g. by X-ray crystallography, the results provide the basis for some hypotheses. Especially the identification of the 2-His-1-carboxylate facial triad allows to postulate distinct steps during the catalytic mechanism of the SOR.

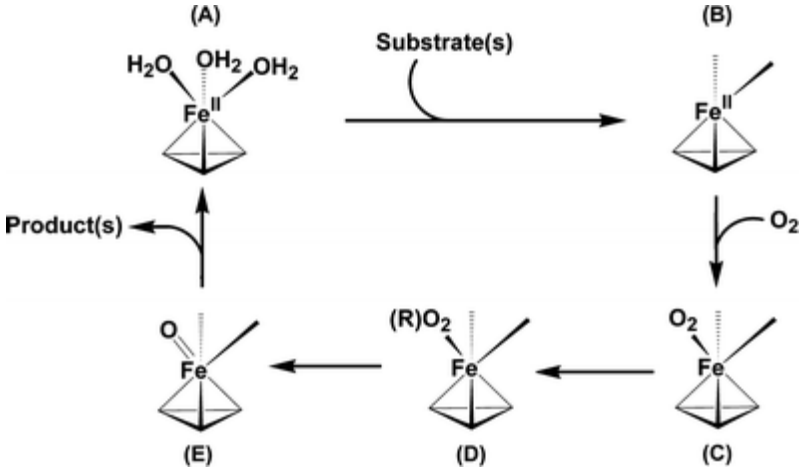


Figure 5.3: General mechanistic scheme proposed for enzymes with the 2-His-1-carboxylate facial triad (Koehntop *et al*., 2005). Refer to text for detailed explanation.

A recent review summarized the current knowledge on enzymes with the 2-His-1-carboxylate facial triad and derived mechanistic features general to all enzyme classes (Fig. 5.3; Koehntop *et al*., 2005). At the beginning of the catalytic cycle, the Fe(II) center is coordinatively saturated and relatively unreactive towards oxygen (Fig 5.3A). Upon substrate binding, the abstraction of one or several water ligands occurs, leading to a coordinatively unsaturated five-coordinated Fe site which is poised to bind O_2 (Fig 5.3B, C). The substrate binds not in all cases directly to the center, but it always

triggers the opening of the sixth coordination site for the attack of dioxygen. This is a mechanistic strategy to prevent a self-inactivation. Common to all of these enzymes is the generation of an Fe(III)-peroxo species upon dioxygen binding (Fig. 5.3D). The next step involves cleavage of the O-O bond. As catalytic steps subsequent to O_2 activation often occur rapidly and only one enzymatic intermediate has been trapped thus far (Fe(IV)-oxo intermediate in TauD; Price *et al.*, 2003), different oxygen species are discussed as intermediates. The intermediate, whatever nature it has, transforms the substrate into product which is released from the active site. The resting form of the enzyme is subsequently regenerated.

Transferred onto the SOR, distinct steps in the catalytic mechanism can be derived (Fig. 5.4). It has to be noted that several cycles of catalysis are necessary for a complete oxidation/reduction of the probably long-chain sulfur substrate, which distinguishes the SOR from the other enzymes with the 2-His-1-carboxylate facial triad.

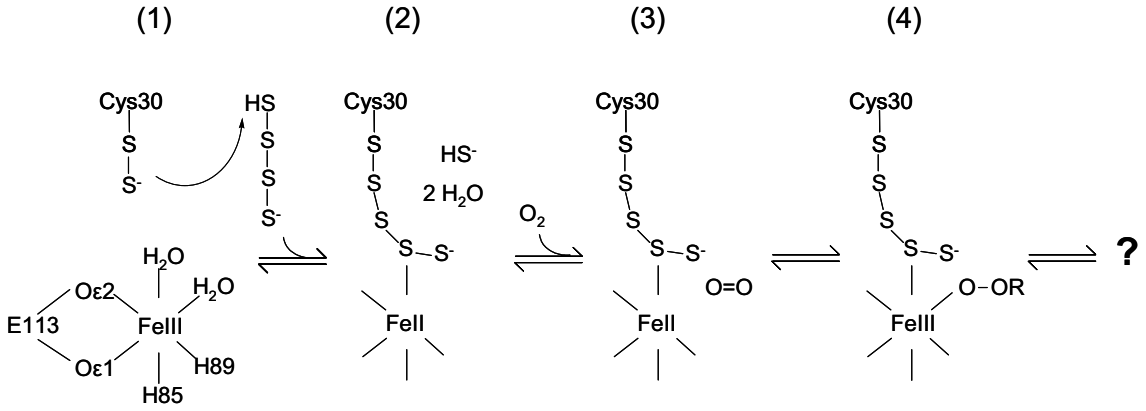


Figure 5.4: Hypothetical proposal for the initial steps of the catalytic mechanism of the SOR. Note that the chain length of the polysulfide species is hypothetical. Refer to text for details.

(1) The resting state of the SOR (Fig. 5.4A). Two coordination sites of Fe are occupied by water molecules. Although isolated as Fe(III), its resting state might be Fe(II), as this is a common theme for enzymes with the 2-His-1-carboxylate facial triad. Support comes from EPR as discussed above. It is not clear whether Cys30 is persulfurated at that stage but it must be in the negatively charged thiolate state.

(2) Covalent binding of a linear polysulfide species occurs at Cys30, probably leading to the formation of one hydrogen sulfide (Fig. 5.4B). Upon binding, the substrate is aligned close to the Fe site, which results in the abstraction of one or both water ligands, leading to a coordinatively unsaturated Fe site. At that stage the center is in the Fe(II) state, eventually by a one-electron transfer from the cysteine-bound polysulfide. As describe

above, both a direct coordination to the Fe(II) site or coordination close to it are described for enzymes with the 2-His-1-carboxylate facial triad. In the case of the SOR, sulfur is the electron donor during catalysis, thus it most probably gets ligated to the site to ease electron transfer.
(3) Dioxygen binds to an open position at the coordinatively unsaturated Fe(II) site (Fig. 5.4C).
(4) One-electron transfer from Fe(II) to dioxygen results in a Fe(III)-peroxo species as the first activated oxygen intermediate (Fig. 5.4D).

Up to here, the proposed steps are supported by experimental findings and by the mechanistic analogy to other 2-His-1-carboxylate facial triad enzymes. Any proposal about subsequent catalytic steps is highly speculative due to the nature of the long-chain sulfur substrate which requires several cycles of catalysis. In addition, the sulfur disproportionation reaction with sulfur being both the electron donor and electron acceptor has to be incorporated in a mechanistic scheme.
For the related sulfur oxygenase from *A. brierleyi* only a moderate incorporation of ^{18}O from $^{18}O_2$ into sulfite has been reported (~11% sulfite ^{18}O-labeled; Emmel *et al*., 1986) and it is reasonable to assume that this also holds true for the *A. ambivalens* enzyme. In consequence, all other oxygen atoms incorporated into sulfite have to be derived from water, albeit no direct experimental proof exists for the *A. ambivalens* SOR. This observation does not point to a direct reaction of sulfur with the peroxo intermediate, because this would require an oxidant capable of ^{18}O exchange with water like it has been shown by $H_2{}^{18}O$ labeling experiments with toluene dioxygenase (Wackett, 2002). A heterolytic cleavage of the O-O bond to generate a OH^- ion and an oxoiron(V) intermediate prior to sulfur oxidation would then be favored. This is supported by the requirement of hydroxyl ions for the non-enzymatic disproportionation of sulfur at elevated temperatures (McCampbell Hamilton, 1991). In mechanistic terms the iron site could then be considered as the initial hydroxyl ion generator for the sulfur disproportionation reaction.
This postulate could only hold true for the first cycle of sulfur conversion, as a complete generation of all hydroxyl ions from dioxygen could not explain the sulfur reductase activity, because of dioxygen being the exclusive electron acceptor. In consequence, alternative mechanisms have to be discussed. The initial oxygenation reaction mediated by the Fe center could result in an "activated" sulfur species which could undergo a disproportionation reaction independent of the Fe site with hydroxyl ions derived from water. Alternatively, the disproportionation reaction could also be mediated by the Fe site, if sulfur could replace oxygen at the Fe center and be reduced instead. This is supported

by the low midpoint potential of the site, which is sufficiently negative to allow a sulfur reduction.

The few mechanistic data that are available so far are not sufficient to develop a complete scheme for the catalytic reaction of the SOR. Considerably more efforts are to be done for its elucidation. But the data collected during this work provide a basis for future experiments.

5.5. THE SUBSTRATE OF THE SOR

Although elemental sulfur in the form of the cyclic α-S_8 ring is used as the *in vitro* substrate of the SOR in the activity assay, for several reasons it is questionable whether α-S_8 is the actual substrate of the enzyme. α-S_8 sulfur is not well suited as substrate because of its low solubility (5 µg/l H_2O at 25°C; Boulegue, 1978) and therefore low availability. The sulfur reductase of *Wolinella succinogenes* has been shown to be a polysulfide reductase (Hedderich *et al*., 1999). As judged from the SOR structure, α-S_8 sulfur can hardly enter the inner cavity and the active sites due to the narrow pores at both entries. In contrast, the more soluble and linear polysulfide species could access the sites more easy.

Indirect support for a linear polysulfide as substrate comes from the cytoplasmic location of the SOR. A transport of α-S_8 sulfur to the cytoplasm is unlikely because it could affect the integrity of the cytoplasmic membrane due to its hydrophobicity. More likely it may be transported as protein- (thiol-) bound polysulfide, which could be released on the cytoplasmic side of the membrane.

5.6. COMPARISON OF THE SOR WITH OTHER SULFUR OXIDIZING ENZYMES

Very limited structural information is available on enzymes involved in the microbial dissimilatory sulfur oxidation pathways. Crystal structures of the thiosulfate oxidoreductase SoxAX (Bamford *et al*., 2002) and the sulfur carrier protein SoxYZ from *Rhodovulvum sulfidophilum* have been reported in recent years (Sauvé *et al*., 2004; the coordinates of SoxYZ are not yet available in public databases). In contrast to the icosatetrameric SOR, SoxAX is a heterodimeric c-type cytochrome composed of two subunits with M_r of 15 kDa (SoxX) and 30 kDa (SoxA, Fig. 5.5). The structural differences are probably based on different functional requirements. The SOR is a cytoplasmic enzyme which disproportiates elemental sulfur to sulfite, thiosulfate and hydrogen sulfide, which subsequently serve as soluble substrates for the enzymes downstream in the sulfur oxidation pathway of *A. ambivalens* (Kletzin *et al*., 2004). In contrast, SoxAX is a

thiosulfate oxidoreductase and part of the periplasmic Sox complex composed of seven different proteins catalyzing the complete oxidation of reduced sulfur compounds like thiosulfate to sulfate without any free intermediates (Friedrich *et al*., 2001).

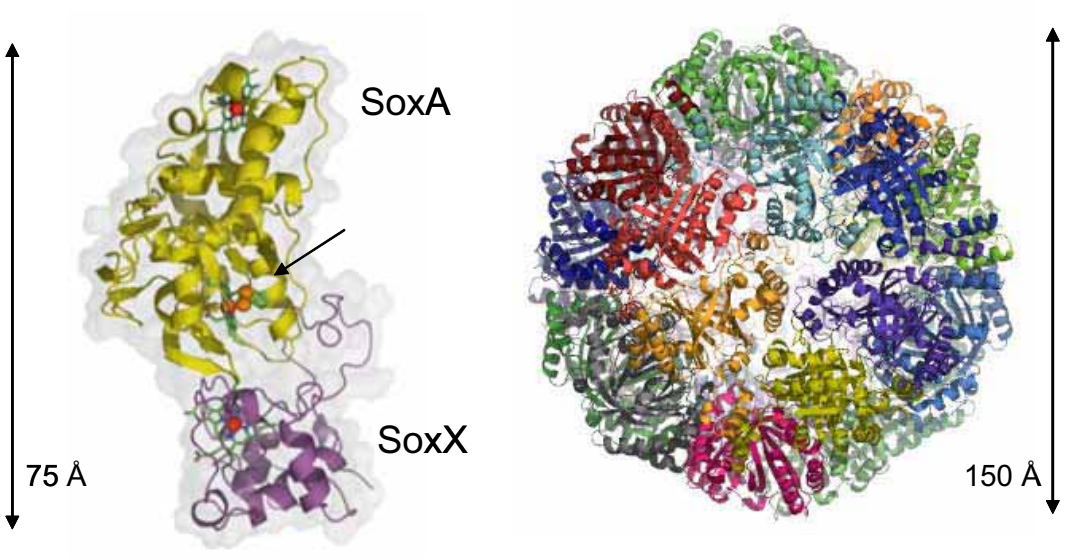


Figure 5.5: The crystal structures of SoxAX (left) and SOR (right) shown as cartoon representations. The two subunits of SoxAX are differently colored. Heme groups are shown as stick representations with their Fe ions colored in red. The arrow indicates the proposed active site with the cysteine residue indicated as spheres.

The SOR and SoxAX also differ in their cofactor content. Whereas the SOR contains a mononuclear non-heme iron site for oxygen activation, SoxAX harbors three heme groups which are reduced upon thiosulfate oxidation and mediate electron transfer to cytochrome c_2 in an oxygen-independent reaction (Bamford *et al*., 2002). The Sox complex contains additional cofactors including heme groups (SoxD), manganese (SoxB) and a molybdenum cofactor (SoxC; Friedrich *et al*., 2001). Siroheme was found in a sulfite oxidoreductase from *Starkeya novella* (Kappler & Dahl, 2001). A flavin cofactor is present in the sulfide:quinone oxidoreductase from *Rhodobacter capsulatus* (Griesbeck *et al*., 2002). No information is thus far available about the cofactors of the bacterial sulfur oxygenases (Rohwerder & Sand, 2003). In summary, a diversity of cofactors is utilized for substrate oxidation and electron transfer.

In contrast, this study together with others highlights a common theme for the sulfur compound-oxidizing enzymes: The presence of cysteines, which act as covalent binding sites of the sulfur substrates.

This work showed that Cys30 is the likely sulfur binding site of the *A. ambivalens* SOR, which is in accordance with other studies on the *A. tengchongensis* SOR (Shen *et al*.,

2005). Similarly, SoxAX contains a cysteine residue in the active site which is postulated to covalently bind thiosulfate prior to the oxidative transfer of latter to SoxYZ (Bamford *et al*., 2002). SoxYZ then serves as a substrate carrier protein during all further steps of catalysis and it was shown that a cysteine is likewise the covalent binding site of thiosulfate (Quentmeier & Friedrich, 2001). Similarly, cysteines are probably involved in covalent binding of sulfide during the catalysis of sulfide:quinone oxidoreductase from *Rhodobacter capsulatus*, as shown by a mutagenesis study (Griesbeck *et al*., 2002).

5.7. CONCLUSIONS AND PERSPECTIVES

The work presented here is the first study focusing on the structure-function relationship of an enzyme involved in the biological oxidation of elemental sulfur. It showed that the SOR adopts a particular architecture to perform its function, i.e. the formation of an inner compartment from where the 24 active sites are accessible. The SOR´s unique sulfur disproportionation reaction appears to be an adaptation to the increased reactivity of sulfur at the high growth temperatures of *A. ambivalens.* The central players during catalysis are Cys30, acting as covalent binding site of the substrate, and a mononuclear non-heme iron center, being the site of the initial oxygen activation.

The SOR will be a valuable model system to study the mechanistic principles underlying enzymatic sulfur oxidation and disproportionation, respectively. Especially the construction of SOR mutants, combined with X-ray crystallography and substrate soaking experiments will be a very promising approach. In addition, reaction intermediates could be identified by spectroscopic methods, like MCD (magnetic circular dichroism), EXAFS (extended X-ray absorption fine structure) and Moessbauer spectroscopy. As large amounts of protein are necessary for these experiments, the heterologous expression of the *sor* gene in *E. coli* and the rapid purification protocol established in the course of this work will be of special value. Furthermore, the SOR could serve as a model system for the study of general biological principles like protein assembly, thermostability or protein folding.

6. MATERIALS AND METHODS

6.1. MATERIALS

6.1.1. CHEMICALS

The chemicals and materials used and their sources are listed below. Chemicals and materials not mentioned were from Merck KG, Darmstadt, Germany or from Sigma, St. Louis, USA, and are mentioned in the context of the respective method.

2-mercaptoethanol	Carl Roth, Karlsruhe, Germany
4x Roti Load reducing/nonreducing	Carl Roth, Karlsruhe, Germany
6-amino caproic acid	Carl Roth, Karlsruhe, Germany
6x Loading Dye	MBI Fermentas, Heidelberg, Germany
Agarose (Sea Kem ME)	BMA, Walkersville, USA
Ampicillin	Roche Diagnostics, Mannheim, Germany
Anhydrotetracycline (AHT)	IBA, Göttingen, Germany
Bovine serum albumine (BSA)	Roche Diagnostics, Mannheim,Germany
Chaps	Serva, Heidelberg, Germany
Chloramphenicol	Carl Roth, Karlsruhe, Germany
Chromatographic columns for FPLC	Amersham Pharmacia, Freiburg, Germany
Coomassie Brilliant Blue G250	Serva, Heidelberg, Germany
Coomassie Brilliant Blue, colloidal	Carl Roth, Karlsruhe, Germany
D-Desthiobiotine	IBA, Göttingen, Germany
Desoxynucleotidtriphosphate	Roche Diagnostics, Mannheim, Germany
Ethanol	Carl Roth, Karlsruhe, Germany
Ethidiumbromide	Carl Roth, Karlsruhe, Germany
Ethylendiamine tetraacetic acid (EDTA)	LS, Darmstadt, Germany
Glycerol	AppliChem, Darmstadt, Germany
Glycine	Serva, Heidelberg, Germany
Hydrochloric acid (HCl)	Carl Roth, Karlsruhe, Germany
Long Ranger gel solution	FMC BioProducts, Rochland, USA
Methanol	Carl Roth, Karlsruhe, Germany
Nitrogen, liquid	Air Liquide, Krefeld, Germany
Rotiphorese Gel 30	Carl Roth, Karlsruhe, Germany
Sodiumdodecylsulfat (SDS)	Carl Roth, Karlsruhe, Germany
Strep-Tactin Gravityflow column (1 ml)	IBA, Göttingen, Germany
Strep-Tactin Superflow column (10 ml)	IBA, Göttingen, Germany
Sulfur, purified	Merck, Darmstadt, Germany
Tetramethylethylendiamine (TEMED)	Serva, Heidelberg, Germany
TPTZ (2,4,6-Tri-(2-pyridyl)-1,3,5-triazine)	Acros, Geel, Belgium
Tri-chloro acetic acid	Riedel-de Haen, Seelze, Germany
Tris-(hydroxymethyl)-aminomethan (Tris)	Carl Roth, Karlsruhe, Germany
Triton X-100	Carl Roth, Karlsruhe, Germany
Tryptone	Oxoid, Wesel, Germany
Tween-20	Serva, Heidelberg, Germany
Urea	Carl Roth, Karlsruhe, Germany
Yeast extract	Difco/Nordwald, New York, USA

6.1.2. MICROORGANISMS

Acidianus ambivalens DSM 3772	DSMZ, Braunschweig, Germany
E. coli BL21-CodonPlus(DE3)-RIL	Stratagene, Heidelberg, Germany
E. coli TOP10F	Stratagene, Heidelberg, Germany
E. coli DH10B	Stratagene, Heidelberg, Germany

6.1.3. PLASMIDS

pASK75	Skerra, 1994
pASK-sor.05	Urich *et al.*, 2004
pASK-sorC30A	Urich *et al.*, 2005b
pASK-sorC30S	"
pASK-sorH85A	"
pASK-sorH89A	"
pASK-sorC100A	"
pASK-sorC100S	"
pASK-sorC103A	"
pASK-sorC103S	"
pASK-sorC100/103A	"
pASK-sorC100/103S	"
pASK-sorE113A	"
pASK-sorE113D	"

6.1.4. ENZYMES AND KITS

Restriction endonucleases	New England Biolabs, Schwalbach, Germany MBI Fermentas, Heidelberg, Germany Stratagene, Heidelberg, Germany
$DAP_{Goldstar}$ Polymerase	Eurogentec, Seraing, Belgium
Taq Polymerase	Invitrogen, Carlsbad, USA
Pfu Polymerase	MBI Fermentas, Heidelberg, Germany
T4 DNA Ligase	Hybaid, Ashford, UK
Alkaline shrimps phosphatase	United States Biochemical, Cleveland, USA
NucleoSpin® Plasmid MiniPrep kit	Machery & Nagel, Düren, Germany
Qiaquick gel extraction kit	Qiagen, Hilden, Germany
SequiTherm EXCEL™ II Sequencing Kit LC	Biozym, Oldendorf, Germany
Ultrafree®-DA	Millipore, Schwalbach, Germany
Nucleobond AX 500 kit	Machery & Nagel, Düren, Germany

6.1.5. SYNTHETIC OLIGONUCLEOTIDES

The synthetic oligonucleotides were from MWG-Biotech, Sigma-Genosys and Invitrogen, all Germany. They were delivered lyophilized and were resuspended in HPLC purified

H_2O and stored at -20°C. The sequencing oligonucleotides were HPLC puridied and stored at -20°C (Tab. 6.1).

Table 6.1: synthetic oligonucleotides. The numbering of the mutation-inducing oligonucleotides refers to the amino acid sequence of the mature protein and not to the codons of the *sor* gene, which differ by one as the N-terminal methionine is post-translationally cleaved off.

Oligonucleotide	Sequence (5'→3')
N- and C-terminal	
sorn_fwd	CGA CAA AAA TCT AGA TAA CGA GGG C
sorc_rev	GTG ACG CCA AGC GCT TTG TTC GTT
Mutation-introducing	
C_{30}A fwd	ACC GAA GGT CGC CAT GGT AAC
C_{30}A rev	GTT ACC ATG GCG ACC TTC GGT
C_{30}S fwd	ACC GAA GGT CTC CAT GGT ACC
C_{30}S rev	GTT ACC ATG GAG ACC TTC GGT
H_{85}A fwd	TTG GAA AGA ACG CTG AAG AAA TGC
H_{85}A rev	GCA TTT CTT CAG CGT CTT TCC CAA
H_{89}A fwd	GAA GAA ATG GCC AGG CAA AAC TGG
H_{89}A rev	CCA GTT TTG CCT GGC CAT TTC TTC
C_{100}A fwd	TTA TTC CAG GCT GCC TAT TCA TGC
C_{100}A rev	GCA TGA ATA GGC TAG CCT GAA TAA
C_{100}S fwd	TTA TTC AGG CTA TCC TAT TCA TGC
C_{100}S rev	GCA TGA ATA GGA TAG CCT GAA TAA
C_{103}A fwd	CTA TGC TAT TCA GCC GCT TCA C
C_{103}A rev	GTG AAG CGG CTG AAT AGC ATA G
C_{103}S fwd	CTA TGC TAT TCA TCC GCT TCA C
C_{103}S rev	GTG AAG CGG ATG AAT AGC ATA G
$C_{100/103}$A fwd	TCA GGC TAG CCT ATT CAG CCG CTT CAC AA
$C_{100/103}$A rev	TTG TGA AGC GGC TGA ATA GG C TAG CCT GA
$C_{100/103}$S fwd	AGG CTA TCC TAT TCA TCC GCT TCA
$C_{100/103}$S rev	TGA AGC GGA TGA ATA GGA TAG CCT
E_{113}A fwd	GAC CCT GGG CGC CAA TTT ATG
E_{113}A rev	CAT AAA TTG GCG CCC AGG GTC
E_{113}D fwd	GAC CCT GGG ATC CAA TTT ATG
E_{113}D rev	CAT AAA TTG GAT CCC AGG GTC
DNA-sequencing[a]	
ask1	TATT TTA CCA CTC CCT ATC AG
ask2	CAT CGC CAT TTG CCG TCT GTT

[a]The sequencing oligonucleotides were labeled 5' with the fluorescence dye IRD 800.

6.1.6. MOLECULAR MASS MARKERS

GeneRuler DNA Ladder Mix	MBI-Fermentas, Heidelberg, Germany
Protein Molecular Weight Marker	MBI-Fermentas, Heidelberg, Germany
MW ND-500 kit	Sigma, St. Louis, USA
HMW Gel filtration calibration kit	Sigma, St. Louis, USA

6.1.7. MEDIA

LB-Medium	10	g	Tryptone
	5	g	Yeast extract
	10	g	NaCl
			ad 1 Liter with H_2O dest., pH 7.0
			Ampicillin: 100 µg/ml Chloramphenicol: 30 µg/ml for solid media: 10 g Agar/l
SOB-Medium	20	g	Tryptone
	5	g	Yeast extract
	0.5	g	NaCl
			ad 1 Liter with H_2O dest.
SOC-Medium	100	ml	SOB-Medium
	1	ml	1 M $MgSO_4$
	1	ml	1 M $MgCl_2$
	1	ml	2 M Glucose (sterile filtered)
Acidianus Medium	1%	w/v	Sulfur
	0.04%	w/v	Yeast extract
	1.5	g	$(NH_4)_2SO_4$
	2.8	g	KH_2PO_4
	0.25	g	$MgSO_4$ x 7 H_2O
	70	mg	$CaCl_2$ x 2 H_2O
	28	mg	$FeSO_4$ x 7 H_2O
	0.5	mg	$Na_2B_4O_7$ x 10 H_2O
	0.36	mg	$MnCl_2$ x 4 H_2O
	440	µg	$ZnSO_4$ x 7 H_2O
	100	µg	$CuCl_2$ x 2 H_2O
	70	µg	$VOSO_4$ x 2 H_2O
	60	µg	Na_2MoO_4 x 2 H_2O
	20	µg	$CoSO_4$ x 7 H_2O
			ad 1 liter with $_{dd}H_2O$; pH 2.5 with H_2SO_4

6.1.8. BUFFERS AND SOLUTIONS

TE-Buffer	20	mM	Tris/HCl, pH 8.0
	1	mM	EDTA
TAE-Buffer	40	mM	Tris
	2	mM	EDTA
			pH 7.5 with acetic acid
10x TBE-Buffer	890	mM	Tris
	889	mM	Borate
	25	mM	EDTA
KPi pH 6.5 (1 M)	107.7	g	K_2HPO_4
	51.82	g	KH_2PO_4
			ad 1 liter with $_{dd}H_2O$
KPi pH 8.0 (1 M)	163.75	g	K_2HPO_4
	8.15	g	KH_2PO_4
			ad 1 liter with $_{dd}H_2O$

6.2. MOLECULAR GENETIC METHODS

6.2.1. SITE-DIRECTED MUTAGENESIS

An expression plasmid had already been constructed previously (Zimmermann, 1998), where the *sor* gene was amplified from genomic *A. ambivalens* DNA using the oligonucleotides SoroN (TAGAC ATCTA GATAA CGAGG GCAAA AAATG CCGAA ACCAT ACGT) and SoroC (AGAAA AAGCT TAAGC GCTTT GTTCG TTTAA ATATT CTC). The product was digested with *Xba*I and *Eco*47III and ligated into the pASK75 plasmid (Skerra, 1994) double digested with the same enzymes resulting in a gene fusion encoding a C-terminal 10 amino acid Strep-tag (Fig. 6.1). Strep-tag is a peptide which binds with high affinity to streptavidin and can therefore be used for purification by affinity chromatography (Skerra, 1994).

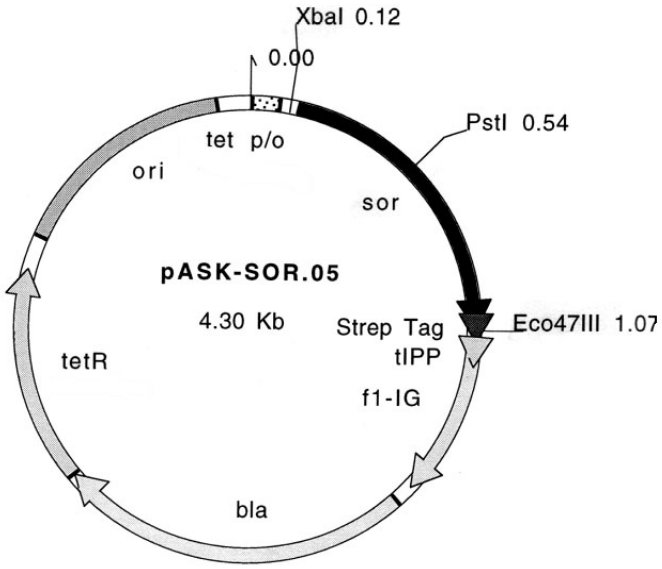


Figure 6.1: Map of the pASK-sor.05 plasmid used for expression of the *sor* gene. It carries the *sor* gene as insert fused to a 3´ sequence coding for a Strep-tag. The expression is under control of the tet promotor/operator (tet p/o) which is repressed by the plasmid-encoded tet repressor (tetR) and can be induced by anhydrotetracycline. The *bla* gene encodes for a β-lactamase used for selection of positive clones.

Site-directed mutants of several SOR codons were constructed according to the overlap extension method (Ho *et al.*, 1989; see Tab. 6.1 for a list of the oligonucleotides). Two separate polymerase chain reactions (PCR) were carried out with the respective flanking and mutation-carrying oligonucleotides using the pASK-sor.05 plasmid as template

(0.25 mM dNTPs, 1.2 U DAP-polymerase, 2 mM $MgCl_2$, 50 pmol oligonucleotides, 1 µl pASK-sor.05 in a total volume of 50 µl; see below for PCR conditions).

	Denaturing	Annealing	Polymerization
1 Cycle	5 min, 94°C	1.30 min, 58°C	1:30 min, 72°C
33 Cycles	0.45 min, 94°C	1.30 min, 58°C	1:30 min, 72°C
1 Cycle	0.45 min, 94°C	1.30 min, 58°C	5 min, 72°C

The PCR products were separated by agarose gel electrophoresis (1 % agarose in TAE buffer; Sambrock *et al.*, 1989), stained with ethidium bromide and the desired fragments were cut out under UV light. The DNA was extracted using the Ultrafree-DA kit (Millipore). Five µl of the respective products were used as template in the 3rd PCR under the same conditions as described above. After agarose gel electrophoresis and staining the desired fragments were isolated and extracted using the Qiaquick gel extraction kit (Quiagen) according to the manufacturers instructions. The fragments were double digested with the restriction endonucleases *Xba*I and *Eco*47III for 2 h at 37°C (Sambrock *et al.*, 1989).

6.2.2. PREPARATION OF THE PASK75 VECTOR

The pASK75 plasmid was transformed into electro-competent *E. coli* DH10B cells (2.4 kV voltage, 25 µF condensator capacity, 200 Ω resistance). The cells were resuspended in 1 ml SOC medium, regenerated for 1 h at 37°C, and afterwards plated on LB_{amp}-agar followed by incubation at 37°C. A 200 ml culture was grown from a single colony. The plasmids were isolated from the harvested cells using the Nucleobond AX 500 kit (Macherey & Nagel) according to the manufacturers instructions. The plasmids were double-digested with *Xba*I and *Eco*47III for 2 h at 37°C (Sambrock *et al.*, 1989) after agarose gel electrophoresis and gel extraction as described above. After dephosphorylation with shrimps alkaline phosphatase the vector was purified by phenol-chloroform extraction and ethanol precipitation (Sambrock *et al.*, 1989) and stored at −20°C in TE buffer.

6.2.3. LIGATION AND ANALYSIS OF TRANSFORMANDS

The ligation of PCR fragments and vector was carried out using the T4 DNA ligase and with PCR product in excess to vector for 2 h at 14°C (Sambrock *et al.*, 1989).
Competent *E. coli* cells were transformed with the ligated plasmids by electroporation (*E. coli* DH10B) or heat shock (*E. coli* Top10F; Sambrock *et al.*, 1989). Positive transformands carrying the desired inserts were detected by plasmid mini prep (Nucleo Spin, Macherey & Nagel) followed by restriction with *Xba*I and *Eco*47III or by colony PCR

using the PCR conditions as described above and the flanking oligonucleotides (Sambrock *et al.*, 1989).

6.2.4. DNA SEQUENCING

Positive constructs were sequenced for the verification of the introduced mutations according to the dideoxy chain termination method (Sanger *et al.*, 1977). The sequencing PCR was done according to the "Cycle Sequencing Protocol" of the SequiTherm EXCEL Sequencing kit (Biozym). The fluorescence-dye labeled oligonucleotides ask1 or ask2 (see Tab. 6.1) were used as primers. They bind 77 base pairs (bp) upstream (ask1) and 78 bp downstream (ask2) of the *sor* gene and allowed the forward and reverse sequencing of the introduced mutations. The reaction mixtures contained the following compounds.

7.2 µl 3.5-fold sequencing buffer
2 pmol oligonucleotide
3 µl DNA
ad 16 µl with $_{dd}H_2O$
1 µl SequiTherm EXCEL II DNA Polymerase (5 U/µl)

4 µl of the reaction mixture were added to 2 µl termination mixture, which contained one dideoxynucleotide besides the four deoxynucleotides. The sequencing PCR was done in a Robocycler Gradient 40 (Stratagene, Heidelberg, Germany) using the conditions noted below.

	Denaturing	Annealing	Polymerization
1 Cycle	3 min, 94°C	1.30 min, 56°C	2 min, 70°C
35 Cycles	0.30 min, 94°C	1.30 min, 56°C	2 min, 70°C
1 Cycle	0.30 min, 94°C	1.30 min, 56°C	7 min, 70°C

3 µl stopping buffer were added to each reaction mixture at the end of the PCR and the samples were stored at -20°C. After denaturing at 90°C the fragments were separated on a denaturing poly acryl amide gel (25 g urea, 21 ml $_{dd}H_2O$, 5 ml 10X TBE buffer, 5 ml Long Ranger acryl amide (50 %), 35 µl TEMED and 250 µl 0.1 g/ml APS) and detected in a Li-Cor DNA Sequencer 4000 apparatus. The retrieved sequences were analyzed for the mutated codons using the programs GAP and FASTA of the GCG Wisconsin package (Accelrys, Cambridge, UK).

6.2.5. HETEROLOGOUS GENE EXPRESSION

E. coli BL 21 Codon plus cells (DE3-RIL) were transformed with the original pASK75 vector and the *sor* gene constructs for gene expression. The strain harbors the "Codon plus" plasmid encoding tRNA genes for the arginine codons AGA and AGG, isoleucine AUA and leucine CUA, which are underrepresented in *E. coli* but often used in AT-rich organisms like *A. ambivalens*, thus these tRNAs are provided to avoid their limitation during protein biosynthesis.

The expression of the *sor* gene was induced by addition of anhydrotetracycline (200 µg/l of culture) to cultures growing in LB medium at 37 ˚C at OD_{600} = 0.8. The cultures were incubated for at least 5 h or overnight after induction. Usually, a strong formation of inclusion bodies within the cells was observed microscopically. The cells were sedimented by centrifugation at 10,800 x g (Beckmann SLA-3000, 20 min) and frozen at -20°C.

6.3. BIOCHEMICAL METHODS

6.3.1. PURIFICATION OF RECOMBINANT SOR FROM *E. COLI*

The harvested *E. coli* cells were resuspended in 10 volumes (v/w) 100 mM KP_i buffer pH 8 containing 1 % Tween 20 and 50 µM EDTA. The cells were disrupted in a Vibrogen V/3 cell mill (Bühler, Tübingen, Germany) by shaking with glass beads for 20 min. After removal of the glass beads by filtration the crude extract was centrifuged (10,800 x g, 20 min, Sorvall SLA-3000). Alternatively, cells were resuspended in 10 volumes (v/w) 50 mM Tris/HCl pH 7.2 and disrupted by a single passage through a French press (0.18 mm nozzle and 135 MPa pressure, Basic-Z, Constant Systems, Warwick, UK). The crude extract was centrifuged as described above.

The supernatant containing the soluble SOR was heat-treated for 15 min at 75˚C to precipitate heat labile *E. coli* proteins and centrifuged (85,000 x g, 30 min, Beckmann 45Ti). The supernatant obtained was dialyzed against 20 mM Tris/HCl pH 7.5 and again centrifuged (85,000 x g, 30 min, Beckmann 45Ti). 100 ml of the dialyzed protein fraction were applied to a DEAE-Sepharose fast-flow column (V_c 60 ml) and eluted at 5 ml/min with an increasing gradient of NaCl in the same buffer. The recombinant SOR eluted at 150 mM NaCl. Active fractions were pooled. The SOR was further purified and sedimented by ultracentrifugation at 235,000 x g for 16 h (Beckmann 45Ti). The pellet containing the SOR was resuspended in 50 mM Tris/HCl pH 7.5 and frozen at -20°C.

In later stages of this work, recombinant SOR was purified by Strep-Tactin affinity chromatography (IBA, Göttingen, Germany). Large amounts of soluble extract were

applied to a Strep-Tactin Superflow column (V_C 10 ml) connected to an FPLC apparatus, whereas smaller volumes of supernatant (< 20 ml) were loaded on 1 ml Strep-Tactin Gravity flow columns. The columns had been equilibrated with 5 V_C buffer W prior to loading of the protein solution. Subsequently, the columns were washed with 5 V_C buffer W. The protein was eluted with 3 V_C buffer E containing desthiobiotin. The columns were regenerated by the application of 3 x 5 V_C buffer R.
A Strep-Tactin protein purification protocol without the metal chelator EDTA in the buffers was applied when the protein samples were subsequently be subjected to the TPTZ assay for iron quantification. It had been observed that the presence of EDTA interfered in the TPTZ assay and that it could not quantitatively be removed by dialysis.

Buffer W: 100 mM Tris/HCl pH 8.0, 150 mM NaCl, 1 mM EDTA
Buffer E: 100 mM Tris/HCl pH 8.0, 150 mM NaCl, 1 mM EDTA, 2.5 mM desthiobiotin
Buffer R: 100 mM Tris/HCl pH 8.0, 150 mM NaCl, 1 mM EDTA, 1 mM hydroxy-azophenyl benzoic acid (HABA)

The pellet of the crude *E. coli* extract containing large amounts of inclusion bodies was washed four times by sonication for 5 min (Branson Sonifier II, Heinemann, Schwäbisch Gmünd, Germany) in washing buffer and centrifugation (10,800 x g, 10 min, Sorvall SLA-3000).

Washing buffer: 100 mM KP_i pH 8, 1 % sodium deoxycholate, 1 M urea, 10 % glycerol, 2 mM EDTA

6.3.2. UNFOLDING AND REFOLDING OF RECOMBINANT SOR FROM INCLUSION BODIES

Washed inclusion bodies (~ 5 g wet mass) were dissolved in 50 ml 8 M urea containing 30 mM DTT. After denaturation at 60˚C for 1 h the solution was centrifuged to remove particular material at 48,000 x g (Sorvall SS34, 20 min). The supernatant containing unfolded SOR was used for refolding.
An initial screen for appropriate refolding conditions was performed aerobically and anaerobically at 4˚C. Unfolded SOR samples were dialyzed aerobically and anaerobically against 50 mM Tris/HCl pH 7.2 containing 2 mM 2-mercaptoethanol and in the presence or absence of 10 µM ferric citrate. Additionally, samples were refolded by a two-step dialysis against 50 mM Tris/HCl pH 7.2 containing 2 M urea in the first and no urea in the second step, respectively. All dialysis steps were done overnight. Subsequently, the samples were centrifuged at 48,000 x g to remove precipitated protein (Sorvall SS34,

20 min). The supernatants containing soluble SOR were subjected to the SOR activity assay. They were kept aerobically at 4°C for up to 8 d to analyze the refolding efficiency.
Finally, refolding of SOR was performed aerobically at 4˚C by a one-step dialysis against 50 mM Tris/HCl pH 7.2 containing 2 mM 2-mercaptoethanol, 100 µM ferric citrate and 2 M urea in the first and no urea in the second step, respectively. Residual ferric citrate and 2−mercaptoethanol were removed from the refolded enzyme preparation in another dialysis step against 50 mM Tris/HCl pH 7.2 prior to centrifugation at 48,000 x g (Sorvall SS34, 20 min).
In a different experiment, unfolded inclusion bodies were dialyzed three times against 8 M urea without ferric citrate to remove residual iron from the samples. The samples were subsequently refolded with 0 or 10 µM ferric citrate as described above and analyzed for activity.

6.3.3. PURIFICATION OF WILD TYPE SOR FROM *A. AMBIVALENS*

Acidianus ambivalens DSM 3772 was grown aerobically according to published procedures (Zillig *et al.*, 1986; Teixeira *et al.*, 1995). Two different protocols were applied for the preparation of wild type SOR.
In protocol I (Scholz, 1999), cells were resuspended in 10 volumes (v/w) 100 mM KP_i buffer pH 8.0 and broken by sonication (Branson sonifier II). After the addition of 50 µg / ml DNase I and incubation for 30 min at 37°C the suspension was centrifuged at 15,000 x g (Sorvall GSA, 20 min) to remove unbroken cells and sulfur particles. The supernatant was centrifuged at 85,000 x g for 1 h (Beckmann 45Ti). 1 ml each of the resulting supernatant was loaded onto a tube with a 10 ml linear 28-40 % (w/w) sucrose density gradient and centrifuged for 48 h at 288,000 x g (Beckmann SW41). 1 ml fractions of the gradient were collected from top to bottom and analyzed for SOR activity. Active fractions were combined and dialyzed against 0.5 M ammonium sulfate, 20 mM Tris/HCl pH 7.5. After centrifugation to remove precipitated proteins the supernatant was loaded on a HiLoad 26/10 phenyl sepharose column equilibrated with the same buffer and eluted at 5 ml/min with a decreasing ammonium sulfate gradient in an FPLC. The SOR eluted at around 50 mM $(NH_4)_2SO_4$ and was pooled according to activity. The active fractions were dialyzed against 20 mM Tris/HCl pH 7.5 and applied to a DEAE-Sepharose fast-flow column (V_c 60 ml) and eluted at 5 ml/min. The SOR eluted at 150 mM NaCl. The active fractions were pooled and concentrated using a Jumbosep ultrafiltration unit (cut off size 30 kDa; Pall Filtron, Bonn, Germany) and stored at -20 °C.
In protocol II, cells of *A. ambivalens* were resuspended in 40 mM Potassium phosphate buffer pH 6.5 and broken at 6000 psi in a SLM Aminco KINO20 French press (Spectronic

Instruments; Rochester, USA). Unbroken cells were removed by centrifugation at 15,000 x g for 20 min. The crude extract was then centrifuged for 8h at 138,000 x g (Beckmann 45Ti). The resulting supernatant (900 ml) was applied to a 330 ml Q-Sepharose fast flow column and eluted at 5 ml/min. The SOR eluted at around 180 mM NaCl and was pooled according to activity. This fraction was dialyzed against 50 mM potassium phosphate buffer pH 7.5 and applied to a second Q-Sepharose High performance column (Vc 60 ml). The active fractions, which eluted at around 130 mM NaCl, were pooled and concentrated by ultrafiltration over a 30 kDa cut-off membrane. The concentrated fraction was applied onto an S-200 gel permeation column (Superdex 200, 3.2 x 60 cm) equilibrated with 50 mM Potassium phosphate buffer pH 7.5, 150 mM NaCl. Active fractions were pooled, concentrated and stored at -20°C.

6.3.4. THE SOR ACTIVITY ASSAY

The SOR activity assay was performed aerobically as previously described (Kletzin, 1989). SOR aliquots (usually 5 µg) were incubated in 1 ml assay buffer (70 mM Tris/HCl pH 7.2, 0.1 % Tween 20, 2 % (w/v) sulfur, dispersed by sonication) at 85°C for different time intervals (usually 0, 4, 8, 12 and 16 min) and the reaction was stopped by cooling on ice. After centrifugation for sulfur sedimentation the products sulfite, thiosulfate and hydrogen sulfide were determined colorimetrically and quantitated by the use of respective calibration curves. The reaction velocity was calculated by the linear increase of product concentrations. It was assumed that sulfite and sulfide are the sole primary products and that thiosulfate resulted from a non-enzymatic reaction between sulfite and sulfur (Kletzin, 1989). One unit (U) of enzyme activity was defined as the formation of 1 µmol of sulfite + thiosulfate (oxygenase activity), respectively 1 µmol hydrogen sulfide (reductase activity) per minute.

6.3.4.1 Hydrogen sulfide determination

The determination is based on the Fe^{3+} catalyzed formation of methylene blue from hydrogen sulfide and N,N-dimethyl-p-phenylenediamine (King & Morris, 1967). The reagents were applied in the following order.

250 µl	zinc acetate 2.6 % (w/v) in $_{dd}H_2O$, 0.1 % (v/v) acetic acid (100%)
350 µl	sample
50 µl	6 % NaOH
125 µl	N,N-dimethyl-p-phenylenediamine 0,1% (w/v) in 5 M HCl
50 µl	11.5 mM $FeCl_3$ in 0.6 N HCl

After 12-24 h the adsorption was determined at 670 nm against H_2O.

6.3.4.2 Thiosulfate determination

The decolorization of methylene blue by thiosulfate was measured. 250 µl of sample were applied to 750 µl methylene blue solution (24 mg/l methylene blue in H_2O, 1:1 mixed with 10 M HCl; Pachmair, 1960). The adsorption at 670 nm was determined after 30 min against H_2O as reference.

6.3.4.3 Sulfite determination

Sulfite reduces fuchsin and the addition of formalin (37 % formaldehyde) leads to the formation of a stable purple compound (Pachmair, 1960). 50 µl reagent (40 mg fuchsine dissolved in 87.5 ml $_{dd}H_2O$ and 12.5 ml H_2SO_4 conc.) and 195 µl $_{dd}H_2O$ were mixed with 250 µl sample. After 5 min incubation 5 µl formalin was added. The adsorption at 570 nm was determined after 90 min against H_2O as reference.

6.3.4.4 K_M determination

Due to the insolubility of elemental sulfur, different detergents (Tween20, sodium deoxycholate, CHAPS) and concentrations (0.1 - 1 %) were tested for their ability to solubilize sulfur at high temperatures and for their interference in the colorimetric assays. Finally, a concentration of 0.5 % (w/v) CHAPS was applied for the K_M and K_{cat} determinations and the assay temperature was lowered to 65 °C to avoid sulfur precipitation. Sulfur concentrations of 624, 312, 156, 78, 39.5, 20 and 4 mM (calculated on the basis of S_o, not S_8) were used. The determinations of the two K_M values for the oxygenase and reductase reactions were performed using the rate of formation of the products thiosulfate and sulfide separately. V_{max} values were calculated from a Lineweaver-Burk plot, whereas the K_M values were derived from manually fitting a curve derived from the Michaelis-Menten formula $v_0 = v_{max} \times [S] / (K_M + [S])$ into the observed velocity values for each substrate concentration using the v_{max} values calculated from the Lineweaver-Burk plots.

6.3.4.5 Inhibitor studies

Enzyme assays for inhibition studies were usually performed with EDTA-washed sulfur prepared as follows: 50 g of sulfur was dispersed in 500 ml enzyme buffer with 10 mM EDTA by sonication for 5 min with a Branson Sonifier (II) followed by centrifugation three times each. EDTA was removed by filtration of the sulfur with 5 l $_{dd}H_2O$. Alternatively, unwashed sulfur or potassium tetrathionate were used as substrates. The following reagents were tested for potentially inhibitory effects: EDTA, $ZnCl_2$, $CaCl_2$, $MgCl_2$, Fe-citrate, $FeCl_3$, $(NH_4)_2Fe(SO_4)_2$, and Na_3citrate. The reagents were added to the enzyme-containing assay mixture. After 30 min incubation at room temperature the pH was readjusted to 7.2 with a solution of 1 M Tris base and the enzymatic reaction was started by raising the temperature to 85 °C. An enzymatic reaction without inhibitor and a non-enzymatic reaction with inhibitor but without SOR were used as controls. The interference of the substances with the colorimetric assays was determined by adding mixtures of sodium sulfite, sodium thiosulfate and sodium sulfide (50 µM each) in enzyme buffer to the colorimetric assay reactions with inhibitor concentrations equal to the ones used in the enzymatic assays.

6.3.5. PROTEIN QUANTIFICATION

The protein concentration was determined by the Bradford (Bradford, 1976) or BCA method (Smith *et al.*, 1985). The protein concentration of pure SOR preparations was determined by measuring the OD_{280} using the calculated absorption coefficient ε = 57,840.

6.3.6. IRON QUANTIFICATION

The iron content of SOR samples was determined by the 2,4,6-tripyridyl-1,3,5-triazine method (TPTZ, Fischer & Price, 1964). TPTZ and Fe^{2+} form a blue complex which can be quantified at 593 nm. A protein solution of desired concentration was filled up to a volume of 800 µl with $_{dd}H_2O$. 100 µl 8 M HCl were added for iron extraction. After 10 min incubation at room temperature the protein was precipitated by the addition of 100 µl 80 % trichloroacetic acid (TCA) and sedimented by centrifugation (20 min, 13,000 rpm, 4°C). 250 µl of 75 % ammonium acetate were added to 800 µl of supernatant for neutralization. Fe^{3+} was reduced to Fe^{2+} by the addition of 100 µl 10 % hydroxylamine hydrochloride. After the addition of 100 µl 4 mM TPTZ solution the resulting blue complex was quantified immediately at 593 nm in a Beckmann DU 640 spectrophotometer.

6.3.7. GEL PERMEATION CHROMATOGRAPHY

200 µl of the purified wild type and recombinant SOR preparations were run over a Superose 6 HR 10/30 gel permeation column (Amersham) for the determination of the apparent molecular masses of the holoenzymes. The column was equilibrated with 20 mM Tris/HCl buffer pH 7.5 and 150 mM NaCl. The proteins were eluted at 0.2 ml/min in the same buffer. Thyroglobulin, ferritin, catalase, aldolase, chicken albumin and carbonic anhydrase (Sigma) were used for calibration.

For the determination of the molecular masses of unfolding intermediates 200 µl each of a 1 mg/ml solution of recombinant SOR were dialyzed against 0-8 M urea in 20 mM KP_i buffer pH 6.5 at room temperature for 48-72 h. The dialyzed protein fractions were applied to a Superose 6 HR 10/30 gel permeation column equilibrated with 20 mM Tris/HCl buffer pH 7.5, 150 mM NaCl not containing any urea and evolved as described above.

6.3.8. SDS-PAGE

The purity of protein samples was assessed by denaturing SDS-PAGE (sodium dodecylsulfate polyacrylamide *g*el electrophoresis; Schägger & Von Jagow, 1987), which separates proteins according to their molecular masses. Several 10 % gels were prepared in a vertical chamber.

50 ml separation gel (10 %)	30 ml stacking gel (4 %)
18.8 ml Rotiphorese Gel 30[a]	3.9 ml Rotiphorese Gel 30[a]
16.6 ml Gel buffer (3x)	7.5 ml Gel buffer (3x)
5.4 ml Glycerol	-
9.2 ml $_{dd}H_2O$	19.5 ml $_{dd}H_2O$
16.3 µl TEMED	24 µl TEMED
162.5 µl APS (10% w/v)	240 µl APS (10% w/v)

[a] Rotiphorese Gel 30 (Roth) is a mixture of acrylamide and bis-acrylamide (37.5:1).

The protein samples were mixed with 4x Roti-Load buffer (Roth) and denatured at 95°C for 5 min prior to the loading of the gel. The electrophoresis was done at ~80 V. The gels were subsequently stained with colloidal coomassie blue (Roth) overnight.

3x Gel buffer: 3 M Tris/HCl, 0.3 % SDS, pH 8.45

10x Anode buffer: 0.2 M Tris/HCl pH 8.9

10x Kathode buffer: 0.1 M Tris/HCl, 0.1 M Tricine, 0.1 % SDS, pH 8.45

Staining solution: 20 ml colloidal coomassie brilliant blue, 60 ml H_2O, 20 ml methanol

6.3.9. BN-PAGE

Non-denaturing Blue Native polyacrylamide gel electrophoresis (BN-PAGE) (Schägger & Von Jagow, 1991) was used for the determination of the apparent molecular mass of the SOR holoenzyme and for the analysis of the unfolding of SOR upon chemical denaturation with urea. The following buffers were used for the electrophoresis:

3x Gel buffer: 150 mM Bis-Tris/HCl, 200 mM ε-aminocaproic acid, pH 7.0 at 4°C
Kathode buffer: 50 mM Tricine, 15 mM Bis-Tris/HCl, +/- 0.002 % (w/v) coomassie blue G 250, pH 7.0 at 4°C
Anode buffer: 50 mM Bis-Tris/HCl pH 7.0 at 4°C
10x Loading buffer: 5 % coomassie blue G 250, 500 mM ε-aminocaproic acid, 100 mM Bis-Tris/HCl, 10 % glycerol, pH 7.0

The gels (14 x 14 x 0.15 cm) were prepared and run in a SE400 vertical electrophoresis chamber (Hoefer Scientific Instruments, San Francisco, USA). Usually, 5 µl of 10x loading buffer were added to 45 µl sample and the mixture was applied to the sample wells of the stacking gel. The gels were run at 4°C. The voltage was set to 100 V until the proteins entered the separation gel and was then increased to 500 V. Initially, a cathode buffer with coomassie blue was used, which was exchanged against a cathode buffer not containing coomassie blue after the migration front of the coomassie blue tracking dye was half way through the separation gel. The electrophoresis was stopped when the migration front of the coomassie blue tracking dye reached the end of the separation gel. The gels were stained with colloidal coomassie blue for visualization of the protein bands.

For the **determination of unfolding intermediates** 200 µl each of a 1 mg/ml solution of recombinant SOR were dialyzed against 0-8 M urea in 20 mM KP_i buffer pH 6.5 at room temperature for 48-72 h. The samples were separated on 4-15 % gradient gels, which were prepared using a gradient former as described below.

	stacking gel	separation gel	
		light solution	heavy solution
Acryl amide	4%	4%	15%
Gel buffer	1x	1x	1x
Glycerol (v/v)	-	-	20%
TEMED (v/v)	0.08%	0.08%	0.08%
APS (w/v)	0.008%	0.008%	0.008%

The **apparent molecular mass** of the SOR was calculated from retardation coefficients of standard proteins after electrophoresis on BN gels of different acryl amide concentrations (5.5, 6.5, 7.5, 8.5, 9.5, 10.5 % acryl amide) applying the Ferguson plot (Ferguson, 1964). The MW ND-500 kit (Sigma) containing lactalbumin, carbonic anhydrase, ovalbumin, bovine serum albumin and urease was used as standard, in addition to catalase, ferritin, and thyreoglobulin (all from Sigma). The relative mobility of the proteins (R_f) was determined by dividing its migration distance from the top of the separating gel to the center of the protein band by the migration distance of the coomassie blue tracking dye. 100 x (log (R_f x 100)) were plotted against the gel concentrations (%) for each standard protein and for the SOR. The negative slopes from these graphs were then plotted against the known molecular masses of the standard proteins. The molecular mass of the SOR was determined from the resulting slope.

6.4. SPECTROSCOPIC METHODS

6.4.1. UV/VISIBLE SPECTROSCOPY

UV/Visible spectra were recorded at room temperature in a spectrophotometer (Beckmann DU 640).

6.4.2. CD SPECTROSCOPY

CD (circular dichroism) spectroscopy was used for analyzing the secondary structural content of SOR. CD spectra were recorded in a Jasco J810 spectropolarimeter. The CD spectra were obtained using a cuvette of 0.2 cm path length at a scanning speed of 50 nm / min and a bandwidth of 2 nm. Scans were collected at 1 nm intervals with a response time of 0.25 s and accumulated 10 times. The CD spectra were corrected by subtracting the spectrum of the buffer solution (20 mM KP_i, pH 6.5). The experimental data in mdeg (millidegrees) were converted to the mean molar ellipticity per residue by applying the following formula:

$$\Theta mrd = \Theta d / 10^3 * (M / (c * l * n_{res}))$$

with the terms included being:

Θmrd = mean molar ellipticity per residue [deg * cm^2 * $dmol^{-1}$ * res^{-1}]

Θd = ellipticity [mdeg]

M = molar mass of protein [g * $dmol^{-1}$]

c = protein concentration [mg * cm^{-3}]

l = pathlength [cm]

n_{res} = number of amino acid residues

The Θ_{mrd} of the spectra between 187 and 250 nm were used to calculate the α-helical and β-sheet content of the native, recombinant and refolded SOR preparations using the programs CONTINLL and SELCON3 included in the CDPro software package and a reference protein set containing 37 proteins (Sreerama & Woody, 2000).

6.4.3. EPR SPECTROSCOPY

Electron paramagnetic resonance (EPR) spectroscopy can detect atoms which contain unpaired (paramagnetic) electrons. EPR spectra of wild type and recombinant SOR were recorded at 4.6 K on a Bruker ESP380 spectrometer, equipped with an ESR 900 continuous flow helium cryostat.

For an EPR-followed redox titration *A. ambivalens* wild type SOR was titrated anaerobically in 50 mM KP_i pH 6.5 by sequential addition of alkaline buffered sodium dithionite and continuously flushing the cell with argon. EPR samples were withdrawn at the desired potential and immediately frozen in liquid nitrogen. The intensity of the g = 4.3 resonance was followed as a function of the reduction potential of the solution. The following compounds were used as redox mediators (40 µM each, E'_o in brackets): 1,2,-naphtoquinone-4-sulphonic acid (215 mV); 1,2-naphtoquinone (180 mV); trimethylhydroquinone (115 mV); phenazine methosulphate (80 mV), phenazine ethosulphate (55 mV), methylene blue (11 mV); indigo-tetrasulphonate (-30 mV); resorufin (-51 mV); indigo-trisulphonate (-70 mV); indigo disulphate (-182 mV); 2,5-hydroxy-p-benzoquinone (-130 mV); 2-hydroxy-1,4-naphtoquinone (-152 mV); antraquinone-2-sulphonate (-225 mV); phenosafranine (-255 mV); safranine (-280 mV); neutral red (−325 mV); benzyl viologen (-359 mV); and methyl viologen (-446 mV). A platinum and a silver/silver chloride electrode were used, calibrated against a saturated quinhydrone solution at pH 7.0. The reduction potentials are quoted in respect to the standard hydrogen electrode. The experimental data were analyzed by Tiago Bandeiras (ITQB, Oeiras, Portugal) using MATLAB (Mathworks, South Natick, Massachusetts).

6.4.4. FLUORESCENCE SPECTROSCOPY

The fluorescence of the aromatic amino acid tryptophane was used as an intrinsic probe to monitor the unfolding of the SOR upon incubation with the denaturants urea and guanidine hydrochloride. The experiments were done by Claudio M. Gomes and Sonia S.

Leal (ITQB, Oeiras, Portugal). Stock solutions of guanidine hydrochloride (GuHCl) and urea for protein unfolding of the soluble recombinant SOR were prepared following published recommendations (Nozaki, 1972) and final molarities were calculated from refractive index measurements. Denaturation curves were determined at 25 °C using wild type SOR in 50 mM Tris/HCl pH 7 buffer at a final concentration of 0.01 mg/ml. The unfolding process was monitored by fluorescence spectroscopy following tryptophane emission (300–450 nm, excitation 292 nm, 1 cm cell) on a Varian Eclipse fluorometer. The spectroscopic data were processed in order to determine the average emission wavelength at each denaturant concentration (Royer, 1995) and the resulting unfolding transitions were analysed using a two-state model, allowing for the calculation of thermodynamic parameters (Shirley, 1995). Direct fits to the experimental unfolding curves were performed using the following expression:

$$Y_{obs} = (Y_U - Y_F * (\exp((\Delta G_U(H_2O) - m * [\text{denaturant}])/RT)))/(1 + (\exp((\Delta G_U(H_2O) - m * [\text{denaturant}]/RT)))$$

where Y_{obs} , Y_U and Y_F are the observed spectroscopic signal, denaturant protein baseline, and folded protein baseline, respectively. The transitions midpoints were calculated as $\Delta G_U(H_2O)/m$ or by direct inspection of the transitions.

6.5. ELECTRON MICROSCOPY

Electron microscopy was done by Reinhard Rachel (Archaeenzentrum, Regensburg, Germany). SOR samples were diluted to a protein concentration of 0.1 mg ml^{-1}, applied to glow-discharged carbon-coated copper grids and negatively stained with 2% uranyl acetate. Electron micrographs were taken using a slow-scan CCD Tietz camera mounted to a Philips CM 12 electron microscope operated at 120 kV. Image analysis by single particle averaging was performed as described (Schurig *et al.*, 1995; Schmidt *et al.*, 1994) using the EM program (Hegerl, 1996) and the SEMPER software (Saxton, 1996).

6.6. SEQUENCE ANALYSIS

The amino acid sequence of the *A. ambivalens* SOR was compared to the public databases using different BLAST tools for the identification of homologues. SOR sequences were retrieved and aligned using the alignment tools PILEUP, CLUSTALW and HMMERALIGN included in the GCG Wisconsin package (Accelrys, Cambridge, UK). A phylogenetic tree was constructed with PILEUP. The PREDICTPROTEIN server and the programs within were used to predict secondary structure elements

(http://cubic.bioc.columbia.edu/predictprotein). The *Acidithiobacillus ferrooxidans* genome was queried at the TIGR web server (http://www.tigr.org). The helix representation was drawn using the program HELICALWHEEL (GCG Wisconsin package).

6.7. CRYSTALLOGRAPHIC METHODS

6.7.1. CRYSTALLIZATION

Recombinant SOR was concentrated to 13.1 mg ml^{-1} in 50 mM Tris/HCl pH 7.5. Initial crystallization attempts were performed at 22°C using Jena Bioscience Screens I, II, IV and V (http://www.jenabioscience.com; 144 conditions) on a Douglas Instruments Oryx 6 robot and applying the microbatch technique with drops of 0.4 µl protein solution and 0.4 µl crystallization solution. Other tries were performed at 22°C using the Hampton Research MembFac Screen (http://www.hamptonresearch.com; 64 conditions) applying the sitting drop vapor diffusion technique with drops of 2 µl protein solution and 2 µl of crystallization solution, which were allowed to equilibrate against 500 µl of crystallization solution in the well reservoir. In addition, the influence of temperature on crystal nucleation was investigated by crystallization tries at 4°C and 32°C. Appropriate cryo protection conditions were established by soaking crystals for 1 min in their crystallization solution complemented with glycerol 25% (v/v) prior to flash-freezing in liquid nitrogen.

6.7.2. PREPARATION OF DERIVATIVES

The Mercury (Hg^{2+}) and gold (Au^{3+}) heavy metal derivatives were prepared by applying 5 mM parahydroxy mercury benzoic acid for 5 d and 10 mM potassium tetra chloro aureat for 1 d at 22°C to crystallization drops containing grown crystals. After washing in a cryoprotectant solution for 1 min crystals were flashfrozen in liquid nitrogen.

Several attempts were undertaken to prepare crystals with inhibitors, substrates, or products of the enzymatic reaction. A zinc complexed crystal was prepared by cocrystallizing the SOR in the presence of 10 mM zinc-acetate for 2 d at 32°C. SOR was cocrystallized with 10 mM iodoacetamide for 2 d. SOR crystals were grown during 4 d in the presence of 1 % (w/v) sulfur and 0.1 % TRITON X-100, with or without a pre-incubation at 70°C (in an attempt to induce a catalytic reaction). Crystals were soaked aerobically for 10 min with anaerobically prepared solutions of sodium thiosulfate, sodium sulfite and sodium sulfide (final concentrations 10 mM). A dithionite-reduced crystal was prepared by adding solid sodium dithionite for 10 min to drops with grown crystals. A solution of polysulfides was prepared by the addition of potassium tetrathionate to a final

concentration of 20 mM to an anaerobic solution of 100 mM Na_2S in Tris/HCl pH 7.5 (Klimmek et al., 1991). Formally, a part of the sulfide was oxidized by tetrathionate to elemental sulfur which reacted with the excess sulfide to give polysulfides. The resulting yellowish solution contained, besides polysulfides, sulfide, thiosulfate and elemental sulfur. Crystals were soaked for 1 d by adding 0.5 µl of this solution to 4 µl drops with grown crystals. After washing in a cryoprotectand solution without the reagents used for soaking, the crystals were flashfrozen in liquid nitrogen.

6.7.3. X-RAY DIFFRACTION DATA COLLECTION, PROCESSING AND ANALYSIS

Several X-ray diffraction data sets were collected at 110 K in house using a rotating anode or with synchrotron radiation at DESY (**De**utsches **E**lektronen **S**ynchrotron, Hamburg, Germany) or at ESRF (**E**uropean **S**ynchrotron **R**adiation **F**acility, Grenoble, France).
The highest resolution native diffraction data set (1.7 Å) was collected at beam line X11 of DESY. Another native diffraction data set was collected at the iron absorption edge (1.732 Å) at beam line BW7-A of DESY. Diffraction data of the Hg and Au derivatives were collected at ESRF beam line ID 14-3. Several diffraction data sets of crystals soaked with SOR inhibitors, substrates or products were collected at ESRF beam line ID 14-1. All diffraction data sets were processed with DENZO and scaled together into sets of observed intensities with SCALEPACK, both programs of the HKL suite (Otwinowski & Minor, 1997). The CCP4 suite (Collaborative Computational Project, 1994) was used with TRUNCATE to calculate observed amplitudes, and with FFT and POLARRFN to calculate the Patterson and the self-rotation Patterson functions of the native diffraction data set.

6.7.4. MIRAS PHASING

The resolution of the phase problem was obtained by Carlos Frazão (ITQB, Oeiras, Portugal) using SOLVE (Terwilliger & Berendzen, 1999). The native, Hg and Au derivatives intensity data sets were scaled together, and the heavy atom sites coordinates were obtained by the automatic analysis of the Patterson functions from their isomorphous differences. A common origin was assigned to both sets of heavy atoms and initial phases were estimated using MIRAS (**m**ultiple **i**somorphous **r**eplacement with **a**nomalous **s**cattering signal).

6.7.5. MODEL BUILDING AND REFINEMENT OF THE NATIVE STRUCTURE

The program ARP/WARP (Perrakis et al., 1999) was used for automated molecule tracing and side-chain model building. Subsequently, a model of one SOR monomer was

manually checked on a graphics workstation using XTALVIEW-XFIT (McRee, 1999) and chain gaps were filled in, leading to a model containing residues 1-307 out of 318 possible. The application of the non-crystallographic symmetry (NCS) operators derived from 12 heavy atom sites of the Hg^{2+} derivative, using LSQKAB from CCP4 (Collaborative Computational Project, 1994), led to the generation of the 6 monomers of SOR in the asymmetric unit, which were fitted into the electron density by their refinement as rigid bodies with REFMAC5 (Murshodov et al, 1999). ARP/WARP with dummy atom positioning was used to obtain the best maps, and one Fe atom per monomer was added manually using XTALVIEW-XFIT. Further refinement was performed with REFMAC5, using loose positional and isotropic thermal NCS restraints, interleaved by manual graphics inspection and improvement of the model using XTALVIEW-XFIT, in particular for discrete alternating side chain conformations. Doubtful side chain conformations were analyzed with the help of averaged sigmaA-weighted $2mIF_oI$ - DIF_cI and mIF_oI - DIF_cI electron density maps, generated by MAPROT (Collaborative Computational Project, 1994) applying rotation/translation matrices from the superposition of the monomers obtained with LSQKAB (Collaborative Computational Project, 1994). Water molecules were added by ARPWATERS (Perrakis et al., 1999) and manually checked based on sigmaA-weighted mIF_oI - DIF_cI maps of at least 3 σ and sensible hydrogen bonding.

An additional atom located at covalent distance of the Sγ atom of Cys30 in every monomer was found after the ARPWARP dummy atoms refinement. A tentative refinement with Cys30 exchanged to Css (cysteine-persulfide) or Cea (cysteine-sulfenic acid) led to acceptable flat sigmaA-weighted mIF_oI - DIF_cI maps. The chemical nature of the extra sulfur atom was confirmed by anomalous difference Fourier maps calculated from the dataset collected at 1.736 Å, where sulfur has a significant anomalous signal (0.7 electrons).

SHELXL (Sheldrick & Schneider, 1997) was used in the final stages of refinement allowing the refinement of the six independent iron sites restrained to their common geometry using SAME/SADI instructions. In addition, coupled alternating side chain conformations occupancies were allowed until the threshold of 0.75/0.25. A model of the 24 subunit holoenzyme was generated with PDBSET (Collaborative Computational Project, 1994) applying the crystallographic symmetry operations of space group I4.

6.7.6. DERIVATIVES: DATA PROCESSING, MODEL REFINEMENT AND ANALYSIS

All derivative datasets were processed with DENZO and scaled together with SCALEPACK, both programs of the HKL suite. They were further processed with programs included in the CCP4.2.2 program suite (Collaborative Computational Project,

1994). SCALEPACK2MTZ was used to convert the scalepack files into MTZ format. Intensities were converted into structure factor amplitudes with TRUNCATE. CAD was used for dataset sorting and merging the derivative and native datasets, which were subsequently scaled with SCALEIT. The final model of the native structure including cysteine persulfide, Fe ions and waters molecules was used as template for the refinement in REFMAC. After a rigid body refinement the models were refined until R values converged and were manually analyzed with XTALVIEW-XFIT. In addition, the mIF_oI - DIF_cI electron density maps were analyzed for regions of electron density not covered by the model with PEAKMAX (Collaborative Computational Project, 1994). Furthermore, averaged $2mIF_oI$ - DIF_cI and mIF_oI - DIF_cI electron density maps were calculated as described above.

6.7.7. MODEL ANALYSIS

The **stereochemical quality** of the models was analyzed using the program PROCHECK (Laskowski et al., 1993). The nature of the **subunit interfaces** was analysed with the Protein-Protein interaction server (http://www.biochem.ucl.ac.uk/bsm/PP/server; Jones & Thornton, 1996). The fractions of **solvent-accessible and buried monomer atoms** in the oligomeric and monomeric state were calculated with AREAIMOL (Collaborative Computational Project, 1994). The **electrostatic potentials and fields** were calculated with MEAD (Bashford, 1997) using a protein and external dielectric constants of 4 and 80, respectively, a temperature of 300 K, an ionic strength of 0 M, and a solvent probe radius of 1.4 Å. The calculations were done for a cubic box with xyz grid dimensions of 321 Å and a grid spacing of 1 Å, respectively. VOIDOO (Kleywegt & Jones, 1994) was used for particle volume calculations and for the **detection and analysis of cavities** in the model, applying the rolling probe function with a probe radius of 1.4 Å. **Searches for structural homologues** were performed using the programs DALI (http://www.ebi.ac.uk/dali; Holm & Sander, 1995), DEJAVU (http://portray.bmc.uu.se/cgi-bin/dejavu/scripts/dejavu.pl; Madsen & Kleywegt, 2002) and CE (http://cl.sdsc.edu/ce.html; Shindyalov & Bourne, 1998) using mainly the default search parameters. The **structural superposition** of the SOR monomer halves and of proteins with a related fold was done with MODELLER6v2 (Sali & Blundell, 1993), a program for protein structure modeling based on spacial restraints. An initial alignment as starting model was manually constructed in which the β-strands were aligned. The α carbon main chain atoms were then automatically fitted with a cutoff distance of 3.5 Å for structurally equivalent positions. **Substrate modeling was done using** XTALVIEW-XFIT by manually placing the substrate molecules in the electron density of the SOR model based on spacial restraints derived from the Van der Waals

distances between protein and substrate. The **figures** were prepared with Pymol (Delano Scientific, California, USA; http://pymol.sourceforge.net).

ABBREVATIONS

°C	degrees Celsius
A	Adenine
ad	filled up
Amp	Ampicillin
APS	Ammonium persulfate
AU	asymmetric unit
BN	Blue native
bp	base pairs
C	Cytosine
CHAPS	3-[(3-Cholamidopropyl)-dimetylammonio]-1-propane-sulfonate
Da	Dalton
$_{dd}H_2O$	double deionized water
ddNTP	Dideoxynucleotide triphosphate
DESY	Deutsches Elektronen Synchrotron
DNA	Dsoxyribonucleic acid
dNTP	Deoxynucleotide triphosphate
DSMZ	Deutsche Sammlung für Mikroorganismen und Zellkulturen
e.g.	for example
EDTA	Ethylene-diamine-tetra acetic acid
ESRF	European Synchrotron Radiation Facility
et al.	and others
F_c	calculated structure factor
F_o	observed structure factor
FPLC	Fast Performance Liquid Chromatography
g	gram
G	Guanine
h	hour
HABA	Hydroxy-azophenyl-benzoic acid
HCl	hydrochloric acid
I	intensity
i.e.	that is
K	Kelvin
kb	kilo bases
kDa	kilo dalton
l	liter
m	meter
M	molar
MAD	Multiple wavelength anomalous diffraction
min	minute
MIRAS	multiple isomorphous replacement with anomalous difference
mU	milli unit
NCS	non-crystallographic symmetry
nt	Nucleotide
OO_{280}	optical density at 280 nm
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
PEG	polyethyleneglycol
r.m.s.(d.)	root mean square (deviation)
rpm	rounds per minute
S^0	elemental sulfur
SDS	sodium dodecyl sulfate
sec	second
SOR	Sulfur oxygenase reductase
T	Thymine
TCA	Trichloro acetic acid
TEMED	Tetramethylethylendiamine

TPTZ	2,4,6-Tripyridyl-s-triazine
Tris	Tris(hydroxymethyl)-aminomenthane
Triton X100	PEG(9-10)p-t-octylphenol
Tween20	PEG(20) Sorbitan monolaurate
U	Unit
UV	Ultra-violet
UV/Vis	Ultra-violet/visible
V	Volt
v/v	volume per volume
v/w	volume per weight
w/v	weight per volume
w/w	weight per weight

LITERATURE

Abramson, J., Smirnova, I., Kasho, V., Verner, G., Kaback, H. R. & Iwata, S. (2003). Structure and mechanism of the lactose permease of *Escherichia coli*. *Science* **301**, 610-615.

Abreu, I. A., Xavier, A. V., LeGall, J., Cabelli, D. E. & Teixeira, M. (2002). Superoxide scavenging by neelaredoxin: dismutation and reduction activities in anaerobes. *J Biol Inorg Chem* **7**, 668-674.

Alberts, B. (1998). The cell as a collection of protein machines: preparing the next generation of molecular biologists. *Cell* **92**, 291-294.

Amend, J. P. & Shock, E. L. (2001). Energetics of overall metabolic reactions of thermophilic and hyperthermophilic Archaea and bacteria. *FEMS Microbiol Rev* **25**, 175-243.

Andersen, O. A., Flatmark, T. & Hough, E. (2002). Crystal structure of the ternary complex of the catalytic domain of human phenylalanine hydroxylase with tetrahydrobiopterin and 3-(2-thienyl)-L-alanine, and its implications for the mechanism of catalysis and substrate activation. *J Mol Biol* **320**, 1095-1108.

Bahadur, R. P., Chakrabarti, P., Rodier, F. & Janin, J. (2003). Dissecting subunit interfaces in homodimeric proteins. *Proteins* **53**, 708-719.

Bamford, V. A., Bruno, S., Rasmussen, T., Appia-Ayme, C., Cheesman, M. R., Berks, B. C. & Hemmings, A. M. (2002). Structural basis for the oxidation of thiosulfate by a sulfur cycle enzyme. *Embo J* **21**, 5599-5610.

Bandeiras, T. M., Salgueiro, C., Kletzin, A., Gomes, C. M. & Teixeira, M. (2002). *Acidianus ambivalens* type-II NADH dehydrogenase: genetic characterisation and identification of the flavin moiety as FMN. *FEBS Lett* **531**, 273-277.

Barlow, D. J. & Thornton, J. M. (1983). Ion pairs in proteins. *J Mol Biol* **168**, 867-885.

Barns, S. M., Jeffries, M. W. & Pace, N. R. (1994). Remarkable achaeal diversity in a Yellowstone National Park hot sprong environment. *Proc Natl Acad Sci* **91**, 1609-1613.

Bashford, D. (1997). An object-oriented programming suite for electrostatic effects in biological molecules. In *Scientific Computing in Object-Oriented Parallel Environments*, pp. 233-240. Edited by Y. Ishikawa, Oldehoeft, R.,R., Reynders, J.,V., Tholburn, M. Berlin: Springer.

Baumeister, W., Walz, J., Zuhl, F. & Seemuller, E. (1998). The proteasome: paradigm of a self-compartmentalizing protease. *Cell* **92**, 367-380.

Beinert, H. (2000). A tribute to sulfur. *Eur J Biochem* **267**, 5657-5664.

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. & Bourne, P. E. (2000). The Protein Data Bank. *Nucleic Acids Res* **28**, 235-242.

Bordo, D., Deriu, D., Colnaghi, R., Carpen, A., Pagani, S. & Bolognesi, M. (2000). The crystal structure of a sulfurtransferase from *Azotobacter vinelandii* highlights the evolutionary relationship between the rhodanese and phosphatase enzyme families. *J Mol Biol* **298**, 691-704.

Boulegue, J. (1978). Solublity of elemental sulfur in water at 298 K. *Phosphorus Sulfur* **5**, 127-128.

Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**, 248-254.

Brock, T. D., Brock, K. M., Belly, R. T. & Weiss, R. L. (1972). Sulfolobus: A new genus of sulfur-oxidizing bacteria living at low pH and high temperature. *Arch Microbiol* **84**, 54-68.

Brüser, T., Lens, P. & Trüper, H. G. (2000). The biological sulfur cycle. In *Environmental Technologies to treat Sulfur Pollutions*. Edited by P. Lens & L. Hulshoff Pol. London: IWA Publishing.

Chan, M. K., Mukund, S., Kletzin, A., Adams, M. W. W. & Rees, D. R. (1995). Structure of a hyperthermophilic Tungstopterin Enzyme, Aldehyde Ferredoxin Oxidoreductase. *Science* **267**, 1463-1469.

Chen, Z. W., Jiang, C. Y., She, Q., Liu, S. J. & Zhou, P. J. (2005). Key role of cysteine residues in catalysis and subcellular localization of sulfur oxygenase-reductase of *Acidianus tengchongensis*. *Appl Environ Microbiol* **71**, 621-628.

Claiborne, A., Yeh, J. I., Mallett, T. C., Luba, J., Crane, E. J., 3rd, Charrier, V. & Parsonage, D. (1999). Protein-sulfenic acids: diverse roles for an unlikely player in enzyme catalysis and redox regulation. *Biochemistry* **38**, 15407-15416.

Cobessi, D., Huang, L. S., Ban, M., Pon, N. G., Daldal, F. & Berry, E. A. (2002). The 2.6 A resolution structure of *Rhodobacter capsulatus* bacterioferritin with metal-free dinuclear site and heme iron in a crystallographic 'special position'. *Acta Crystallogr D Biol Crystallogr* **58**, 29-38.

Coelho, A. V., Macedo, S., Matias, P. M., Thompson, A. W., LeGall, J. & Carrondo, M. A. (2001). Structure determination of bacterioferritin from *Desulfovibrio desulfuricans* by the MAD method at the Fe K-edge. *Acta Crystallogr D Biol Crystallogr* **57**, 326-329.

Collaborative Computational Project, N. (1994). The CCP4 Suite: Programs for Protein Crystallography. *Acta Crystallogr D Biol Crystallogr* **50**, 760-763.

Costas, M., Mehn, M. P., Jensen, M. P. & Que, L., Jr. (2004). Dioxygen activation at mononuclear nonheme iron active sites: enzymes, models, and intermediates. *Chem Rev* **104**, 939-986.

Dahl, C., Engels, S., Pott-Sperling, A. S., Schulte, A., Sander, J., Lubbe, Y., Deuster, O. & Brune, D. C. (2005). Novel genes of the dsr gene cluster and evidence for close interaction of Dsr proteins during sulfur oxidation in the phototrophic sulfur bacterium *Allochromatium vinosum*. *J Bacteriol* **187**, 1392-1404.

Dietmann, S. & Holm, L. (2001). Identification of homology in protein structure classification. *Nat Struct Biol* **8**, 953-957.

Doolittle, R. F. (1995). The multiplicity of domains in proteins. *Annu Rev Biochem* **64**, 287-314.

Drenth, J. (1999). *Principles of protein X-ray crystallography*, 2nd edn. Heidelberg: Springer-Verlag.

Dutta, S. & Berman, H. M. (2005). Large macromolecular complexes in the protein data bank: a status report. *Structure (Camb)* **13**, 381-388.

Emmel, T., Sand, W., König, W. A. and Bock, E. (1986). Evidence for the existence of a sulphur oxygenase in *Sulfolobus brierleyi*. *J Gen Microbiol* **132**, 3415-3420.

Ferguson, K. A. (1964). Starch-Gel Electrophoresis--Application To The Classification Of Pituitary Proteins And Polypeptides. *Metabolism* **13**, 985-1002.

Finster, K., Liesack, W. & Thamdrup, B. (1998). Elemental sulfur and thiosulfate disproportionation by *Desulfocapsa sulfoexigens* sp. nov., a new anaerobic bacterium isolated from marine surface sediment. *Appl Environ Microbiol* **64**, 119-125.

Fischer, D. S. & Price, D. C. (1964). Simple serum iron method using new sensitive chromogen tripyridyl-s-triazine. *Clin Chem*, 21-25.

Friedrich, C. G. (1998). Physiology and genetics of sulfur-oxidizing bacteria. *Adv Microb Physiol* **39**, 235-289.

Friedrich, C. G., Quentmeier, A., Bardischewsky, F., Rother, D., Kraft, R., Kostka, S. & Prinz, H. (2000). Novel genes coding for lithotrophic sulfur oxidation of *Paracoccus pantotrophus* GB17. *J Bacteriol* **182**, 4677-4687.

Friedrich, C. G., Rother, D., Bardischewsky, F., Quentmeier, A. & Fischer, J. (2001). Oxidation of reduced inorganic sulfur compounds by bacteria: emergence of a common mechanism? *Appl Environ Microbiol* **67**, 2873-2882.

Frolow, F., Kalb, A. J. & Yariv, J. (1994). Structure of a unique twofold symmetric haem-binding site. *Nat Struct Biol* **1**, 453-460.

Fuchs, T., Huber, H., Burggraf, S. & Stetter, K. O. (1996). 16S rDNA-based phylogeny of the archaeal order *Sulfolobales* and reclassification of *Desulfurolobus ambivalens* as *Acidianus ambivalens* comb. nov. *System Appl Microbiol* **19**, 56-60.

Gomes, C. M. & Teixeira, M. (1998). Ambineela, an unusual blue protein isolated from the archaeon *Acidianus ambivalens*. *Biochem Biophys Res Commun* **249**, 23-25.

Gomes, C. M., Backgren, C., Teixeira, M., Puustinen, A., Verkhovskaya, M. L., Wikstrom, M. & Verkhovsky, M. I. (2001a). Heme-copper oxidases with modified D- and K-pathways are yet efficient proton pumps. *FEBS Lett* **497**, 159-164.

Gomes, C. M., Bandeiras, T. M. & Teixeira, M. (2001b). A new type-II NADH dehydrogenase from the archaeon *Acidianus ambivalens*: characterization and in vitro reconstitution of the respiratory chain. *J Bioenerg Biomembr* **33**, 1-8.

Griesbeck, C., Schutz, M., Schodl, T., Bathe, S., Nausch, L., Mederer, N., Vielreicher, M. & Hauska, G. (2002). Mechanism of sulfide-quinone reductase investigated using site-directed mutagenesis and sulfur analysis. *Biochemistry* **41**, 11552-11565.

Hagedoorn, P. L., Schmidt, P. P., Andersson, K. K., Hagen, W. R., Flatmark, T. & Martinez, A. (2001). The effect of substrate, dihydrobiopterin, and dopamine on the EPR

spectroscopic properties and the midpoint potential of the catalytic iron in recombinant human phenylalanine hydroxylase. *J Biol Chem* **276**, 22850-22856.

Hahn, T. (2002). International Tables for Crystallography. Volume A: Space-group symmetry. Heidelberg: Springer Verlag.

He, Z., Li, Y., Zhou, P. & Liu, S. (2000). Cloning and heterologous expression of a sulfur oxygenase/reductase gene from the thermoacidophilic archaeon *Acidianus* sp. S5 in *Escherichia coli*. *FEMS Microbiol Lett* **193**, 217-221.

He, Z. G., Zhong, H. & Li, Y. (2004). *Acidianus tengchongensis* sp. nov., a new species of acidothermophilic archaeon isolated from an acidothermal spring. *Curr Microbiol* **48**, 159-163.

Hedderich, R., Klimmek, O., Kröger, A., Dirmeier, R., Keller, M., Stetter, K. O. (1999). Anaerobic respiration with elemental sulfur and with sulfides. *FEMS Microbiol Rev* **22**, 353-381.

Hegerl, R. (1996). The EM Program Package: A Platform for Image Processing in Biological Electron Microscopy. *J Struct Biol* **116**, 30-34.

Hegg, E. L. & Que, L., Jr. (1997). The 2-His-1-carboxylate facial triad--an emerging structural motif in mononuclear non-heme iron(II) enzymes. *Eur J Biochem* **250**, 625-629.

Ho, S. N., Hunt, H. D., Horton, R. M., Pullen, J. K. & Pease, L. R. (1989). Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene* **77**, 51-59.

Höcker, B., Beismann-Driemeyer, S., Hettwer, S., Lustig, A. & Sterner, R. (2001). Dissection of a (betaalpha)8-barrel enzyme into two folded halves. *Nat Struct Biol* **8**, 32-36.

Holm, L. & Sander, C. (1995). Dali: a network tool for protein structure comparison. *Trends Biochem Sci* **20**, 478-480.

Iciek, M. & Wlodek, L. (2001). Biosynthesis and biological properties of compounds containing highly reactive, reduced sulfane sulfur. *Pol J Pharmacol* **53**, 215-225.

Jaenicke, R. & Böhm, G. (1998). The stability of proteins in extreme environments. *Curr Opin Struct Biol* **8**, 738-748.

Jones, S. & Thornton, J. M. (1996). Principles of protein-protein interactions. *Proc Natl Acad Sci U S A* **93**, 13-20.

Kaiser, J. T., Clausen, T., Bourenkow, G. P., Bartunik, H. D., Steinbacher, S. & Huber, R. (2000). Crystal structure of a NifS-like protein from *Thermotoga maritima*: implications for iron sulphur cluster assembly. *J Mol Biol* **297**, 451-464.

Kambampati, R. & Lauhon, C. T. (1999). IscS is a sulfurtransferase for the in vitro biosynthesis of 4-thiouridine in *Escherichia coli* tRNA. *Biochemistry* **38**, 16561-16568.

Kantardjieff, K. A. & Rupp, B. (2003). Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein-nucleic acid complex crystals. *Protein Sci* **12**, 1865-1871.

Kappler, U. & Dahl, C. (2001). Enzymology and molecular biology of prokaryotic sulfite oxidation. *FEMS Microbiol Lett* **203**, 1-9.

Karlsson, A., Parales, J. V., Parales, R. E., Gibson, D. T., Eklund, H. & Ramaswamy, S. (2003). Crystal structure of naphthalene dioxygenase: side-on binding of dioxygen to iron. *Science* **299**, 1039-1042.

Karshikoff, A. & Ladenstein, R. (2001). Ion pairs and the thermotolerance of proteins from hyperthermophiles: a "traffic rule" for hot roads. *Trends Biochem Sci* **26**, 550-556.

Kelly, D. P. (1979). The sulphur cycle: definitions, mechanisms and dynamics. *Ciba Found Symp*, 3-18.

Kelly, D. P. (1982). Biochemistry of the chemolithotrophic oxidation of inorganic sulphur. *Philos Trans R Soc Lond B Biol Sci* **298**, 499-528.

Kelly, D. P. (1985). Physiology of the thiobacilli: elucidating the sulphur oxidation pathway. *Microbiol Sci* **2**, 105-109.

Kelly, D. P. & Wood, A. P. (1994). Enzymes involved in the microbial oxidation of thiosulfate and polythionates. *Meth Enzymol* **243**, 501-510.

Kelly, D. P., Shergill, J. K., Lu, W. P. & Wood, A. P. (1997). Oxidative metabolism of inorganic sulfur compounds by bacteria. *Antonie Van Leeuwenhoek* **71**, 95-107.

Kim, B. S., Lee, C. S., Yun, C. Y., Yeo, S. M., Park, W. M. & Kim, H. R. (2001). Characterization and immunological analysis of ferritin from the hemolymph of *Galleria mellonella*. *Comp Biochem Physiol A Mol Integr Physiol* **129**, 501-509.

King, T. E. & Morris, R. (1967). Determination of acid-labile sulfide and sulfhydryl groups. *Meth Enzymol* **10**, 634-641.

Kletzin, A. (1989). Coupled enzymatic production of sulfite, thiosulfate, and hydrogen sulfide from sulfur: purification and properties of a sulfur oxygenase reductase from the facultatively anaerobic archaebacterium *Desulfurolobus ambivalens*. *J Bacteriol* **171**, 1638-1643.

Kletzin, A. (1992). Molecular characterization of the *sor* gene, which encodes the sulfur oxygenase reductase of the thermoacidophilic archaeon *Desulfurolobus ambivalens*. *J Bacteriol* **174**, 5854-5859.

Kletzin, A. (1994). Sulfur oxidation and reduction in Archaea: Sulfur oxygenase/-reductase and hydrogenases from the extremely thermophilic and facultatively anaerobic Archaeon *Desulfurolobus ambivalens*. *System Appl Microbiol* **16**, 534-543.

Kletzin, A., Urich, T., Muller, F., Bandeiras, T. M. & Gomes, C. M. (2004). Dissimilatory oxidation and reduction of elemental sulfur in thermophilic archaea. *J Bioenerg Biomembr* **36**, 77-91.

Kleywegt, G. J. & Jones, T. A. (1994). Detection, delineation, measurement and display of cavities in macromolecular structures. *Acta Crystallogr D Biol Crystallogr* **50**, 178-185.

Klimmek, O., Stein, T., Pisa, R., Simon, J. & Kroger, A. (1999). The single cysteine residue of the Sud protein is required for its function as a polysulfide-sulfur transferase in *Wolinella succinogenes*. *Eur J Biochem* **263**, 79-84.

Klimmek, O., Kroger, A., Steudel, R., Holdt, G. (1991). Growth of *Wolinella succinogenes* with polysulphide as terminal acceptor of phosphorylative eletron transport. *Arch Microbiol* **155**, 177-182.

Koehntop, K. D., Emerson, J. P. & Que, L., Jr. (2005). The 2-His-1-carboxylate facial triad: a versatile platform for dioxygen activation by mononuclear non-heme iron(II) enzymes. *J Biol Inorg Chem*.

Lang, D., Thoma, R., Henn-Sax, M., Sterner, R. & Wilmanns, M. (2000). Structural evidence for evolution of the beta/alpha barrel scaffold by gene duplication and fusion. *Science* **289**, 1546-1550.

Lange, S. J. & Que, L., Jr. (1998). Oxygen activating nonheme iron enzymes. *Curr Opin Chem Biol* **2**, 159-172.

Laska, S., Lottspeich, F. & Kletzin, A. (2003). Membrane-bound hydrogenase and sulfur reductase of the hyperthermophilic and acidophilic archaeon *Acidianus ambivalens*. *Microbiology* **149**, 2357-2371.

Laskowski, R. A., Moss, D. S. & Thornton, J. M. (1993). Main-chain bond lengths and bond angles in protein structures. *J Mol Biol* **231**, 1049-1067.

Lawson, D. M., Artymiuk, P. J., Yewdall, S. J. & other authors (1991). Solving the structure of human H ferritin by genetically engineering intermolecular crystal contacts. *Nature* **349**, 541-544.

Lin, Y. J., Dancea, F., Lohr, F., Klimmek, O., Pfeiffer-Marek, S., Nilges, M., Wienk, H., Kroger, A. & Ruterjans, H. (2004). Solution structure of the 30 kDa polysulfide-sulfur transferase homodimer from *Wolinella succinogenes*. *Biochemistry* **43**, 1418-1424.

Lombardi, A. T. & Garcia Junior, O. (1999). An evaluation into the potential of biological processing for the removal of metals from sewage sludges. *Crit Rev Microbiol* **25**, 275-288.

Lupas, A., Flanagan, J. M., Tamura, T. & Baumeister, W. (1997). Self-compartmentalizing proteases. *Trends Biochem Sci* **22**, 399-404.

Macedo, S., Romao, C. V., Mitchell, E. & other authors (2003). The nature of the di-iron site in the bacterioferritin from *Desulfovibrio desulfuricans*. *Nat Struct Biol* **10**, 285-290.

Madsen, D., Kleywegt G., J., (2002). Interactive motif and fold recognition in protein structures. *J Appl Cryst* **35**, 137-139.

Matthews, B. W. (1968). Solvent content of protein crystals. *J Mol Biol* **33**, 491-497.

McCampbell Hamilton, E. (1991). The chemistry of low valent sulfur compounds in the sulfur-water system: University of Minnesota.

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

Middelburg, J. J. (2000). The geochemical sulfur cycle. In *Environmental Technologies to treat Sulfur Pollutions*. Edited by P. Lens & L. Hulshoff Pol. London: IWA Publishing.

Müller, F. H., Bandeiras, T. M., Urich, T., Teixeira, M., Gomes, C. M. & Kletzin, A. (2004). Coupling of the pathway of sulphur oxidation to dioxygen reduction: characterization of a novel membrane-bound thiosulphate:quinone oxidoreductase. *Mol Microbiol* **53**, 1147-1160.

Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* **53**, 240-255.

Nakamura, K., Nakamura, M., Yoshikawa, H. & Amano, Y. (2001). Purification and properties of thiosulfate dehydrogenase from *Acidithiobacillus thiooxidans* JCM7814. *Biosci Biotechnol Biochem* **65**, 102-108.

Nelson, M. (1988). Catecholate Complexes of Ferric Soybean Lipoxygenase 1. *Biochemistry* **27**, 4273-4278.

Nozaki, Y. (1972). The preparation of guanidine hydrochloride. *Methods Enzymol* **26**, 43-50.

Olson, G. J., Brierley, J. A. & Brierley, C. L. (2003). Bioleaching review part B: progress in bioleaching: applications of microbial processes by the minerals industries. *Appl Microbiol Biotechnol* **63**, 249-257.

Otwinowski, Z. & Minor, M. (1997). Processing of X-ray Diffraction Data Collected in Oscillation Mode. In *Macromolecular Crystallography, part A*, pp. 307-236. Edited by C. W. J. Carter & R. M. Sweet: Academic Press.

Pachmair, F. (1960). Vorkommen und Bestimmung von Schwefelverbindungen in Mineralwasser. Munich: Justus-Maximilian Universität.

Palenchar, P. M., Buck, C. J., Cheng, H., Larson, T. J. & Mueller, E. G. (2000). Evidence that ThiI, an enzyme shared between thiamin and 4-thiouridine biosynthesis, may be a sulfurtransferase that proceeds through a persulfide intermediate. *J Biol Chem* **275**, 8283-8286.

Perrakis, A., Morris, R. & Lamzin, V. S. (1999). Automated protein model building combined with iterative structure refinement. *Nat Struct Biol* **6**, 458-463.

Ploegman, J. H., Drent, G., Kalk, K. H., Hol, W. G., Heinrikson, R. L., Keim, P., Weng, L. & Russell, J. (1978). The covalent and tertiary structure of bovine liver rhodanese. *Nature* **273**, 124-129.

Price, J. C., Barr, E. W., Tirupati, B., Bollinger, J. M., Jr. & Krebs, C. (2003). The first direct characterization of a high-valent iron intermediate in the reaction of an alpha-ketoglutarate-dependent dioxygenase: a high-spin FeIV complex in taurine/alpha-ketoglutarate dioxygenase (TauD) from Escherichia coli. *Biochemistry* **42**, 7497-7508.

Pronk, J. T., Meulenberg, R., Hazeu, W., Bos, P. & Kuenen, J. G. (1990). Oxidation of reduced inorganic sulfur compounds by acidophilic thiobacilli. *FEMS Microbiol Rev* **75**, 293-306.

Pushkaev, A. V., Koo, L. & Ortiz de Montellano, P. R. (2002). Aromatic stacking as a determinant of thermal stability of CYP119 from *Sulfolobus solfataricus*. *Arch Biochem Biophys* **409**, 52-58.

Que, L., Jr. & Ho, R. Y. (1996). Dioxygen Activation by Enzymes with Mononuclear Non-Heme Iron Active Sites. *Chem Rev* **96**, 2607-2624.

Que, L., Jr. (2004). The oxo/peroxo debate: a nonheme iron perspective. *J Biol Inorg Chem* **9**, 684-690.

Quentmeier, A. & Friedrich, C. G. (2001). The cysteine residue of the SoxY protein as the active site of protein-bound sulfur oxidation of Paracoccus pantotrophus GB17. *FEBS Lett* **503**, 168-172.

Ratnayake, D. B., Wai, S. N., Shi, Y., Amako, K., Nakayama, H. & Nakayama, K. (2000). Ferritin from the obligate anaerobe *Porphyromonas gingivalis*: purification, gene cloning and mutant studies. *Microbiology* **146 (Pt 5)**, 1119-1127.

Rohwerder, T., Gehrke, T., Kinzler, K. & Sand, W. (2003). Bioleaching review part A: progress in bioleaching: fundamentals and mechanisms of bacterial metal sulfide oxidation. *Appl Microbiol Biotechnol* **63**, 239-248.

Rohwerder, T. & Sand, W. (2003). The sulfane sulfur of persulfides is the actual substrate of the sulfur-oxidizing enzymes from *Acidithiobacillus* and *Acidiphilium* spp. *Microbiology* **149**, 1699-1710.

Rother, D., Henrich, H. J., Quentmeier, A., Bardischewsky, F. & Friedrich, C. G. (2001). Novel genes of the sox gene cluster, mutagenesis of the flavoprotein SoxF, and evidence for a general sulfur-oxidizing system in *Paracoccus pantotrophus* GB17. *J Bacteriol* **183**, 4499-4508.

Roy, A. B. & Trudinger, P. A. (1970). The biochemistry of inorganic compounds of sulphur.

Roy, S., Gupta, S., Das, S., Sekar, K., Chatterji, D. & Vijayan, M. (2003). Crystallization and preliminary X-ray diffraction analysis of *Mycobacterium smegmatis* Dps. *Acta Crystallogr D Biol Crystallogr* **59**, 2254-2256.

Royer, C. A. (1995). In *Protein stability and Folding: theory and practice*, pp. 65-89. Edited by B. A. Shirley. New Jersey: Humana Press Inc.

Sali, A. & Blundell, T. L. (1993). Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol* **234**, 779-815.

Sali, A. & Chiu, W. (2005). Macromolecular assemblies highlighted. *Structure (Camb)* **13**, 339-341.

Sambrock, J., Frisch, E. F.; Maniatis, T. (1989). *Molecular Cloning. A laboratory manual.* Cold Spring Harbour: Cold Spring Harbour Laboratory Press.

Sauvé, V., Bruno, S., Hemmings, A. & Berks, B. (2004). Structure of a sulphur carrier protein involved in the oxidation of inorganic sulphur compounds. In *EUROBIC 7 - 7th European Biological Inorganic Chemistry Conference*, pp. P-236. Edited by B. Lippert. Garmisch-Partenkirchen, Germany.

Saxton, W. O. (1996). Semper: Distortion Compensation, Selective Averaging, 3-D Reconstruction, and Transfer Function Correction in a Highly Programmable System. *J Struct Biol* **116**, 230-236.

Schagger, H. & von Jagow, G. (1987). Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal Biochem* **166**, 368-379.

Schägger, H., Aquila, H. & Von Jagow, G. (1988). Coomassie blue-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for direct visualization of polypeptides during electrophoresis. *Anal Biochem* **173**, 201-205.

Schägger, H. & von Jagow, G. (1991). Blue native electrophoresis for isolation of membrane protein complexes in enzymatically active form. *Anal Biochem* **199**, 223-231.

Schmidt, M., Rutkat, K., Rachel, R., Pfeifer, G., Jaenicke, R., Viitanen, P., Lorimer, G. & Buchner, J. (1994). Symmetric complexes of GroE chaperonins as part of the functional cycle. *Science* **265**, 656-659.

Scholz, C. (1999). Schwefel-Oxygenasen/-Reduktasen in extrem thermophilen Prokaryonten. In *Institute of Microbiology and Genetics*. Darmstadt: Darmstadt University of Technology.

Schurig, H., Rutkat, K., Rachel, R. & Jaenicke, R. (1995). Octameric enolase from the hyperthermophilic bacterium *Thermotoga maritima*: purification, characterization, and image processing. *Protein Sci* **4**, 228-236.

Sciara, G., Kendrew, S. G., Miele, A. E., Marsh, N. G., Federici, L., Malatesta, F., Schimperna, G., Savino, C. & Vallone, B. (2003). The structure of ActVA-Orf6, a novel type of monooxygenase involved in actinorhodin biosynthesis. *Embo J* **22**, 205-215.

Sheldrick, G. M. & Schneider, T. R. (1997). *SHELXL: High Resolution Refinement*. Orlando, Florida: Academic Press.

Shindyalov, I. N. & Bourne, P. E. (1998). Protein structure alignment by incremental combinatorial extension (CE) of the optimal path. *Protein Eng* **11**, 739-747.

Shirley, B. A. (1995). In *Protein stability and Folding: theory and practice*, pp. 177-190. Edited by B. A. Shirley. New Jersey: Humana Press Inc.

Silver, M. & Lundgren, D. G. (1968). Sulfur-oxidizing enzyme of *Ferrobacillus ferrooxidans* (*Thiobacillus ferrooxidans*). *Can J Biochem* **46**, 457-461.

Singer, G. & Hickey, D. A. (2003). Thermophilic prokaryotes have characteristic patterns of codon usage, amino acid composition and nucleotide content. *Gene* **317**, 39-47.

Skerra, A. (1994). Use of the tetracycline promoter for the tightly regulated production of a murine antibody fragment in *Escherichia coli*. *Gene* **151**, 131-135.

Smith, P. K., Krohn, R. I., Hermanson, G. T. & other authors (1985). Measurement of protein using bicinchoninic acid. *Anal Biochem* **150**, 76-85.

Spallarossa, A., Forlani, F., Carpen, A., Armirotti, A., Pagani, S., Bolognesi, M. & Bordo, D. (2004). The "rhodanese" fold and catalytic mechanism of 3-mercaptopyruvate sulfurtransferases: crystal structure of SseA from *Escherichia coli*. *J Mol Biol* **335**, 583-593.

Sreerama, N. & Woody, R. W. (2000). Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. *Anal Biochem* **287**, 252-260.

Stetter, K. O. (1996). Hyperthermophilic prokaryotes. *FEMS Microbiol Rev* **18**, 149-158.

Steudel, R. (2000). The chemical sulfur cycle. In *Environmental Technologies to treat Sulfur Pollutions*. Edited by P. Lens & L. Hulshoff Pol. London: IWA Publishing.

Sun, C. W., Chen, Z. W., He, Z. G., Zhou, P. J. & Liu, S. J. (2003). Purification and properties of the sulfur oxygenase/reductase from the acidothermophilic archaeon, *Acidianus* strain S5. *Extremophiles* **7**, 131-134.

Suzuki, I. (1965a). Oxidation of elemental sulfur by an enzyme system of *Thiobacillus thiooxidans*. *Biochim Biophys Acta* **104**, 359-371.

Suzuki, I. (1965b). Incorporation of atmospheric oxygen-18 into thiosulfate by the sulfur-oxidizing enzyme of *Thiobacillus thiooxidans*. *Biochim Biophys Acta* **110**, 97-101.

Suzuki, I. & Silver, M. (1966). The initial product and properties of the sulfur-oxidizing enzyme of thiobacilli. *Biochim Biophys Acta* **122**, 22-33.

Suzuki, I. (1999). Oxidation of inorganic sulfur compounds: Chemical and enzymatic reactions. *Can J Microbiol* **45**, 97-105.

Tano, T. & Imai, K. (1968). Physiological studies on thiobacilli. Part II: The metabolism of colloidal sulfur by the cell-free enzyme system of *Thiobacillus thiooxidans*. *Agri Biol Chem* **32**, 51-54.

Teixeira, M., Batista, R., Campos, A. P., Gomes, C., Mendes, J., Pacheco, I., Anemuller, S. & Hagen, W. R. (1995). A seven-iron ferredoxin from the thermoacidophilic archaeon *Desulfurolobus ambivalens*. *Eur J Biochem* **227**, 322-327.

Terwilliger, T. C. & Berendzen, J. (1999). Automated MAD and MIR structure solution. *Acta Crystallogr D Biol Crystallogr* **55 (Pt 4)**, 849-861.

Terwilliger, T. C. (2003). SOLVE and RESOLVE: automated structure solution and density modification. *Methods Enzymol* **374**, 22-37.

Thauer, R. K., Jungermann, K. & Decker, K. (1977). Energy conservation in chemotrophic anaerobic bacteria. *Bacteriol Rev* **41**, 100-180.

Theil, E. C. (2001). Ferritin. In *Handbook of Metalloproteins*, pp. 771-781. Edited by A. Messerschmidt, R. Huber, T. Poulos & K. Wieghardt: Wiley, UK.

Thurl, S., Witke, W., Buhrow, I. & Schafer, W. (1986). Quinones from archaebacteria, II. Different types of quinones from sulphur-dependent archaebacteria. *Biol Chem Hoppe Seyler* **367**, 191-197.

Tonello, F., Dundon, W. G., Satin, B., Molinari, M., Tognon, G., Grandi, G., Del Giudice, G., Rappuoli, R. & Montecucco, C. (1999). The *Helicobacter pylori* neutrophil-activating protein is an iron-binding protein with dodecameric structure. *Mol Microbiol* **34**, 238-246.

Urich, T., Bandeiras, T. M., Leal, S. S. & other authors (2004). The sulphur oxygenase reductase from *Acidianus ambivalens* is a multimeric protein containing a low-potential mononuclear non-haem iron centre. *Biochem J* **381**, 137-146.

Urich, T., Coelho, R., Kletzin, A. & Frazao, C. (2005a). The sulfur oxygenase reductase from *Acidianus ambivalens* is an icosatetramer as shown by crystallization and Patterson analysis. *Biochim Biophys Acta* **1747**, 267-270.

Urich, T., Kroke, A., Bauer, C., Seyfarth, K., Reuff, M. & Kletzin, A. (2005b). Identification of the core active site of the sulfur oxygenase reductase from *Acidianus ambivalens. FEMS letters* **in review**.

Vance, C. K. & Miller, A. F. (2001). Novel insights into the basis for *Escherichia coli* superoxide dismutase's metal ion specificity from Mn-substituted FeSOD and its very high E(m). *Biochemistry* **40**, 13079-13087.

Wackett, L., P., (2002). Mechanism and applications of Rieske non-heme iron dioxygenases. *Enzyme Microb Technol* **31**, 577-587.

Wang, G., Guo, R., Bartlam, M., Yang, H., Xue, H., Liu, Y., Huang, L. & Rao, Z. (2003). Crystal structure of a DNA binding protein from the hyperthermophilic euryarchaeon *Methanococcus jannaschii*. *Protein Sci* **12**, 2815-2822.

Woese, C. R. (2000). Interpreting the universal phylogenetic tree. *Proc Natl Acad Sci U S A* **97**, 8392-8396.

Yuvaniyama, P., Agar, J. N., Cash, V. L., Johnson, M. K. & Dean, D. R. (2000). NifS-directed assembly of a transient [2Fe-2S] cluster within the NifU protein. *Proc Natl Acad Sci U S A* **97**, 599-604.

Zillig, W., Yeats, S., Holz, I., Bock, A., Gropp, F., Rettenberger, M. & Lutz, S. (1985). Plasmid-related anaerobic autotrophy of the novel archaebacterium *Sulfolobus ambivalens*. *Nature* **313**, 789-791.

Zillig, W., Yeats, S., Holz, I., Böck, A., Rettenberger, M., Gropp, F. and Simon, G. (1986). *Desulfurolobus ambivalens* gen. nov., sp. nov., an autotrophic archaebacterium facultatively oxidizing and reducing sulfur. *System Appl Microbiol* **8**, 197-203.

Zimmermann, P. (1998). Biochemische und molekularbiologische Untersuchungen an der Schwefel Oxygenase/Reduktase aus *Acidianus ambivalens*. Darmstadt: Institute of Microbiology and Genetics.

Zimmermann, P., Laska, S. & Kletzin, A. (1999). Two modes of sulfite oxidation in the extremely thermophilic and acidophilic archaeon *Acidianus ambivalens*. *Arch Microbiol* **172**, 76-82.