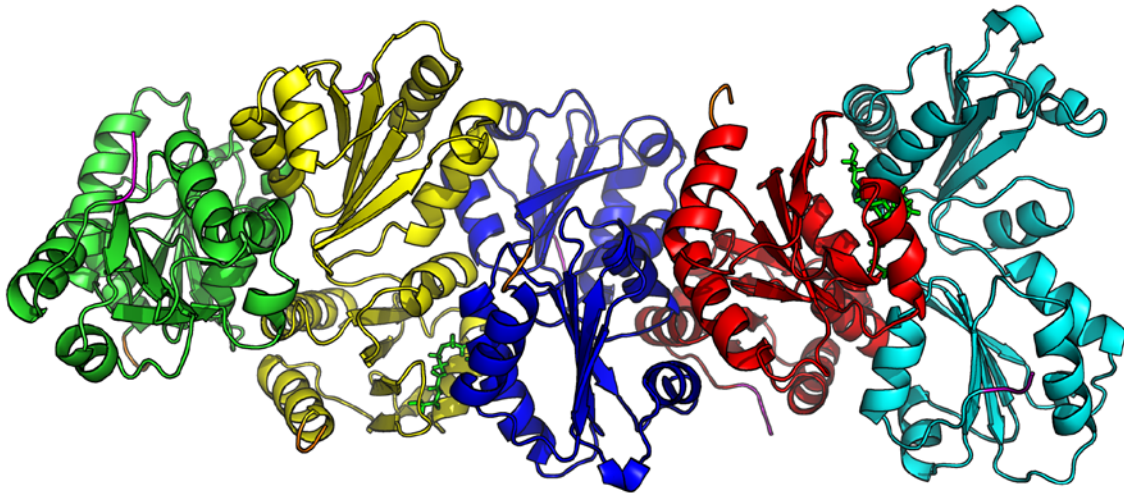




Molecular Insights into the Protein Disulfide Isomerase Family



Molekulare Mechanismen der Protein Disulfid Isomerase Familie

Doctoral Thesis for a doctoral degree
at the Graduate School of Life Sciences,
Julius-Maximilians-Universität Würzburg
Section: Biomedicine

Submitted by:
Franz-Xaver Wilhelm Kober
born in Bad Kissingen

Submission: Würzburg 2012

*Scientists often confuse people
with a fascination of the boring
with an interest for the inconspicuous
and with a passion for immaterial.*

Submitted on / eingereicht am:

.....

Members of PhD student's committee/ Promotionskomitee:

Chairperson:

Primary Supervisor: Prof. Dr. Hermann Schindelin (Würzburg)

Supervisor (second): Prof. Dr. Franz-Xaver Schmid (Bayreuth)

Supervisor (third): Prof. Dr. Alexander Buchberger (Würzburg)

Supervisor (fourth): Prof. Dr. James C. Bardwell (Ann Arbor, MI)

Date of Public Defense / Verteidigung am:

Date of Receipt of Certificates / Zeugnis erhalten am:

A Acknowledgements

This thesis – and the project it describes – represents over four years of my life and is the finale to about ten years spent at the university learning, working and researching. Up to this point is also the most comprehensive work I have completed and there are many people that have enabled and supported me – professionally or personally – for which I am truly grateful.

Hermann Schindelin my supervisor who made all this possible. He approached me with the interesting PDI project when I was indecisive and gently nudged me towards my PhD. There were difficult times as I was definitely not an easy student, but he always remained patient and supported me with his time and work. No matter what, he always set an example for scientific conduct and accuracy.

James Bardwell enabled me to spend more than a year in the United States and take my own first steps into the world of research. He supported me a lot during Diploma Thesis and I could always discuss science with him. He also introduced me to **Jonathan Pan**, who was a PhD student himself back when I did my Diploma thesis. Jon showed me the ropes concerning the day to day life in the lab. He also became an awesome friend and introduced me to the American way of life.

Franz-Xaver Schmid, Gregory Harms and Alexander Buchberger agreed become a part of my supervisory committee. Each of them gave me their time, listening to my hypothesis and problems. They always made sound suggestions and discussion with them greatly helped me to develop new ideas and approach my project in a better way.

Wolfgang Kölmel is the man with the crystallizing hands. He was an awesome intern and he definitely deserves his spot as a PhD student. Without him, the structure of ERp27 would not have been solved as it is the product of his internship.

Hans Maric and Carolyn Delto are the yin and yang of the office life. Both helped me immensely with their innate talents and their friendship. Be it that I needed criticism when I was overeager, sympathy after a failure, heated discussions to fortify ideas or encouragement when I was in doubt.

Uwe Dietzel and Wilko Rauert both were great companions and made the office, the lab and the life of a PhD student more enjoyable.

Antje Schaefer, Jochen Kuper, Petra Hänzelmann und Ingrid Tessmer – together they form the post-doc oracle that omnisciently answers virtually every question concerning research methods. Each of them in their own way not only patiently taught me the practical aspect of research but also helped me to become a better scientist.

Sara Leonhardt always managed to motivate me, cheer me up when I was frustrated and always provided a reason to go home. Her support was invaluable – not only on a personal level. Discussing research and science in general over dinner, being a sound board for each other's ideas improved me as a scientist. She simply was – and still is – the best friend, love and partner I could have ever wished for.

Ulrike Kober, my mum, is a very unique person. She was the one always pushing me to overcome my weak points. While it helped me immensely and I am – in hindsight – grateful for it, I have not always shown appreciation. So I want to say “Thank you! For Everything.”

Max, my brother, has to be mentioned for his valiant effort distracting me with awesome and fun times spent together.

Franz Kober, my father, was the one who fed my curiosity in you years. Our drive to school was oftentimes more informative than the day of lessons that followed. I hope that you remember those times and derive strength from it.

Bjoern, Christian & Eva and Claudi & Felix and the rest of my friends have to be mentioned. For I know that I'm not an easy person to be around and I'm grateful for people that think of me and that I can consider friends.

B Summary - English

Upon synthesis, nascent polypeptide chains are subject to major rearrangements of their side chains to obtain an energetically more favorable conformation in a process called folding. About one third of all cellular proteins pass through the secretory pathway and undergo oxidative folding in the endoplasmic reticulum (ER). During oxidative folding, the conformational rearrangements are accompanied by the formation of disulfide bonds – covalent bonds between cysteine side chains that form upon oxidation. Protein disulfide isomerase (PDI) assists in the folding of substrates by catalyzing the oxidation of pairs of cysteine residues and the isomerization of disulfide bonds as well as by acting as chaperones. In addition to PDI itself, a family of related ER-resident proteins has formed. All PDI family members share the thioredoxin fold in at least one of their domains and exhibit a subset of the PDI activities. Despite many studies, the role of most PDI family members remains unclear.

The project presented in this thesis was aimed to establish tools for the biochemical characterization of single members of the PDI family and their role in the folding process. A combination of fluorescence based assays was developed to selectively study single functions of PDI family members and relate their properties of either catalysis of oxidation or catalysis of isomerization or chaperone activity to the rest of the protein family. A binding assay using isothermal titration calorimetry (ITC) was established to complement the activity assays. Using ITC we could show for the first time that members of the PDI family can distinguish between folded and unfolded proteins selectively binding the latter. The unique information provided by this method also revealed a two-site binding of unfolded proteins by PDI itself.

In addition to the functional characterization, experiments were conducted to further investigate the oligomeric state of PDI. We could show that the equilibrium between structurally different states of PDI is heavily influenced by the redox state of the protein and its environment. This new data could help to further our understanding of the interplay between oxidases like PDI and their regenerative enzymes like Ero1, which may be governed by structural changes in response to the change in redox status. Another structural approach was the screening of all investigated PDI family members for suitable crystallization conditions. As a result of this screening we could obtain protein crystals of human ERp27 and were able to solve the structure of this protein

with X-ray crystallography. The structure gives insight into the mechanisms of substrate binding domains within the PDI family and helps to understand the interaction of ERp27 with the redox active ERp57. In collaboration with the group of Heike Hermanns we could further show the physiological importance of this interaction under oxidative stress.

In conclusion, the project presented in this thesis provides novel tools for an extensive analysis of the activities of single PDI family members as well as a useful set of methods to characterize novel oxidoreductases and chaperones. The initial results obtained with the our novel methods are very promising. At the same time, the structural approach of this project could successfully solve the structure of a PDI family member and give information about the interplay within the PDI family.

C Zusammenfassung - Deutsch

Neu gebildete Polypeptidketten arrangieren schon direkt nach ihrer Synthese die Seitenketten ihrer Aminosäuren um, damit sie einen energetisch günstigeren Zustand erlangen. Diesen Umlagerungsprozess nennt man Proteinfaltung. Schätzungsweise ein Drittel aller zellulären Proteine werden über den sekretorischen Transportweg geschleust und durchlaufen die oxidative Proteinfaltung im endoplasmatischen Retikulum (ER). Während der oxidativen Faltung werden zusätzlich zur Umlagerung von Seitenketten auch Disulfidbrücken gebildet. Dies sind kovalente Bindungen zwischen Zystein-Seitenketten durch Oxidation entstehen. Protein Disulfid Isomerase (PDI) unterstützt die Faltung von Proteinen im ER indem es die Oxidation zweier Zystein-Seitenketten katalysiert. Neben der Oxidation katalysiert PDI ebenfalls die Isomerisierung von fehlverknüpften Disulfidbrücken und wirkt als Chaperon der Aggregation entgegen. Im Laufe der Evolution hat sich zusätzlich zu PDI eine Gruppe verwandter ER-lokalisierter Proteine gebildet. Diese Mitglieder der PDI-Familie weisen alle das Thioredoxin-Faltungsmotiv in mindestens einer ihrer Domänen auf und besitzen mindestens eine der drei PDI-Aktivitäten. Trotz eingehender Untersuchung ist die Rolle der meisten dieser PDI Familienmitglieder weiterhin unklar.

Im Rahmen des Projekts, welches dieser Dissertation zugrunde liegt, wurden Methoden zur biochemischen Charakterisierung einzelner Mitglieder der PDI-Familie, und deren Rolle im Faltungsprozess, entwickelt. Eine Kombination von Fluoreszenzexperimenten wurde etabliert mit der selektiv einzelne Aktivitäten von Faltungshelfern analysiert und qualitativ in die PDI-Familie eingeordnet werden können. Diese fluoreszenzbasierten Methoden wurden durch isothermale Titrationkalorimetrie (ITC) ergänzt. Mit ITC konnten wir als Erste zeigen, dass PDI-Familienmitglieder gefaltete von ungefalteten Proteinen unterscheiden können und letztere selektiv binden. Die zusätzlichen Informationen, die in einem ITC-Experiment gewonnen wurden, zeigten, dass PDI mit Substraten mit Hilfe von zwei unterschiedlichen Bindungsstellen interagiert.

Neben der funktionellen Analyse der PDI-Familie wurde Experimente durchgeführt um den oligomeren Zustand von PDI näher zu untersuchen. Wir konnten zeigen, dass das Gleichgewicht zwischen strukturell verschiedenen Zuständen entscheidend vom Redox-Status von PDI abhängt.

Diese neuen Daten werfen ein neues Licht auf die Interaktion zwischen Oxidasen wie PDI und ihren regenerativen Enzymen wie Ero1. Diese Interaktion könnte sehr wohl durch strukturelle Veränderungen, ausgelöst durch Redox-Reaktionen, reguliert werden. Als weiteren strukturellen Ansatz zur Erforschung der PDI-Familie wurden alle verwendeten Familienmitglieder auf aussichtsreiche Kristallisationsbedingungen hin untersucht. Durch dieses Screening konnte ERp27 kristallisiert und seine Struktur durch Röntgenkristallografie aufgeklärt werden. Die so gewonnene Struktur gibt Aufschluss über die Mechanismen der Substratbindung in der PDI-Familie und hilft ebenfalls dabei, die Interaktion zwischen ERp27 und dem redoxaktiven ERp57 besser zu verstehen. Auf Grund dieser Daten konnten wir gemeinsam mit der Gruppe von Heike Hermanns die physiologische Bedeutung dieser Interaktion bei oxidativem Stress aufzeigen.

Zusammenfassend konnten im Rahmen dieses Projektes neue Werkzeuge zur eingehenden Analyse der PDI-Familie etabliert werden, welche auch zur Charakterisierung neuer Oxidoreduktasen und Chaperone verwendet werden können. Die ersten Ergebnisse die mit Hilfe dieser neuen Methoden gewonnen werden konnten sind vielversprechen. Gleichzeitig konnten wir mit ERp27 die Struktur eines weiteren PDI-Familienmitgliedes lösen und so weitere Einblicke in das komplexe Netzwerk der PDI-Familie gewinnen.

D Table of contents

A	Acknowledgements.....	- 2 -
B	Summary - English	- 4 -
C	Zusammenfassung - Deutsch	- 6 -
D	Table of contents.....	- 8 -
1.	Introduction.....	- 10 -
1.1	On proteins, structures and energy landscapes	- 10 -
2.	Materials & Methods	- 28 -
2.1	Molecular Biology	- 28 -
2.1.1	Separation of DNA molecules by size.....	- 28 -
2.1.2	Extraction of DNA from agarose gels	- 28 -
2.1.3	Purification of plasmid DNA.....	- 29 -
2.1.4	Digestion with restriction endonucleases.....	- 29 -
2.1.5	Polymerase chain reaction and site directed mutagenesis	- 30 -
2.1.6	Sequencing of DNA.....	- 31 -
2.1.7	Cultivation and storage of <i>E. coli</i> strains.....	- 31 -
2.1.8	Transformation of <i>E. coli</i>	- 32 -
2.1.9	Large scale lysis of <i>E. coli</i> cells.....	- 33 -
2.2	Bioinformatics.....	- 35 -
2.2.1	UniProtKB Database.....	- 35 -
2.2.2	ExPasy Tools and ProtParam.....	- 35 -
2.2.3	ClustalW	- 35 -
2.2.4	GeneRunner	- 36 -
2.2.5	Pymol	- 36 -
2.2.6	NCBI database and tools.....	- 36 -
2.2.7	Mosflm / XDS.....	- 37 -
2.2.8	CCP4 program suite.....	- 37 -
2.2.9	Phenix program suite	- 37 -
2.2.10	HydroPro.....	- 38 -
2.3	Protein biochemistry	- 39 -
2.3.1	Determination of protein concentration	- 39 -
2.3.2	Separation of proteins by SDS-polyacrylamide gel electrophoresis.....	- 39 -
2.3.3	Coomassie staining of Polyacrylamide gels	- 40 -

2.3.4. Protein Purification Protocols	- 41 -
2.3.5 Reduced carboxymethylation of model substrates	- 46 -
2.3.6 Preparation of apo-riboflavin binding protein	- 46 -
2.3.7 Preparation of di-eosine-gluthatione (Di-E-GSSG).....	- 47 -
2.3.8 Fluorescence self-quenching (FSQ) assay	- 47 -
2.3.9 Riboflavin quenching assay	- 48 -
2.3.10 Isothermal titration calorimetry (ITC)	- 49 -
2.3.11 Circular dichroism spectroscopy (CD spectroscopy)	- 50 -
2.3.12 Dynamic Light Scattering (DLS).....	- 51 -
2.3.13 Crystallization of proteins.....	- 52 -
2.3.13 Diffraction experiments	- 53 -
2.3.14 Data Processing.....	- 54 -
2.3.15 Structure solution	- 55 -
2.4 Materials	- 58 -
2.4.1 List of primers	- 58 -
2.4.2 List of expression vectors	- 61 -
2.4.3 Equipment	- 62 -
2.4.4 Consumables	- 63 -
3.4.5 Chemicals and Solvents	- 63 -
3. Results & Discussion	- 65 -
3.1 The multistate-equilibrium of yeast PDI.....	- 65 -
3.2 Crystallization of PDI family members	- 79 -
3.3 Fluorescence-based assays for activity measurements of PDI family proteins	- 102 -
3.4 Substrate recognition of the PDI family	- 122 -
4. Synopsis & Outlook.....	- 140 -
5. References.....	- 143 -
E List of abbreviations.....	- 151 -
F List of Figures	- 153 -
G List of Publications and Congress contributions	- 155 -
H Curriculum Vitae	- 156 -
Affidavit.....	- 158 -

1. Introduction

1.1 On proteins, structures and energy landscapes

Proteins are the key to life. An organism's genetic information for the most part either specifies the amino acid sequence of proteins or regulates the spatial or temporal distribution of protein production. Proteins are extremely versatile and are utilized for functions as different as structural support, movement, recognition, coordination of cellular events and exchange of information between cells. With very few exceptions, every biochemical reaction within a cell is catalyzed by a protein. No matter the task in an organism, one will find proteins involved in it. In the face of this multitude and variety of tasks, it comes as a surprise that proteins themselves are assembled from only twenty different building blocks: the proteogenic amino acids.

By linking together amino acids in different combinations, proteins with diverse biochemical properties and functions are generated. Proteins binding transition states of reactants for example catalyze the corresponding reactions by lowering the activation energy. Proteins are assembled at the ribosome as polypeptide chains. These strings of amino acids are defined by the sequence of its residues, also known as the primary structure. *In vivo*, polypeptide chains do not keep this thread-like form. Instead, at a first level of higher order structure neighboring amino acids interact to form three-dimensional patterns, also known as the secondary structure. These patterned groups in turn interact with each other forming a complex three-dimensional tertiary structure. In many cases protein molecules become subunits in a stable protein complex. Such an arrangement is called the quaternary structure. This process of rearranging the relative orientation of residues to each other from the primary structure into a stable tertiary structure is called protein folding. The tertiary structure of a protein is more important in determining a protein's function than the exact sequence of amino acids, as the tertiary structure determines the spatial arrangement of certain key residues. Independent studies [1-3] have found very similar structures, when investigating proteins which perform the same function (Fig. 1.1).

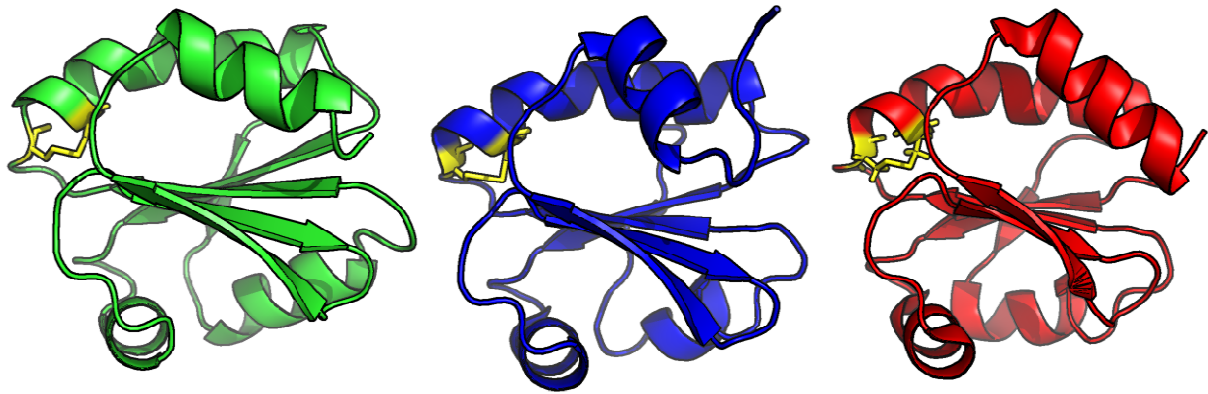


Figure 1.1 – Structure and function

Models of thioredoxins from *Homo sapiens* (green – 1ERU), *Mycobacterium tuberculosis* (blue – 2I1U) and *Chlamydomonas reinhardtii* (red – 1DBY) were generated from coordinate files retrieved from the protein data base. Each protein performs the task of reducing protein disulfide bonds and all three proteins adopt the thioredoxin consisting of a central β -sheet which is flanked by two helices on each side. In all three proteins the active site (highlighted in yellow stick model) is located at the same position at the beginning of a helix.

Despite originating from completely unrelated life forms, proteins performing an identical task usually share similar structures. On the other hand, two protein molecules might share the same amino acid sequence, but while one fulfils a vital role in the cell, the second may even be harmful to the organism due to a different fold. Multiple medical conditions are known to be caused by incorrect folding. Alzheimer's disease is probably the most prominent example of an illness caused by the misfolding and subsequent aggregation of two proteins, amyloid-beta and tau, in the brain. [4] The build-up of protein aggregation – visible as amyloid plaques - finally leads to the pruning of neuronal connections and eventually cell death. Another publicly well-known disease caused by protein misfolding is the Creutzfeldt-Jakob disease, which was covered extensively in the media in the 1990ies. Here prions – protein molecules with an alternative stable fold – act as chaperones convert their correctly folded forms into the prion fold [5]. Eventually, as the concentration of the misfolded prion form rises, the prions aggregate causing apoptosis. The difference in secondary and tertiary structure alone renders one a functional part of the cell and the other an infectious and pathogenic agent. These examples highlight the importance of protein folding to transform a newly translated polypeptide chain into a functional protein.

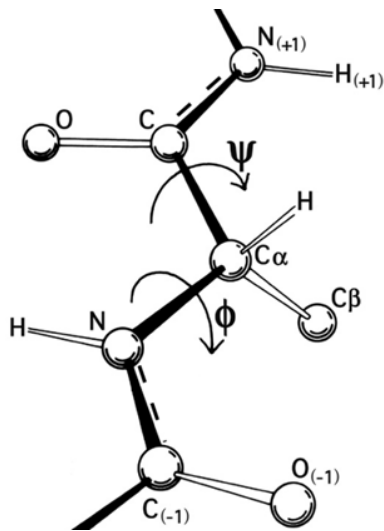


Figure 1.2 – Schematic of a peptide backbone

The bonds involving the C_{α} atom are free to rotate. Thus the orientation of the amino acid in relation to its neighbors is defined by the dihedral angles ψ and ϕ . Due to the partial double bond character (dashed line) the peptide bond is fixed and cannot rotate. While both bonds can rotate freely, not all combinations of angles are possible due to steric hindrance and unfavorable bond angles. The favored and disfavored combinations for amino acids are displayed in a Ramachandran plot.

(Modified from Wikimedia Commons:
Protein_backbone_PhiPsiOmega_drawing.jpg)

A protein is able to fold because two bonds per residue within the polypeptide backbone are flexible and can rotate (see Figure 1.2). This flexibility allows the peptide chain to adopt many different conformations. Yet, as shown by X-ray diffraction analysis of crystallized proteins, most proteins only adopt a limited number of folded, native conformations, and single domain proteins usually only have one single native conformation [6]. At a conference in 1969 Cyrus Levinthal [7, 8] presented a calculation for the case of a small protein with 100 amino acids. Even if each amino acid of the hypothetical model is limited to only three different conformations and the interconversion of structures takes place with the speed of the fastest known biochemical reaction ($10^{13}/s$, the reaction velocity for the hydration of carbon dioxide), the average time needed for folding would be 1.6×10^{27} years as there would be 3^{100} or 5×10^{47} possible structures. The problem of the folding time of a small protein with limited conformations exceeding the age of the universe was termed the *Levinthal paradox*. In contrast, Barnase (*Bacillus amylofaciensis* Rnase, 110 amino acids) has been shown experimentally to fold with a half time of 50 ms [9]. This discrepancy between theoretical calculation and experimental observation shows that protein folding cannot succeed by randomly trying out every possible conformation. Anfinsen and coworkers could show [10] using RNaseA as a model, that proteins which were unfolded using chaotrope agents such as urea will reacquire their native conformation when the denaturing agent is removed. This refolding occurs in the absence of cellular factors demonstrating that the primary structure contains all relevant information for folding. Despite more than fifty years of research since those initial experiments,

the prediction of a structure *ab-initio* from just its sequence of amino acids cannot be done with certainty.

Any transition between two states depends on the energy gradient between those states. As the conformational change of a protein within the cell happens at constant temperature and pressure, the Gibbs free energy is used to describe the thermodynamic potential of a conformation. Folding is driven by the fact that an unfolded protein has a higher free energy than the folded state, which is stabilized by – among others – hydrogen bonds and hydrophobic interactions [11]. The previously discussed Levinthal paradox is based on the assumption that only the final native conformation possesses a lower free energy and that all non-native states are equal in their thermodynamic potential. The paradox breaks down, when the different free energies are attributed to folding intermediates [12]. Today the predominant model for protein folding is the energy landscape model. As two amino acids are brought into close proximity by their shared

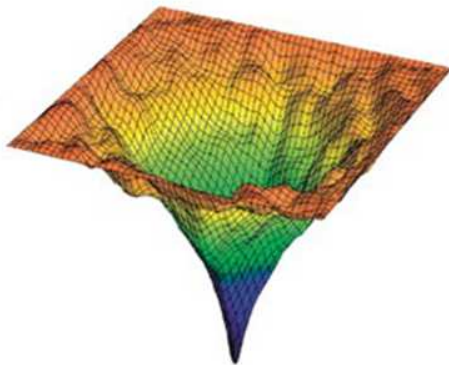


Figure 1.3 – The ‘energy funnel’

3D model of a hypothetical energy landscape of a protein undergoing protein folding. The Gibbs’ free energy is plotted on the vertical axis. Proteins with energetically unfavorable conformations (red regions in the catchment area) move towards more lower energy states through a series of intermediate conformations. The blue center point with the lowest free energy corresponds to the native conformation of the protein.

peptide bond, their bonds will rotate in a way that optimizes the interaction between them. Thus, there are many concurrent micro-folding processes as each amino acid accommodates to its micro environment, thus lowering the global free energy of the macro molecule, albeit at the level of $k_B T$ [13]. In parallel there is also a similar process to minimize the free energy between non-neighboring residues. Between those two often competing processes, some states are generated with an even lower free energy, from which, through rearrangement of larger segments of the protein, even more energetically favorable states emerge. When plotting all possible conformational states against their free energy, the energy landscape for a given sequence is generated. As this would require an n -dimensional plot (with n being the number of amino acids) and it is currently beyond our possibilities to accurately determine the free energy of fleeting

folding intermediates, the landscape is usually visualized in the form of an energy funnel with the native conformation at its center (see Figure 1.3 - Taken from [14])

Such a funnel directs all possible conformations within its catchment area towards the central conformation with the lowest free energy within a short time. Experimental work with random and designed amino acid sequences show that such a smooth funnel with a single minimum is the exception rather than the rule [15, 16]. This stems from the fact that polypeptides are an inherently frustrated system, meaning that it is not possible for each functional group to attain its optimal environment. As an example, the hydrogen-bonded interactions of the peptide backbone favor the continuation of an α -helix, as the end of a helix contains unsatisfied hydrogen bonds. Yet a helix must be broken, as the hydrophobic interactions of apolar side chains require that the polypeptide chain folds back onto itself. This thermodynamic frustration often causes sequences to contain special amino acids to initiate or terminate helices, where side chains either provide additional hydrogen bonds or, in the case of proline at the start of helices, do not feature a nitrogen atom with hydrogen bond potential to reduce the frustration [17]. Spatial constraints are an additional source of frustration. Often the ideal thermodynamic arrangement is impossible due to the fact that polypeptides are a chain and single residues cannot move independently. This geographic frustration results in local minima that are similar in their free energy and structure, but are divided by a high energy barrier [12] and are thus not interconvertible. This leads to the formation of folding pathways, where few well-folded and energetically favorable intermediates compete for less folded intermediates. Depending on the topography of such a rough energy landscape, local minima can either lead on towards the folded state, or can represent a kinetic trap (as it has been mapped out for bovine pancreatic trypsin inhibitor, BPTI [18]). Statistical models such as those proposed by Bryngelson and Wolynes [19] postulate, that most local energy minima are way points on the path towards a native conformation as the folding molecule minimizes frustration.

Besides the roughness of an energy landscape, its 'depth' – the difference in free energy (ΔG) between the unfolded and the fully folded state – is also a very important characteristic of a protein. The amount of this free energy difference is directly related to the stability of a protein [20]. Stable proteins remain in their native conformation over a wide range of temperature, pH or

ionic strength. Naturally proteins have to balance competing needs. While they have to be stable enough to attain and retain their functional conformation at physiological conditions, very stable enzymes would be unable to undergo functionally relevant conformational changes needed e.g. for substrate binding. Also the function of a protein itself usually reduces stability, as binding interfaces for interaction partners or substrates as well as catalytic residues need to be present at a higher free energy level to thermodynamically favor the substrate-bound state [21]. With all those constraints on protein stability it is not surprising that enzymes are usually very well adapted to the physiological conditions inside the cell. This also highlights the importance of cellular homeostasis, as many proteins depend on a stable constitution of the solvent in order to remain in their native state.

In conclusion, the shape and structure are the defining features of a protein. Made up from a limited variety of amino acids, proteins fold along the gradient of free energy, gradually optimizing the environment of their functional groups. A defined fold is a necessity, but protein stability is usually not the primary feature, a given sequence is selected throughout the course of evolution. Therefore the energy landscape is rarely a smooth funnel towards a single minimum as shown in Figure 1.3. Experimental work on β -barrel folds [22, 23] revealed that even with sequences designed to fold smoothly, the energy landscape is rather rough and contains many local energy minima. This roughness results from geometric constraints and competing but mutually exclusive tendencies to saturate hydrogen bonds and avoid solvent contact by apolar groups. In such a rough energy landscape, unfolded polypeptides are forced into defined pathways usually leading towards the native state. Once folded, the energy barriers surrounding the native state keep a protein folded and functional. Thus, proteins rely on a gradient of free energy not only to fold properly, but also to remain folded. As the energy landscape is shaped not only by the sequence, but also by environmental factors such as temperature, pH and ionic strength among others, cellular homeostasis is very important to keep proteins functional.

1.2 Redox chemistry and its application to proteins

As previously discussed, proteins fold because their native state is favored energetically. The free energy difference between the unfolded and the folded state dictates the stability of a protein, which is correlated with the ability to retain its 3D structure under stress. This stress can be induced through a rise in temperature, a shift in pH or ionic strength, or a change in the solvent content, like the addition of ethanol or chaotropic agents. While proteins within the cell benefit from a stable and controlled environment maintained in the cytosol, there are many functional proteins on the outside of the cell wall or plasma membrane. These proteins need to keep a functional conformation under a wide variety of conditions that the cell can only influence to a limited degree. Therefore proteins working on the outside usually contain stabilizing features that raise energy barriers in order to lock a peptide chain in its functional state.

The energy gain through the optimal arrangement of residues is relatively small, usually on the order of $k_B T$ [24], which corresponds to 4.11×10^{-21} J per residue or 2.48 kJ/mol at standard conditions. On the other hand, the binding enthalpy for a covalent bond between atoms ranges between 146 kJ/mol and 945 kJ/mol [25], so with the introduction of covalent bonds between residues, a structure can be stabilized significantly. The most prominent bond utilized for protein stability is the disulfide bond. Cysteine residues carry a terminal free thiol group that is susceptible to oxidation under physiological conditions. Two cysteine residues with their thiol groups in proximity can transfer electrons to a suitable acceptor and form a disulfide bond – a covalent linkage between their side chains. Once formed, it requires either extremely high energies or a suitable redox reaction to break such a bond. Thus, disulfide bonds raise energy barriers to trap a protein in its native conformation. Due to their nature as a redox sensitive bond, they are reversible, they will only form once the protein is in an oxidative environment and they can be removed through reduction with a suitable electron donor.

The redox potential within the cell is regulated to keep the proteins in the proper redox state. Eukaryotic cells generally maintain a reducing environment in the cytoplasm with regard to their proteins. To keep the cysteines of a protein reduced, a buffer system of the oxidized and reduced forms of the tri-peptide glutathione is used. Upon spontaneous oxidation of protein cysteines, the resulting disulfide bond is reduced by the enzyme thioredoxin with electrons supplied by reduced glutathione which itself becomes oxidized and forms a disulfide bond. The dimeric glutathione is

then reduced by the enzyme glutathione reductase which uses the reductant NADH generated by the catabolic branch of the metabolism to regenerate the reduced glutathione. Therefore proteins localized in the cytosol, the nucleus or other reducing environments have evolved to be functional with their cysteines in the reduced state.

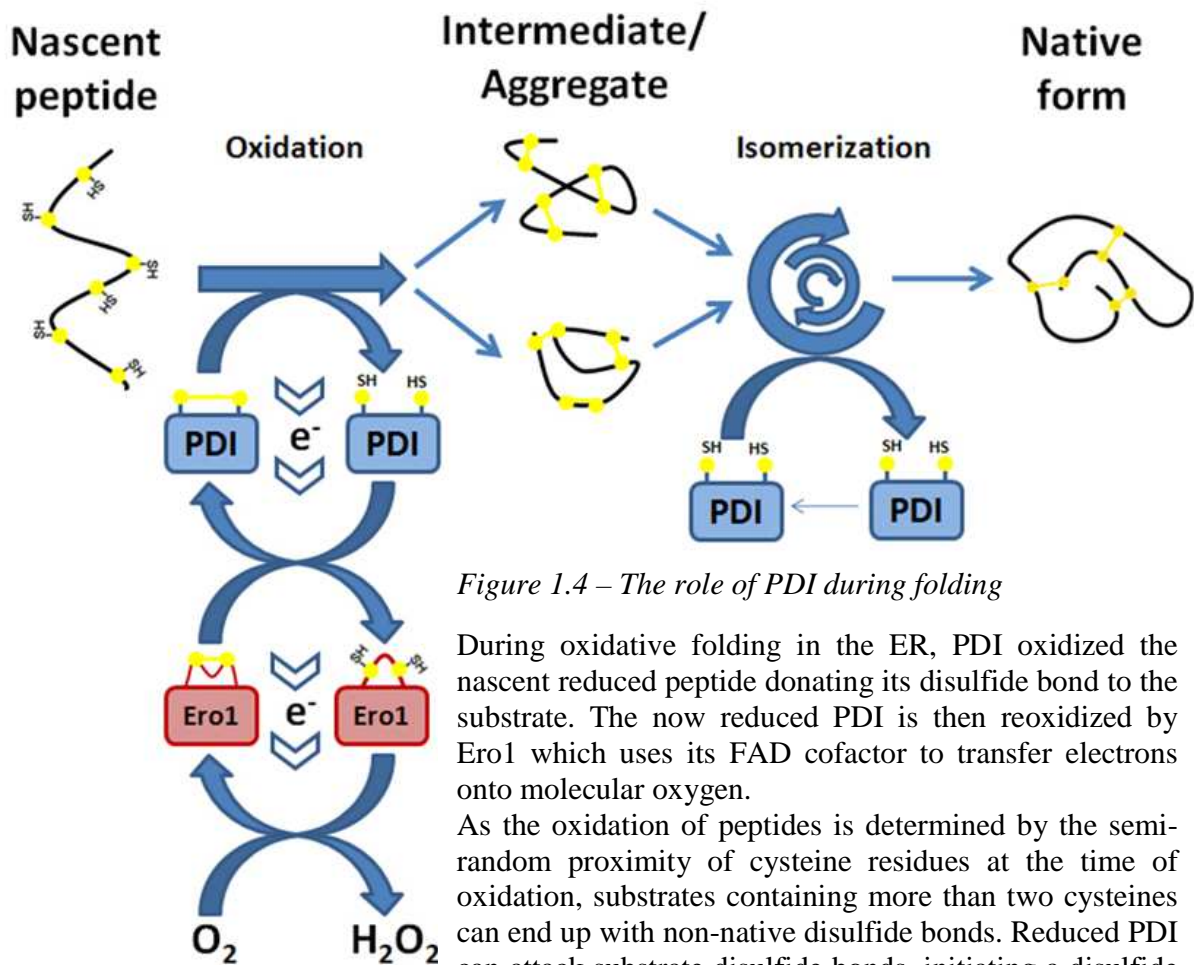


Figure 1.4 – The role of PDI during folding

During oxidative folding in the ER, PDI oxidized the nascent reduced peptide donating its disulfide bond to the substrate. The now reduced PDI is then reoxidized by Ero1 which uses its FAD cofactor to transfer electrons onto molecular oxygen.

As the oxidation of peptides is determined by the semi-random proximity of cysteine residues at the time of oxidation, substrates containing more than two cysteines can end up with non-native disulfide bonds. Reduced PDI can attack substrate disulfide bonds, initiating a disulfide rearrangement in the substrate.

In contrast to the cell interior, the environment of cells is usually oxidizing with regard to proteins due to the presence of molecular oxygen. This means that extracellular proteins or the extracellular portions of membrane spanning proteins usually feature disulfide bonds in their functional conformation. In order to have properly disulfide bonded proteins ready for the export to the cell surface, there must be cellular compartments where maturing proteins can be oxidized and processed. In prokaryotes the periplasm – the space between the inner and the outer membrane – contains such an oxidizing environment. In eukaryotes, the endoplasmic reticulum

(ER) is the most important site for oxidative folding. Proteins destined for an oxidative environment, be it the cell surface, a mitochondrion or the ER itself, are translated by ribosomes directly bound to the ER membrane (rough ER) that separates the cytoplasm from the ER lumen [26]. In the ER, nascent proteins are often glycosylated cotranslationally and then oxidized by protein disulfide isomerase (PDI). PDI transfers an active site disulfide bond to a folding substrate and is in turn reoxidized by the actions of Ero1 [27], which generates disulfide bonds *de novo*, utilizing a FAD cofactor to directly transfer electrons onto molecular oxygen resulting in the generation of hydrogen peroxide (Fig. 1.4). In general, the oxidation of folding proteins and *de novo* disulfide bond generation are separated. PDI oxidizes substrates in both the plant [28] and animal kingdom and is subsequently reoxidized by ER oxidoreductases such as Ero1p in yeast [27], AERO1 in plants [29] or Ero1-L α in mammals [30] that funnel the electrons to final acceptors such as molecular oxygen. In bacteria proteins are oxidized by DsbA[31], which in turn gets reoxidized by DsbB [32]. The hydrogen peroxide is further reduced to water by peroxiredoxins[33], which also can generate disulfide bonds and transfers them onto PDI [34, 35].

The redox potential is the chemical property that decides whether a compound placed in an environment with other redox active compounds is present in its oxidized or reduced form. For a given redox couple such as oxidized and reduced glutathione or cysteine and cystine (the disulfide linked form of two cysteines), the redox potential is determined experimentally by measuring the electrochemical potential towards the standard reaction $2\text{H}^+ + 2\text{e}^- \rightarrow \text{H}_2$ at a pH of 7, which is assigned the redox potential of 0 V. The experimentally measured redox potential for glutathione, if present in equal concentrations in its reduced and oxidized form, is -240 mV [36]. In general, electrons will flow from lower to higher

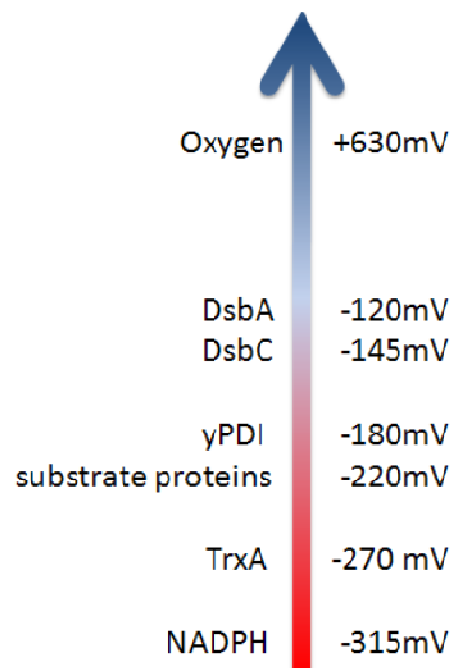


Figure 1.5 - Standard redox potentials of selected cellular redox active proteins and compounds.

Electrons are transferred along the gradient with molecular oxygen as terminal electron acceptor in aerobic organisms.

electron potentials. So, if for example the cellular reductant NADPH with a potential of -315 mV is added to glutathione, electrons will be transferred from NADPH to glutathione, oxidizing the former and reducing the latter. The larger the difference between the redox potentials the more complete will each partner become oxidized and reduced, respectively.

In the cell, redox catalysts create a micro-environment by the arrangement of appropriate side chains in the vicinity of the active site, to modify the redox potential of the catalytic cysteines. Reductases such as TrxA are able to reduce their redox potential to values of -270 mV [37] while on the other hand oxidases such as DsbA increase their redox potential to -120 mV [38] in order to carry out their respective reaction more efficiently. PDI – being an isomerase in addition to its oxidase functionality – keeps its two active sites in a moderately oxidizing range at -175/190 mV [39] in order to be able to partly reduce target proteins at all during the initial step of the isomerization reaction.

This work will focus on the redox biochemistry in the endoplasmic reticulum. Approximately one third of all proteins are destined for an oxidative environment and are synthesized by ER membrane associated ribosomes, which directly release the nascent polypeptide chain into the lumen of the ER [40]. In the ER a chain of redox partners funnels the electrons of the reduced nascent cysteines onto molecular oxygen, forming disulfide bonds in the protein undergoing folding during this process. The formation of disulfide bonds is a kinetic challenge, as their formation is orders of magnitude slower than the rapid folding processes [31, 41-43]. Yet the disulfide bonds due to their covalent nature provide a significant increase in protein stability enabling proteins to retain their functional conformation over wider ranges of pH, ionic strength, solvent composition and redox environment. Thus, a network of enzymes has evolved to accelerate the oxidation of cysteines and to provide assistance in forming the correct disulfide linkages.

1.3 Assistants in folding: The protein disulfide isomerase family

In the lumen of the endoplasmic reticulum – a network of membranes attached to the nucleus – nascent proteins undergo folding in an oxidative environment. Compared to the formation of

secondary and tertiary structures which develop within milliseconds [41], the formation of covalent disulfide bonds *in vitro* is considerably slower with rates on the timescale of minutes or even hours [42, 43]. In comparison, the reported rate for the *in vivo* oxidation of proteins is generally faster, occurring within seconds [31]. The reason for the discrepancy between *in vitro* and *in vivo* oxidation rates is the presence of folding catalysts within the cell that assist in the oxidation and folding of maturing proteins. Those folding catalysts exhibit three major functions in order to accelerate oxidative refolding. For one, disulfide bonds need to be inserted into folding substrates. Oxidases – such as the *Escherichia coli* enzyme DsbA [31]– donate their active site disulfide bond to

Number of Cysteines	Possible combinations
2	1
4	3
6	15
8	105
10	945
14	135135
20	654729075

Table 1.1: Increased number of cysteines exponentially increases the complexity of oxidative folding.

substrates. The now-reduced oxidases get in turn oxidized by specialized regenerating proteins – such as DsbB [32] – that are capable of generating disulfide bonds *de novo*, passing the electrons over intermediates onto molecular oxygen. With the combination of oxidases and an oxidizing environment in the periplasm (prokaryotes) or the endoplasmic reticulum (eukaryotes), respectively, proteins undergoing folding are oxidized upon translocation quickly and efficiently. As oxidases usually also catalyze the reverse reaction – the reduction of disulfide bonds given a reducing environment – they are also often called oxidoreductases. While rapid oxidation is all that is needed for small proteins to fold within a sensible timeframe, larger proteins are facing an additional challenge. As the number of possible disulfide bond combinations increases exponentially with the numbers of cysteines in a substrate (see Table 1.1), and many substrates follow distinct and complex oxidation patterns [44-46], proteins may form non-native disulfide bonds during folding. Therefore folding catalysts have to provide a second essential function. They attack non-native disulfides, creating a short-lived intermolecular disulfide bond. This results in a reduced cysteine in the folding protein, which can in turn attack other disulfide bonds in the vicinity starting an isomerization of bonds. This intramolecular rearrangement continues until the intermolecular disulfide is attacked, either by a cysteine of the rearranging substrate, or by the second active site cysteine of the isomerase. In order to be available for either oxidation or isomerization, misfolded proteins have to be present in a soluble form. Yet, most un- or

misfolded proteins tend to aggregate and precipitate out of solution, minimizing the surface area available for interactions with oxidases or isomerases. To overcome this, many folding catalysts also exhibit a chaperone function. This means they selectively bind non-natively folded proteins, sequestering them away from other misfolded proteins in to avoid aggregation. Periodically chaperones release their bound substrate in order to make it gradually available for processing by other folding catalysts.

The very first protein found to accelerate the refolding of a denatured substrate was an enzyme later identified as PDI [47]. In refolding assays, the activity of the analyzed ribonuclease recovered significantly faster when PDI was added to the denatured protein undergoing refolding. PDI increased the rate of protein oxidation in a glutathione buffer, proving itself to be an oxidase. PDI also increased the rate of reactivation of scrambled ribonuclease, which is fully oxidized, but not functional as its cysteines are connected in a non-native way. In order to regain their functionality, such scrambled proteins need to isomerize their disulfide bonds, rearranging the way their cysteines are connected. Thus PDI was shown to also act as an isomerase. In later studies, PDI was also shown to act as a chaperone [48, 49], binding model substrates in a transient manner. Whether this is a byproduct of its substrate-binding capabilities or an independent function recently became a topic of discussion as anti-chaperone activity has been tied to PDI-protein interactions as well [48, 50]. Exhibiting three important functionalities PDI has become one of the most studied folding catalysts. In the course of this thesis, the focus will be on PDI's role during oxidative folding. However, PDI also moonlights in other activities. The most prominent is its involvement in the assembly of the enzyme complexes prolyl-4-hydroxylase [51] and microsomal triacylglycerol transfer protein (MTTP) [52]. In those cases PDI is an integral structural component, but is not involved in catalysis and does not require its active site. Several studies have reported that PDI is also located outside of the ER (for a review see [53]). Up to date, it is still unclear, how PDI is able to escape the ER retention mechanism and what role it plays outside of the ER. Yet it has been implied into medically relevant processes [54, 55] in those unusual locations and is currently under intense investigation.

1.3.1 The architecture of protein disulfide isomerase (PDIA1)

Human PDI is a 57 kDa protein with a main chain comprised of 507 amino acids. Secondary structure predictions and protein characterization revealed early that the protein consists of four domains, each adopting the thioredoxin fold. The thioredoxin fold – named after the reductase thioredoxin A (TrxA) – is a common structural feature in many oxidoreductases. The core of the thioredoxin fold is characterized by a secondary structure sequence of β_1 - α_1 - β_2 - α_2 - β_3 - β_4 - α_3 (see Fig. 1.6 and [56]) forming a central four stranded mixed β -sheet flanked by three α -helices. Often, the fold is expanded N-terminally by an additional β -strand and α -helix each to β_0 - α_0 - β_1 - α_1 - β_2 - α_2 - β_3 - β_4 - α_3 [57]. The active site consists of a CXXC motif which is located at the beginning of the first α -helix (α_1). This allows the active site cysteines to form a disulfide as their thiols are placed in close spatial proximity. The two residues between the active site cysteines have been shown to greatly influence the redox potential [58]. Also the neighboring residues of the active site are highly conserved and thus form an extended active site motif of WCXXCK [59] most likely being involved in forming a suitable microenvironment for catalysis, for substrate binding or both. Further away in the sequence but in structural proximity is the so called cis-loop proline. This proline [60] and especially the N-terminally adjacent residue [61] have been shown to be important for the catalytic activity of the active site.

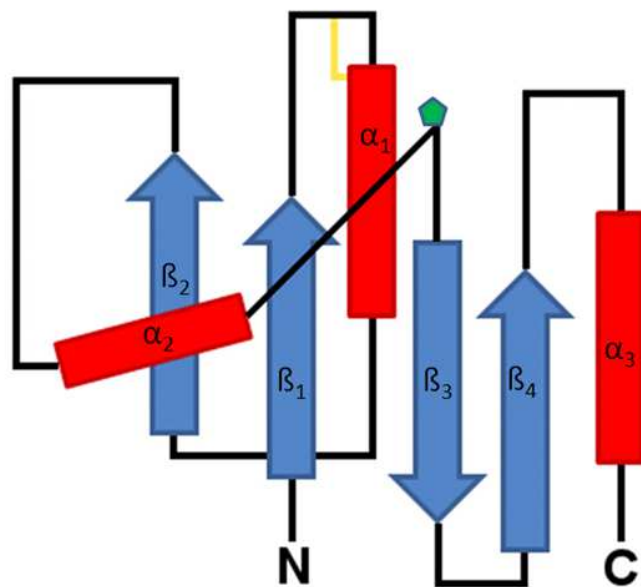


Figure 1.6 – Schematic of the thioredoxin fold

The canonical thioredoxin fold consists of four beta strands (blue arrows) forming a mixed beta sheet which is flanked by three alpha helices (red bars). The active site (marked in yellow) is located at the N-terminus of the first alpha helix. The green pentagon denotes the location of the cis-proline, an important residue modifying the properties of the active site.

PDI, like many other related oxidoreductases, consists only of thioredoxin fold domains, though not all have an active site. Throughout the PDI family, there are three types of thioredoxin fold domains. The *a* domains contain a redox sensitive active site with active site cysteines. These domains are usually directly involved in the redox reaction performed by the enzyme. Domains devoid of an active site are referred to as *b* domains. These domains may or may not contain a rudimentary extended active site motif, but they never contain active site cysteines. Among the *b* domains, there is a subgroup called *b'* domains. These domains share a structural feature of exposed hydrophobic residues or – if the structure is unknown – align with

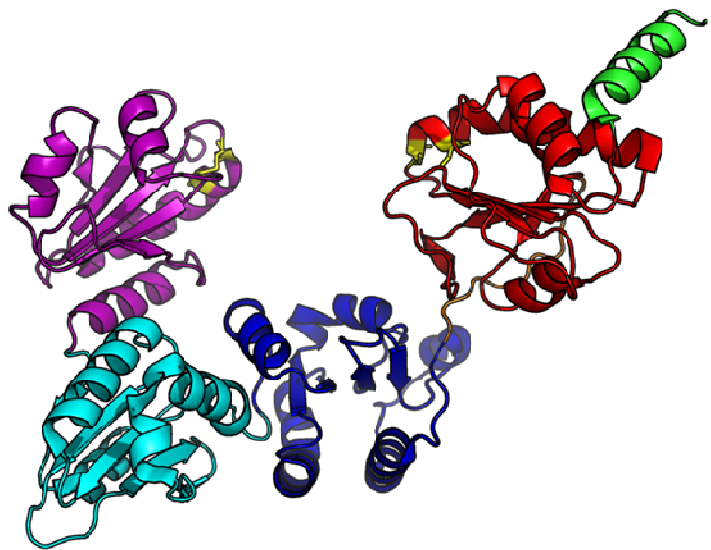


Figure 1.7 – Crystal structure of yeast PDI

Generated from structure 2B5E deposited at the RCSB. The four thioredoxin fold domains *a* (magenta), *b* (cyan), *b'* (blue) and *a'* (red) are arranged in the shape of a 'U' with the *b* domains forming the bottom and the *a* domains the sides. The active sites of the *a* domains are marked in yellow. The orange loop marks the *x*-linker, which has been suggested to regulate substrate binding. The highly charged and acidic *c* domain (green) contains the C-terminal ER retention signal (KDEL), which is located beyond the residues visualized in the structure.

known *b'* domains. Together, these residues form a cleft or groove of hydrophobic residues where unfolded substrate proteins have been shown to bind [62]. The thioredoxin domain architecture of PDI is *a*-*b*-*b'*-*a*. In addition to the thioredoxin domains, two more structural features could be detected in a secondary structure prediction of PDI. For one, the C terminus forms an additional helix – rich in acidic residues – not belonging to the thioredoxin motif. At the very end of this negatively charged helix, the canonical ER retention motif KDEL is located, but apart from ER retention, no functional role could be assigned up to this point to the C terminus. Despite being only a single helix, the C terminal extension is generally referred to as the *c* domain of PDI. The second structural feature is a very long linker between the *b'* domain and the C terminal *a* domain. This extended linker region is often added to the domain

architecture as an 'x' due to its larger size. So the general architecture of PDI is often described as a-b-b'-x-a'-c, with the a' domain being the a domain adjacent to the b' domain.

The first picture of a full length PDI molecule at atomic resolution came in 2006, when the X-ray crystal structure of yeast PDI (also referred to as PDI_A1) was solved [63]. In the crystal structure, the four domains are arranged in the shape of a twisted 'U' (see Fig 1.7). At the bottom of the 'U' the b and b' domains form a base. The interface between those two domains is extensive and therefore there is little flexibility between the b and b' domains. In contrast the two a domains are connected to their respective b domains through flexible linkers, allowing for larger movements of the active site domains in relation to the rigid base. In the originally published structure, both a-type domains are positioned to have their active sites facing each other. Yet, a second structure [64] (3BOA) showed the same molecule, but with its a' domain in a completely different orientation facing outwards. This alternate conformation highlights the flexibility of the a-type domains in their orientation in relation to the b-b' core.

The crystal structure of yeast PDI not only provided a high resolution image of the domain arrangement, but also gave insights into features and possible functions of the domain types. The

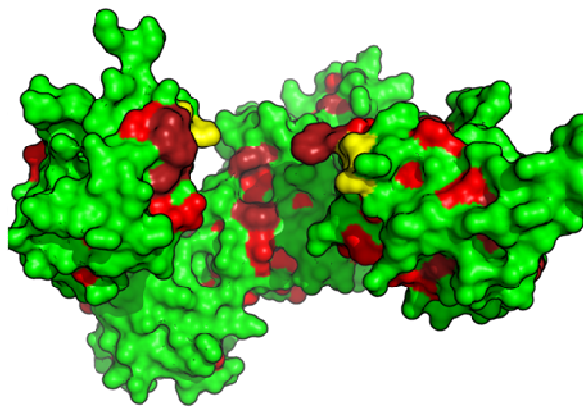


Figure 1.8 – Surface hydrophobicity of PDI

Surface representation of yeast PDI (green) with hydrophobic residues (Ala, Leu, Ile, Val, Pro, Met, Trp and Phe) colored in red. Active sites are marked in yellow. The hydrophobic cleft is clearly visible on the b' domain between the active sites.

b-b' domain core contains an unusual amount of solvent exposed hydrophobic residues. These residues are all facing towards the a domains, increasing the hydrophobicity on the base of the 'U'. In addition, the b' domain contains a cleft which opens towards the a-type domains and which is very hydrophobic in nature (see Fig. 1.8). This cleft is rather narrow not offering enough space for domains or even a single helix, but is suitable for the binding of unstructured peptide chains. Unfolded or misfolded proteins – the potential substrates of PDI and PDI family members – generally have a higher content of exposed hydrophobic surface, which usually leads to aggregation of

misfolded protein. Potential substrates also contain an above average amount of regions, where amino acids form neither α -helices nor β -sheets. These random coil or loop regions can fit into the hydrophobic cleft of the b' domain. With its capability to bind unstructured and hydrophobic regions or other proteins, the hydrophobic base formed by the b-b' domains appears to be an ideal candidate for substrate binding and recognition [65].

The peptide binding capability was confirmed by a nuclear magnetic resonance (NMR) structure of the isolated b' domain of human PDI and the associated 'x' linker region [66] that normally connects the b' domain with the C-terminal a domain. In this structure, the hydrophobic region of the flexible linker binds into the hydrophobic pocket of the b' domain, suggesting a regulatory function of the 'x' linker and further supporting the hypothesis that the b' domain – especially its hydrophobic cleft – represents the core of the substrate binding site. In the initial X-ray structure of full-length yeast PDI, another feature also pointed towards the substrate binding role of the b-b' domain core; there were crystal contacts between two yeast PDI molecules from different asymmetric units, placing the b domain of one molecule between the active site domains of the other. The resulting model shows nicely, how domain-sized substrates might position themselves via the hydrophobic surface of the b-b' core to interact with the active sites.

The crystal structure also advanced our understanding of the active site domains. They are positioned with the active sites facing each other over the substrate binding site, hence suggesting a cooperative action. The active site domains of yeast PDI are very homologous to other active site domains of this family of oxidoreductases, with a highly conserved consensus motif. It also becomes clear, how key residues of the active site domains determine the redox potential of the active site. The most important factor are the residues between the active site cysteines. As shown *in vitro* by *Wunderlich & Glockshuber* [59] and *in vivo* by *Quan et al.*[58], mutations in this position can radically alter the redox potential, turning a reductase like TrxA into an oxidase able to replace DsbA in the periplasm of *E. coli*. There are also other key residues, supposed to modulate the redox potential such as the side chains of the cis-proline loop. This loop, located between helix two and strand three, comes into close proximity of the active site and was shown to be important for the modulation of the enzymatic activity. The full-length structure of PDI revealed the presence of a buried pair of charged and polar residues. Located

within the hydrophobic interior of the domain, these residues approach the active site from the back and are hypothesized to stabilize the reduced state and to assist in the deprotonation of an active site cysteine. In this way, the buried charged polar pair can simultaneously lower the redox potential and increase the reactivity of the reduced state. Unfortunately, the influence of this residue pair could not yet be conclusively proven.

Despite all the information gained through the determination of the structure and a multitude of biochemical studies, many questions still remain regarding the function of PDI. The second crystal structure for example revealed large movements occurring within the molecule, which are necessary for the activity of the protein. Large scale conformational changes could also explain, how PDI might adapt to the enormous size range of potential substrates, and how the interaction with regenerating enzymes is regulated. Also PDI has evolved into a whole family in higher eukaryotes. While PDI1 is still regarded as the most important enzyme in the oxidative folding process in the ER, the role of most PDI family members still remains unknown.

1.3.2. The PDI protein family

PDI is ubiquitous throughout the eukaryotic realm. While PDI itself fulfills a critical part in the oxidative folding process and is lethal when knocked out, there is also a group of proteins with thioredoxin fold domains resident in the ER called the PDI protein family. In general there are more PDI like proteins expressed by an organism, the more complex the organism is. In humans there are six genes belonging to the PDI gene family (PDIA1 through PDIA6) encoding for the proteins PDI, PDIp, ERp57, ERp72, PDIr and P5. Their encoded proteins all have multiple thioredoxin fold domains, contain at least one b-type domain and at least one a-type domain with a CGHC as the active site motif. Besides this core group, ERp18, ERp27, ERp28, ERp44, ERp46 and ERdj5 are usually also included into the PDI family. The members of the latter group are all ER resident proteins containing at least one thioredoxin fold domain, though they might lack either a-type domains or b-type domains or carry completely different active site motifs. Additionally a group of transmembrane oxidoreductases (TMX) are also usually included, extending the family to more than 15 related in higher eukaryotes [67]. These proteins are located in the ER membrane and contain at least one thioredoxin fold domain with an active site.

In this large – and still growing – group of proteins, the roles for most individual family members still remain unknown. Apart from PDI itself, the proteins ERp57, ERp27 and P5 were in the focus of this project.

ERp57 is one of the most extensively studied PDI family members. It shares its domain composition (a-b-b'-a'-c) and more than 20% sequence identity with PDI as well as the overall 'twisted U' architecture as seen in its crystal structure [68]. ERp57 catalyzes disulfide bond formation of a glycoprotein subset of substrates and requires the lectins calnexin and calreticulin to facilitate substrate-binding.

Human P5 consists of two a-type and one b-type domain and is reported to have isomerase and chaperone activity, albeit with lower efficiency than PDI [69]. A second report also showed *in vitro* disulfide exchange activity [70], but its involvement in the ER folding process remains largely unexplored compared to the reports on its activities outside of the ER.

ERp27 is only comprised of two b-type domains and completely devoid of an active site. It can bind unfolded peptides like Δ somatostatin as well as ERp57 [71] and might act as a mediator for ERp57 to interact with substrates not recognized by calnexin or calreticulin.

Despite sharing structural features and belonging to the same protein family, these three proteins have been shown to behave completely different. The knowledge currently available about the roles and interactions of the PDI family members is very limited and to gain a comprehensive understanding of the folding machinery processing a wide variety of substrates in an efficient and timely manner, much remains to be done. The projects detailed within this thesis attempt to increase our knowledge about mechanistic aspects of the PDI family.

2. Materials & Methods

2.1 Molecular Biology

This section encompasses all methods used to:

- a) analyze and manipulate DNA
- b) transform *E. coli* cells with DNA and
- c) cultivate *E. coli* cells.

2.1.1 Separation of DNA molecules by size

In order to analyze fragments of deoxyribonucleic acid (DNA), they were separated by size via electrophoresis in an agarose gel. To prepare a gel, agarose was boiled in TAE running buffer. The amount of agarose varied between 0.8% (w/v) and 1.4% (w/v) according to the size of the

TAE buffer	
4.84 g	Tris
1.14 ml	Acetic Acid
2.00 ml	0.5M EDTA
add to 1L	dd H ₂ O
adjust to pH 8.5	

Contents of TAE buffer

DNA fragments to be analyzed. After cooling down the gel slowly to temperatures below 60°C without letting it solidify, ethidium bromide was added to a final concentration of 0.5 µg per ml. The liquid agarose solution was then poured into a gel scaffold, allowing it to solidify. One part of 6x concentrated DNA sample buffer was added to five parts of DNA solution and mixed before pipetting the sample into the wells of the agarose gel. The gel was immersed into TEA buffer and subjected to electrophoresis at 120 V for 45 min. To detect the fragment pattern of the separated DNA, the gel was subjected to UV light. The ethidium bromide within the gel intercalates into the DNA double helix and shows fluorescence if excited with light around 290 nm wavelength. For documentations, the fluorescent image was recorded using the Biorad GelDoc documentation system

2.1.2 Extraction of DNA from agarose gels

To retrieve fragments from an agarose gel after analysis (see 2.1.1) the QIAquick gel extraction kit was used. First the bands of interest were excised using a fine scalpel under UV light from a

UV illuminator on the lowest setting. After each excision, the blade was carefully cleaned to avoid cross contamination. The target DNA was extracted from the agarose gel fragments using the NucleoSpin extract kit by Macherey-Nagel. In this protocol the gel is first dissolved using a urea buffer. Next, the DNA is bound to an anion exchange column, washed with ethanol and eluted using either TE buffer (10 mM Tris pH8.0, 1 mM EDTA) or ddH₂O. The exact procedure is detailed in the kit protocol.

2.1.3 Purification of plasmid DNA

Plasmid DNA was extracted and purified from *E. coli* cell cultures using the Macherey-Nagel NucleoBond miniprep kit. Cells from a 5 mL overnight culture of *E. coli* were harvested by centrifugation (Eppendorf centrifuge 5417R, 25,000 x g, 3 minutes, 4 °C). The medium was discarded and the cells resuspended in 250 µL S1 resuspension buffer. Cells were subsequently subjected to alkaline lysis by the addition of 250 µL S2 lysis buffer containing sodium dodecyl sulfate and sodium hydroxide followed by a 5 minute incubation at room temperature. As soon as the solution cleared, 350 µL S3 neutralization buffer was added to move the pH to the neutral range and precipitate the SDS together with cell debris and membranes. In order to separate the precipitate from the solution, the samples were centrifuged (Eppendorf centrifuge 5417R, 25,000 x g, 10 minutes, 4 °C) and the supernatant removed carefully to avoid contamination with precipitate. The supernatant was loaded on an anion exchange column provided with the kit, washed with ethanol and eluted using either TE buffer (10 mM Tris pH 8.0, 1 mM EDTA) or ddH₂O. The exact procedure followed is detailed in the kit protocol.

2.1.4 Digestion with restriction endonucleases

Restriction endonucleases (also called restriction enzymes) are proteins that cleave the sugar-phosphate-backbone of double stranded DNA molecules. These breaks are introduced only at target sites that are highly specific to each restriction enzyme. Due to their sequence specificity,

they are used to cut DNA-molecules in a predicted pattern, producing fragments that can be recombined later.

A typical reaction with restriction enzymes contained:

3.0 μ l	10x reaction buffer*
1.0 μ l	restriction endonuclease (10u/ μ l)
x μ l	target DNA – 1 μ g of DNA in total
Add to 30 μ l	autoclaved distilled water

* - *The buffer was specified and supplied by New England Biolabs together with the enzyme.*

The reaction mix was incubated at the appropriate temperature as specified by the supplier for 1 h. The reaction was stopped either by heat inactivation (if applicable), or by separation of the reagents and the DNA by agarose gel electrophoresis (see 2.1.1).

2.1.5 Polymerase chain reaction and site directed mutagenesis

The polymerase chain reaction (PCR) is a powerful tool to amplify DNA sequences with a known start and stop site. It is based on an *in-vitro* replication, where DNA-polymerase is provided with a template, a primer and activated nucleotides (dNTP). To duplicate a target sequence, it is first heated up to melt double stranded DNA, the temperature is then reduced, to enable the primer to bind. As soon as a primed single strand is available, the polymerase starts to duplicate the sequence following the primer, yielding two double stranded DNA molecules – one from each strand of the original. By iteration, the region between the two primers are nearly doubled on each cycled. Due to the exponential growth, even a very small amount of template DNA is sufficient to generate large quantities of the piece that is located between the two primers.

As the primers get incorporated into the product, this method can also be used to introduce changes into a target DNA sequence. Primers used for site directed mutagenesis are designed to match with the sequence upstream and downstream of the mutation, but contain a mismatch at the site of mutagenesis. When the PCR is carried out, DNA molecules containing the sequence of the primer will vastly outnumber the ones containing the original sequence.

A typical PCR reaction contained:

x µl	template DNA – total of 200ng
1.0 µl	each Primer (concentration = 50µmol)
1.0 µl	dNTP (dATP, dCTP, dGTP, TTP – 10mM)
10.0 µl	5x Pfu reaction buffer
1.0 µl	Pfu polymerase
add to 50µl	autoclaved and distilled water

The reaction was then subjected to the following PCR protocol using an Eppendorf Mastercycler EPgradient S PCR cycler:

Cycle name	Temp.	Duration	Number of cycles
Initial denaturation	95°C	5 min	1
Denaturation	95°C	30 sec	25
Primer annealing	55°C	45 sec	
Elongation	65°C	1 min/1kb*	
Final elongation	65°C	7 min	1

* - Elongation time depends on the length of the expected product.

2.1.6 Sequencing of DNA

Sequences of DNA were determined by the company SeqLab (Goettingen, Germany) from samples of plasmid DNA using dideoxy sequencing based on the method originally published by Sanger et al. [72].

2.1.7 Cultivation and storage of *E. coli* strains

For general usage *E. coli* was grown in LB-medium (20 g/L LB powder in H₂O) under permanent selection with one or more of the following antibiotics:

Antibiotic	Final concentration
Ampicilin	200 µg/ml
Tetracycline	200 µg/ml

Chloramphenicol	34 µg/ml
Kanamycin	200 µg/ml

A standard overnight culture contained 5 ml LB media supplemented with the appropriate antibiotic in a 15 mL test tube. It was inoculated by touching a pre-existing culture or plate with a sterile micropipette tip and inserting this tip briefly into the fresh culture medium.

For protein expression, pre-cultures were grown in 200 mL of LB medium supplemented with the appropriate antibiotics in a 500 mL shaker flask. These cultures were inoculated from single colonies of freshly transformed *E. coli* cells. Main cultures of 2.5 L LB with the appropriate antibiotics in 5 L shaker flasks were inoculated by transferring 20 mL of pre-culture grown to an optical density at 600nm (OD₆₀₀) of 1.5.

For long-term storage of *E. coli* strains, 500 µl of a grown overnight culture were rapidly mixed with 1 ml of 20% (v/v) autoclaved glycerol and stored at -80°C.

2.1.8 Transformation of *E. coli*

Preparation of competent cells

The desired strain was streaked out and grown on an LB agar plate supplemented with antibiotics according to the strain resistances. 200 ml of LB medium were inoculated from a single colony of this LB plate and incubated shaking at 200 rpm and 37 °C. During the incubation, the optical density (OD₆₀₀) was monitored. Cells were grown into the early log phase (OD₆₀₀ = 0.3-0.5) and then harvested in sterile 50 mL falcon tubes by centrifugation (Eppendorf Centrifuge, 4000 x g, 10 minutes, 4 °C). The supernatant was discarded and the pellet was kept on ice. All subsequent steps were carried out at 4 °C or on ice. The pellet was resuspended in 15 ml of ice cold sterile filtered TFB1 buffer (contents see table below) and incubated on ice for 90 minutes. The cells were then collected by another centrifugation step (Eppendorf Centrifuge, 2000 x g, 20 minutes, 4 °C). After discarding the supernatant, the pellet was resuspended in 2 mL of ice cold sterile filtered TFB2 buffer (contents see table below) and aliquoted into 50 µl aliquots, which were flash frozen in liquid nitrogen and stored at -80 °C.

TFB1 buffer	TFB2 buffer
30 mM potassium acetate pH 5.0	10 mM MOPS pH 6.8
100 mM RbCl	10 mM RbCl
50 mM MnCl ₂	75 mM CaCl ₂
10 mM CaCl ₂	15% glycerol
15% glycerol	

Transformation of competent cells

Plasmids were inserted into *E. coli* through heat shock transformation. For the transformation a 50-100 µL aliquot of competent cells was removed from storage at -80°C and thawed on ice. 50 µg of DNA were added to the thawed cells and the mixture was incubated on ice for 20 minutes. The cells were then transferred from the ice to a 42 °C heating block for one minute and then put back on ice. In order to increase the rate of transformation, 600 µL of LB medium devoid of antibiotics and for an incubation period for 20 minutes at 37°C could be added as an optional step. The transformed cells were spread on an LB agar plate containing appropriate antibiotics and incubated overnight at 37 °C.

2.1.9 Large scale lysis of *E. coli* cells

For protein purification purposes cultivated cells have to be harvested and lysed. Large scale lysis was performed with either an e615 cell disruptor (Constant Cell Disruption Systems LTD) or a M-110P microfluidizer (Microfluidics). Both systems subject cells to an extreme pressure through application of Bernoulli's principle. In both cases a cell suspension is pumped through a very small nozzle with high pressure. As the liquid passes the nozzle it is accelerated due to an decrease in diameter. The increase in velocity corresponds to an increase in kinetic energy which in turn has to result in a decrease of potential energy through a decrease in pressure in order to be in accordance with the principle of conservation of energy. Bernoulli's principle relates the increase of velocity with a decrease in pressure. The very sudden change in pressure causes the cells to rupture and lyse. The procedure for each system was similar, with the exact handling as described in the manual. First the system was rinsed with the same buffer that was used to resuspend the cells. Then the cell suspension was passed twice through the system that was set to

a pump pressure of 1.5 kbar. After the second passage, the system was cleaned by rinsing with water and 20% ethanol (which was used to collapse and remove foam that formed during lysis), followed by rinsing with a solution of 0.5 M sodium hydroxide in water. In the end the system was rinsed multiple times with water to remove traces of sodium hydroxide to avoid corrosion.

2.2 Bioinformatics

This section encompasses commonly used scientific programs, data bases and web bases tools. Office programs for writing, spread sheet calculus and the design of graphics and presentations are interchangeable and therefore not listed.

2.2.1 UniProtKB Database

This data base located at <http://www.uniprot.org/> was used to gain information about amino acid sequences, post translational modifications, alternative names and DNA sequences (via links to the NCBI data base). PDI family members were searched using their common name (e.g. ERp57) and the organism *Homo sapiens*. The accession number from each entry was stored to use for other web tools such as the Expsy web tools.

2.2.2 Expsy Tools and ProtParam

Using the accession number gathered from UniProt, a wide array of tools could be applied. The most comprehensive collection of those programs are the ExpsyTools located at <http://expasy.org/> which use the amino acid sequence for prediction of protein properties. The program most commonly used within this thesis is ProtParam (<http://web.expasy.org/protparam/>). This tool was used to calculate the molar absorbtion coefficients, the theoretical PI and the molecular weight from the amino acid sequence. These parameters were then used to determine the concentration of each protein in solution.

2.2.3 ClustalW

ClustalW is an algorithm to align DNA and amino acid sequences and is employed by many web based alignment services and offline alignment tools. The most commonly used application of

ClustalW was the web service of genome.jp (<http://www.genome.jp/tools/clustalw/>) for alignment of amino acid sequences.

2.2.4 GeneRunner

GeneRunner is a freeware program for Windows based computers to edit, store and manage DNA sequences. It was used to plan the cloning of new expression vectors and to design primers from DNA sequences. Its DNA alignment feature was used to compare sequencing results to the source sequences in order to spot mutations. It was also used for *in silicio* restriction digests of sequences.

2.2.5 Pymol

Pymol is a powerful tool to visualize 3D-models generated from .coordinate files such as those deposited in the proteins database (PDB - www.rcsb.org). In these files the coordinates of all atoms are denoted, determined through either NMR or X-ray crystallography. Pymol was used under an academic license to create figures used for presentations, publications and this thesis.

2.2.6 NCBI database and tools

The National Center for Biotechnology Information (NCBI) houses the largest publicly available data bases on DNA sequences, ribonucleic acid (RNA) sequences and amino acid sequences. Its service PubMed contains an extensive literature database for the field of life science and related topics. In addition to data base searches, the NCBI hosts web servers for many alignment and database tools such as the various BLAST tools. All these services can be accessed free of charge (<http://www.ncbi.nlm.nih.gov/>).

2.2.7 Mosflm / XDS

Mosflm (<http://www.mrc-lmb.cam.ac.uk/harry/mosflm/>) and XDS (<http://homes.mpimf-heidelberg.mpg.de/~kabsch/xds/>) are programs used to process the image files obtained during an X-ray diffraction experiment. The experimental data is searched for patterns coinciding with the predicted patterns from known crystallographic space groups and each spot visible in the X-ray diffraction data is indexed and its intensity is integrated. The processed data can then be utilized to solve the structure using for example the CCP4 suite.

2.2.8 CCP4 program suite

CCP4 is a compilation of a wide range of computational tools for structure determination of macromolecules by X-ray crystallography. It encompasses many scripts for the conversion of file types and the calculation of parameters used for structure determination. The programs most prominently used were Scala for the scaling of the intensity values, MolRep to solve the phase problem by molecular replacement, Phaser to solve the phase problem by single anomalous wavelength diffraction (SAD) and Buccaneer for automated model building. More information about this program suite can be obtained under <http://www.ccp4.ac.uk/>

2.2.9 Phenix program suite

Phenix is a compilation of crystallography software similar to CCP4. Phenix offers programs to solve structures by molecular replacement as well as single (or multiple) wavelength anomalous dispersion (SAD or MAD respectively). The most prominently used tool from the Phenix suite is phenix.refine for the model refinement of ERp27. More information about this program suite can be obtained under <http://www.phenix-online.org/>

2.2.10 HydroPro

This freeware program was used to calculate the radii of gyration from a structure in the PDB format. Calculations were done based on an atomic level shell model (calculation mode 1) with 6 passes and a radius sigma ranging from 1 Å to 2 Å. HydroPro is available at:

<http://leonardo.inf.um.es/macromol/programs/hydropro/hydropro.htm>

2.3 Protein biochemistry

This section encompasses the expression and purification of proteins as well as their use and handling in assays and crystallization.

2.3.1 Determination of protein concentration

The amount of purified protein in solution was measured by UV-spectroscopy using the PeqLab Nanodrop ND-1000 spectrophotometer. To calculate the protein concentration of a solution, an absorbance spectrum of the wavelength range 220-350 nm was recorded corrected for a blank against the solution devoid of protein. Absorbance and protein concentration are directly related via the extinction coefficient which was calculated from the amino acid sequence using the web-based ExPASy tool ProtParam.

2.3.2 Separation of proteins by SDS-polyacrylamide gel electrophoresis

Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) is a technique to separate proteins according to their molecular weight based on the protocol of *Laemmli et al.* [73]. Proteins move within an electric field according to their charge. As an excess of SDS is added to the sample, the negatively charged dodecyl sulfate anion binds to the protein covering it with negative charges. As the number of dodecyl sulfate ions is very large compared to the number of charged groups of a protein at a pH of 8.8, the inherent charge of the protein becomes insignificant. Thus, when subjected to an SDS-PAGE, all proteins migrate towards the anode of the electrophoresis chamber with their migration speed only determined by their size as they are slowed by the meshwork of polyacrylamide fibres.

Polyacrylamide gels were prepared using the casting devices of the Mini Protean II system from BioRad. A front glass plate and a back glass plate were carefully cleaned and inserted into casting frames. The separating gel solution was prepared and filled into the glass frames. The acrylamide solution was overlaid with isopropanol to create a clear and level surface and then incubated for 45 minutes allowing the solution to polymerize and solidify. After polymerization, the isopropanol was decanted and stacking gel solution was poured on top of the solid separating

gel and a comb was inserted into the still liquid stacking gel. After another 45 minute incubation, the stacking gel was solidified and the gels with their glass frames were removed from the casting frames. For storage, the gels were wrapped in wet paper towels and stored at 4 °C.

Reagent	Stacking Gel	Separating Gel (12%)	Separating Gel (15%)
30% Polyacrylamide*	17 mL	40 mL	50 mL
1M Tris pH 6.8	12.5 mL	none	none
1M Tris pH 8.8	none	25 mL	25 mL
10% SDS in H ₂ O	1 mL	1 mL	1 mL
TEMED solution [#]	100 µL	40 µL	40 µL
10% APS in H ₂ O [#]	1 mL	1 mL	1 mL
ddH ₂ O	Add to 100 mL	Add to 100 mL	Add to 100 mL

*The used polyacrylamide solution had an acrylamide:bis-acrylamide ratio of 37.5:1

[#]Marked items were added last and only shortly before casting the gels

To prepare samples for an SDS-PAGE, proteins were fully reduced and denatured by heating them for 5 min at 95 °C in sample buffer. Precast gels (12 or 15%) were installed into the gel chamber and immersed into SDS running buffer. The boiled samples were then applied on the gel and subjected to electrophoresis at 180 V, 300 mA for 60-120 min until the dye buffer front had nearly traversed the gel completely.

1x SDS running buffer		5x sample buffer	
14.4 g	Glycine	5 g	SDS
2.0 g	SDS	30.5 g	Glycerol
2.9 g	Tris pH 8.8	15 ml	1M Tris/HCl pH 7.0
add to 1L	dd H ₂ O	25 µl	Bromphenol Blue
		2.5 ml	1M DTT
		add to 50 ml	dd H ₂ O

Contents of buffers used for SDS polyacrylamide gel electrophoresis

2.3.3 Coomassie staining of Polyacrylamide gels

After running an SDS-PAGE, gels were removed from their glass frame and stained with Coomassie Brilliant Blue. For staining, gels were treated first with staining solution, then with destaining solution. In each solution, the gels were heated in a microwave oven for 1 min and then incubated for 20 min while slowly shaking (40 rpm). Gels were kept in destaining solution

for short term storage. For long term storage, gels were inserted into a drying frame between two sheets of cellophane and dried under hot air for 2 h.

Staining and destaining solutions:

Reagent	Staining	Destaining
Acetic Acid (glacial)	100 ml	50 ml
Ethanol	200 ml	100 ml
Coomassie Brilliant Blue R-250	2000 mg	
dd H ₂ O	add to 1 l	add to 1 l

2.3.4. Protein Purification Protocols

Within this project three different protocols were employed to purify all proteins. All PDI family member proteins were purified from *E. coli* using either a His-Tag or an Intein-Tag system. Riboflavin binding protein (RfbP) was purified without tag directly from chicken eggs.

Purification of PDI family proteins using the intein expression system

This purification protocol was modified from *Tian et al.* [63] and used for every protein expressed with the intein purification system. An *E. coli* expression strain (BL21DE3 or BL21DE3/RIL codon augmented) was transformed with a pYTB12 expression construct carrying the protein of interest as a fusion protein with the intein tag. *E. coli* cells were grown on an LB plate supplemented with the antibiotics ampicillin (resistance conferred by the pYTB12 vector) and in case of BL21DE3/RIL cells with chloramphenicol. After incubation at 37 °C overnight, single colonies were picked and transferred into a starter culture of 200 mL LB medium (500 mL shaker flask, supplemented with antibiotics). This starter culture was incubated overnight while shaking at 200 rpm for 14-18 h at 37 °C. The starter culture should reach an OD₆₀₀ ≥ 2.0 in order to be used for inoculation of the main culture. 25 mL of starter culture were added to 2.5 L of LB medium (5 L shaker flask, supplemented with antibiotics) and incubated while shaking at 200 rpm and 37 °C. The OD₆₀₀ of the main culture was monitored and upon

reaching a value of 0.6 (usually after 3.5 h), protein production was induced by the addition of IPTG to a final concentration of 1 mM. The induced culture was then incubated while shaking at 200 rpm and 15 °C for 16-20 h. During this incubation step the main culture should reach at least an $OD_{600} \geq 1.3$. Cells were harvested by centrifugation (Beckman J-X26P centrifuge with a JLA8.100 rotor, 10,000 x g, 8 minutes, 4 °C) and weighted after removal of the LB medium. The cells were resuspended in 5 mL of lysis buffer (500 mM NaCl, 20 mM Tris, pH 8.5) per g of cell pellet by pipetting the cell suspension up and down. Next the cells were subjected to lysis with one of the cell disruption systems ‘French Press’, ‘CellDisruptor’ or ‘MicroFluidizer’. The lysate was cleared of cellular debris by centrifugation (Beckmann J-X26P centrifuge, JLA16.250 rotor, 35,000 x g, ≥ 40 minutes, 4 °C). Usually during this centrifugation step, the chitin beads were prepared. Per gram of cell pellet, 2 mL of Biorad Chitin Beads (equal to 4 mL of slurry) were loaded into an empty column resulting in a column volume of 60-90 mL. The storage solution was drained and the column equilibrated with 5-10 column volumes (CV) of low salt buffer (100 mM NaCl, 20 mM Tris, pH 8.5). The low salt buffer was drained from the column and the cleared lysate was added to the column and incubated for at least two hours at room temperature (while rocking or slowly rotating) with the chitin beads. After incubation the column was washed with at least 5 CV of low salt buffer. The column was put into high salt buffer (1 M NaCl, 20 mM Tris, pH 8.5) and washed with 20 CV of high salt buffer. After re-equilibration with at least 5 CV of low salt buffer, the column was drained of buffer, filled with low salt buffer containing 50 mM of fresh and reduced dithiothreitol (DTT) and incubated rocking overnight (at least 16 h) at room temperature for redox-induced intein cleavage.

After cleavage, the protein was eluted from the column by draining the cleavage buffer and the addition of 1-2 CV low salt buffer. The eluate was loaded onto a MonoQ 10/100 GL anion exchange column (GE Healthcare) using an ÅktaPurifier FPLC system (GE Healthcare) and the column was eluted using filtrated and degassed low salt buffer and high salt buffer as buffer A and B, respectively.

For the elution the following profile was used:

Step	1	2	3	4	5
Elution volume in CV	0→5	5→10	10→25	25→30	30→35
Buffer composition in % B	0→0	0→20	20→40	40→100	100→100
Fractionation size in mL	0	12	4	8	12

During the elution, the absorbance at 280 nm and 260 nm was monitored and PDI family members eluted around 25-35% of buffer B in step 3 usually with multiple peaks corresponding most likely to different conformations exposing a different amount of charged residues at the protein surface. During the first purification of a protein, fractions with a mean absorbance of more than 200 mAU were checked for protein content and purity via SDS-PAGE. Fractions containing reasonably pure (>80%) protein of the expected size were pooled and concentrated to a final volume of 5 mL (with a maximum concentration of 50 mg/mL) using Centricon concentrators with an appropriate MWCO depending on the molecular weight of the protein (see Table). The pooled and concentrated sample was loaded onto a Superdex 200 26/60 HiLoad size exchange chromatography column (GE Healthcare) pre-equilibrated with 1.2 CV of sample buffer (250 mM NaCl, 20 mM Tris pH 8.5) and eluted with 1.2 CV of sample buffer using an ÅktaPurifier FPLC system while monitoring the absorbance at 280 nm and 260 nm. Starting at an elution volume of 120 mL, the elution was sampled into 4 mL fractions. PDI family members usually eluted in one or two peaks at an elution volume that is earlier than expected for a protein of their size. Fractions with a mean absorbance over 250 mAU were checked for purity and all fractions with a purity >95% were pooled and concentrated using a Centricon concentrator (MWCO: see table). The concentrated protein was either used directly for experiments or divided into 200 µL aliquots. These aliquots were flash frozen in liquid nitrogen and stored at -80 °C.

MW of sample	MWCO
4 kD – 15 kD	3 kD
15 kD – 45 kD	10 kD
45 kD or more	30 kD

Purification of PDI family proteins using a His-tag for purification.

This purification protocol was used for all proteins expressed from the modified pET23 vector kindly provided by Lloyd Ruddock from the University of Oulu, Finland. The codon augmented *E. coli* expression strain BL21DE3/RIL was transformed with a pYTB12 expression construct carrying the protein of interest. *E. coli* cells were grown on an LB plate supplemented with the antibiotics ampicillin (resistance conferred by the pYTB12 vector) and chloramphenicol (resistance inherent to the RIL expression strain). After incubation at 37 °C overnight, single colonies were picked and transferred into a starter culture of 200 mL LB medium (500 mL

shaker flask, supplemented with antibiotics). This starter culture was incubated overnight while shaking at 200 rpm for 14-18 h at 37°C. The starter culture should reach an $OD_{600} \geq 2.0$ in order to be used for inoculation of the main culture. 25 mL of starter culture were added to 2.5 L of LB medium (5 L shaker flask, supplemented with antibiotics) and incubated while shaking at 200 rpm and 37 °C. The OD_{600} of the main culture was monitored and upon reaching a value of 0.6 (usually after 3.5 h), protein production was induced by the addition of IPTG to a final concentration of 0.5 mM. The induced culture was incubated at 15°C for 16-20h while shaking at 200 rpm. During this incubation step the main culture should reach at least an $OD_{600} \geq 1.3$. Cells were harvested by centrifugation (Beckman J-X26P with a JLA8.100 rotor, 10,000 x g, 8 minutes, 4 °C) and weighed after removal of the LB medium. The cells were resuspended in 5 mL of Buffer A (1 M NaCl, 20 mM Tris, 10 mM Imidazol, pH 8.5) per g of cell pellet by pipetting the cells up and down. Next the cells were subjected to lysis with one of the cell disruption systems 'French Press', 'CellDisruptor' or 'MicroFluidizer'. The lysate was cleared of cellular debris by centrifugation (Beckmann J-X26P centrifuge with a JLA16.250 rotor, 35,000 x g, ≥ 40 minutes, 4°C). The cleared lysate was supplemented with PMSF (~1 mg per 20 mL of lysate) and loaded onto a HisTrap column (Novagen) using a peristaltic pump. The loaded column was attached to an an ÅktaPurifier FPLC system and washed extensively with buffer A at a flowrate of 4 mL/min until the monitored absorbance at 280 nm became stable (<1 mAU absorbance change per mL). The protein was then eluted with 40% of buffer B (1 M NaCl, 20 mM Tris, 300 mM Imidazol, pH 8.5) at a reduced flow rate of 0.8 mL/min. Fractionation was started as soon as the monitored absorbance at 280 nm exceeded 500 mAU and the following 10 mL were directly loaded onto a Superdex 200 26/60 HiLoad size exchange chromatography column pre-equilibrated with 1.2 CV of sample buffer (250 mM NaCl, 20 mM Tris pH 8.5). The column was eluted with 1.2 CV of sample buffer using an ÅktaPurifier FPLC system while monitoring the absorbance at 280 nm and 260 nm. Starting at an elution volume of 120 mL, the elution was sampled into 4 mL fractions. PDI family members usually eluted in one or two peaks at an elution volume that is earlier than expected for a protein of their size. Fractions with a mean absorbance over 250 mAU were subjected to SDS-PAGE and all fractions with a purity $>95\%$ and the correct size were pooled and concentrated using a Centricon concentrator. The concentrated protein was either used directly for experiments or divided into 200 μ L aliquots. These aliquots were flash frozen in liquid nitrogen and stored at -80 °C.

Purification of riboflavin binding protein (RfbP) from chicken eggs

In order to obtain correctly disulfide linked and glycosylated protein, riboflavin binding protein (RfbP) was purified directly from chicken eggs. Approximately 35 eggs purchased at a local grocery store were opened and separated to obtain one liter of egg white, discarding the yolk. The egg white was diluted using an equal volume of 100 mM sodium acetate buffer, pH 4.3. After centrifugation (Beckmann J-X26P with a JLA8.100 rotor, 10,000 x g, 15 minutes, 4 °C) the supernatant was loaded onto a XK16/60 column packed with FastFlow DEAE Sepharose (GE Healthcare) anion exchange chromatography material pre-equilibrated with buffer A (100 mM sodium acetate, pH 4.3). After loading the sample, the column was attached to an ÄktaPurifier FPLC system and was washed and eluted using a linear gradient of buffer A and buffer B (150 mM NaCl, 100 mM sodium acetate, pH 4.3) with the following profile:

Step	1	2
Elution volume in CV	0→5	5→25
Buffer composition in %B	0→0	0→100
Fractionation size in mL	0	5

The absorbance of the bound riboflavin at 445 nm was monitored to pool all RfbP-containing fractions. For the first purification, all protein containing fractions were checked by SDS-PAGE and the pure (>98%) fractions were concentrated and subjected to size exclusion chromatography. For the size exclusion chromatography, the sample was loaded onto a Superdex 200 26/60 HiLoad size exchange chromatography column pre-equilibrated with 1.2 CV of sample buffer (250 mM NaCl, 20 mM Tris, pH 8.5). The column was eluted with 1.2 CV of sample buffer using an ÄktaPurifier FPLC system while monitoring the absorbance at 280 nm, 260 nm and 445 nm. Starting at an elution volume of 120 mL, the elution was sampled into 4 mL fractions and fractions with an absorbance at 445 nm of more than 50 mAU were checked on a SDS-PAGE for size and purity. As all the protein bound to the DEAE sepharose column was RfbP with a very high purity (>98%), the size exclusion chromatography was omitted in all following purifications. The RfbP was dialyzed twice against 50 mM ammonium formate buffer (pH 8) and lyophilized for storage at -80 °C.

2.3.5 Reduced carboxymethylation of model substrates

In order to permanently unfold model substrate proteins such as RfbP or RNase A for ITC binding experiments, all cysteine residues were modified with iodoacetic acid (IAA) to prevent any refolding reaction. For the modification, the usually lyophilized substrate protein was reconstituted in modification buffer (100 mM Tris, pH 8.5). Solid guanidium chloride as well as a solution of 1 M DTT in modification buffer were added to the reconstituted protein to a final concentration of 50 mM DTT and 6 M guanidium chloride. The protein was incubated under these denaturing conditions at 55 °C for 40 minutes. After the denaturation period, the protein solution was allowed to cool to room temperature, before 1 M IAA (dissolved in 1 M NaOH) was added to a final concentration of 150 mM and the reaction was incubated at room temperature for 45 minutes in the dark. The carboxymethylation reaction was quenched by the addition of an excess of DTT (1 M DTT in modification buffer to a final concentration of 200 mM). The modified protein was subsequently dialyzed against 5 L of ITC buffer (250 mM NaCl, 2 mM EDTA, 20 mM Tris, pH 8.5) overnight at 4 °C with a 3 kD MWCO.

2.3.6 Preparation of apo-riboflavin binding protein

In its native conformation RfbP binds riboflavin with high affinity ($K_D=1\text{nM}$) [74]. As the absorbance of bound riboflavin interferes with the absorbance of the 2-nitro-5-thiobenzoate (NTB²⁻) chromophore used in Ellman's assay, the riboflavin has to be removed before native RfbP can be used in this assay. Also the apo-RfbP can be used for a titration curve of riboflavin fluorescence as apo-RfbP stoichiometrically removes riboflavin from the solution while quenching its fluorescence. To generate apo-RfbP, RfbP was reconstituted in ITC buffer (250 mM NaCl, 2 mM EDTA, 20 mM Tris, pH 8.5) and was dialyzed thrice for at least 2 h each against 2 L of 6 M urea in ITC buffer at 4°C with a 3 kD MWCO dialysis membrane. After removal of riboflavin, the urea was removed by dialyzing twice for at least 2 h against 5 L of ITC buffer at 4°C with a 3 kD MWCO membrane.

2.3.7 Preparation of di-eosine-glutathione (Di-E-GSSG)

Oxidized glutathione (GSSG) was modified with the eosine fluorophore to generate the substrate used in the fluorescence self-quenching assay (FSQ assay), according to a procedure which was based on the protocol published by *Raturi et al.* [75] with modifications. For this purpose GSSG was dissolved in conjugation buffer (100 mM potassium phosphate buffer, pH 8.5, 2 mM EDTA) to a final concentration of 1 mM. To this solution solid eosin isothiocyanate was added to a final concentration of 10 mM and incubated in the dark at room temperature overnight (at least 8 h). A SephadexG25 size exclusion chromatography column (Biorad Econo, 5 mL) pre-equilibrated in FSQ reaction buffer (100 mM sodium citrate, pH 7.0, 2 mM EDTA) was used in order to remove both excess eosin isothiocyanate and incompletely conjugated GSSG. The size exclusion chromatography column was loaded with the overnight reaction and run by gravity flow with FSQ reaction buffer. The eluate was divided into 0.5 mL fractions and fractions were tested for an increase in fluorescence before and after reduction with 10 mM DTT. Fluorescence readings were taken with a Fluoromax-4 (excitation at 525 nm, emission at 545 nm, bandpass of 2 nm, 10 measurements of 0.5 seconds each). All di-E-GSSG fractions that showed fluorescence_{reduced} to

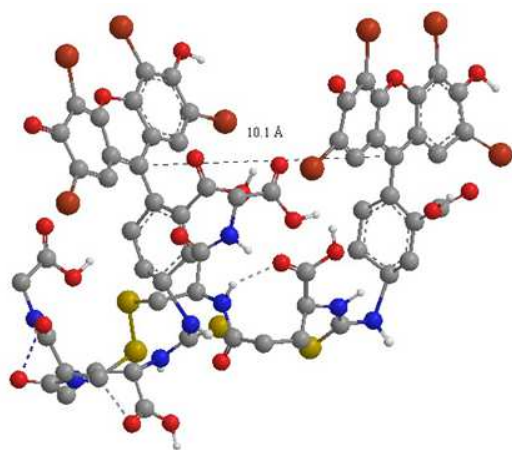


Figure 2.1 – Di-Eosin-Glutathione

Hypothetical structure of oxidized glutathione modified with Eosin groups. In the oxidized form the two eosin moieties (triple ring structures visible in the top half) are kept in proximity.

fluorescence_{oxidized} ratio of equal to or greater than 10 were pooled. The concentration of di-E-GSSG was determined by the absorbance of E-GSH at 525 nm with a molar extinction coefficient of $\epsilon=88,000 \text{ M}^{-1}\text{cm}^{-1}$. If needed, di-E-GSSG can be concentrated by dialysis against solid PEG 5,000 at 4 °C in the dark with a 500 dalton MWCO membrane.

2.3.8 Fluorescence self-quenching (FSQ) assay

The FSQ assay based on the publication by *Raturi et al.* [75] was employed in an attempt to characterize the oxidoreductase function of PDI

family members. This assay is based upon the fluorescence properties of di-E-GSSG. In its oxidized form, the two eosine moieties of the molecule are kept in close proximity (see Figure 2.1.) by the disulfide bond of the oxidized glutathione, thus quenching the fluorescence signal of each other. Upon reduction the spatial restriction and thus the quenching is removed and the fluorescence signal increases. In the original publication, the assay was performed in a phosphate assay buffer (100 mM potassium phosphate buffer, pH 7.0, 2 mM EDTA) with DTT as reductant in the reaction. As it is very hard to exactly reproduce DTT concentrations and DTT is unstable in aqueous solution and prone to air oxidation, it was replaced by TCEP as a reductant. This allowed for a large batch of aliquoted stock solution to be prepared in advance which could then be frozen for storage. As TCEP is unstable in phosphate buffer the buffer substance was replaced creating a new FSQ assay buffer (100 mM sodium citrate, pH 7.0, 2 mM EDTA). For a standard curve FSQ substrate was measured at different concentrations in the oxidized and reduced form using the Fluoromax-4 fluorimeter with the excitation wavelength set to 525 nm (bandwidth = 1 nm) and emission at 545 nm (bandwidth = 1 nm). After establishing a standard curve of a FSQ substrate batch, kinetic experiments were set up using the following composition:

Parameter	Setting	Parameter	Setting
Temperature	25 °C	Integration time	50 ms
Interval time	50 ms	Total time	10 seconds
Excitation wavelength	520 nm	Emission wavelength	540 nm
Excitation bandpass	1 nm	Emission bandpass	1 nm

The experiment was started by the addition of TCEP and the fluorescence signal was recorded using the kinetic record mode (measurement frequency = 5 Hz). The data of the first 10 seconds of the measurement were fitted to a linear equation. The slope of the straight line was plotted against the concentration of the FSQ substrate and this plot was fitted against the Michaelis-Menten equation.

2.3.9 Riboflavin quenching assay

This refolding assay was published first by Rancy&Thrope [76] and uses the fluorescence properties of riboflavin to monitor refolding. Riboflavin binding protein (RfbP) is used to store

riboflavin (vitamin B₂) in high quantities for example in the eggs of birds. The protein binds riboflavin with high affinity (K_D of 1 nM) in a one to one stoichiometry. Free riboflavin in solution fluoresces at an excitation of 450 nm, yielding an emission at 530 nm. Upon binding of riboflavin by apo-RfbP, this fluorescence is quenched. Thus, the fluorescence can be used to observe the refolding of denatured RfbP, as only the correctly folded protein is able to bind riboflavin and to quench its fluorescence. RfbP has 18 cysteines forming nine disulfide bonds and is therefore an excellent model substrate for oxidative folding.

The experiment was set up by recording a baseline of a sample containing only riboflavin and refolding buffer (50 mM potassium phosphate buffer pH 7.0, 2 mM EDTA, 1 mM GSSG, 0.5 mM GSH). The experiment was then started by the addition of reduced and denatured RfbP devoid of riboflavin. The fluorescence was recorded in regular intervals using a Fluoromax-4 fluorescence spectrometer with the following settings:

Parameter	Setting	Parameter	Setting
Temperature	25 °C	Integration time	300 ms
Interval time	10 seconds	Total time	54000 seconds
Excitation wavelength	450 nm	Emission wavelength	525 nm
Excitation bandpass	2 nm	Emission bandpass	3 nm

2.3.10 Isothermal titration calorimetry (ITC)

In preparation for ITC experiments the proteins to be analyzed – usually a member of the PDI family and substrate – were dialyzed against 5 L of ITC buffer (250 mM NaCl, 2 mM EDTA, 20 mM Tris, pH 8.5) overnight at 4 °C with a 3 kD MWCO dialysis tubing. A 500 mL sample was taken from the 5 L of buffer used in the dialysis, filtered and degassed using a Whatman Membrane filter (pore size = 0.2 µm). The dialyzed proteins were centrifuged (Eppendorf centrifuge 5417R, 25,000 x g, 4 °C, 30 minutes) to remove aggregates and precipitated protein. After centrifugation the concentration of the proteins was calculated by measuring the absorbance at 280 nm with a Nanodrop spectrophotometer blanked against the filtered dialysis buffer using the extinction coefficients calculated by the web-based ExPasy tool ProtParam and the amino acid sequence of the protein. The proteins were diluted to the correct concentrations

(standard concentrations: 1 mM for each PDI family member in the syringe and 50 μ M for the substrate in the cell) with the filtered sample buffer. Aliquots of the proteins sufficient for one experiment (450-500 μ L for the syringe, 1.8-2.0 mL for the cell) as well as 2 mL of filtered dialysis buffer were then thermally equilibrated and degassed for at least 10 minutes, using the Microcal Vacuum station set to a temperature 1 $^{\circ}$ C below the temperature planned for the ITC experiment. During this time period, the Microcal VP-ITC used in this project was prepared for the experiment. Each of the syringes, the sample cell and the reference cell were extensively rinsed using at least three times their volume capacity of ddH₂O followed by at least three times their volume capacity of filtered dialysis buffer. The degassed and equilibrated proteins were filled into the syringe (300 μ L – air bubble free) and the reference cell (1.65 mL), respectively, and the degassed and equilibrated buffer was filled into the reference cell (1.65 mL). The standard settings for ITC experiments measuring the interactions between PDI family members and model substrates were:

Parameter	Setting	Parameter	Setting
Temperature	25 $^{\circ}$ C	Injection volume	10 μ L
Reference Power	15 μ cal/sec	Injection speed	300 μ L/min
Syringe concentration	1 mM	Injection spacing	6 minutes
Cell concentration	50 μ M	Injection delay	3 minutes
Number of injections	30	Initial injection	5 μ L

Deviations from those conditions are noted in the Results section of this thesis.

2.3.11 Circular dichroism spectroscopy (CD spectroscopy)

In order to verify the unfolding of reduced carboxymethylated substrate proteins, or to exclude the influence of reduction or mutation on the folding status of PDI family members, CD spectroscopy was performed. All buffer substances and denaturants interfering with CD spectroscopy were removed by dialysis against CD buffer (50 mM potassium phosphate buffer, pH 8.0). In case of Tris buffer, it is also possible to minimize its interference with this technique by dilution of the buffer below 3 mM. The protein sample was diluted to an A_{S280nm} of 1.0 with either CD buffer (for dialyzed samples) or ddH₂O (for non-dialyzed samples). The JASCO J-810

spectropolarimeter was purged with nitrogen and the sample cell set to a temperature of 20°C. The sample was filled into the sample cuvette and inserted into the sample cell. For test scans, the following parameters were used:

Parameter	Setting	Parameter	Setting
Scan type	Continuous Scan	Scan Speed	100 nm/minute
Scan range	185-260 nm	Bandwidth	2 nm
Number of scans	1		

The test runs were terminated as soon as the voltage needed for a sufficient signal exceeded 600 V. The sample was continuously diluted between runs until either the voltage never exceeded 600 V, or until the maximum absolute value of the circular dichroism signal fell below 20 mdeg/cm. With the protein concentration thus adjusted the measurement data scans were taken with the following parameters:

Parameter	Setting	Parameter	Setting
Scan type	Continuous Scan	Scan Speed	30 nm/minute
Scan range	185*-260 nm	Bandwidth	2 nm
Number of scans	10		

*value chosen as low as possible with the voltage during a scan not exceeding 600 V

For final data evaluation the signal was normalized against the protein concentration used in the final scan.

2.3.12 Dynamic Light Scattering (DLS)

Dynamic light scattering, also known as photon correlation spectroscopy, is a spectroscopic method to determine the volume distribution of particles in solution. A monochromatic beam of light is directed at a solution of particles. The wavelength of the scattered light differs from the input light due to the Doppler shift as it interacts with particles moving in solution due to Brownian motion. The magnitude of the shift is related to the speed of the Brownian motion which in turn is related to the diffusion constant and the size of the particle. A detector registers the wavelengths of scattered light perpendicular to the incident laser light. From the distribution of wavelength deviations, the program calculates a distribution profile of volumes in solution

under the assumption that each particle has a spherical shape. Additionally, molecular weights are calculated based on the determined volumes.

For the experiment, the sample solution is adjusted to an approximate protein concentration of 0.5 mg/ml. A quartz cuvette is filled with 80 μ l of sample and inserted into a DynaPro Titan DLS photometer (Wyatt Technology). A single experiment consists of a series of measurements with the following parameters:

Parameter	Setting	Parameter	Setting
Acquisition time	10 sec	Solvent Model	PBS
Number of Acquisitions	20	Refraction Index*	1.333
Laser Power	100%	Viscosity*	1.019 cp
Temperature	10 °C	Cauchy Coefficient*	3119 nm ²

* These values are given by the solvent model 'PBS'

The data is automatically evaluated by the Dynamics V6 software supplied with the equipment and presented as a distribution of molecular radii and molecular weights.

2.3.13 Crystallization of proteins

In order to determine the three dimensional structure of a protein at atomic resolution through X-ray crystallography, single crystals of the investigated proteins are needed. As up to this date there is no known method to determine crystallization conditions of a protein *ab initio*, proteins are screened in many test setups for suitable crystallization conditions. Once promising conditions are identified, they are verified in the fine-screening process and then – should they yield reproducible crystals – optimized. Crystallization is a special form of precipitation and therefore a protein solution is slowly brought to the solubility limit. As the protein becomes supersaturated, the protein begins to form ordered protein-protein-contacts to reduce solvent exposure and thus the need for solvation. While being inside a supersaturated solution, these nuclei of precipitation will extend and can – if forming slowly enough – form the regular lattices of a protein crystal.

Within the scope of this project, the predominant crystallization technique was the vapor diffusion crystallization. Here protein and precipitant are mixed at lower concentrations and

placed in a closed system together with a precipitant solution of higher concentration and several orders of magnitude greater volume. Due to its greater volume, this reservoir precipitant solution will determine the vapor pressure and the sample drop will lose solvent until its osmolarity is equal to the osmolarity of the reservoir solution. This loss in solvent slowly increases the concentration of both protein and precipitant in the sample drop, potentially leading to supersaturation and crystallization of the sample.

The initial screening was done using sitting drop vapor diffusion crystallization. Here the sample drop sits on a shelf above the reservoir solution in a sealed chamber. This was the method of choice as it can be automatized using the Honeybee 963 Cartesian dispensing system, which uses only very small amounts of protein solution (0.3 $\mu\text{L}/\text{drop}$) to create a large number of crystallization experiments. The first step in crystallization screening was to test the protein in standardized screens on 96-well crystallization plates. The crystallization plates were subsequently examined in 24 h intervals for the formation of crystals or crystalline precipitation. Promising conditions were varied in fine screens using the hanging drop vapor diffusion crystallization technique, where the sample drop is suspended from a cover slide over the reservoir solution, increasing the drop surface. This method is set up by hand and uses larger amounts of protein to yield larger crystals.

2.3.13 Diffraction experiments

Protein crystals obtained from the crystallization process outlined above were removed from the sample drop using a mounted CryoLoop (Hampton Research). In order to preserve the protein crystals, they are kept in liquid nitrogen for storage and are cooled to $-180\text{ }^{\circ}\text{C}$ with cold nitrogen gas during the experiment. As the formation of ice crystals within the solvent channels of the crystal can damage or destroy the crystal lattice during the freezing process, the crystals were first transferred into a suitable cryo-solution. This cryo-solution usually consists of the so-called ‘mother liquor’ – the precipitant solution present in the sample drop – supplemented with cryo-protectants. Commonly used cryo-protectants are glycerol, ethylene glycol, PEG 400, methoxypolyethylene glycol 350 (MPG), sugars or oil. The crystals were allowed to equilibrate in the cryo solution for 2-5 minutes before being flash frozen in ice-free liquid nitrogen.

For the initial diffraction experiments, crystals were mounted on the goniometer of the in-house X-ray system. X-rays generated by a MicroMax-007HF X-ray generator (Rigaku) were targeted onto the crystal and the diffraction pattern was recorded using an R-Axis HTC imaging plate detector. For the initial estimation of crystal quality, the following settings were used to save beam time:

Parameter	Setting
Detector distance	200 mm
Number of Images	2
Φ angle	0.5°
Interval angle	90°
Exposure time	10 minutes

Crystals diffracting to a resolution of 3 \AA were then used to record a complete data set using either the home source or synchrotron beam lines. All data sets which were used to solve the structure of ERp27 were taken at the BESSY synchrotron with the following settings:

Parameter	Setting (native ERp27)	Setting (SeMet ERp27)
Detector distance	222.41 mm	249.42 mm
Number of Images	500	720
Φ angle	1.0°	0.5°
Interval angle	1.0°	0.5°
Exposure time	5 sec	2 sec
Wavelength	0.91841 \AA	0.97977 \AA

2.3.14 Data Processing

A diffraction experiment generates data in the form of diffraction images. Each image corresponds to the X-ray scattering of the crystal at a particular orientation. Each image contains a significant number (usually more than 50) of observed reflections and the data sets used to solve and refine the structure contained 500 and 720 images respectively. In order to gain the structure factor amplitudes used for structure determination, the data needs to be processed. Data processing was done with the programs Mosflm or XDS consisting of four steps in each case. First the images are indexed. Based on a small subset of the data, the agreement of the observed reflections with one of the 14 Bravais lattices is calculated. The objective of this step is to

suggest possible space groups, to determine initial unit cell parameters of the crystal and to assign the Miller indices to each reflection. After a preliminary space group has been assigned, the intensity of each reflection is calculated in the integration step. The program measures the intensity of each observed reflection and corrects for the local background. Due to space group symmetry, the very good signal-to-noise ratio of strong reflections can be used to correctly identify and integrate weak reflections. After the intensity of all reflections has been measured, the intensities are scaled and merged. A single structure factor can give rise to multiple observed reflections due to symmetry and data set redundancy. During the course of an experiment at a synchrotron, the beam intensity typically declines. This leads to weaker measured intensities in later images. To counteract this decline in signal, reflection intensities are scaled before all intensities belonging to one structure factor are averaged. As a result of scaling and merging, a single averaged intensity is calculated for each structure factor. After this step, R_{sym} is calculated from the agreement between the separate observations of a reflection and its symmetry mates. Using this parameter the quality of the data set can be judged with good data sets resulting in R_{sym} below 5% while bad data sets exceed an R_{sym} of 10%. The final step of the data processing is the calculation of the structure factor amplitudes from the scaled and merged intensities. The calculated structure factor amplitudes can then be used to solve the structure.

2.3.15 Structure solution

During a diffraction experiment, only the position and intensity of the scattered X-rays are recorded. These measurements give rise to the structure factor amplitudes. The information contained in the phase of the scattered X-rays cannot be recorded by current detectors and is therefore lost. This lost information is at the core of the phase problem of crystallography. A phase has to be assigned to each structure factor in addition to its amplitude in order to be able to calculate an electron density map. Multiple methods have been developed in order to obtain the lost information. Within this project, the single wavelength anomalous dispersion (SAD) and molecular replacement methods have been used.

For the SAD approach a data set is collected from a crystal containing an anomalous scatterer. Selenium is often employed, as seleno-methionine can be incorporated into proteins during expression and thus modified proteins often from crystals which are identical to the native

protein. All chemical elements, when excited at their element specific absorption edge, give rise to anomalous scattering. Due to the absorption of radiation, the phase of the scattered X-rays shifts uniquely. This leads to differing intensities in reflections related by Friedel's Law, which should have equal intensities in the absence of anomalous scatterers. This difference contains the phase information. Through comparison of the Friedel pairs $I(h,k,l)$ and $I(-h,-l,-k)$, an initial phase can be assigned to the structure factor if sufficient anomalous signal is available to accurately determine the phases. An electron density map can be calculated using the determined phases and amplitudes of the structure factors.

For molecular replacement, phase information from already existing structures is used to solve the phase problem. The pre-existing template structure should have high structural homology with the structure to be solved. The Patterson function is calculated from both the template and the experimentally determined structure factor amplitudes. The resulting Patterson maps are then matched first by rotation of the search template in three dimensions, followed by translation in three dimensions. After the most suitable position and orientation of the template has been found, an electron density map is calculated using the experimentally determined structure factor amplitudes and the phases of the template.

An electron density map gives the distribution of electrons – and thereby the position of each atom – in the three dimensions. Using the initial electron density map, a model is built in accordance with the electron densities as well as with the pre-existing knowledge about the molecule such as amino acid sequence and stereochemical parameters such as bond lengths and bond angles. This initial model building is often automated using software tools from the CCP4 or Phenix program suites. Within this project Buccaneer from the CCP4 suite was used for the initial model building. With the initial model the refinement process can be started. During refinement, the model parameters (i.e. coordinates, B-factors) are used to improve the accuracy of the determined phases and therefore the electron density map. The improved electron density map allows for a more detailed model which in turn will lead to an even more accurate electron density map. This model improvement is usually done manually in Coot using the pre-existing knowledge about a protein. During refinement, the R-factor is calculated to judge the agreement of the model with the experimental data. The R-factor gives the deviation of the calculated structure factor amplitudes from the experimentally observed structure factor amplitudes. Thus

lower R-factors indicate a higher model quality. As the model quality increases, it is important that the number of model parameters should not exceed the number of observed reflections in order to avoid model bias. For validation of the model, the free R-factor (R_{free}) is calculated [77]. During each refinement cycle, the same random fraction of data (usually 5-10%) is not refined and the R_{free} is calculated from this unrefined data. If for a new model only the R-factor decreases while the R_{free} does not improve, the model could be over fitted and the refinement step should be reevaluated. After multiple steps of successful refinement, the final model should yield an R-factor below 25% and an R_{free} below 30%.

The quality of the final model is validated by comparing model parameters like bond lengths and angles with the ideal values from literature [78]. As these parameters are often constrained to the ideal values during refinement, the orientation of amino acids is used as independent criterion of model accuracy. In a Ramachandran plot, the Φ and Ψ angles of each amino acid are plotted. For each amino acid there are favorable regions on the plot, unfavorable but allowed regions and disallowed regions defined in the literature [79]. In the Ramachandran statistics the vast majority (>95 %) of amino acids should fall into the favored region with only a small portion in unfavorable or disallowed regions. It is possible for amino acids to attain conformations from the disallowed regions of the plot, but those amino acids should have well defined electron density and an environment that can explain the unusual conformation.

2.4 Materials

2.4.1 List of primers

#	Name	Sequence 5` to 3`	Usage
1	CGHC->SGHS a` for	GTACTATGCCCCATGGTCTGGTCACTC TAAGAGATTGGCCCAAC	Mutagenesis of yeast PDI
2	CGHC->SGHS a` rev	GTTGGGGCCAATCTCTTAGAGTGACC AGACCATGGGGCATAGTAC	Mutagenesis of yeast PDI
3	Q87K for	CATTACCTTGGCCAAGATCGACTGTA C	Mutagenesis of yeast PDI
4	Q87K rev	GTACAGTCGATCTTGGCCAAGGTAAT G	Mutagenesis of yeast PDI
5	L400E for	GGACGTTCTTGTGAGTACTATGCCCC	Mutagenesis of yeast PDI
6	L400E rev	GGGGCATAGTACTCAACAAGAACGTC C	Mutagenesis of yeast PDI
7	Seq -20/6490 for	GTTTTTGCTTGGATCCCAGG	Sequencing of yeast PDI
8	seq 575/7085 for	CCCTCCGCCATGGACGAGCC	Sequencing of yeast PDI
9	seq 555/7065 rev	GGCTCGTCCATGGCGGAGGG	Sequencing of yeast PDI
10	seq 1038/7548 for	CCTCCCAATCGTGAAGTCC	Sequencing of yeast PDI
11	seq 1018/7528 rev	GGACTTCACGATTGGGGAGG	Sequencing of yeast PDI
12	seq 1433/8043 rev	CAGCCGGATCCCCTTCTGC	Sequencing of yeast PDI
13	yPDI GLY242 for	GCATGCCATATGGGTGAAATCGACGG TTCCG	Mutagenesis of yeast PDI
14	yPDI GLY242 rev	CAGAATTC AACCAAAGTAGGGCAAGG CTTCC	Mutagenesis of yeast PDI
15	yPDI GLN140 for	GCATGCCATATGCAACCGGCTGTGCG CG	Mutagenesis of yeast PDI
16	yPDI GLN140 rev	GTGAATTCATTGGCTTTGCTTGATCAT GAATTGG	Mutagenesis of yeast PDI
17	yPDI VAL365 for	GCATGCCATATGGTGAAGTCCAAGA GATCTTCG	Mutagenesis of yeast PDI
18	yPDI VAL365 rev	GCATGAATTCACACGATTGGGGAGGC ATCAC	Mutagenesis of yeast PDI
19	yPDI start for	GCTGGTCATATGCAACAAGAGG	PCR amplification
20	yPDI term rev	CGAGTCAGAATTACAATTCATCGTGA ATGG	PCR amplification
21	yeast htm1 5'	GCATACTAGTGTGCTGCTTATGGGT GCTTCTAG	PCR amplification
22	yeast htm1 3'	GCATCTCGAGTCATCATAAATAAAT AAGTTGATGATGG	PCR amplification
23	yeast htm1 c term min 5'	GCATCATATGAACAACCTGTCTACTC GTCGCACTTGG	PCR amplification
24	yeast htm1 c term	CGATCTCGAGTCATCATGGGATCAGG	PCR amplification

	min 3'	CACGTGCTAAATTGC	
25	human calreticulin 5'	CGATCATATGGAGCCTGCCGTCTACTT CAAGGAGC	PCR amplification
26	human calreticulin 3'	CGATCTCGAGTCACTACAGCTCGTCCT TGG	PCR amplification
27	human calnexin 5'	GCATCATATGCATGATGGACATGATG ATGATGTG	PCR amplification
28	human calnexin 3'	CGATCTCGAGTCATCACTCTCTTCGTG GCTTTCTG	PCR amplification
29	human ERp57 5'	CGATCCATGGCCGACGTGCTAGAACT CACGG	PCR amplification
30	human ERp57 3'	GCATCTCGAGTCATTAGAGATCCTCCT GTGCC	PCR amplification
31	yeast htm1 639for	See Constantin Braun's Diploma Thesis	Sequencing of yeast htm1
32	yeast htm1 1032for	See Constantin Braun's Diploma Thesis	Sequencing of yeast htm1
33	hPDI cluster 1 for	GACGTGGCTTTTGTATGCGGCAGCAA CGTCTTTGTGGAG	Entropy reduction mutagenesis
34	hPDI cluster 1 rev	CTCCACAAAGACGTTTGCTGCCGCAT CAAAAGCCACGTC	Entropy reduction mutagenesis
35	hPDI cluster 2 for	GAGTTCTTTGGCCTGGCGGCGGCAGC GTGCCCGGCCGTGCGCC	Entropy reduction mutagenesis
36	hPDI cluster 2 rev	GGCGCACGGCCGGGCACGCTGCCGCC GCCAGGCCAAAGAACTC	Entropy reduction mutagenesis
37	hPDI cluster 3 for	CGCCTCATCACCTGGCGGCGGCGAT GACCAAGTACAAG	Entropy reduction mutagenesis
38	hPDI cluster 3 rev	CTTGTACTIONGGTCATCGCCGCCGCCAG GGTGATGAGGCG	Entropy reduction mutagenesis
39	hPDI cluster 4 for	CAAGTACAAGCCGAAGCGGCAGCGC TGACGGCAGAGAGG	Entropy reduction mutagenesis
40	hPDI cluster 4 rev	CCTCTTGCCGTCAGCGCTGCCGCTTC GGGCTTGTACTTG	Entropy reduction mutagenesis
41	yPDI cluster 1 for	TACGAAGAAGCCAGGCAGCAGCTGC TGCGGCAGCCGATGCTGACGCTG	Entropy reduction mutagenesis
42	yPDI cluster 1 rev	CAGCGTCAGCATCGGCTGCCGCAGCA GCTGCTGCCTGGGCTTCTTCGTA	Entropy reduction mutagenesis
43	yPDI cluster 2 for	TTATACCCAGGTGGTGC GGCCCGCCGC ATCTGTTGTGTACCAA	Entropy reduction mutagenesis
44	yPDI cluster 2 rev	TTGGTACACAACAGATGCGGCGGCCG CACCACCTGGGTATAA	Entropy reduction mutagenesis
45	yPDI cluster 3 for	GGCAACTTGAACATGGCGGCAGCATT CCCTCTATTTGCC	Entropy reduction mutagenesis
46	yPDI cluster 3 rev	GGCAAATAGAGGGAATGCTGCCGCCA TGTTCAAGTTGCC	Entropy reduction mutagenesis

47	a forw	CACCACCATATGGACGCCCCGAGGA GGAGGACC	human PDI truncations
48	a rev	GAATTCGGATCCTTACGGGCCCCGTGC GCTTCTCAGC	human PDI truncations
49	a' forw	CACCACCATATGAAGCAGCCTGTCAA GGTGCTTG	human PDI truncations
50	b forw	CACCACCATATGGCTGCCACCACCCT GCCTGACG	human PDI truncations
51	b rev	GAATTCGGATCCTTACAGCTGGTTGTG TTTGATAAAGTCC	human PDI truncations
52	b' forw	CACCACCATATGCCCTTGTCATCGAG TTCACCG	human PDI truncations
53	b' rev	GAATTCGGATCCTTACTTGATTTGCC CTCCAGGAACG	human PDI truncations
54	c rev	GAATTCGGATCCTTACAGTTCATCTTT CACAGC	human PDI truncations
55	pET23 mod forw	TAAGGATCCGAATTCCTAG	Sequencing of plasmids provided by L. Ruddock
56	pET23 mod rev	CATATGGTGGTGATGGTG	Sequencing of plasmids provided by L. Ruddock

The suffix 'for' or 'forw' denominates the primer binding 5' to the sequence to be amplified, mutated or sequenced. The suffix 'rev' denominates the primer binding 3' to the target sequence.

2.4.2 List of expression vectors

#	Vector	Antibiotic	Insert	Origin	Tag	Position	Linker
1	pTYB12	Amp	yeast PDI1	G.Tian	CBD	N-term	Intein
3	pTYB12	Amp	yeast PDI1 SSSS	XK1	CBD	N-term	Intein
5	pTYB12	Amp	human PDI	G.Tian	CBD	N-term	Intein
7	pTYB12	Amp	chicken ERp57	G.Tian	CBD	N-term	Intein
9	pTYB12	Amp	yeast PDI SSCC	XK1	CBD	N-term	Intein
11	pTYB12	Amp	yeast PDI CCSS	XK1	CBD	N-term	Intein
13	pTYB12	Amp	yeast PDI 18AA C-trunc #3	G.Tian	CBD	N-term	Intein
14	pTYB12	Amp	yeast PDI 28AA C-trunc #4	G.Tian	CBD	N-term	Intein
15	pTYB12	Amp	yeast PDI abb'a' c-del #5	G.Tian	CBD	N-term	Intein
16	pTYB12	Amp	yeast PDI b'a'c #6	G.Tian	CBD	N-term	Intein
17	pTYB12	Amp	yeast PDI bb'a'c #7	G.Tian	CBD	N-term	Intein
18	pTYB12	Amp	yeast PDI abb' #8	G.Tian	CBD	N-term	Intein
19	pTYB12	Amp	yeast PDI ab #9	G.Tian	CBD	N-term	Intein
20	pTYB12	Amp	yeast PDI abb'c #10	G.Tian	CBD	N-term	Intein
21	pTYB12	Amp	yeast PDI C64A #11	G.Tian	CBD	N-term	Intein
22	pTYB12	Amp	yeast PDI C409A #12	G.Tian	CBD	N-term	Intein
23	pTYB12	Amp	yeast PDI P106A #13	G.Tian	CBD	N-term	Intein
24	pTYB12	Amp	yeast PDI P451A #14	G.Tian	CBD	N-term	Intein
25	pTYB12	Amp	yeast PDI P106A + P451A #15	G.Tian	CBD	N-term	Intein
26	pTYB12	Amp	yeast PDI C64S + C409S #18	G.Tian	CBD	N-term	Intein
27	pTYB12	Amp	yeast PDI C64A + C409A #19	G.Tian	CBD	N-term	Intein
28	pET23_X	Amp	human ERp27	L. Ruddock	His (x6)	N-term	none
29	pET23_X	Amp	human ERp46	L. Ruddock	His (x6)	N-term	none
30	pET23_X	Amp	human ERp18	L. Ruddock	His (x6)	N-term	none
31	pET23_X	Amp	human ERp44	L. Ruddock	His (x6)	N-term	none
32	pET23_X	Amp	human PDIp	L. Ruddock	His (x6)	N-term	none
33	pET23_X	Amp	human PDIr	L. Ruddock	His (x6)	N-term	none
34	pET23_X	Amp	human P5	L. Ruddock	His (x6)	N-term	none
35	pET23_X	Amp	human ERp57	L. Ruddock	His (x6)	N-term	none

All vectors carried the Lac promoter which is induced by the addition of isopropyl β -D-1-thiogalactopyranoside (IPTG) to a final concentration as indicated in the Methods section.

2.4.3 Equipment

Type	Model	Supplier
Autoclave	Systec V-150	Systec
Balance 0.5-500 g	XS 6002S Dual Range	Mettler Toledo
Balance 0-1 g	XS 105 Dual Range	Mettler Toledo
CD-Spectropolarimeter	J-810	Jasco
Cell disruptor	M-110P	Microfluidics
	E615	Constant Cell Disruption Systems LTD
Centrifuge 0-1000 mL	Avanti J-26 XP	Beckmann Coulter
Centrifuge 0-2 mL	5417 R	Eppendorf
Centrifuge 0-5 0mL:	5810 R	Eppendorf
Column body (FPLC)	XK 16	GE Healthcare
Column body (gravity flow)	Econo-Column 1,5 x 15 cm	Biorad
Column body (gravity flow)	Econo-Column 2,5 x 20 cm	Biorad
Crystal mounting	CryoLoop	Hampton Research
Crystallization robot	HoneyBee 963	Zinsser Analytik
Dish washer	Professional G7883 CD	Miele
Dynamic Light Scattering Spectrophotometer	DynaPro Titan	Wyatt Technology
Fluorescencespectrometer	Flouromax4	Horiba Jobin Yvon
FPLC system	Äkta purifier	GE Healthcare
FPLC: Anion exchange column	MonoQ 10/100 GL	GE Healthcare
FPLC: Size exclusion chromatography column	HiLoad 26/60 Superdex 200 prep grade	GE Healthcare
FPLC: Nickel affinity column	Ni-MAC Cartridge 5 mL	Merck (Novagen)
Gelelectrophoresis chamber	Mini-Protean 3-cell	Biorad
Ice machine	Eismaschine 94774	Ziegra Eismaschinen
Illumination table	P265.1	Roth
Incubator	Type B15	Thermo Electron Corp.
Liquid culture incubator	ISF-1-W	Kühner
Liquid handling robot	Lissy	Zinsser Analytik
Magnetic stirrer	MR 3002	Heidolph
Microscope	SteREO Discovery.V12	Zeiss
Microscope: Camera	AxioCam MRC	Zeiss
Microscope: Light source	KL 2500 LCD	Zeiss
PCR Cycler	Mastercycler EPgradient S	Eppendorf
pH-electrode	BlueLine 14pH	Schott
Pipettes	pipet lite	Rainin
Pipetboy	pipetus	Hirschmann Laborgeraete

Power supply	PowerPac Basic	Biorad
Spectrophotometer	Bio-Photometer	Eppendorf
Spectrophotometer	Nanodrop ND 1000	Peqlab
Spectrophotometer	UV-Vis	Agilent
Thermoblock	Rotilabo-Block Heater 250	Roth
Thermomixer	comfort	Eppendorf
X-ray detector	R-AXIS HTC	Rigaku
X-ray generator	MicroMax-007HF	Rigaku
X-ray optics	VariMax HF	Osmic Inc.

2.4.4 Consumables

Type	Model	Supplier
24-well plate	SuperClear	Crystalgen
96-well plate	ClearPlate Halfarea MB	Greiner
Chitin Beads	S6651L	New England Biolabs
Concentrator 10 mL	Vivaspin 6	Sartorius stedim Biotech
Concentrator 30 mL	Vivaspin 20	Sartorius stedim Biotech
Disposable cuvettes	UVette	Eppendorf
DEAE Sepharose	DFE100	Sigma-Aldrich
Dialysis cassettes	Slide-A-Lyzer	Pierce Biotechnology
Dialysis tubing	Spectra/Por	SpectrumLabs.Com
Micro reaction tubes	Safe-Lock 1.5 mL / 2 mL	Eppendorf
Pipette tips	10 µl, 200 µl, 1000 µl	Rainin
Reaction tubes	Greiner tube 15 ml / 50 ml	Greiner

3.4.5 Chemicals and Solvents

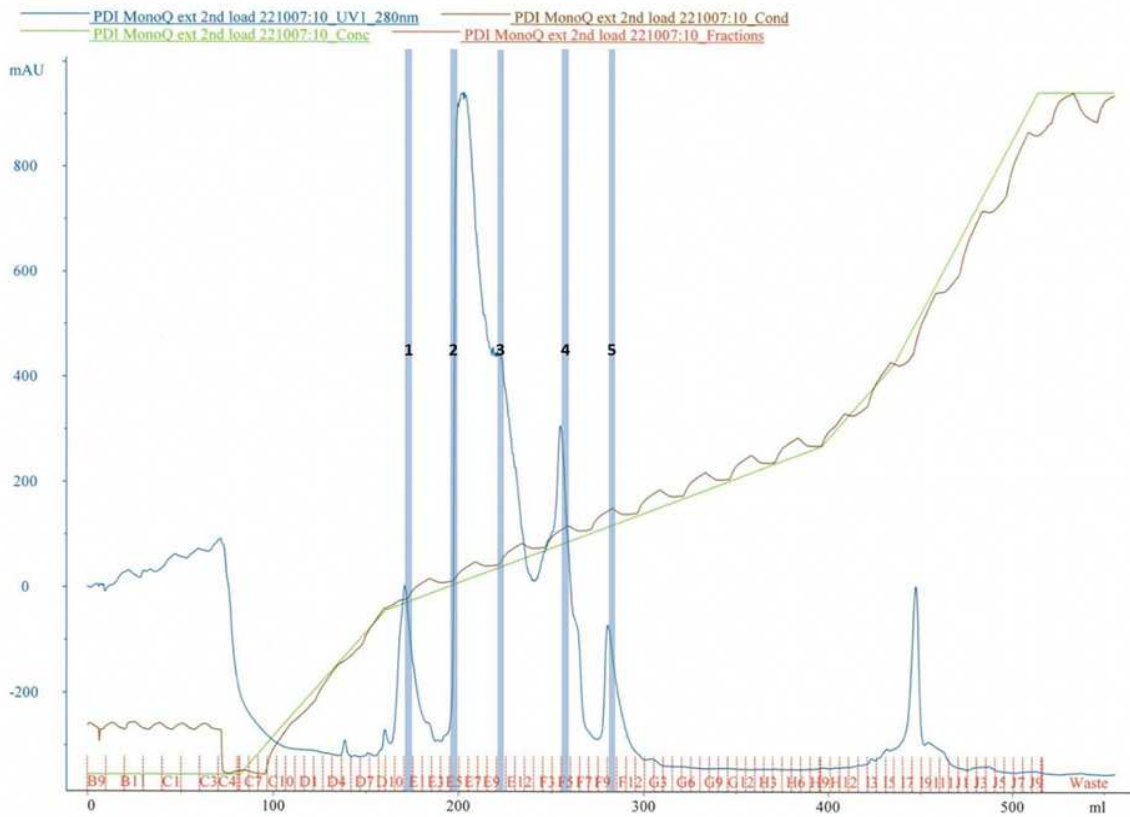
Chemical	Catalogue ID	Supplier
2-Propanol (Isopropanol)	6752.4	Roth
Acetic acid (glacial)	7332.2	Roth
Acetonitrile	AE70.2	Roth
Ammonia	CP17.1	Roth
Ammonium carbonate	09832	Fluka
Ammonium peroxodisulfate	09913	Fluka
Ammonium sulfate	9318.1	Roth
Ampicillin (sodium salt)	K029.2	Roth
Barium Chloride	11760	Fluka
Chloramphenicol	3886.3	Roth
Chymotrypsin	C4129	Sigma-Aldrich

Coomassie Brilliant Blue R25	3862.2	Roth
Dithiothreitol	6908.4	Roth
Ethanol (absolute)	9065.2	Roth
Ethanol (denatured)	K928.4	Roth
Formic acid	399388	Sigma-Aldrich
Glycerol	A3552	AppliChem
Guanidinium chloride	0037.1	Roth
Hydrochloric acid	4623.2	Roth
Hydrogen peroxide	8070.4	Roth
Imidazole	3899.4	Roth
Iodo acetic acid	I6806	Sigma-Aldrich
L-Glutathione (oxidized)	G4376	Sigma-Aldrich
L-Glutathione (reduced)	A2084	AppliChem
Magnesium chloride	HN03.3	Roth
MES	69892	Sigma-Aldrich
Methanol	4627.5	Roth
Nickel sulfate	T111.1	Roth
Papain	8933.1	Roth
Phenylmethylsulfonyl fluoride	6367.2	Roth
Poly ethylene glycol 10,000	81280	Fluka
Poly ethylene glycol 3,000	81227	Fluka
Poly ethylene glycol 400	81350	Fluka
Poly ethylene glycol 5,000 monomethyl ether	81323	Sigma-Aldrich
Potassium chloride	HN02.3	Roth
Potassium hexacyano ferrate (III)	60300	Fluka
Potassium hydroxide	6751	Roth
Potassium phosphate dibasic	P749.3	Roth
Potassium phosphate monobasic	P9791	Sigma-Aldrich
Potassium phosphate tribasic	60494	Fluka
Riboflavin	R7649	Sigma-Aldrich
Sodium acetate	3580.1	Roth
Sodium azide	K305.1	Roth
Sodium chloride	3957.2	Roth
Sodium dodecylsulfate	2326.2	Roth
Sodium hydroxide	6771.1	Roth
Trichloroacetic acid	91228	Fluka
Trifluoroacetic acid	302031	Sigma-Aldrich
TRIS (2-Amino-2-hydroxymethyl-propane-1,3-diol)	4855.3	Roth
Tris(2-carboxyethyl)phosphine	HN95.2	Roth
Trypsin	T1426	Sigma-Aldrich
Urea	2317.2	Roth

3. Results & Discussion

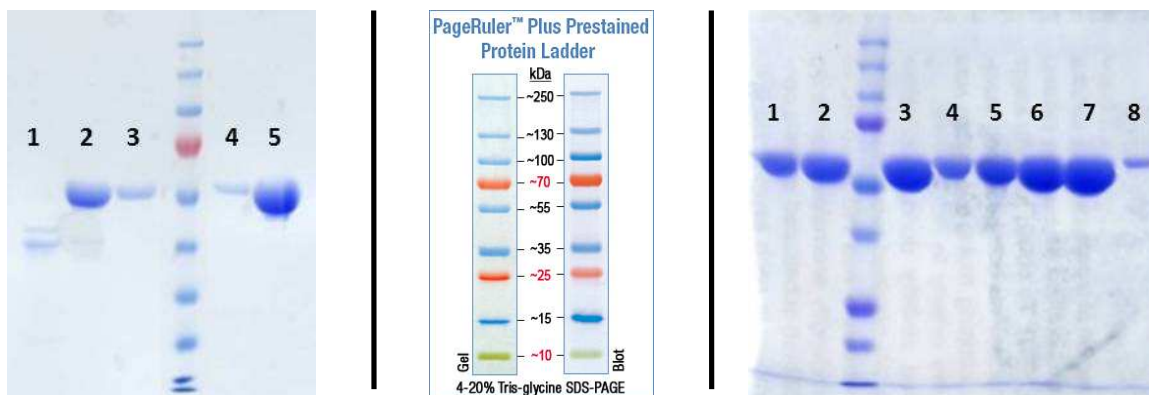
3.1 The multistate-equilibrium of yeast PDI

One of the first tasks of this project was to purify recombinant yeast PDI for biochemical experiments and crystallization attempts. For purification the *E. coli* XK2 strain was used together with a pTYB12-PDI vector allowing for the inducible expression of an intein-PDI-fusion protein. The purification was carried out as described in the Methods section of this document utilizing a three-step purification strategy consisting of an affinity chromatography with chitin beads, an anion exchange chromatography using a MonoQ column and a final size exclusion chromatography using a Superdex200 column. During the purification process protein heterogeneity could be detected on the the anion exchange column. As can be seen from the elution profile monitoring the absorbance at 280 nm (Figure 3.1), protein eluted from the column at concentrations of 260 mM, 290 mM, 310 mM and 335 mM of sodium chloride with a possible additional species eluting at 295 mM sodium chloride forming a shoulder of the second peak. When samples from each absorbance peak were subjected to SDS-PAGE analysis (Figure 3.2. A), the peaks 2-5 contained only a single pure protein band running at a height of 57 kD while the first peak contained a possible degradation product. So the anion exchange chromatography revealed the existence of multiple isoforms of yeast PDI, apparently differing in its surface charge. This indicates either multiple different conformations that expose a different amount of charged residues to the solvent, or the existence of multiple oligomeric forms of PDI that bury different amounts of charged residues in the interface, or a combination of both. All peaks containing pure yeast PDI were subsequently pooled and concentrated. The absorbance profile of a size exclusion chromatography of the pure yeast PDI sample clearly shows two distinct peaks each (Figure 3.3). When characterized with SDS-PAGE both peaks contained one pure protein running at the same height indicating equal size (Figure 3.2. B). The elution volume during size exclusion chromatography is determined by the molecular weight, the oligomeric state and the molecular shape of the sample. The expected elution volume for a globular 54 kD protein on a Superdex 200 XK26/60 column is 220 mL. With elution volumes of 154 mL and 182 mL for the two species, the smaller species is already eluting considerably earlier than expected for its size. For the remainder of this thesis, the earlier eluting species will be referred to as the dimer while



(Top) Figure 3.1 – Anion exchange chromatography during the purification of yeast PDI. Multiple protein species elute over a range of 250-400 mM salt. A sample was taken from each peak as indicated. The last peak eluting around 450mL was ignored as its absorbance at 260 nm was greater than the absorbance at 280 nm.

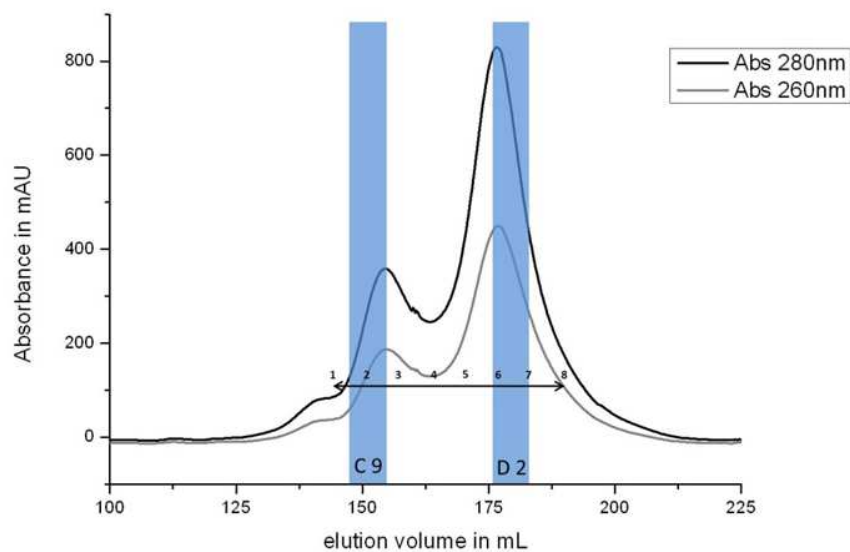
(Below) Figure 3.2 – SDS gel electrophoresis of protein samples from the PDI purification. Left panel: Fractions taken from the anion exchange chromatography (Fig. 3.1). Middle panel: Pre-stained protein ladder (taken from www.invitrogen.com). Right panel: Fractions taken from the size exclusion chromatography (Fig. 3.3).



the species eluting later will be referred to as a monomer since a monomer/dimer equilibrium would explain the elution behavior with two peaks.

In order to study the monomer-dimer equilibrium, samples were taken from both the monomer and the dimer species (as indicated in Fig. 3.3). After an 18 h incubation at 4°C both samples were subjected to analytical size exclusion chromatography on a Superdex200 10/100 GL. Apart from the protein amount, both samples eluted in an identical fashion (Fig. 3.4) indicating that after 18 h both samples had re-equilibrated to the same state. Interestingly, the monomer-dimer-equilibrium is the same despite the concentration difference. So as a follow-up, a dilution series with a dilution factor of three was created from the monomer sample. After 18 h of equilibration, each sample was separated by analytical size exclusion chromatography. Each absorbance profile at 280 nm was analyzed using the UNICORN software to automatically detect peaks and calculate the corresponding peak areas. As can be seen from both the elution profiles and the integrated peak areas (Fig. 3.5), the ratio of monomer to dimer was more or less constant with an average factor of 1.96 ± 0.18 . Only the highest concentration differed significantly, most likely due to the higher overlapping area between the two peaks (factor changes to 2.03 ± 0.12 when the highest concentration is excluded). These results show that the equilibrium between both molecular species is independent of the protein concentration. To further characterize the dynamics of the monomer-dimer-equilibrium, a time course experiment was performed. Pooled fractions of yeast PDI were separated once again and a sample taken from a dimer fraction was injected onto the analytical size exclusion column at different time points after elution (Figure 3.6). Over time the monomer was formed from the dimer and after three hours the equilibration between both forms was nearly complete. To investigate the physiological relevance of the size exclusion experiments, a glutathione redox buffer was added to a sample and, as can be observed on the chromatogram of the size exclusion chromatography run, the glutathione buffer greatly accelerated re-equilibration (Fig. 3.6).

Since the glutathione buffer accelerated the re-equilibration and PDI being a redox-active enzyme, the influence of the redox status on the monomer/dimer distribution was investigated. Samples of both PDI species were kept under oxidizing conditions through the addition of the inorganic oxidant potassium ferricyanide at a concentration of 10 mM immediately after the separation. After incubation for 18 h at 4°C both species were analyzed by analytical size

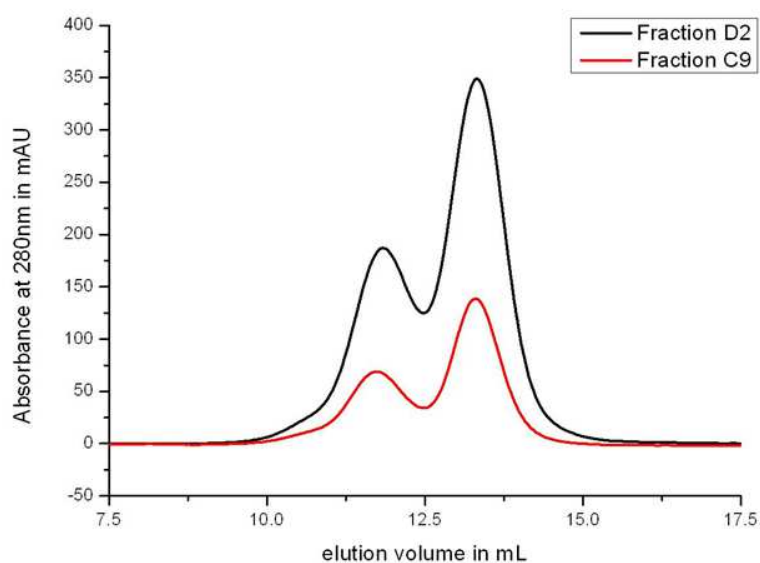


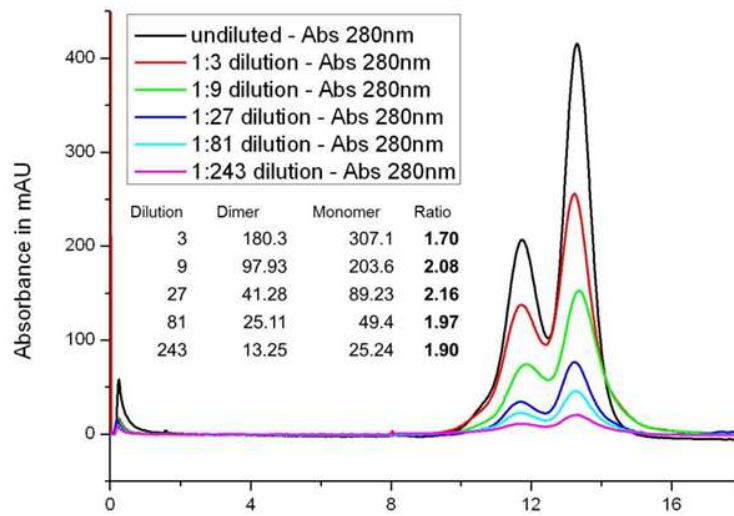
(Top) Figure 3.3 – Size exclusion chromatography of the pooled fractions 2-5 from Fig. 3.1

Despite uniformity in the SDS PAGE analysis (Fig. 3.2A), at least two molecular species could be observed when the sample was loaded onto a preparative Superdex200 HiLoad SD26/60 column. The absorbance at 280 nm yielded maxima at 154.7 mL and 177.2 mL, corresponding to apparent molecular weights of 188.8 kD and 95.5 kD, respectively.

(Below) Figure 3.4 – Analytical Size Exclusion Chromatography

Fractions corresponding to both peaks from the previous size exclusion chromatography step were incubated at 4°C overnight and injected onto a Superdex200 10/100 GL column. Again, the absorbance profile shows two major peaks (11.9 mL and 13.4 mL) for both fractions. For further experiments, fractions were taken from the ‘D2’ run as indicated.



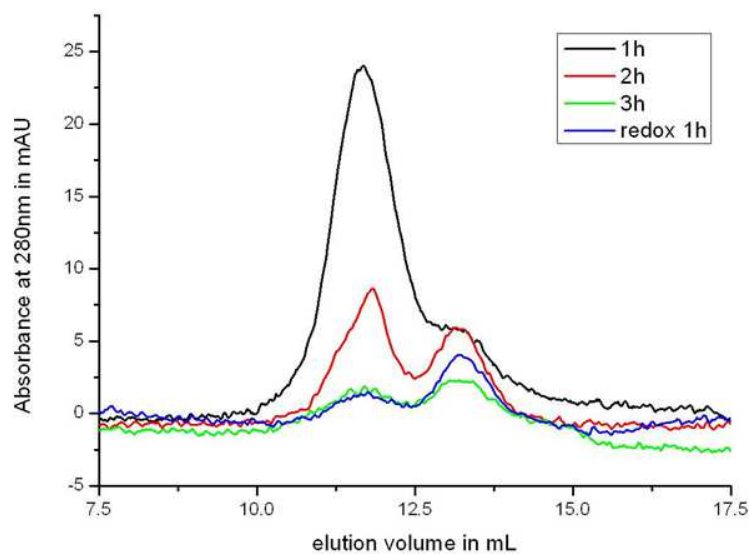


(Top) Figure 3.5 – Equilibrium formation at different concentrations

A sample taken from the 177 mL peak (Figure 3.3 – fraction D2) was diluted with sample buffer as indicated and incubated overnight at 4°C. The undiluted sample had a concentration of 270 μM. The dilution samples were analyzed on a Superdex200 10/100GL column injecting increasing volumes in order to keep the absorbance in a reasonable range for analysis resulting in volumes of: 50 μL (undiluted & 1:3); 100 μL (1:9 & 1:27); 200 μL (1:81 & 1:243)

(Below) Figure 3.6 – Time course of the equilibrium formation

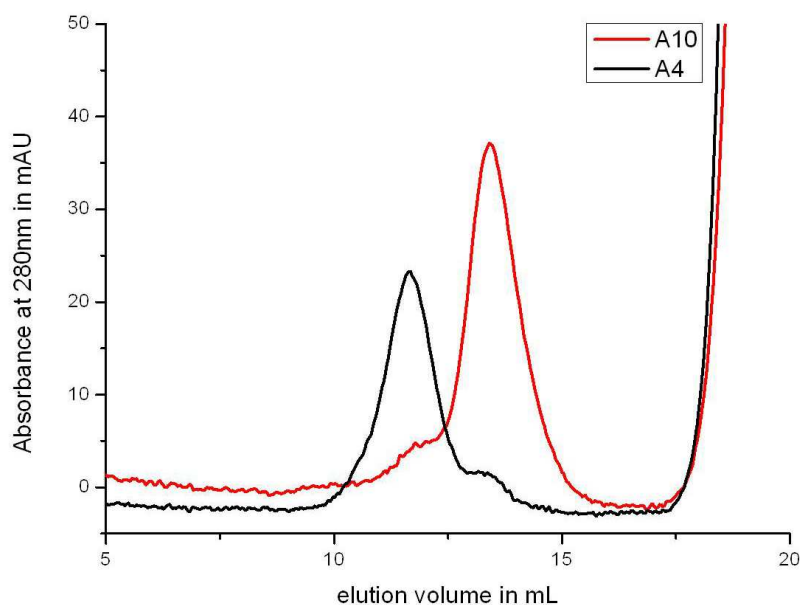
A sample eluting earlier on the analytical size exclusion chromatography (Figure 3.4 – 11.2-11.7 mL) was incubated for the indicated time at 4°C. ‘Redox’ indicated the addition of oxidized and reduced glutathione to a final concentration of 1 mM and 250 μM, respectively, prior to incubation.



exclusion chromatography. Despite the long incubation period, the dimeric species still eluted predominantly as a dimer (>95%) while the monomeric species eluted primarily as a monomer (>95%) as can be seen in Figure 3.7. In contrast, when the two PDI species were incubated with the reductant DTT (10 mM, 18 h, 4°C), both samples eluted as a monomer when subjected to size exclusion chromatography (Fig. 3.8). So while the addition of an oxidant prevents re-equilibration and locks each species in its current state, the addition of a reductant forces both species into the monomeric form.

As the initial anion exchange chromatography had shown the existence of differently charged groups, a charged-based separation of the PDI species was attempted. Purified yeast PDI was transferred into a sample buffer supplemented with 0.1% hydrogen peroxide and containing only 100 mM sodium chloride. This sample was then loaded onto a small (2 mL) Resource Q anion exchange column equilibrated with buffer A (20 mM Tris pH 8.5, 100 mM NaCl, 0.1% hydrogen peroxide). After washing with five column volumes of buffer A, the sample was eluted using a linear gradient over 60 column volumes to 100% of buffer B (20 mM Tris, 1 M NaCl, 0.1% hydrogen peroxide). As oxidants will damage the column matrix over time, the column was washed extensively immediately after the elution. PDI eluted in two major peaks corresponding at 250 mM and 350 mM of salt, respectively (Figure 3.9). Each of those peaks was pooled and concentrated separately and then subjected to size exclusion chromatography. As can be seen in Figure 3.10 the species eluting at 250 mM salt from the anion exchange column corresponds to the monomeric form, while the species eluting at 350 mM salt corresponds to the dimeric form. Thus it could be demonstrated that yeast PDI can be present in at least two distinct states. One state is either dimeric or very elongated and exposes more negatively charged residues on the surface. The other state is either monomeric or less elongated and possesses a surface with a reduced number of negatively charged residues.

In the literature, there have been many older studies mentioning and investigating the dual state behavior of PDI when subjected to size exclusion chromatography. Depending on the study, the two states have been interpreted to be full-length PDI and a truncated form [80], a monomer and a dimer [81] or a dimer and a tetramer [82]. A more recent publication attributed the occurrence of a monomer-dimer-equilibrium to the presence of Zn^{2+} cations [83]. The zinc is supposed to coordinate the cysteine thiols present in the b'domain of two different protein disulfide

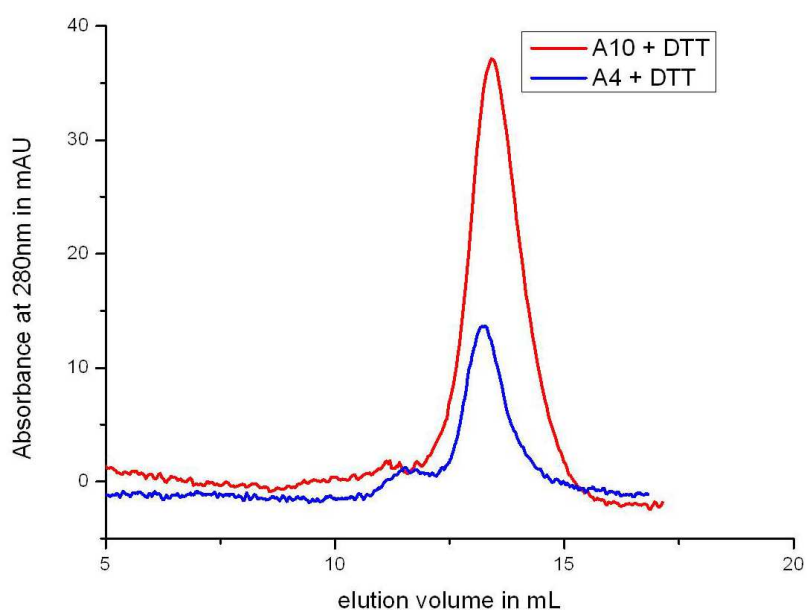


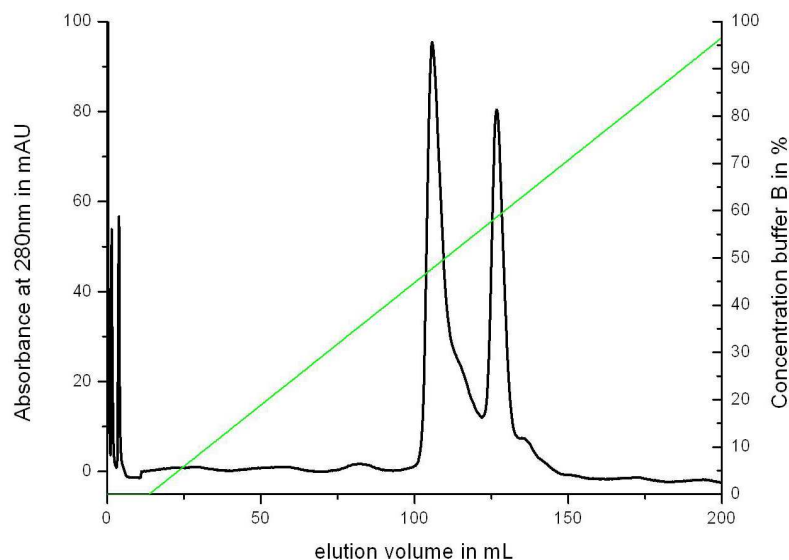
(Top) Figure 3.7 – Arresting the equilibration under oxidizing conditions

Using fractions taken from the analytical separation of the two PDI species (see Fig. 3.4), samples were supplemented with the inorganic and thiol-free oxidant potassium ferricyanide ($K_3[Fe(CN)_6]$) to a final concentration of 20 mM. Under this oxidizing condition, the samples were incubated at 4°C overnight and then analyzed on a Superdex200 10/100GL analytical size exclusion column.

(Below) Figure 3.8 – Effects of reductant on the equilibrium

Using fractions taken from the analytical separation of the two PDI species (see Fig. 3.4), samples were supplemented with the reductant dithiothreitol (DTT) to a final concentration of 20 mM. Under this reducing condition, the samples were incubated at 4°C overnight and then analyzed on a Superdex200 10/100GL analytical size exclusion column.



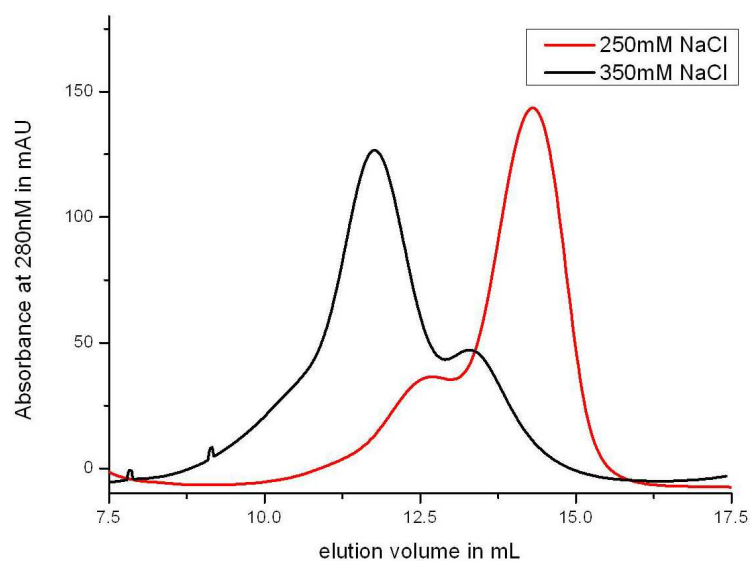


(Top) Figure 3.9 – Separation of PDI species by surface charge

Yeast PDI was loaded on a Resource Q 2mL anion exchange chromatography column and eluted using a linear gradient of salt (Buffer A: 100 mM NaCl, 20 mM Tris pH 8.5; Buffer B: 500 mM NaCl, 20 mM Tris pH 8.5) and a slow flow rate of 0.5 mL/min. Two distinct species eluted at approximately 250 mM NaCl and 350 mM NaCl, respectively. Fractions were taken from the absorbance peaks, concentrated using a Microcon concentration device (MWCO: 10 kD) and analyzed by size exclusion chromatography.

(Below) Figure 3.10 – Size exclusion analysis of charge-separated samples

As describes above, samples eluting at different salt concentrations during the anion exchange chromatography were injected onto a Superdex200 10/100GL after a concentration step. The two species differed significantly in their elution profile.



isomerase molecules to create a non-covalently linked dimer. So while the existence of multiple states has been described, the relevance of this observation remains contested. The existence of a truncation can also be excluded as both observed forms can interconvert into each other. When the elution volumes of both PDI species were compared to a calibration curve of globular proteins of known molecular weight, the two species revealed apparent molecular weights of 95.5 kD and 188.8 kD respectively. A dimer of yeast PDI has a calculated molecular weight of 114 kD and thus the observed apparent weight of the smaller species is too small for a dimer, even if the dimer of PDI would assume a perfectly globular shape. As there are no reports of an interaction of PDI with size exclusion chromatography materials, it makes the hypothesis of a dimer-tetramer-equilibrium highly unlikely. The currently favored hypothesis of a zinc dependent dimerization can be excluded, as it has been shown to rely on a pair of cysteine residues located in the b' domain of PDI. These two cysteines, while not involved in catalysis, were suggested to coordinate a Zn^{2+} ion with the cysteines of a reduced active site of another PDI molecule [83]. However, protein disulfide isomerase from *S. cerevisiae* used in these experiments does not contain this crucial pair of amino acids. The only cysteine residues apart from the active site are found in the a domain where they form a structural disulfide bond as could be seen in the crystal structure and by Ellman's assay [63]. The same crystal structure was solved as a tetramer, but with very small protein-protein-contacts so that the tetrameric state was concluded to be a crystallization artifact. However, the second crystal structure of yeast PDI was solved as a dimer (Fig 3.10) and in this case the two PDI molecules shared approximately 2700 Å of interface surface. The interface of this dimer occludes both the hydrophobic cleft of the b' domain and involves the surface around the active site of the a' domain. Therefore, it was suggested that the dimeric state represents an inactive or storage form [64].

The monomer-dimer-equilibrium is therefore the favored hypothesis, as both the dimer and the monomer have been shown to exist *in vivo* and *in vitro*. The apparent weight derived from the size exclusion chromatography is not in disagreement with a monomer and a dimer, as the protein deviates from a spherical form and becomes elongated and thus yields an increased apparent molecular mass. This is in line with results obtained by sedimentation velocity experiments in the literature [84] that determined an axial ratio of 5.7 for a monomer of rat PDI. As it has been shown, the formation of the monomer-dimer-equilibrium is dependent on the redox environment of the proteins. With two redox active cysteine pairs present in PDI, the

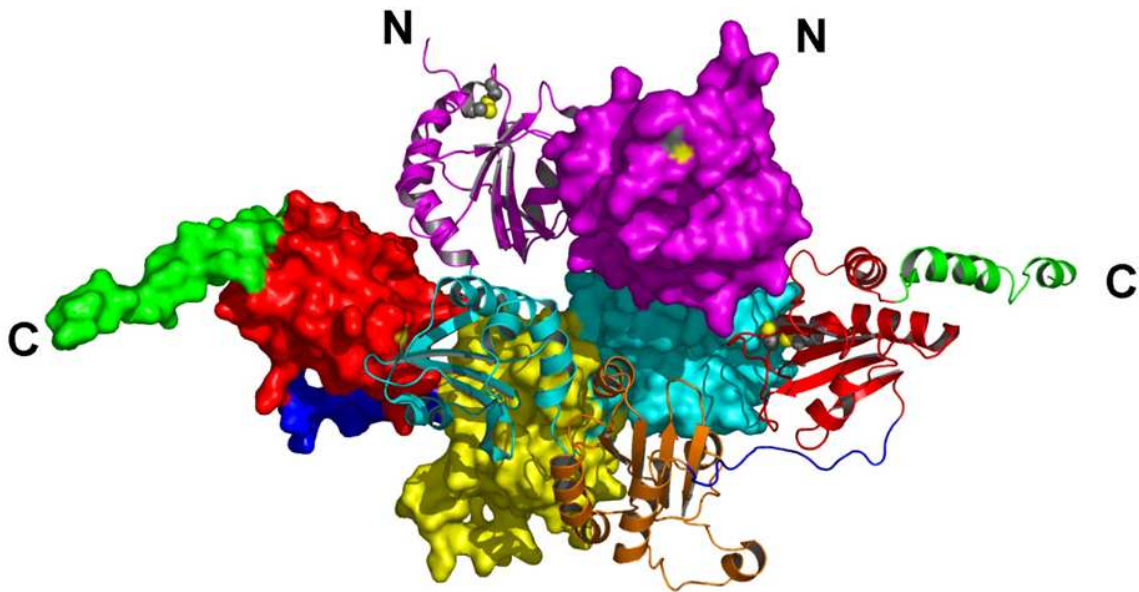


Figure 3.11 – Structure of a yeast PDI dimer

Crystals grown at 22°C diffracted in a different space group (C222₁ compared to the I4 of the earlier structure). In this structure PDI forms a dimer with an interface of 2700Å. Most of this surface is split evenly between the b, b' and a' domain. While the active site of both a domains are still solvent exposed and therefore potentially functional, the active site of the a' domains are buried into the outward surface of the b domain of the respective partner molecule. In addition to the removal of the a' domain active site from the solvent, the access to the hydrophobic cleft of the b' domain is also occluded by the interaction.

redox dependency suggests a disulfide linked dimer. It has been shown [85], that the redox potential of the a domain active site (-188 mV) differs from the redox potential of the a' domain active site (-152 mV). Also the crystal structure of yeast PDI shows the a' domain in a reduced state while the active site of the a domain is mainly in the oxidized state with only 20% being reduced [63]. This is the redox equilibrium PDI reaches after reduction in the absence of further reductants. Thus, when two PDI molecules engage each other, the reduced active site of the a' domain can attack the disulfide bond of the a domain of the second molecule forming an intermolecular disulfide bond. The most likely candidate for such a nucleophilic attack is the N terminal cysteine of the active site as it has been shown [86-88] that the pK_A of this residue is reduced, and it is therefore more likely to form the reactive thiolate. Under the experimental condition we see rather slow interconversion taking several hours to reach equilibrium. Yet, under physiological conditions this should be a short lived link, as it leaves two reduced

cysteines in close proximity ready to attack the intermolecular bond for another rearrangement. This would trigger an attack by the C terminal cysteine of the active site in the same fashion as it disengages PDI from an isomerization substrate. This is in line with the results of the equilibration in the presence of a redox buffer where additional free thiols result in a greater chance of a nucleophilic attack on the intermolecular disulfide. Under reducing conditions, no disulfide can form and thus the complete protein population is present in the monomeric state. In contrast, oxidation would not shift the population to the dimeric state. With all active sites in the oxidized form, there is no reduced cysteine present to initiate a disulfide exchange between molecules and thus all the monomers will stay monomers. Likewise, dimers will remain in the dimeric state as the two free cysteines which would otherwise attack the intermolecular disulfide are in close proximity and thus are likely to form an additional disulfide. With no free thiols present in the solution, no interconversion can take place just as observed experimentally. The only problem with the monomer-dimer-hypothesis is the static monomer dimer ratio over a large range of PDI concentration. In standard monomer dimer dynamics as well as in the outlined model, an increase in concentration should result in an increase in chance encounters of two PDI molecules in solution and thus to an increase in the dimer over the monomer. Respectively a decrease in PDI concentration should result in an increase of the monomer over the dimer. This dimer would also have to adopt a radically different conformation as the one solved by X-ray crystallography (Fig. 3.11 – taken from [64]) as it needs a direct contact of the active sites of two different molecules.

It is also possible to explain the results through two different protein conformations. The different species on the size exclusion chromatography could be the result of different shapes instead of different sizes. The difference in surface charge that enables the separation through anion exchange chromatography could be a result of different intramolecular interfaces between the domains of PDI exposing and burying different charged residues in their two conformations. In prior publications, PDI has already been shown to be a flexible protein prone to large movements of its domains [64], and, recently, conformational changes of a single domain have been documented [89] which extend to the x-linker and the b' domain. Different conformations of a monomer would explain the similar ratios of both species over a wide concentration range as the distribution of two conformations would not necessarily be concentration-dependent. The active site of the thioredoxin fold is located at the beginning of an α -helix and the active site

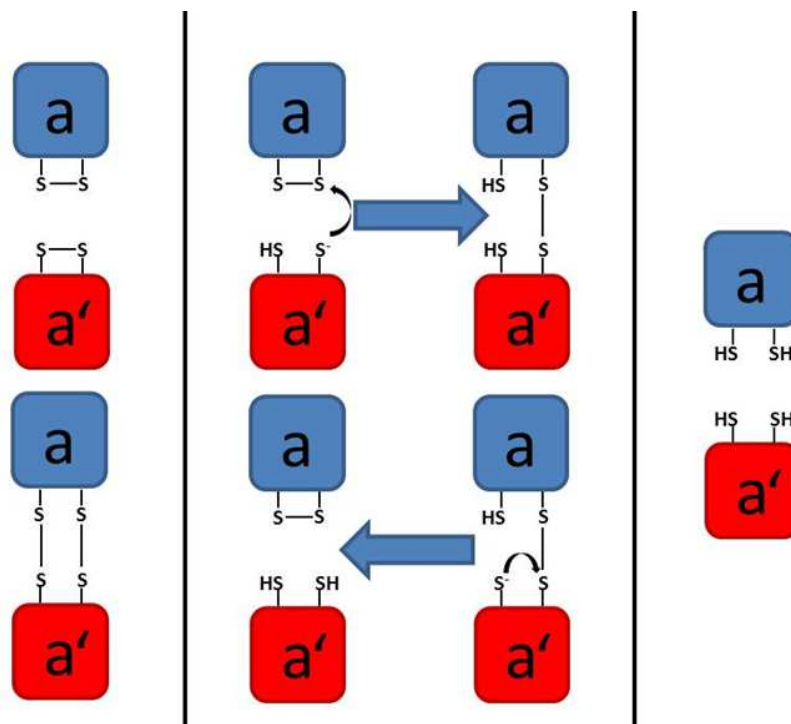


Figure 3.12 – Suggested model for the monomer-dimer-equilibrium.

Middle panel: The active sites of two different domains (blue and red, respectively) can interact with each other due to their different redox states. A cysteine in the thiolate form performs a nucleophilic attack on the disulfide bonded cysteine of the a domain. Likewise, the dimer can be disbanded by another nucleophilic attack, thus restoring both active sites to their monomeric form.

Right panel: Under reducing conditions, all active sites are present in the reduced state and therefore no linkage between two active site domains can occur.

Left panel: Under oxidatizing conditions, there are no free thiols present which could initiate a disulfide exchange reaction. Therefore the separated form stays separated (top), while the disulfide-bonded species stays linked (bottom).

disulfide bridges the helix and the preceding loop. Thus, it is possible that a change in redox status of the active site results in a conformational change of the corresponding active site domain. Such a conformational change of a single domain could be translated via the domain interfaces to the whole protein, resulting in an altered shape and thus the apparent molecular weight on a size exclusion chromatography differs. It has been shown that the redox status actually changes the conformation of the a' domain of human PDI [89], thereby controlling its activity as a chaperone. Attributing the different apparent size to conformational changes has the great advantage that it explains why the ratio of the two species stays constant over a large concentration range. Conformational changes within a single molecule are not influenced by its

concentration (excluding dimerization). Yet there are also inherent problems with this hypothesis. First of all a monomer running at nearly four times the apparent molecular weight has so far not been described in the literature. The elution volume in a size exchange chromatography has been shown to correlate with the radius of gyration of a molecule [90]. This property depends on both the size and the shape of a given sample. In order to estimate the potential range of the radii of gyration for a given molecular weight, I compared the globular carbonic anhydrase (29kD) with a stretch of collagen triple helix. While the radius of gyration for the carbonic anhydrase has been reported at 2.08 ± 0.03 nm [91], the radius of a 30 kD stretch of collagen triple helix was calculated to be 9.7 nm. The calculation was done with HydroPro and based on the available structure 1CAG, of which four were fused head to tail to yield a 30 kD fragment. Thus while it is possible for monomers to run at more than four times their actual weight in a size exclusion chromatography, it would require an extremely stretched conformation. Also the observed rate of equilibration is very slow, hinting at the formation or breaking of covalent bonds. Finally, a change in redox state should shift PDI from one form to the other. While this is the case for a treatment with reductant, an oxidization of the protein does not shift every protein to the second state but only inhibits the formation of the equilibrium.

It might be possible for the two active sites of one protein to interact with each other via the same mechanism described for the monomer-dimer-hypothesis. In this case the behavior under different redox conditions can be explained by the disulfide exchange between the active sites. Also the constant ratio of both species at varying concentrations is not in disagreement, as the two species would indeed be two different conformations of the protein. It would also explain the extreme shifts in the apparent weight as the disulfide bonded state forces PDI to adopt a compact ring-like arrangement with its four domains, while the non-bonded state could be much more elongated. This compact arrangement would be similar in shape to the existing crystal structure of yeast PDI where both active sites are facing each other. For the first structure of yeast PDI (2B5E), the radius of gyration was calculated to be 3.04 nm, which is larger than the radius of the 66 kD bovine serum albumin (BSA - 2.99 nm [91]). Thus, even the ring-like arrangement would most likely run at the observed molecular weight. However, it is unclear, if the protein is flexible enough to allow both active site domains to interact without sterical hindrance and conformational strain. A mechanism excluding intermolecular interactions would also be

required as otherwise an increase in concentration would increase the chance of active sites interacting with each other, again resulting in a shift in the ratio of both species.

In conclusion, the two different species are both PDI and are most likely a result of a disulfide exchange between its active sites. With its dependency on the redox environment it is likely that the equilibrium between both species has a regulatory role to either control the activity of PDI, or to change the substrate selectivity. However, further experiments are required to determine the mechanisms underlying this dynamic. With the separation of both species by charge and under oxidizing conditions it should be possible to isolate and lock both species and subject them to mass spectrometry for a determination of the molecular mass. Unfortunately, it will be extremely difficult to assess the catalytic activities of each species as the isolation requires a strongly oxidizing environment which would interfere with the redox activity assays. The chaperone activity, however, could still be tested using appropriate assays like the refolding assay utilizing glyceraldehyde-3-phosphate-dehydrogenase (GAPDH) [92].

3.2 Crystallization of PDI family members

At the time this thesis was written, only two structures of PDI family had been solved without the truncation of domains. The first structure published was protein disulfide isomerase from *S. cerevisiae* [63]. It was previously known from secondary structure prediction and structural studies of single domains that this protein consists of four thioredoxin fold domains two of which contain an active site. While PDI had been suggested to be linear [83], the crystal structure showed the arrangement of these four domains in the shape of a twisted 'U', with the two active sites facing each other over the hydrophobic surface of the b' domain. This non-linear shape was maintained, despite large-scale domain movements, in a second structure of the yeast enzyme [64]. The second protein solved in full-length is ERp57. This protein could be crystallized in complex with tapasin and showed a similar 'U' shape as the first PDI structure, even in complex with a substrate [68]. As it is still unclear, how PDI family member differ from each other, new structures are greatly anticipated.

As there are two different structures of yeast PDI, each showing a different conformation although the crystals were grown with the same precipitant, yeast PDI was crystallized to screen for additional conformations. The published crystallization conditions for full-length yeast PDI were reported to be 28-30% PEG 2000 MME, 200-500 mM magnesium chloride and 100 mM sodium cacodylate buffer (pH=6.5) using hanging drop vapor diffusion with 1 μ L of protein (30 mg/mL) and 1 μ L of precipitant [63]. Thus, hanging drop crystallization experiments were set up with the precipitant concentration extended from 26 to 32% PEG 2000 MME, 100-600 mM magnesium chloride at 4°C and 20°C. While most conditions yielded a clear solution or a separation into two liquid phases, several setups resulted in protein crystals. Unfortunately, the resulting crystals were not very large and most were not uniform (Figure 3.12 A-C). In testing at the in-house X-ray generator the crystals diffracted only weakly. When diffraction experiments were carried out at the BESSY synchrotron in Berlin, the crystals only diffracted up to a resolution of 6 Å. In an effort to increase crystal size and uniformity, a second screen was performed with a smaller range of precipitant conditions and selected additives as instructed by Dr. Geng Tian. Indeed the crystal size increased markedly, especially with barium chloride as an additive (Figure 3.12 D). Although the crystals were still not uniform, they had uniform parts that could be selectively exposed at the synchrotron. Despite the increase in crystal quality, these

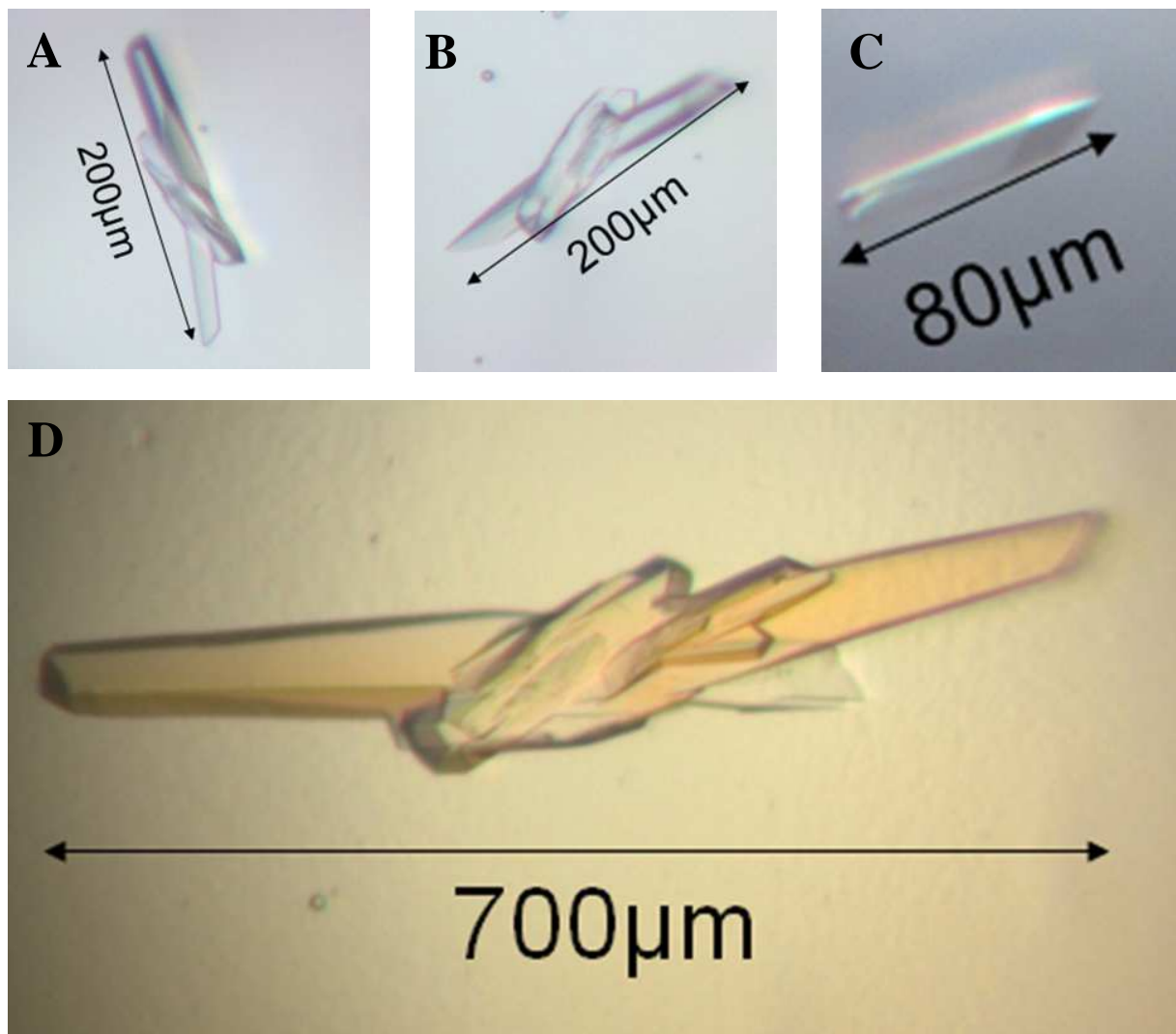


Figure 3.13 – Protein crystals of PDI from S. cerevisiae

A+B: Crystals grown from 30% PEG 2000; 300 mM MgCl₂; 30 mg/mL protein; 1 µL protein + 1 µL precipitant after five weeks at 22°C.

C: A mostly single crystals grown from 30% PEG 2000; 200 mM MgCl₂; 30 mg/mL protein; 1µL protein + 1µL precipitant after five weeks at 22°C.

D: Larger crystals with clean ends appeared after three months at 22°C from 28% PEG 2000; 300 mM MgCl₂; 10 mM BaCl₂; 30 mg/mL protein; 1 µL protein + 1 µL precipitant.

crystals only diffracted up to 4.8 Å. Due to the extremely long growth period of more than three months, the crystallization of yeast PDI was abandoned in favor of other not yet crystallized family members.

Together with yeast PDI, PDI from *H. sapiens* and ERp57 from *Gallus gallus* were set up for crystallization. Due to the similarity of both proteins to yeast PDI, the screening was performed with the honeybee pipetting robot in a sitting drop setup at a similar protein concentration of 500 µM. Unfortunately, the initial screening including the Index screen, the Wizard Screens 1-5, the Crystal Clear screen and the Nextal PEG screen yielded no lead conditions. With the optimization of the purification of human PDI, larger quantities of protein became available and the crystallization attempts were resumed. A rule of thumb for the correct protein concentration in crystallization is a ratio of approximately 50% precipitation in the Index crystallization screen conditions after 24 h. When the Index screen was performed at a concentration of 40 mg/mL human PDI (700 µM), only 22% (21 out of 96) conditions yielded a precipitate. The concentration of protein was therefore increased to 85 mg/mL (1.5 mM), but still only 38% (37 out of 96) of the tested conditions yielded precipitation. At a concentration of 300 mg/mL (5.25 mM) the Centricon concentration device was unable to further concentrate the protein solution, yet human PDI still remained in solution. When the Index screen was performed at this concentration, the 50% rule was met with 52% of the conditions (50 out of 96) yielding precipitates, however, no protein crystals were observed. Therefore, additional crystallization screens were performed in a sitting drop setup with the honeybee crystallization robot: Optimix 1/2/3/5/PEG, Protein Complex screen and Nextal PEG. After four weeks of observation, a single condition with potential microcrystals could be observed in the Optimix 5 screen (Figure 3.13 A). Starting from this condition, a hanging drop fine screen was set up to find suitable conditions for optimization. In the initial crystallization range (100 mM MES pH5.5-6.5; 12-18% PEG 5000 MME; 50-300 mM CaCl₂) several crystalline objects were observed (Figure 3.13 B+C). Unfortunately, these crystals could not be tested, as the PEG precipitant itself falls out of solution and forms a thick rubbery layer in which the crystals are embedded. Additional fine screens were performed around this condition with no improvement, even when set up as sitting drop vapor diffusion crystallization or microcapillary diffusion crystallization. Also the Additive screen from Hampton Research yielded no improved results.

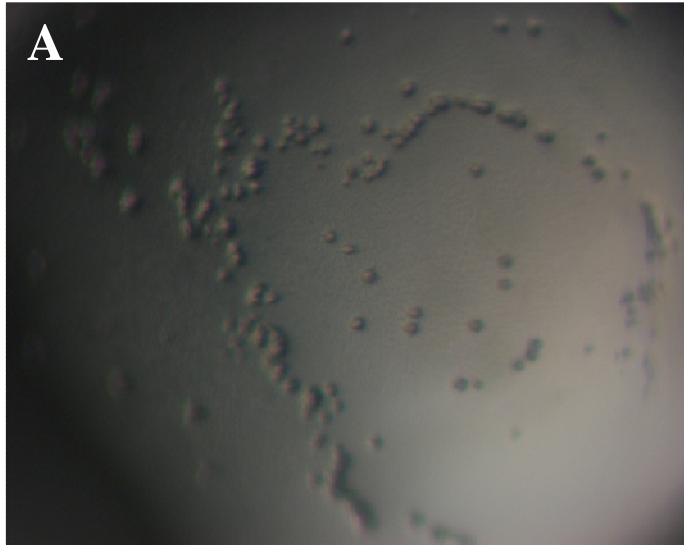


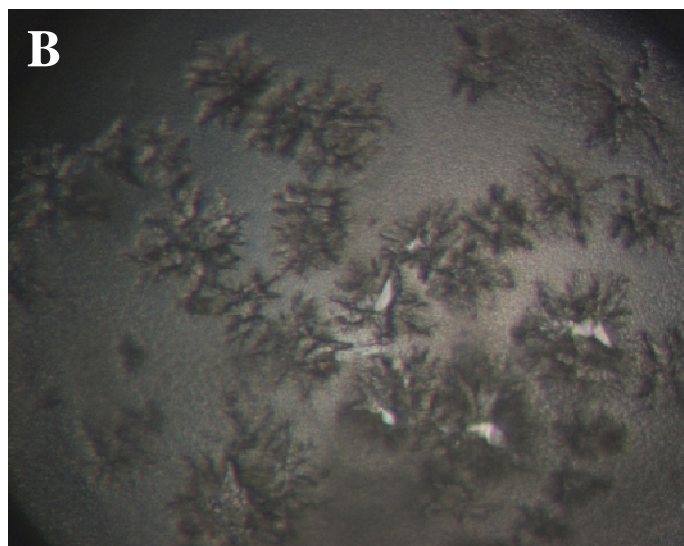
Figure 3.14 – Crystallization of human protein disulfide isomerase

A: Initial condition with crystalline objects

Optimix 5 – E9

100 mM	MES pH 5.5
150 mM	CaCl ₂
15%	PEG 5000 MME

Protein concentration: 300 mg/mL
 0.3 μL protein + 0.3 μL precipitant
 Four weeks of growth

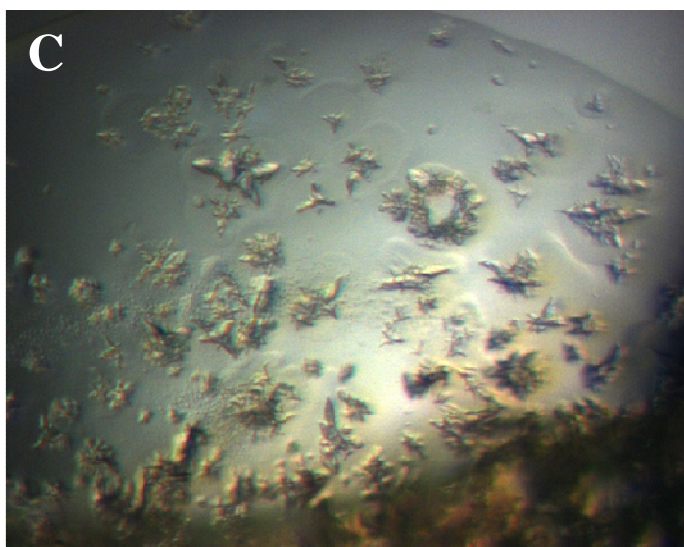


B: Crystals in fine screen 1

Fine screen – A1

100 mM	MES pH 5.5
50 mM	CaCl ₂
12%	PEG 5000 MME

Protein concentration: 300 mg/mL
 1 μL protein + 1 μL precipitant
 Three weeks of growth



C: Crystals in fine screen 2

Fine screen 2 – A5

100 mM	MES pH 5.5
250 mM	CaCl ₂
12%	PEG 5000 MME

Protein concentration: 300 mg/mL
 2 μL protein + 2 μL precipitant
 Six weeks of growth

With all further efforts leading into dead ends, protein modification was considered. Protein-protein-interfaces are a necessity for the formation of proteins crystals. In an ordered crystal, the protein molecules form a regular lattice as they are oriented in the same way by crystal contacts, which usually involve surface contacts between protein molecules. It has been shown that the large and charged side chains of glutamate and lysine are both predominantly present on the surface [93], yet at the same time they are disfavored in protein-protein-interfaces [94]. Therefore replacing these charged and flexible residues with alanine – which has a small and low-entropy side chain but does not introduce an increase in main chain flexibility like glycine – can establish surface areas suitable for crystal contacts [95]. Indeed this mutational approach could be shown to increase the likelihood of crystallization and the quality of the structural data obtained [96]. A surface entropy reduction analysis was performed with the sequence of human PDI and the folding model of yeast PDI as known from the crystal structure. With this input, the web-based service SERp (Surface Entropy Reduction prediction located at <http://services.mbi.ucla.edu/SER/>) suggested several regions for mutation (Fig. 3.14). This type of surface mutation remains an option for further crystallization attempts, but due to time constraints the crystallization of human PDI within the scope of this project was discontinued.

Proposed Mutations:



Figure 3.15 – Proposed mutations for surface entropy reduction

With the sequence of human PDI as input, the algorithm proposed several regions for mutation. Displayed above are the two top results with a final score of 7.0 or greater. These residues are calculated to increase the surface entropy of human PDI and are not very conserved. In an alignment with the sequence of yeast PDI, the side chains of the corresponding residues were located at the protein surface.

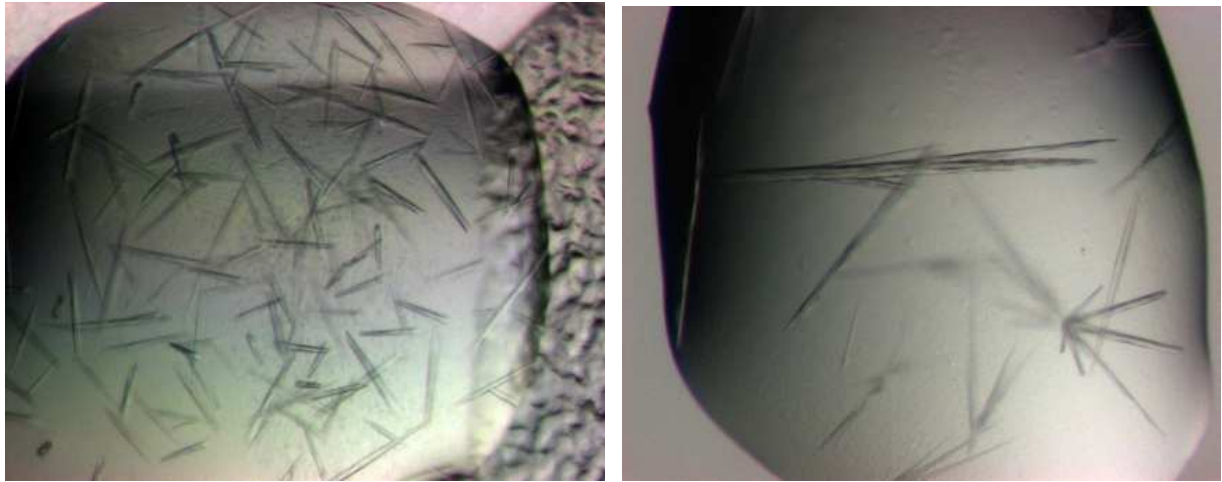


Figure 3.16 – ERp27 crystallization

A: Crystal Clear screen G2

100 mM MES pH=6.5; 200 mM ammonium sulfate; 30% PEG 5000 MME

B: Optimix 1screen E7

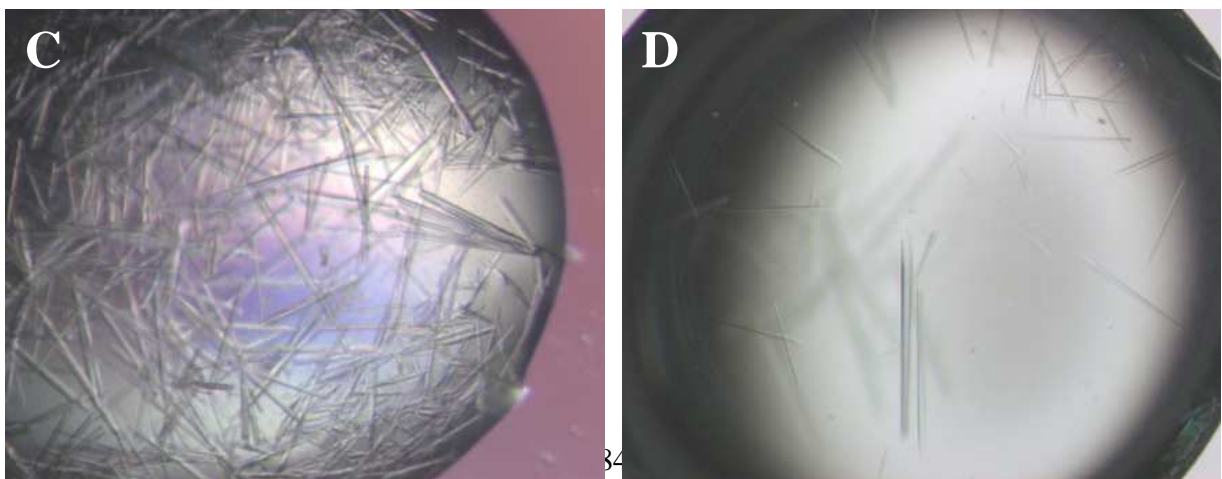
100 mM Tris pH=8.5; 100 mM MgCl₂; 25% PEG3350

C: Fine screen 1 B5 (see next page)

100 mM Tris pH=8.5; 50 mM MgCl₂; 25% PEG 3350

D: Fine screen 2 with micro seeding

100 mM Tris pH=8.5; 150 mM MgCl₂; 22.5% PEG 3350



Within the context of this project, large protein amounts of other human PDI family members were generated. The initial crystallization screens Index, Wizard 1+2, the Crystal Clear screen and the Optimix screens 1-3 were performed with PDIp, ERp46, ERp27, ERp18 and PDIr as sitting drop vapor diffusion setups at 20°C with typical protein concentrations of slightly above 1 mM. Out of all those crystallization attempts, only ERp27 produced crystals during the screening process. ERp27, a protein predicted to consist of two thioredoxin fold domains but devoid of active site cysteines, crystallized initially in three conditions (Fig 3.15 A+B). All conditions were at near neutral pH (6.5 and 8.5) with polyethylene glycol (PEG 3350 and PEG 5000 MME) as precipitant. Fine screens were designed around both the Optimix1 E7 condition (0-200 mM MgCl₂, 20-30% PEG 3350, 100 mM Tris pH 8.5) and the Crystal Clear G2 condition (100-400 mM ammonium sulfate, 25-35% PEG5000 MME, 100 mM MES pH 6.5). Each fine screen was set up as hanging drop vapor diffusion experiments with 1 μL of 1.2 mM ERp27 and 1 μL of precipitant and kept at 20°C. While no crystallization could be observed for conditions derived from the Crystal Clear G2 condition, the Optimix1 E7 condition translated well into the hanging drop method with many conditions yielding crystals. Under these conditions, ERp27 formed many yet rather thin needles which were mostly grown together or were twinned (Fig. 3.15 C). A finer second screen which varied protein concentrations and drop size gave only slight improvements. The crystals were extracted from the crystallization tray with a nylon loop and incubated in mother liquor supplemented with 15% glycerol, MPD or PEG 400 as cryo protectants for three to five minutes before flash cooling them in liquid nitrogen. Unfortunately, only about 20% of the extracted crystals yielded observable diffraction when tested (as outlined in 2.3.13). As all of the diffracting crystals had been soaked in cryo solution containing glycerol, the other two cryo protectants were disregarded in further trials with ERp27. While the crystals showed diffraction in accordance with crystalline protein, the resolution at the in-house X-ray generator was limited to 8.5 Å. In addition, no clear spots could be detected, instead the crystal diffracted with elongated smeared signals that overlapped each other (Fig. 3.16 A). When the crystals were tested at the ESRF synchrotron in Grenoble, a slightly better resolution could be achieved, but the problem with the smeared spots remained (Fig. 3.16 B). The screening process was extended to the additive screen, but the crystal quality only improved marginally with a higher fraction of the crystals giving diffraction and a slightly improved resolution.

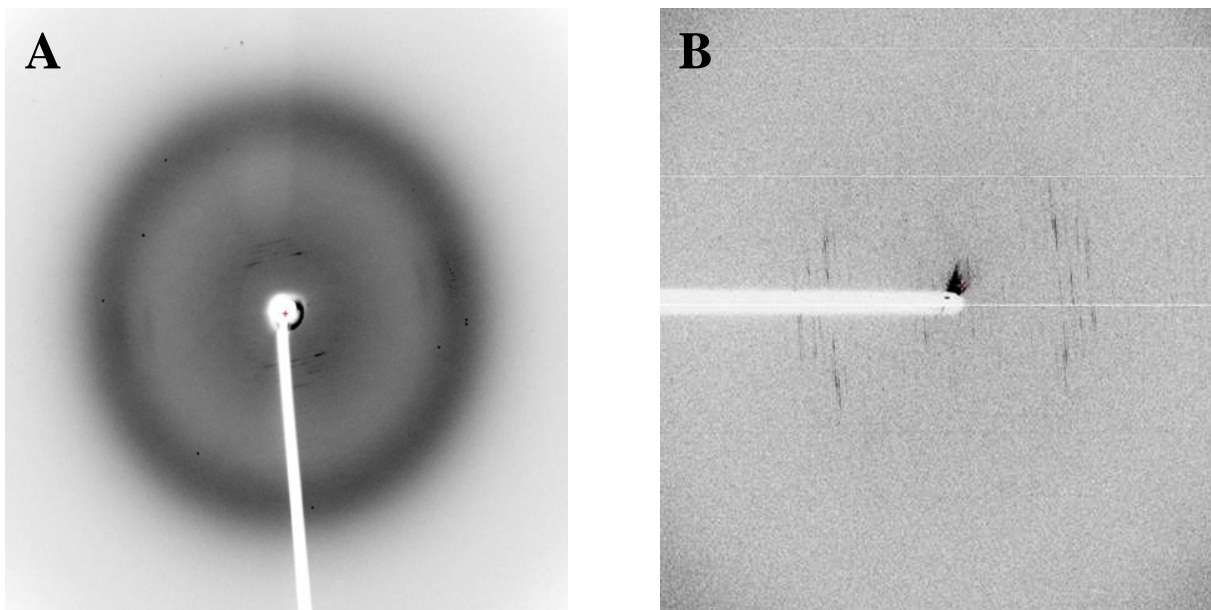


Figure 3.17 – Initial diffraction experiments with ERp27 crystals

A: Test diffraction at the in-house X-ray generator (see 2.3.13 for details)

B: Test diffraction at the ESRF synchrotron facility

Exposure time: 1 sec; Phi angle: 0.5° ; Detector distance: 200 mm; Wavelength 1 \AA

As ERp27 showed promising initial results, crystallization was continued together with the intern Wolfgang Koelmel. To further optimize the crystallization conditions protein stability was tested in different buffers. Yet neither a change in pH nor different salt concentrations could improve ERp27 stability as tested with a ThermoFluor assay. ERp27 also did not show a preference for any specific cations or anions. Crystallization at different temperatures (4°C and 37°C) did not improve the crystal quality over crystallization at 20°C . However, seeding using the crystals obtained in previous crystallization trials resulted in improved crystals. For seeding, 4-8 large crystals were extracted – preferably selecting needles that were not grown together with neighboring crystals. The crystals were transferred into a $200 \mu\text{L}$ PCR tube containing $20 \mu\text{L}$ of mother liquor. Next the tube was placed into a sonication bath and the crystals were subjected to sonication for 60 seconds. This solution was used for seeding in hanging drop crystallization setups, by touching the liquid with a pipette tip and subsequently touching the crystallization drop, transferring shards of ERp27 crystals as seed.

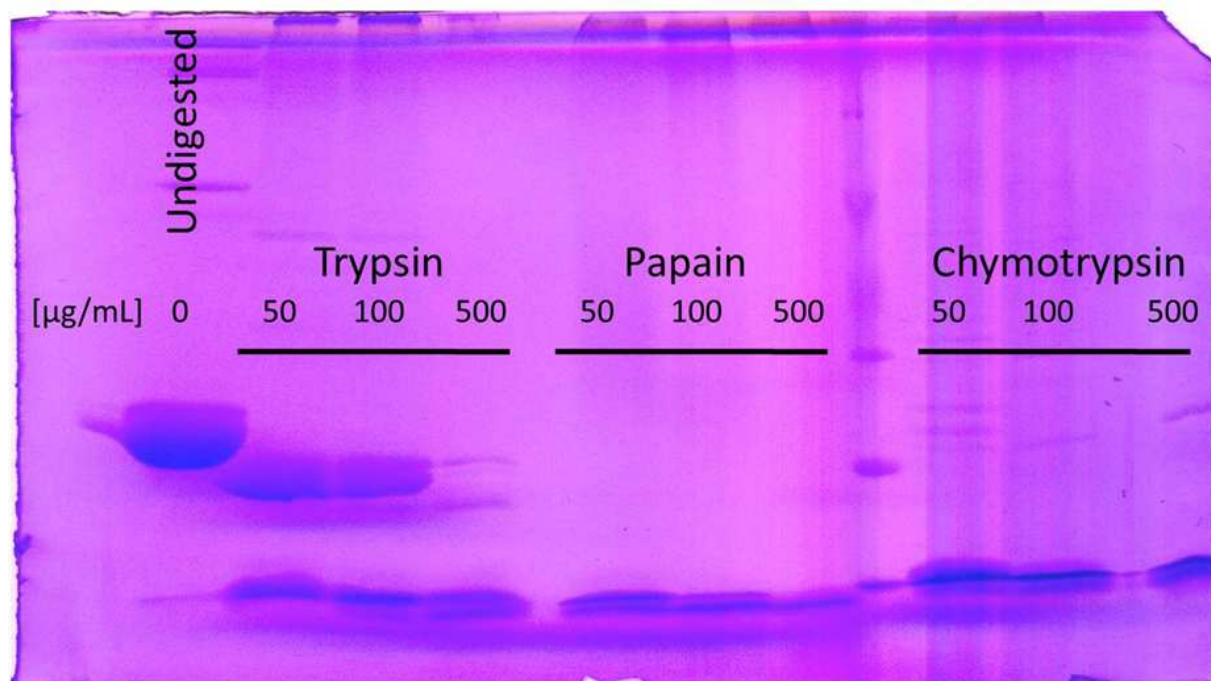


Figure 3.18 – Limited proteolysis of ERp27

ERp27 at a final concentration of 500 μ M was incubated for 20 minutes with the indicated proteases at three different concentrations. The left most lane shows undigested ERp27 as a control. While both chymotrypsin and papain rapidly degraded the protein, trypsin at lower concentrations only trimmed approximately 3 kD from ERp27.

Using this technique, the crystal number per crystallization drop decreased while simultaneously the crystal size and quality increased (Fig. 3.15 D). Unfortunately, the diffraction did not improve with the seeding and no data set could be collected. As the crystal needles were often fanning out after reaching a medium size and even the rare single crystals displayed frayed ends and growth defects, it seemed likely that flexible parts of ERp27 were disrupting the formation of a uniform crystal. To test this hypothesis, we performed limited proteolysis to trim disordered regions. ERp27, at a final concentration of 500 μ M, was incubated with varying amounts of three proteases, trypsin, chymotrypsin and papain, for 20 minutes at room temperature. The digests were analyzed by SDS PAGE (Figure 3.17). Both papain and chymotrypsin heavily degraded ERp27, leaving only a 10 kD fragment which possibly corresponds to one of the tightly folded thioredoxin domains. In contrast, a digest with trypsin initially removes only small parts of approximately 3 kD in size from the protein. At higher concentrations and over time, this stable fragment of approximately 24kD was further degraded yielding the same 10 kD fragment

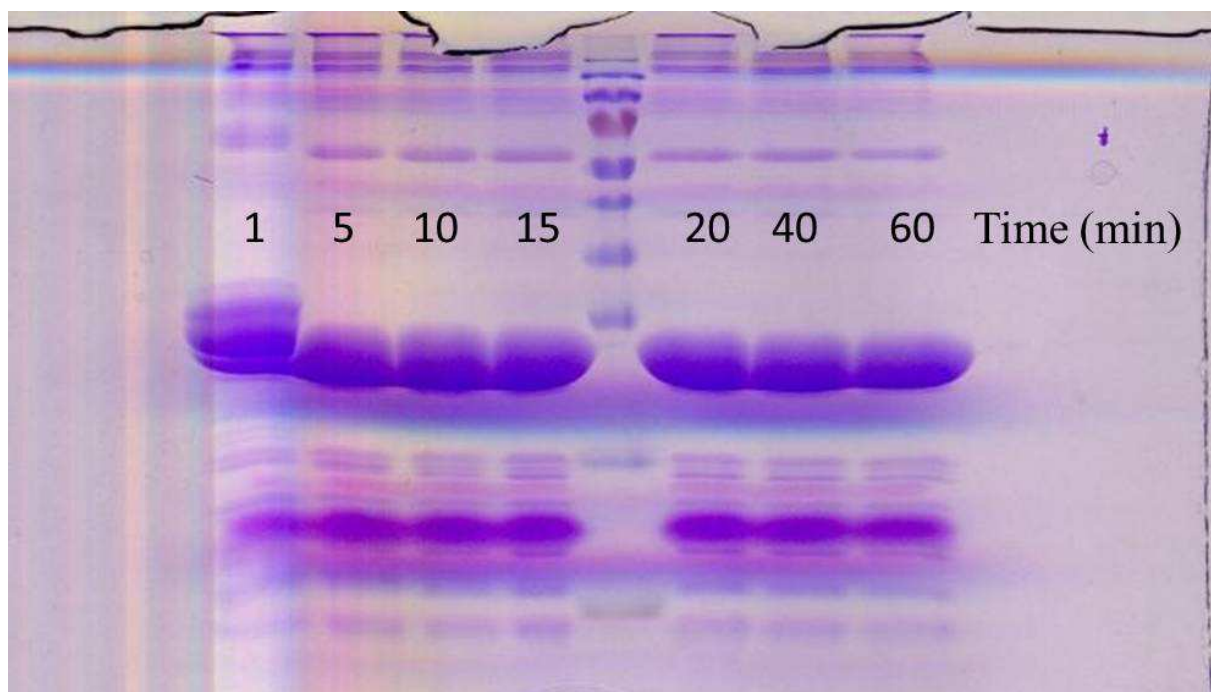


Figure 3.19 – Time course of ERp27 proteolysis by trypsin

ERp27 at a final concentration of 500 nM was incubated with 50 $\mu\text{g}/\text{mL}$ trypsin for the indicated times at room temperature. At this low concentration, the digestion time is of little consequence as the stable fragment is very resistant to further digestion. For the final digestion protocol the incubation time was set to 20 minutes.

as papain. The stable 24 kD fragment had a large enough size to encompass both thioredoxin domains. The digestion time was therefore optimized with trypsin at a concentration which only yields this fragment (see Fig. 3.18). After a large scale digest of ERp27 the tryptic fragment was purified using preparative size exclusion chromatography (HiLoad 26/60 Superdex 200 prep grade; GE Healthcare). As expected, the digested fragment eluted slightly later ($\Delta V_e=5$ mL) than undigested ERp27 (see Fig 3.19) with no observable impurities. Using this fragment, the fine screen established for ERp27 was set up. The fragment crystallized in the same conditions but yielded larger crystals of a higher quality (Fig 3.20 A). These crystals diffracted to a resolution of 3.5 \AA when tested at the in-house X-ray generator (Fig 3.20 B) and a complete data set (Exposure time: 10 minutes per frame; Phi-angle: 0.5 $^\circ$; 360 images) was collected. Initial indexing suggested the monoclinic space group C2, however, the data could not be scaled in this space group and hence the crystal were assigned to the triclinic space group P1. Unfortunately, it became apparent during data analysis, that the seemingly single crystal contained more than one

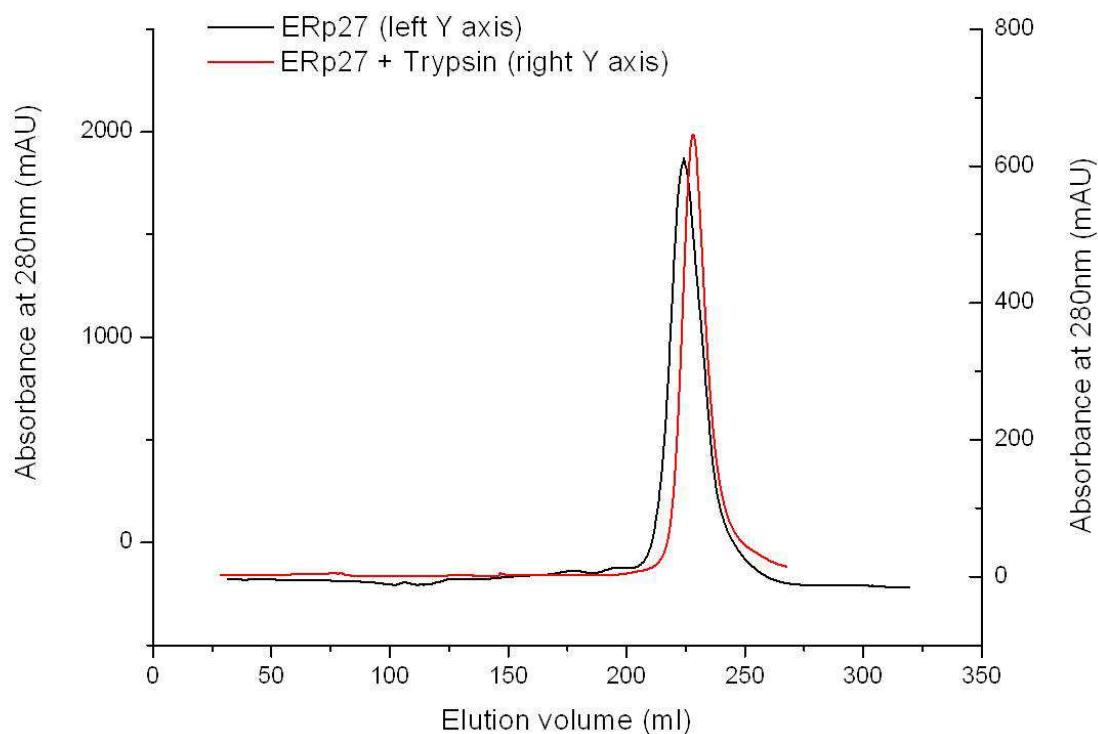


Figure 3.20 – Purification of the tryptic fragment by size exclusion chromatography

Comparison of the elution profile of full length (black, left Y axis) ERp27 and the same protein partially digested with 50 $\mu\text{g}/\text{mL}$ of trypsin (red, right Y axis). The fragment generated by proteolytic digestion elutes 5 mL later than the full-length protein and appears as a single species. As only one third of the full-length protein was subjected to proteolysis, each curve was scaled to its own axis to allow for a better comparison.

lattice which all diffracted simultaneously. After several cycles of reproduction and testing, uniform crystals were found that in testing showed only signs of a single crystal lattice. These crystals were sent to the BESSY synchrotron facility in Berlin, where they diffracted to a resolution of 2.6-2.8 \AA . Two complete data sets were recorded (Exposure time: 1 second per frame; Phi-angle: 0.5° ; 720 images). The images were indexed with iMOSFLM in the space group P1 and scaled with SCALA from the CCP4 program suite. During indexing and scaling it became obvious, that each dataset contained data from multiple slightly misaligned crystal lattices. Using MOLREP and Phaser (as implemented in the CCP4 package) it was attempted to solve the phase problem by molecular replacement. Numerous models were tested in molecular replacement calculations: the existing NMR structure of the N-terminal thioredoxin fold domain

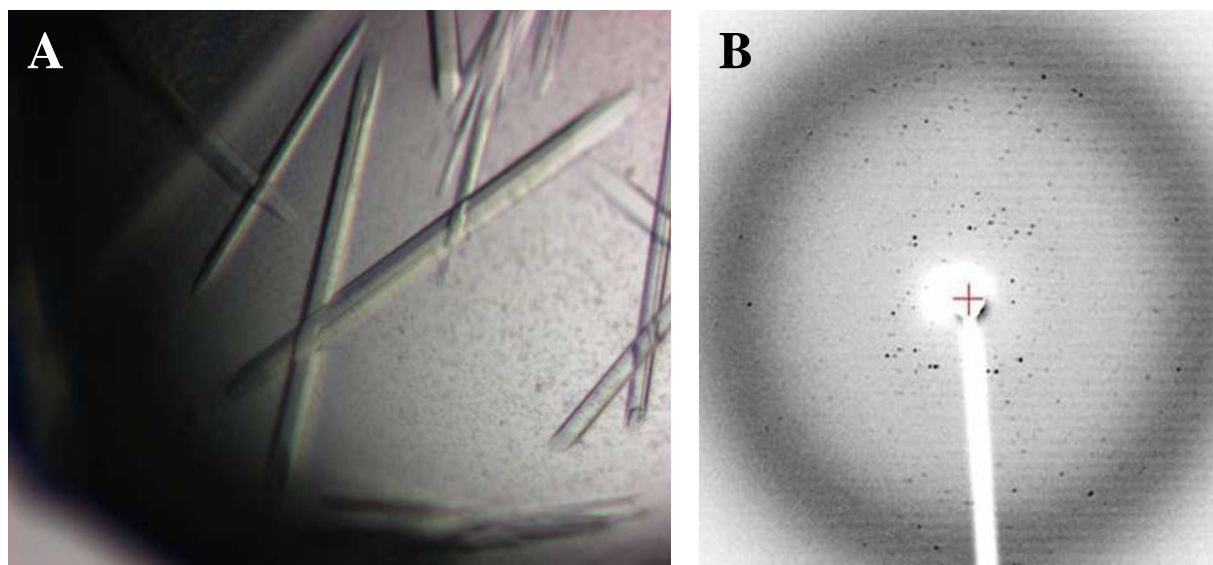


Figure 3.21 – Crystallization of a tryptic fragment of ERp27

- A: Shortened ERp27 crystallizes under similar conditions as the full-length variant, but yields larger and more uniform crystals. Yet there are still growth defects visible as a single needle apparently consists of multiple smaller parallel needles.
- B: Test diffraction taken at the in-house X-ray generator (Exposure time: 20 minutes; Phi angle 0.5°; Detector distance: 240mm)

of ERp27 and the bb' domain fragments of yeast PDI, human PDI and human ERp57. Yet despite this multitude of molecular replacement templates, no suitable matches could be made. Unable to solve the structure by molecular replacement it was attempted to soak ERp27 crystals with heavy metal atoms to solve the phase problem by multiple isomorphous replacement. However, when ERp27 crystals were placed in mother liquor containing heavy metal salts, they shattered and dissolved within a minute. Therefore seleno-methionine substituted ERp27 was expressed using M9 minimal medium supplemented at induction with seleno-methionine. The purification of seleno-methionine ERp27 was analogous to the purification of regular ERp27, as was the protocol for the tryptic digestion. The seleno-methionine protein crystallized under the same conditions as regular ERp27, although the resulting crystals diffracted X-rays only poorly. The crystallization conditions were optimized by Wolfgang Koelmel by screening for different cations. By replacing magnesium chloride with sodium chloride, he was able to create selenium-containing crystals that diffracted up to a resolution of 2.8 Å at the BESSY synchrotron in Berlin. Multiple data sets were recorded at a wavelength of 1 Å, which is very close to the

Native dataset	
Spacegroup	P1
Unit cell dimensions (Å, °)	a = 57.83 b = 68.36 c = 86.78 α = 70.45 β = 88.16 γ = 64.96
Unique reflections	54,973
Resolution limits (Å)	57.8–2.2
Completeness (highest shell)	96.2 (95.4)
Multiplicity	5.4 (5.4)
R _{sym} / R _{pim} (highest shell)	0.113 (0.836) / 0.063 (0.394)
Mean <l/σl> (highest shell)	11.5 (2.0)
SeMet dataset	
Spacegroup	P1
Unit cell dimensions (Å, °)	a = 61.04 b = 63.94 c = 105.07 α = 90.59 β = 106.81 γ = 118.65
Unique reflections	31,863
Resolution limits (Å)	44.9–2.8
Completeness (highest shell)	98.5 (98.5)
Multiplicity	3.9 (3.9)
R _{sym} / R _{pim} (highest shell)	0.176 (0.700) / 0.103 (0.408)
Mean <l/σl> (highest shell)	7.5 (2.0)
Phasing statistics	
Number of sites	24
Figure of merit	0.371
Refinement	
Resolution limits (Å)	41.7–2.2
Number of working/test reflections	53,561/1,371
Number of protein/solvent atoms	8,770/738
Wilson B-factor (Å ²)	30.1
Overall average B-factor (Å ²)	42.5
Average B-factor individual domains (Å ²)	b domains: 43.7±1.8 b' domains: 41.4±4.8
R factor (R _{free})	0.170 (0.217)
Coordinate error (Å)	0.29
RMS deviations from ideal values in	
Bond lengths (Å)	0.007
Bond angles (°)	1.070
Dihedral angles (°)	15.639
Planar groups (Å)	0.005
Ramachandran statistics	97.61/2.11/0.28

Table 3.1 – Crystallographic statistics for the structure of ERp27

Modification of a table generated by Wolfgang Kölmel.

absorption edge of selenium at 0.9795 Å. As before each data set contained the diffraction patterns of multiple crystal lattices. Using selected images, Dr. Jochen Kuper was able to index and subsequently process the intensities arising from the major crystal lattice. With the program XDS the images were processed and scaled. The ShelxC/D/E platform allowed solving the heavy atom substructure, resulting in the identification of 24 Se sites out of a total of 30 sites. The resulting sites were input into PhaserEP to solve the phase problem using single wavelength anomalous dispersion (SAD) resulting in a figure of merit (FOM) of 0.371. The resulting electron density map was promising, showing the positions of helices and a few β -strands. In order to improve the accuracy of the determined phases the density modification program Parrot was used for solvent flattening with a solvent content of 45 % and NCS averaging improving the figure of merit to 0.611. The resulting electron density map was detailed enough to place four copies of the NMR structure of the N-terminal domain into the electron densities. This allowed for automated model building using Buccaneer. The resulting model, as expected, contained five chains of ERp27 in the asymmetric unit, which were nearly complete with chain lengths varying between 217 and 220 amino acids. This model was used for further refinement with Phenix and was manually modified by Wolfgang Koelmel with Coot in further rounds of refinement, resulting in the final model as shown in Figure 3.21 with an R factor of 17% ($R_{\text{free}} = 21.7$). The final crystallographic statistics of the model are given in Table 3.1.

The final model contains five molecules (designated A through E) of ERp27 in the asymmetric unit (Fig. 3.21). The protein molecules form stacks where one molecule is rotated by 72° in comparison to its neighbors. There are a total of five possible orientations that are iterated throughout the crystal in the same order (C→A→B→D→E→C). Crystal contacts within a stack are extensive, with interfaces of approximately 1050 \AA^2 shared by neighboring molecules of the same stack. The hydrophobic cleft – the suggested substrate binding site – is part of this intra-stack interface (Fig. 3.21 A – green PEG molecules). In contrast, the interfaces between molecules of different stacks are much smaller, ranging from 60 \AA^2 to 320 \AA^2 with each molecule being in 7-9 interactions with a total interface area of $\sim 550 \text{ \AA}^2$ on average. Aside from their orientation, the five chains found in the asymmetric unit are very similar. When superimposed onto each other (Fig. 3.22) low pairwise root mean square (rms) deviations varying between 0.337 \AA and 0.684 \AA (mean: 0.504 \AA) were observed, however, it should be pointed out that ncs restraints with target values of 0.5 \AA were incorporated during refinement.

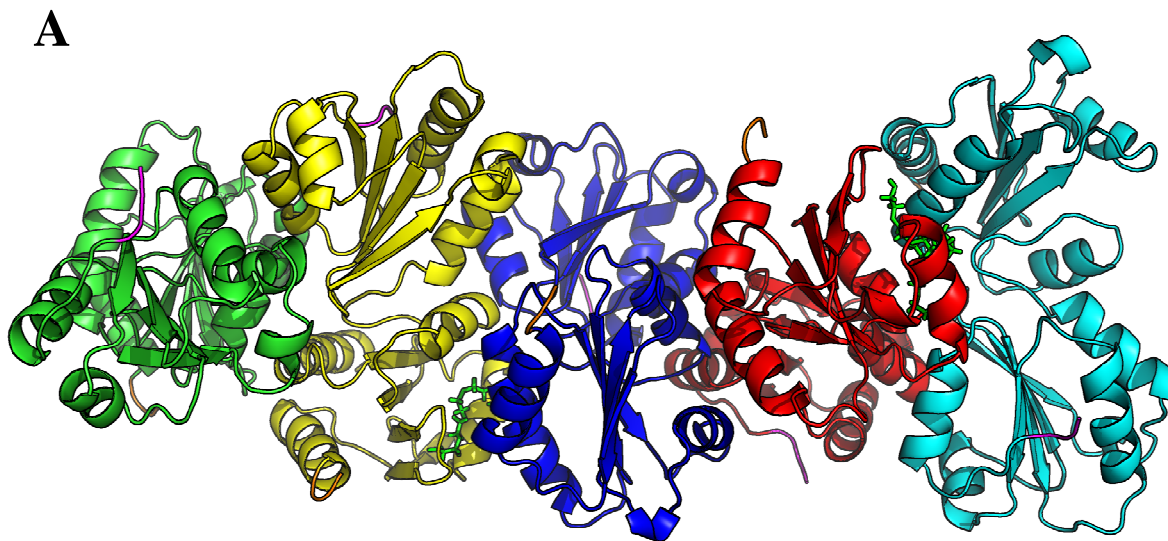


Figure 3.22 – Structure of ERp27- Crystal content

(A) Five molecules make up the asymmetric unit of the ERp27 crystals. The crystal lattice is comprised of five ERp27 molecules (ribbon models colored in blue, cyan, yellow, red and green for the A/B/C/D/E chains respectively) stacked upon each other with a clockwise rotation of 72° between each molecule. Two chains have bound parts of a polyethyleneglycol molecule (green stick model) in their hydrophobic pocket. For better orientation, the N-termini are colored in magenta and the C-termini are colored in orange. (B) Asymmetric unit containing two partial stacks (same coloring). While the crystal contacts within a stack are extensive ($\sim 1000 \text{ \AA}^2$), the contact area between stacks is rather limited ($< 320 \text{ \AA}^2$). This explains the needle shaped morphology of ERp27 crystals along with their tendency to split into parallel needles, causing multiple similar lattices to be present in one crystal.

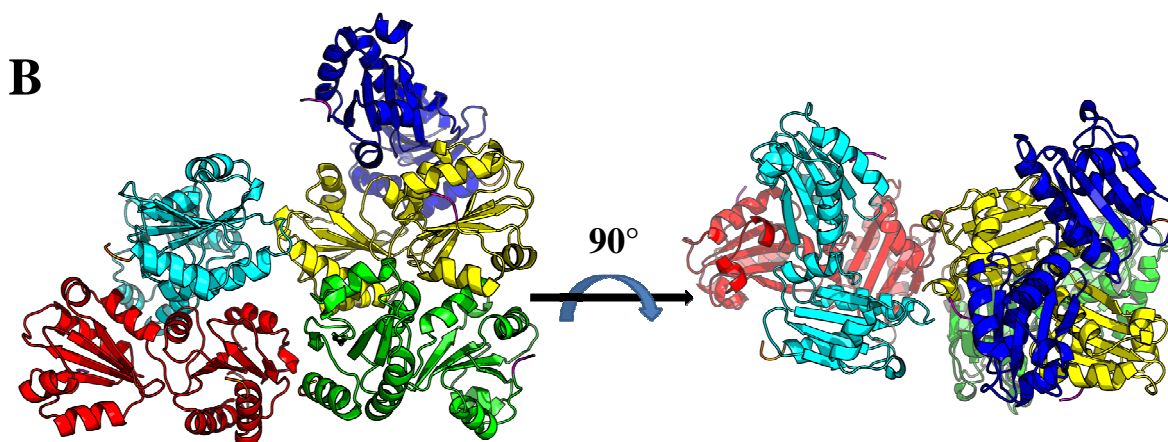
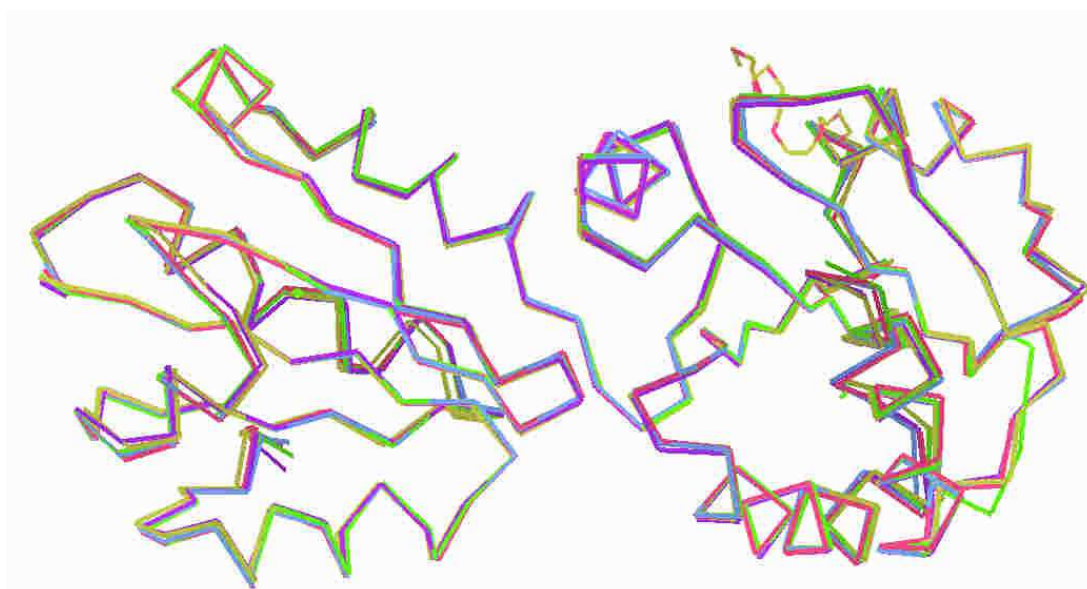


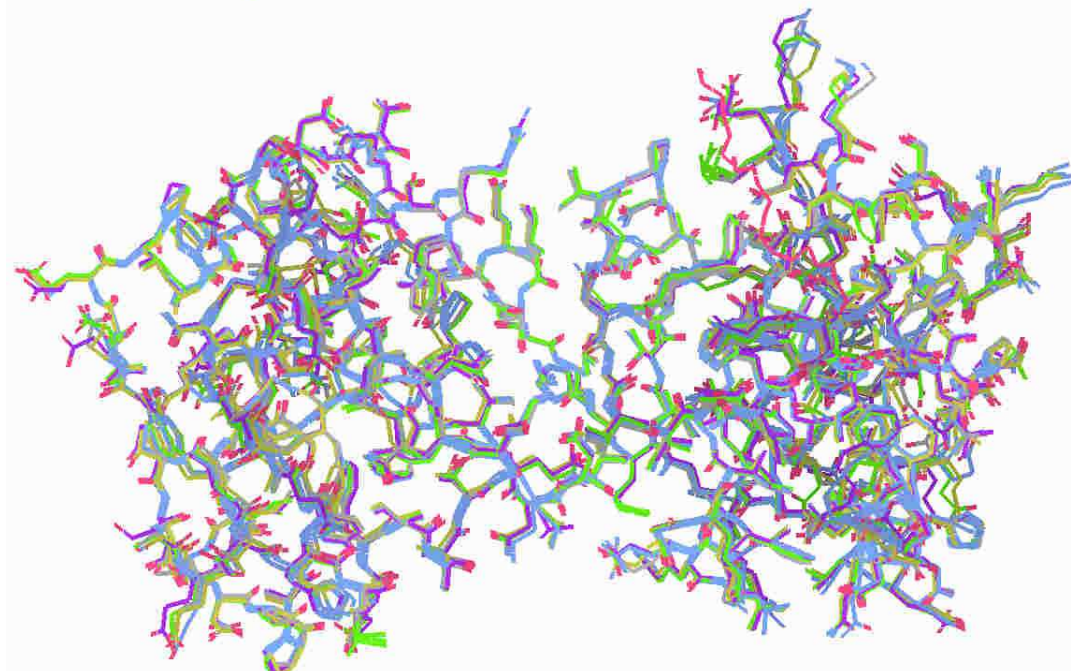
Figure 3.23 – Superposition of the five molecules in the asymmetric unit

Five molecules make up the asymmetric unit of the ERp27 crystals. (A) Superposition of the peptide backbone. All five molecules were superimposed with the secondary-structure-matching (SSM) function of Coot. The backbone atoms of all chains match except for one deviation in the C chain (green, bottom right), where a crystal contact forces a loop inward. (B) The same superposition with a representation of the side chains reveals a very good match for the majority of residues. Only selected amino acids exhibit different conformations, mostly at the surface due to crystal contacts.

A



B



```

      10          20          30          40          50          60
MEAAPSRFMF LLFLLTCELA AEVAAEVEKS SDGPGAAQEP TWLTDVPAAM EFIAATEVAV
      70          80          90          100         110         120
IGFFQDLEIP AVPILHSMVQ KFPGVSEFGIS TDSEVLTHYN ITGNTICLFR LVDNEQLNLE
      130         140         150         160         170         180
DEDIESIDAT KLSRFIEINS LHMVTEYNPV TVIGLFNSVI QIHLLIMNK ASPEYEENMH
      190         200         210         220         230         240
RYQKAARKLFQ GKILFILVDS GMKENGKVIS FFKLKESQLP ALAIYQTLDD EWDTLPTAEV
      250         260         270
SVEHVQNFCD GFLSGKLLKE NRESEKTPK VEL

```

Figure 3.24 – Sequence of ERp27

The open reading frame of ERp27 translates into 273 amino acids (Uni-Prot accession number: **Q96DN0**). The first 25 amino acids (shaded grey) represent the signal sequence for the import into the ER and were not contained within the expression construct used. After purification, ERp27 was subjected to limited proteolysis with trypsin, yielding a fragment approximately 3 kD smaller. Analysis of potential trypsin cleavage sites (red letters) shows that Ser 30 as well as Lys 256 have to be included in the digestion product. The structure (shaded in green) accounts for 219 of the 227 amino acids of the crystallized fragment.

Only molecule C showed a single deviation in the C- α -backbone, where a crystal contact causes a minor deformation of a loop. The side chains are also positioned in a very similar fashion with only minor deviations for surface residues most likely involved in crystal contacts. Each chain in the model is significantly shorter (217 to 220 residues) than the 248 amino acids for full-length human ERp27 after removal of the signal sequence as expected due to the proteolysis with trypsin. Taking into account the cleavage sites of trypsin after Lys and Arg residues, only the first eight amino acids presumably present in the digestion product are not resolved in the structure (Fig 3.23), while the C-terminus terminates with a lysine which apparently constitutes the C-terminal cleavage site. The amino acids at the N-terminus are most likely disordered.

As expected from previous secondary structure analysis in the literature [71], each single ERp27 molecule consists of two domains. Each domain is folded into the canonical thioredoxin fold with the secondary structure sequence of β_1 - α_1 - β_2 - α_2 - β_3 - β_4 - α_3 (Fig 3.24). The N-terminal domain extends this fold with a single alpha helix before and after this motif (α_0 - β_1 - α_1 - β_2 - α_2 - β_3 - β_4 - α_3 - α_4). The C-terminal domain also possesses an α_0 helix but lacks the α_4 helix. As seen from the

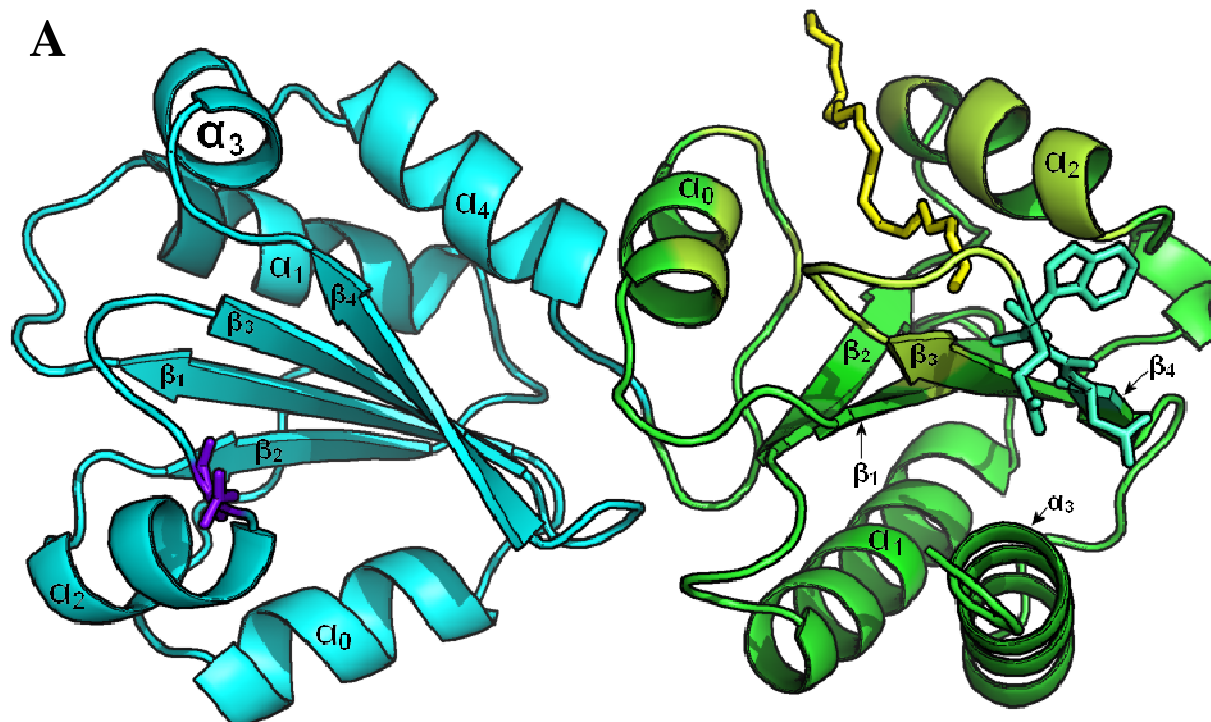


Figure 3.25 – Structure of ERp27 – Single molecule

Model of the D chain (red molecule Figure 3.21). Each molecule consists of two domains (cyan and green). Both domains fold in accordance with the canonical thioredoxin fold motif (β_1 - α_1 - β_2 - α_2 - β_3 - β_4 - α_3) with additional helices (α_0 and α_4) flanking the central fold. The N-terminal domain contains a putative glycosylation site (purple – stick model). The C-terminal domain of the D chain has a PEG molecule (yellow – stick model) bound to its hydrophobic cleft (colored limon) and contains the ERp57 interaction site (turquoise – stick model).

sequence, ERp27 lacks the active site motif (CXXC) of the thioredoxin family. The structure shows that the the sole two cysteine residues present are positioned in such a way that their side chains face into the interior of the protein without any surface accessibility. This further confirms a lack of redox activity which is consistent with the alignment of ERp27's domains with the bb' domains of other PDI family members.

Within this project it has been shown (see 3.4) that ERp27 can bind proteins that are permanently unfolded through reductive carboxymethylation. It is also known from crosslinking experiments, that ERp27 can bind the unstructured peptide Δ -somatostatin [71]. Based on the Δ -somatostatin binding and an alignment of ERp27 with the bb' domain it has been hypothesized that the C-terminal domain of ERp27 – like PDI's b' domain – contains a hydrophobic cleft for substrate

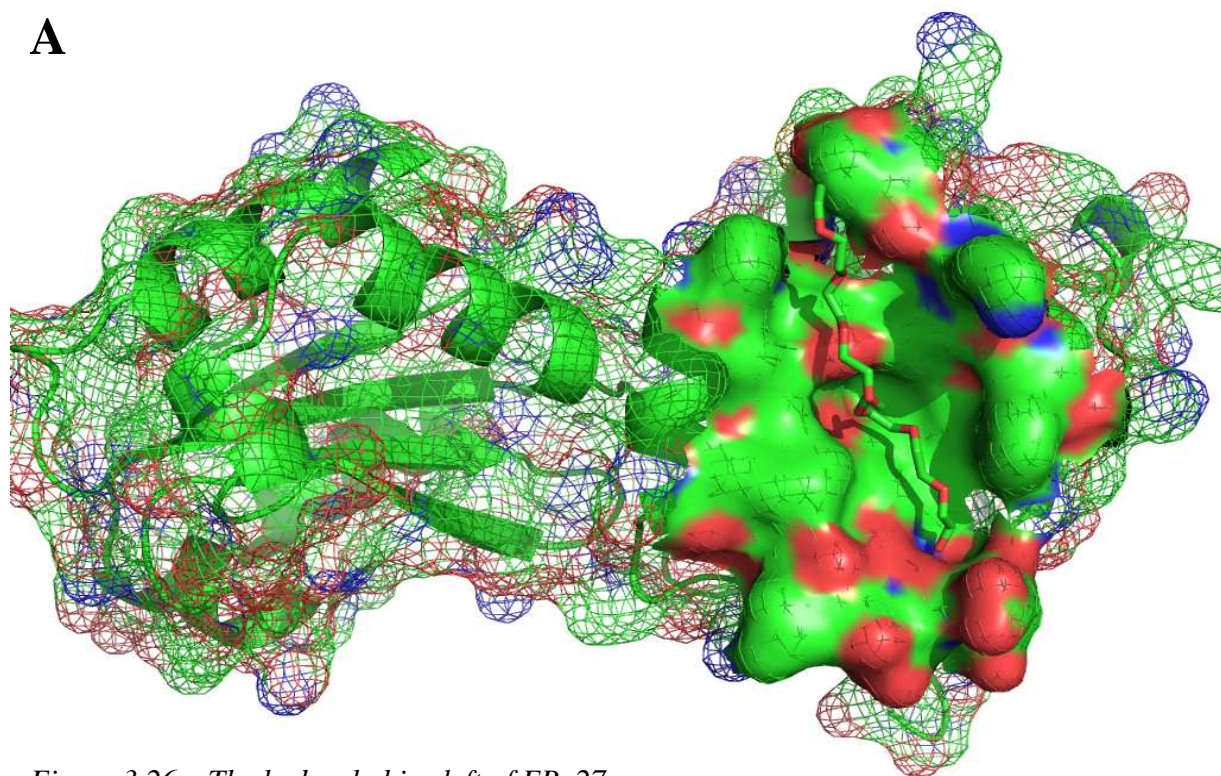
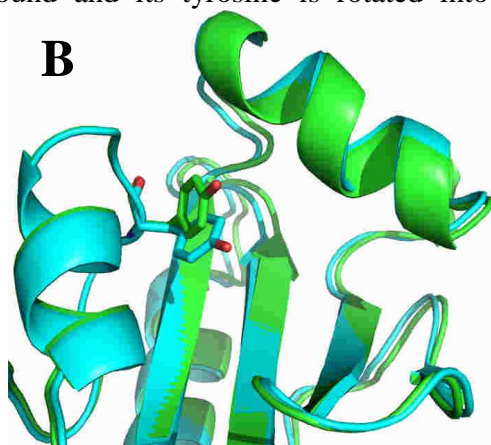
A

Figure 3.26 – The hydrophobic cleft of ERp27

(A) Surface representation of the ERp27 D molecule with a bound PEG. The surface of the protein is represented as a mesh with oxygen atoms in red, nitrogen atoms in blue and carbon atoms in green. The hydrophobic cleft is shown in the solid surface representation (coloring as mesh). A PEG molecule from the crystallization mother liquor is bound to this groove, which is lined mainly by hydrophobic, apolar residues. This structural feature can be found in many PDI family members and has been shown to facilitate substrate binding. In ERp27 tyrosine 147 extends its hydroxyl group into the floor of the cleft, disrupting the hydrophobic character. In the overlay of all five chains, the orientation in this tyrosine is flexible. The deviant A chain also has a molecule of PEG bound and its tyrosine is rotated into a conformation that no longer intrudes the hydrophobic cleft. (B) Close-up overlay of the hydrophobic cleft between the A chain (cyan) and the D chain (green). The PEG molecule of both models has been removed. The crucial tyrosine 147 – represented as stick model – rotates its hydroxyl group away from the surface extending the hydrophobic cleft from 9.5 Å to almost 14 Å.

B

interaction. This hypothesis is further supported by the results of isothermal titration calorimetry analysis (see 3.4). In these experiments ERp27 interacts with unfolded proteins in the same way as the corresponding fragment of human PDI. The structure confirms the presence of a hydrophobic cleft in the C-terminal domain of ERp27. The helices α_0 and α_2 together with the β_3 - β_4 -loop form a groove in the surface of the C-terminal domain (Fig. 3.25). The bottom of this groove is lined with hydrophobic side chains originating from the central β -sheet. In two of the five chains of the structure, this groove is filled with a fragment of the much longer polyethyleneglycol (PEG) molecule used as precipitant. This binding illustrates how unfolded linear polypeptide chains may bind to this site. In the D chain displayed in Figure 3.23, the flexible PEG molecule gets distorted by a ridge that runs through the bottom of the hydrophobic cleft, limiting its length to 9.5 Å. This ridge is formed by the hydroxyl group of Tyr 147, which is the most prominent polar group at the bottom of this otherwise apolar cleft. In the superposition of all five chains, however, the side chain of Tyr 147 in the A chain deviates from the consensus position (Fig. 3.25 B). The A chain is the second chain that has a PEG molecule bound and in this case the side chain of Tyr 147 has rotated to an angle where the hydroxyl group is no longer interrupting the hydrophobic cleft to the same extent. This could be a sign of induced fit upon substrate-binding, where the interaction of the unfolded polypeptide induces a conformational change to open up the hydrophobic cleft to a larger degree. Yet, judging from the superposition, the conformational flexibility of Tyr 147 seems to be independent of the general conformation of the surrounding residues. Thus, the hydroxyl group could initially be in place to act as a hydrogen bond donor or acceptor, but could easily be displaced in case of very hydrophobic substrate to make room for additional hydrophobic residues. So with the Tyr 147 hydroxyl in the raised position, ERp27 binds two to three hydrophobic side chains followed by a contact with the peptide backbone. In contrast, with the tyrosine rotated out of the way, the hydrophobic cleft extends to a length of approximately 16 Å, which would be enough to accommodate two additional hydrophobic side chains. Thus, the flexibility of Tyr 147 appears to be a mechanism to offer an energetically more favorable binding site to a broader substrate population. A co-crystal structure with a peptide substrate like Δ -somatostatin would provide additional insights as to which ERp27 residues make contact with features of the substrate and what groups in the substrate are bound specifically.

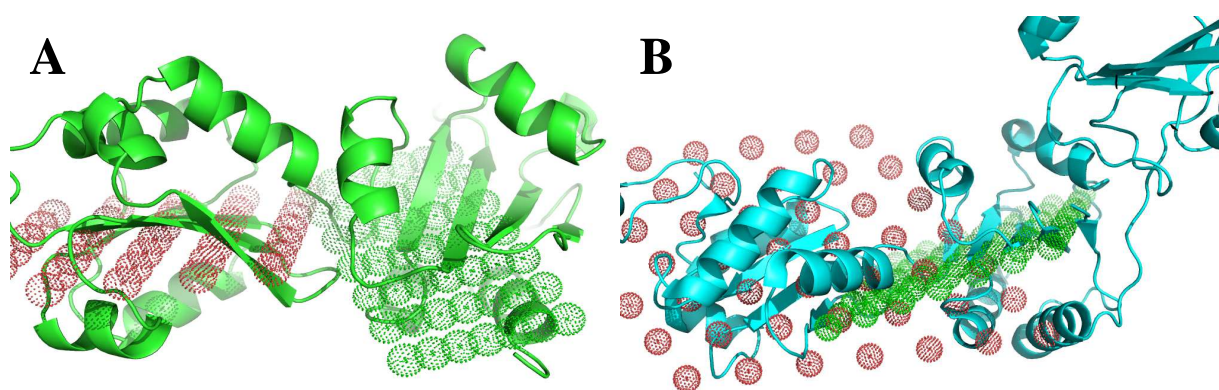


Figure 3.27 – Planar regression analysis of the domain orientation in the PDI family

In order to analyze the relative orientation of the b and b' domains with respect to each other, the planes defined by the central β -sheets of each domain from multiple PDI family members were calculated, using the coordinates of the main chain atoms as basis for a planar regression analysis. From each residue participating in the central β -sheet, the coordinates of the C α atom as well as the carbon and nitrogen of the peptide bond were included into the data set. The resulting parameters defining each plane (see table below) were then used to calculate reference points for each plane which were included into the structure of the respective thioredoxin domain. (A) Ribbon model of ERp27 (green) with the reference planes for the N-terminal b domain (red dots) and for the C-terminal b' domain (green dots). Both planes intersect at a shallow angle of 25.7°. (B) Ribbon model of the bb'a fragment of human PDI (pdb: 3UEM) with the reference planes for the b domain (red dots) and the b' domain (green dots). The angle between the two planes is much larger with 55° in the case of human PDI. In each case, the calculated planes fit very well with the central β -sheet, when superimposed.

Protein	Parameters – b domain	Parameters – b' domain	Interplanar angle
ERp27	$z_0 = -67.1$; $a = 0.46$; $b = 1.19$	$z_0 = -41.8$; $a = -0.13$; $b = 0.77$	25.71°
PDI (3UEM)	$z_0 = -23.8$; $a = 0.84$; $b = -1.26$	$z_0 = -3.91$; $a = -0.57$; $b = -0.70$	55.00°
PDI (2B5E)	$z_0 = -25.0$; $a = -0.10$; $b = 0.38$	$z_0 = -51.1$; $a = 0.88$; $b = -0.28$	56.32°
ERp57 (2H8L)	$z_0 = 77.8$; $a = -0.19$; $b = 0.07$	$z_0 = 134$; $a = -2.10$; $b = 0.85$	54.85°
ERp72 (3EC3)	$z_0 = 16.3$; $a = 0.38$; $b = 0.05$	$z_0 = 33.1$; $a = -0.37$; $b = -0.80$	54.94°

The role of the N-terminal domain is not readily discernable from the structure. This domain is homologous to the b-domain in the four-domain-members of the PDI family with sequence identity levels ranging from 9.2% (ERp44) to 23.3% (PDI_A1). So far only a structural role has been suggested for these domains [97]. Usually the b domain together with the substrate-binding b' domain forms a rigid base. The catalytically active thioredoxin domains (a domains) are flexibly linked to this two-domain core. The b domain has been suggested to correctly space the catalytic domains N- and C-terminal from the bb' core as well as to ensure the correct orientation of N-terminal active sites in relation to the substrate-binding site located in the b' domain.

In an overlay of structures of human PDI, yeast PDI, ERp57 and ERp72, the orientation of the b domain to the b' domain appears to be highly conserved [97]. Even in the three domain protein ERp44 only minor deviations could be observed. In contrast, the N and C terminal domains of ERp27 appear to adopt a very different orientation towards each other. As an overlay provides only qualitative information about similarity, I quantified the orientation of the two thioredoxin domains using the central β -sheet of each domain as a reference. Using a planar regression analysis of the coordinates of all backbone atoms that are part of the central β -sheet, a reference plane was assigned to each domain (Fig. 3.26 A). As can be seen, both planes intersect each other at a shallow angle of 25.7° . When the same regression analysis was performed for PDI (human and yeast), ERp57 and ERp72, the orientation of the two domains is clearly different (Fig 3.26 B) with the regression planes encompassing angles between 54.9° and 56.3° . These results clearly show a common orientation among the larger PDI family members in their bb' core, while ERp27 deviates significantly from this arrangement. This could be due to a novel function inherent in the N-terminal domain, but no such function could be deduced from either structure or function of the protein. It thus stands to reason, that the lack of catalytically active domains removed the orientation constraint from the domain interface allowing it to change during the course of evolution. The function of the N-terminal domain in ERp27 is most likely one of keeping the C-terminal domain in solution. This is supported by the results from the group of Lloyd Ruddock [71], showing a sharp decrease in protein solubility, when the N-terminal domain has been removed. All PDI family members show extremely high solubility in-vitro. This is only sensible as the primary folding catalysts of the cells have to keep hydrophobic substrates in solution during their interaction while acting in a highly crowded cellular compartment with extremely high total protein concentrations. Thus, the N terminal domain of ERp27 keeps the functionally critical C terminal domain soluble, enabling the protein as a whole to fulfill its role. It seems likely that the evolutionary pressure for high solubility has also modified the domain interface resulting in the observed change in orientation.

Another interesting feature surrounding the substrate binding activity of ERp27 is the interaction between ERp27 and the PDI family member ERp57. ERp27, like the known ERp57-interaction-partners calnexin and calreticulin, contains the EWD motif that is recognized by ERp57. This

motif is located in the β_3 - β_4 -loop which closes of the hydrophobic cleft. Initial MNR-studies on the ERp27-ERp57 interaction [71] reported a distinct movement of the involved tryptophan during interactions of ERp27 with both Δ -somatostatin and ERp57. Thus, the interaction with ERp57 initiates a conformational change in the β_3 - β_4 -loop which is involved in substrate-binding. It is unclear, if the conformational change has an influence on the affinity of ERp27 towards an unfolded substrate. ITC experiments with ERp27, ERp57 and Δ -somatostatin could provide insights into a potential modulation of substrate recognition of ERp27 through ERp57 interaction. Taking the properties of both PDI family members into account, such a modulation would make sense. ERp57 is a very potent redox catalyst, but lacks a general substrate binding activity as shown in the literature [98] as well as within this project (see 3.4). It has been shown that ERp57's activity in protein refolding is critically dependent on the interaction with calnexin and calreticulin [98, 99]. These two interaction partners belong to the protein family of lectins that selectively bind substrates carrying N-glycans, thereby limiting the refolding activity of ERp57 to glycoproteins. ERp27 on the other hand only contains a substrate binding site and is devoid of catalytic activity on its own. ERp27 could be shown to bind both glycosylated and non-glycosylated proteins and could therefore be an adaptor protein to enable ERp57 to act on non-glycosylated substrates. This could be a way to cope with a high substrate load or with an accumulation of misfolded proteins during stress situations. A potential upregulation of ERp27 during chemically-induced ER stress is currently investigated by quantitative real-time PCR in cooperation with the group of Heike Hermanns. The results will hopefully further our understanding of the actions of ERp27 and its interactions with the PDI family during the unfolded protein response (UPR) of a cell.

3.3 Fluorescence-based assays for activity measurements of PDI family proteins

In order to understand the web of functions that – as a whole – results in the oxidative folding of proteins in the endoplasmic reticulum it is very useful to compare the activity of different PDI family members to each other. However, due to the presence of multiple activities in a singular molecule – such as oxidase, isomerase and chaperone activity in the case of PDI – it is very difficult to measure the kinetic parameters of a single activity. Yet the knowledge of those kinetic parameters would enable a direct comparison of proteins within the PDI family in regard to other family members, thus possibly suggesting specific roles for each. With the exact determination of kinetic parameters such as rate and binding constants one could also characterize mutations and structural features more exactly. Over the decades of research done on PDI and related proteins many assays have been used to estimate the activity. RNase refolding assays observe the refolding of either reduced and unfolded or scrambled (randomly oxidized) RNase [63, 100-102]. The non-native RNase refolds in the presence of a redox buffer and, at various time points, samples of the refolding RNase are assayed for activity. This widely used assay has several drawbacks. For one it does not measure the activity of PDI directly, thus greatly decreasing the accuracy. The assay only registers completely refolded molecules and does not distinguish between folding intermediates at different stages of maturation. It also does not account for the difference in oxidation versus isomerization. While this is appropriate for an approximation of the overall activity of an enzyme, the assay cannot be used to determine kinetic parameters for selected activities. Another often applied assay is based on the precipitation of redox sensitive insulin [103]. Here, insulin is reduced in the presence and absence of a disulfide exchange catalyst by directly observing the progress of the reduction through the turbidity – and thus the absorbance – of the solution. While this assay can be used to determine the kinetic parameters of the reduction reaction, the assay is rather insensitive and has difficulties to measure low activities. For the chaperone activity, multiple assays have been established. Unfortunately, different assays cast a conflicting picture on the chaperone activity of PDI. While the refolding of the disulfide free protein GAPDH (D-Glyceraldehyde-3-phosphate dehydrogenase) [92] is assisted by the addition of PDI, the binding of unfolded cholera toxin [48] suggest an anti-chaperone or chaperone activity depending on the redox state.

For the in depth characterization of PDI activity, a combination of three assays were tested. The first assay was initially established by *Raturi et al.* [75] to measure the catalysis of redox reactions. It is based on the fluorescent properties of a molecule of oxidized glutathione (GSSG) labeled by a conjugation of eosin to its two N-termini. In the oxidized state, the eosin moieties are in close proximity and thus mutually quench their fluorescence. Upon reduction of glutathione, two independent molecules are created and the quenching is abolished. The resulting rise in fluorescence can be monitored in real time and even allows for detection of nanomolar concentrations of PDI [104], which is a far more sensitive and accurate quantification of the reduction compared to the insulin turbidity assay. As the substrate for reduction is a low molecular weight molecule, the reaction rate observed in this fluorescence self-quenching assay (FSQ assay) is most likely uninfluenced by additional substrate binding and recognition sites and only dependent on the active site. The information gained from the FSQ assay allows for a better analysis of the reduction of macromolecules.

A second suitable assay was proposed by *Stefan Lorenz* (University of Bayreuth) using a domain with a redox sensitive fold [105]. The inserted flap (IF) domain of the phage protein BGX3 is domain inserted into a loop in another protein. The fold of the domain is dependent on the close spatial proximity of its termini and thus the isolated IF domain is unfolded. When cysteine residues are introduced at both the N- and C-terminus, the folding of the IF domain becomes dependent on its redox status. The folding of this domain is a rapid process that can be monitored through the single tryptophan residue which changes its fluorescence as it gets sequestered on the inside of the domain away from the solvent. With the help of this substrate it is possible to observe the catalysis of both the oxidation and the reduction by an enzyme. As the substrate is a macromolecule, the reaction rate is influenced by both the rate of catalysis and the substrate binding of an enzyme. Using the IF domain in an assay would allow to compare oxidase and reductase capabilities to test the hypothesis that they are closely linked. Together with the data from the FSQ assay it could also be used to determine the substrate binding properties of PDI family members.

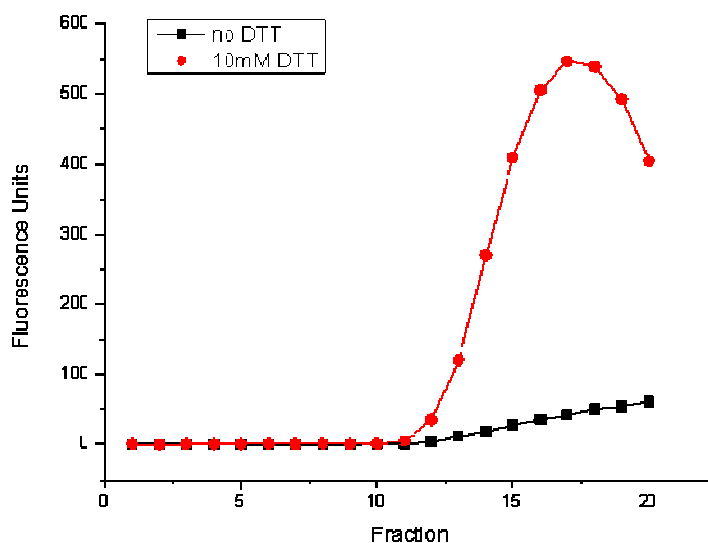


Figure 3.28 – Fluorescence measurements during di-E-GSSG purification

The eluate from the Sephadex 25G chromatography was fractionated and samples drawn from each fraction. After incubation with either buffer (black squares) or 10 mM DTT (red circles) the fluorescence was measured. As the Fluoromax-4 was not yet purchased at the time of the experiment, the measurements were made using a BioRad Fluorostar with an excitation filter of 520 nm (5 nm bandpass) and an emission filter of 555 nm (10 nm bandpass).

The third and final assay is a protein refolding assay that can be used to investigate the isomerase activity of PDI family members. As described above, there are several drawbacks using the refolding of RNase. Also RNase contains only eight cysteine residues which allow for 105 different SS-bond combinations. This means that even in randomly oxidized RNase there is a residual activity of ~1% [47]. All those problems can be sidestepped by the use of riboflavin binding Protein (RfbP) to assay refolding as described by *Rancy & Thorpe* [76]. Riboflavin is bound with a very high affinity ($K_D=1$ nM) by native apo-RfbP [74] and upon binding, the fluorescence of riboflavin is abolished. During refolding reactions, this allows for a direct measure of the number of completely refolded substrates as this is directly linked to a reduction in fluorescence. With 18 cysteines present in RfbP, it provides an isomerization challenge of far greater complexity as it results in more than 34 million potential combinations of disulfide bonds. Just like the RNase refolding, RfbP can be used both as a reduced and a scrambled substrate and is therefore suited to characterize the isomerase function of a protein especially in

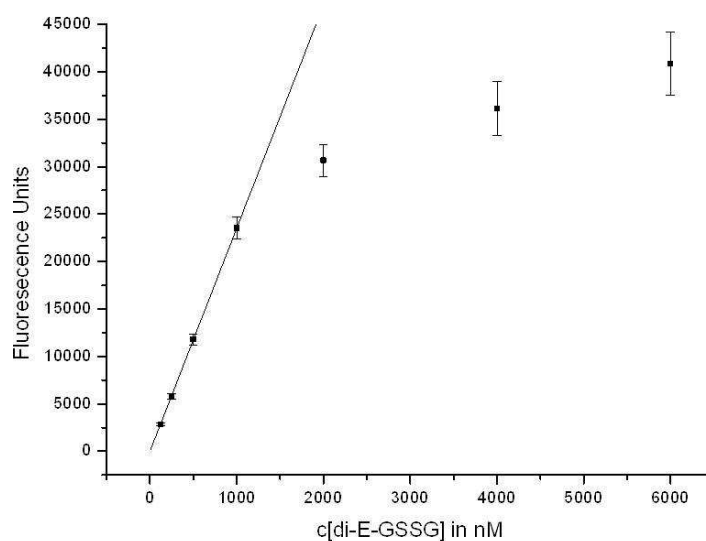


Figure 3.29 – Fluorescence standard curve of di-E-GSSG

Samples of di-E-GSSG at concentrations of 125 μM to 6 mM were reduced with 10 mM DTT for 60 minutes at room temperature. After incubation, the samples were examined for fluorescence using a BioRad Fluostar with an excitation filter of 520 nm (5 nm bandpass) and an emission filter of 555 nm (10 nm bandpass). Each data point is the average of three independently prepared samples. The first four data points were fitted to a linear equation to gain the conversion factor between fluorescence signal and the concentration of fluorescent product. A concentration of E-GSH of 2 μM yields a fluorescence signal of 23400 units.

combination with the other two assays, thus providing information about substrate binding and oxidation or reduction respectively.

The di-eosin-glutathione (di-E-GSSG) substrate used in the FSQ assay was prepared as described in the methods section (see 2.3.7). The final step in the substrate preparation is to separate the conjugated di-E-GSSG from the non-conjugated educts glutathione and eosin isothiocyanate as well as from the glutathione molecules with only a single eosin moiety. This was done using a column packed with Sephadex 25G as chromatography material which is usually used for size exclusion chromatography. For the initial purification, a column with a volume of 5 mL was packed using a peristaltic pump. The reaction mix containing the di-E-GSSG was layered onto the surface of the column. After the substrate had entered the column overlaid with FSQ reaction buffer (see 2.3.7) the eluate was collected in 0.5 mL fractions, but only fractions 11 and onwards

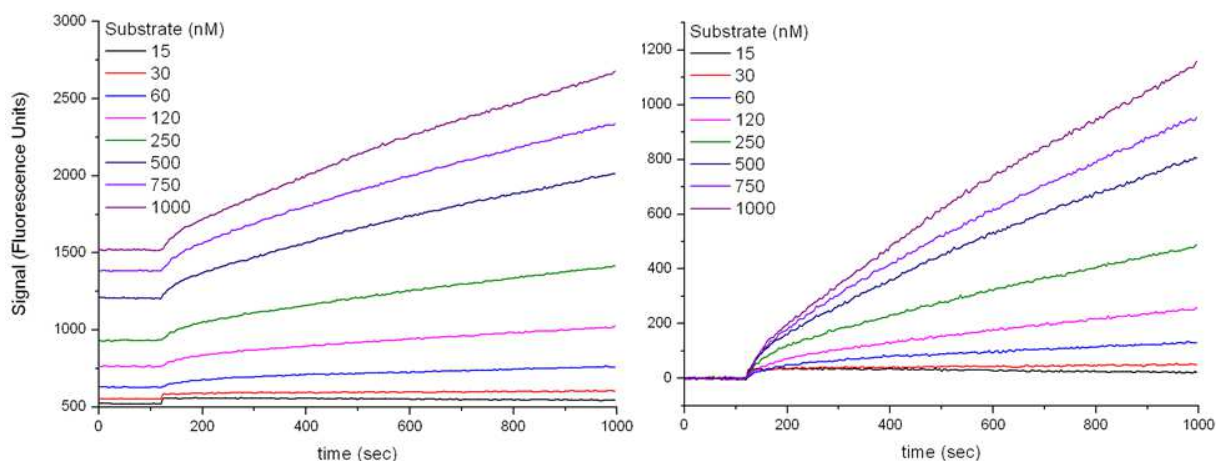


Figure 3.30 – Reduction of di-E-GSSG by DTT

The fluorescence of di-E-GSSG in concentrations ranging from 15 nM to 1 μ M was measured simultaneously in a Fluostar plate reader. For the first 120 seconds only the base fluorescence of the di-E-GSSG in its oxidized form was recorded. At the 120 second mark, DTT was added to a final concentration of 15 μ M. The raw data (left panel) shows a variable base line depending on the di-E-GSSG concentration. For better comparison, the signal of the first 60 seconds of each sample was averaged and then subtracted from each data point of this sample to exclude the background fluorescence (right panel). After the addition of DTT the fluorescence increases with the rate of the increase dependent on the substrate concentration.

showed signs of color. From each fraction two samples of 10 μ L were drawn and incubated with either reaction buffer or reaction buffer containing 10 mM DTT. After an incubation period of 60 minutes at room temperature in the dark, the samples were tested for fluorescence. As suggested by the coloring, fluorescence could only be observed for fractions 11 and later (Fig. 3.3.1). This suggests that there is an interaction between the Sephadex 25G material and the di-E-GSSG, because in a normal size exclusion chromatography every component of the loaded sample should be eluted after one column volume. Yet the first fraction containing any eosin eluted after 1.1 column volumes. As visible in the fluorescence plot, the fluorescence of the reduced eluate peaks at fraction 17. In contrast the fluorescence of the non-reduced fluorescence increases steadily over the monitored range. While di-E-GSSG fluoresces at the examined range, it does so with a very poor efficiency compared to free eosin or the reduced mono-eosin glutathione (E-GSH) [75]. Therefore the fluorescence observed from the non-reduced samples most likely corresponds to single conjugated glutathione carrying only one eosin moiety and is therefore not subject to quenching. When comparing the ratio of the signals given by the reduced and oxidized

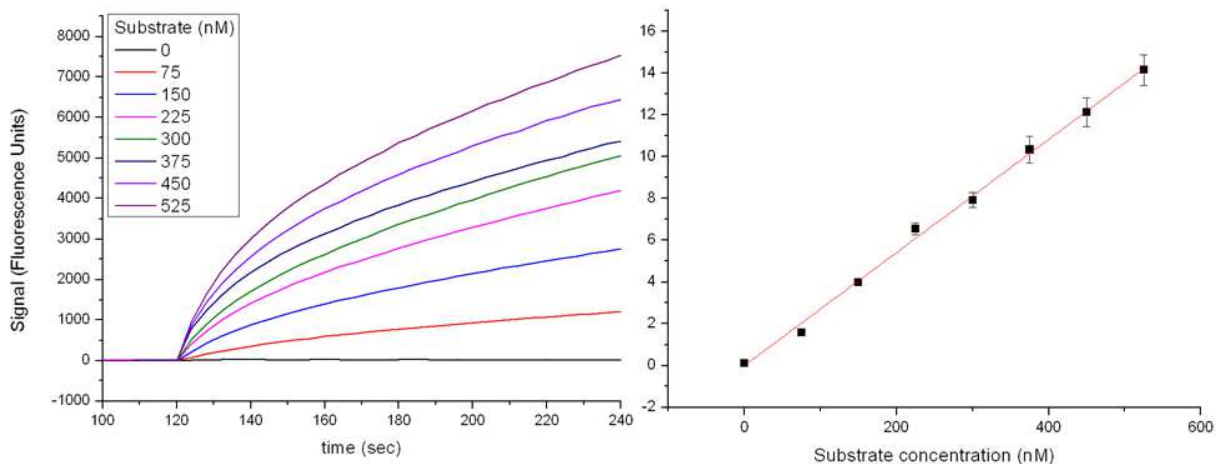


Figure 3.31 – Reduction of di-E-GSSG catalyzed by yeast PDI

The fluorescence of di-E-GSSG at concentrations ranging from 0 nM to 525 nM was measured simultaneously in a Fluostar plate reader. For the first 120 seconds only the base fluorescence of the di-E-GSSG in its oxidized form was recorded. At the 120 second mark, DTT was added to a final concentration of 15 μ M. At the same time yeast PDI in FSQ assay buffer was added to a final concentration of 50 nM through a different injector. The baseline corrected signal (left panel) increases over time with the rate depending on the di-E-GSSG concentration. To determine the reaction rate, the data points of the first 20 seconds after the addition of DTT and PDI were fitted to a linear equation, the slope was plotted against the substrate concentration (right panel) and converted from the change in signal (Δ FU/sec) to the quantity of formed E-GSH (pM/sec) after subtracting the background reduction in the absence of enzyme. The data points were fitted by linear regression with weighted errors and a fixed intercept at $y=0$ yielding a slope of $2.5 \times 10^{-5} \text{ sec}^{-1}$ with an R^2 of 0.999.

samples from each fraction, the maximum ratio of reduced to oxidized signals (signal_{RED} to signal_{OX}) was 14.71 for fraction 15. This was much lower than the ratio of 70 given in the literature. When pooling the fractions with a fluorescence ratio of greater than 10, the final concentration of di-E-GSSG as determined by the absorbance at 525 nm was 150 nM. Following optimization of the di-E-GSSG purification to include a larger Sephadex column (120x10 mm) the purity of the substrate could be improved to yield fluorescence ratios of up to 60.

Using a dialysis tube with a molecular weight cut off of 500 Da, the substrate could be concentrated by dialysis against PEG 5000 to a concentration of up to 8 μ M di-E-GSSG. Using this substrate a standard curve was generated to relate the fluorescence intensity to the quantity of the product E-GSH. When samples in a concentration range between 125 nM and 6 μ M were

reduced by the addition of 10 mM DTT, the increase in fluorescence was dependent on the concentration of di-E-GSSG in a linear fashion up to 1 μ M (Fig. 3.3.2). At higher concentrations, the fluorescence increased at a slower rate, most likely due to inner and outer filter effects.

With these results the FSQ assay had to be restricted to a substrate concentration of 1 mM or lower. The first kinetic experiment was the reduction of substrate by 15 μ M of DTT. Samples of di-E-GSSG in sample buffer were prepared with a final concentration between 20 nM and 1 μ M. Using the BioRad Fluostar, a fluorescence plate reader, the fluorescence was recorded for 2 minutes, then DTT was injected into the well. As expected, the fluorescence increases upon addition of the reductant (Fig. 3.3.3). As each sample gave a different base line signal prior to the addition of DTT, the data were normalized to exclude the background signal. The normalized data showed an increase in fluorescence after DTT addition, which is dependent on the concentration of the substrate present. Interestingly, the signal of the two samples with the lowest concentration of di-E-GSSG showed a slightly different behavior than the other six signals. Instead of an asymptotic growth seen for higher concentrations, the signal suddenly increases upon DTT addition for the 15 nM and 30 nM samples. After this initial surge, the signal stays constant (30 nM) or even decreases in a linear fashion (15 nM). The decrease can be attributed to minor bleaching of the fluorophor, which in case of the 30 nM sample gets presumably compensated by fluorescence of the newly generated E-GSH. The reasons for the jump in the signal, however, remain unclear, especially as it restricted to only two samples. The six samples with higher substrate concentration do not show this jump, but instead an initially linear increase in signal. However, both the bleaching effect as well as the signal anomaly is small (less than 30 units) compared to the increase in fluorescence upon reduction (120-1150 units for 30-1000 nM respectively).

Using this background reduction rate as a blank, experiments were performed with yeast PDI as reduction catalyst. In this setup, after a baseline measurement both DTT and enzyme were added to the reaction. The reductant and the enzyme were not pre-incubated together but added separately. In comparison to the rate of reduction by 15 μ M DTT alone, the rate was greatly increased in the presence of 50 nM yeast PDI (wild type). At 250 nM di-E-GSSG, a concentration of 15 μ M DTT in the absence of enzyme increased the fluorescence signal of 250 nM di-E-GSSG by 121.5 fluorescent units, corresponding to an increase in E-GSH concentration

by 10.3 pM. In the presence of 20 nM yeast PDI, the same condition results in a fluorescence increase of 3295 fluorescent units corresponding to an increase in E-GSH concentration by 563 pM (Fig. 3.3.4 – left side). In order to characterize the catalysis of the reduction by PDI, the rate of reduction was determined from a linear fit of the initial 20 seconds of the reaction. As the concentration of di-E-GSSG during the initial seconds of the reaction is close to the known initial concentration, the reaction rate can be set in relation to the substrate concentration. However, for the samples with a di-E-GSSG concentration of 375 nM and higher the increase in the fluorescence signal had already started to leave the linear initial phase, leading to a worse fit and therefore an increased error in the slope (Fig 3.3.4 – right panel). This is the result of the restrictions of the initially used Fluostar fluorescent plate reader that only enables measurements at intervals of 4 seconds or more. Reducing the period for the linear range would therefore lead to an even greater uncertainty as less than five data points would be fitted. Despite this shortcoming of the initially used fluorometer, it is apparent that there is a linear relationship between the reaction rate and the substrate concentration. This would be in line with a pseudo-first order reaction as the concentration of DTT greatly exceeds the concentration of the substrate and can therefore be approximated to remain constant throughout the reaction. However, the number of thiol groups available for reduction is not significantly changed by the addition of 50 nM PDI to 15 μ M DTT. Therefore a pseudo-first order model cannot explain the increase in reaction velocity.

The other possible explanation is that we are observing the low concentration range of Michaelis-Menten-like kinetics. When the concentration of the substrate is significantly lower than the K_D of the enzyme, the change in the reaction rate is dependent on the change of the substrate concentration in a linear fashion. Further exploration of the Michaelis-Menten model enables at least estimates of the kinetic parameter K_M and v_{max} . The Michaelis-Menten equation (eq.3.3.1) describes the catalysis of a reaction in relation to the concentration of the substrate [S] and the properties K_M and v_{max} of the enzyme. Plotting the rate of the reaction against the concentration of the substrate yields the Michaelis-Menten plot (Fig. 3.3.5), a hyperbolic curve that asymptotically approaches a maximum reaction rate v_{max} as its upper boundary and that reaches $v=v_{max}/2$ at a concentration that is called the Michaelis-Menten constant K_M which is

$$\text{Eq.3.3.1} \quad v = v_{max} \times \frac{[S]}{[S] + K_M}$$

$$\text{Eq.3.3.2} \quad \alpha = \frac{v}{[S]}; \rightarrow v = \alpha \times [S]$$

$$\text{Eq.3.3.3} \quad \alpha \times [S] = v_{max} \times \frac{[S]}{[S] + K_M}; \xrightarrow{\div [S]} \alpha = \frac{v_{max}}{[S] + K_M}$$

$$\text{Eq.3.3.4} \quad \alpha = \frac{v_{max}}{K_M}$$

$$\text{Eq.3.3.5} \quad v_{lin} = v_{max} \times \frac{[S]}{K_M}$$

$$\text{Eq.3.3.6} \quad \frac{v_{lin} - v}{v} = \frac{v_{lin}}{v} - 1 = \frac{v_{max} \times \frac{[S]}{[S] + K_M}}{v_{max} \times \frac{[S]}{K_M}} - 1 = \frac{K_M}{[S]}$$

specific for an enzyme in a given reaction environment. We can see that for concentrations that are very small compared to K_M there is a linear relation between the concentration of the substrate and the reaction rate (eq.3.3.2) which are connected by the linearity factor α . Rearranging eq.3.3.2 for the reaction rate we can now substitute the reaction rate v in the Michaelis-Menten equation (eq.3.3.1) for eq.3.3.2, which can be transformed to yield a relation between α and the kinetic parameters $[S]$, K_M and v_{max} (eq.3.3.3). As this relation only holds true for the linear part of the Michaelis-Menten model, when K_M greatly exceeds $[S]$ or $[S]$ approaches zero, we can approximate eq.3.3.3 by removing the $[S]$ term. This yields eq.3.3.4, a direct relation between the slope of the near-linear part of the Michaelis-Menten equation and the two enzyme specific parameters K_M and v_{max} . Applying this to our data for wild type yeast PDI we can estimate the ratio v_{max}/K_M to $2.5 \times 10^{-5} \text{ sec}^{-1}$. This in itself is not useful without an estimate of either K_M or v_{max} . Yet we can also give a range for our Michaelis-Menten constant. For that we have to look at the deviation of our data from the linear plot. At the highest concentration used in this experiment series, the error from our linear fit amounts to 5%. So we can say that the

difference between a linear fit (eq.3.3.5) and the Michaelis-Menten-model (eq.3.3.1) is limited to 5% at a concentration of 525 nM. When rearranging the equations for the linear model and the Michaelis-Menten kinetic to yield the difference between those two in relation to the value of the Michaelis-Menten constant we arrive at a very simple relationship (eq.3.3.6). The deviation of the Michaelis-Menten model from its initial linear slope is equal to the ratio of the substrate concentration and the Michaelis-Menten constant. The maximum deviation derived from our data of 5.2% at a substrate concentration of 525 nM would put the Michaelis-Menten constant at a value of 10 μM or higher. This in turn means that the maximum reaction speed is in the range of 250 pM/sec or higher.

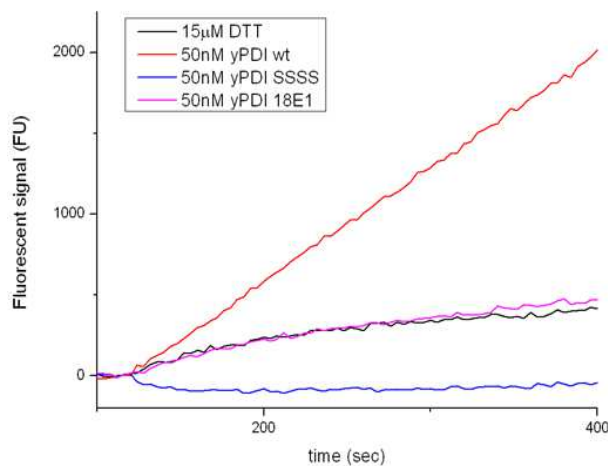


Figure 3.32 – Comparison of the wild-type, SSSS and 18E1 variants of yeast PDI

The reduction of Di-E-GSSG was monitored at a concentration of 100 nM in a Fluostar plate reader. The experimental procedures are as described for Fig 3.3.4 and the baseline corrected graphs are shown at the right. While the wild type accelerates the reduction, the 18E1 mutant has no effect and the SSSS variant even retards the reaction.

In order to use the FSQ assay for comparison of different redox catalysts it has to be ascertained that the increase in reduction rate is caused by the catalytic activity of the active sites. Therefore wild type PDI and a mutant with all four active site cysteines replaced by serine (yPDI SSSS) were directly compared. In addition, another mutant of reduced activity – yeast PDI 18E1 – was tested. The yeast PDI 18E1 contains four additional cysteines that were engineered to form two disulfide bonds linking the a domain with the b domain and the a' domain with the b' domain, respectively. As could be shown [64] the disulfides form as intended and restrict the mobility of the a domains with respect to the bb' domain core. When those three variants of PDI were tested, only the wild type accelerated the reduction of di-E-GSSG (Fig. 3.3.5). The addition of PDI SSSS actually retarded the reduction by DTT. This could be caused by an interaction of the substrate with the hydrophobic cleft, removing part of the substrate from the solution and

quenching the residual fluorescence of both di-E-GSSG and E-GSH impurities. This is also supported by the fact that the fluorescence signal drops below the baseline which was recorded prior to the addition of enzyme and DTT. In contrast, the 18E1 mutant of PDI does not influence the reduction significantly. The lack of activity in either direction points towards both active sites being locked in a conformation where they are inaccessible for the substrate. In addition, the hydrophobic surface that might interact with the substrate in case of the SSSS mutant is covered, rendering the PDI 18E1 completely inactive. Only PDI shown to be catalytically active in other assays accelerates the reduction of di-E-GSSG, while the inactive mutants either do not influence or even impair the reduction.

As a next step, a comparison between human PDI, yeast PDI and ERp57 (from *Gallus gallus*) was started, but was paused due to the introduction of a new fluorimeter. The data collection on the Biorad Fluostar is restricted by multiple limitations of the plate reader itself such as the minimum interval between measurements and the lack of a monochromator to generate or read the exact excitation or emission wavelengths, respectively. The Fluoromax4 fluorimeter allows for steadier and more exact measurements and was therefore used as soon as it was available. After generating standard curves, for the new system the comparison between the two PDI variants and chicken ERp57 was resumed in a concentration range of 100 nM to 2 μ M. As expected both variants of PDI were very similar in their signal. In contrast the addition of ERp57 resulting in a very large increase of the reaction speed (Fig 3.3.6). When fitting the initial reaction rate, the reaction velocity was non-linear even for the first 10 seconds. A further reduction of the time frame used for a linear fit would reduce the number data points used in a fit to a level where the noise inherent in single data points has a great effect on the result of the fit. Despite those problems, the initial reaction velocity was determined by linear fit using a minimum of ten data points. Plotting the reaction velocity against the substrate concentration, the data for chicken ERp57 followed the Michaelis-Menten model rather well (Fig 3.3.7). In contrast, the data of both human and yeast PDI were not following Michaelis-Menten kinetic and could not be fitted well with a hyperbola. Both proteins showed a sigmoidal plot of reaction velocity versus substrate concentration as described by the Hill equation for cooperative ligand binding [106].

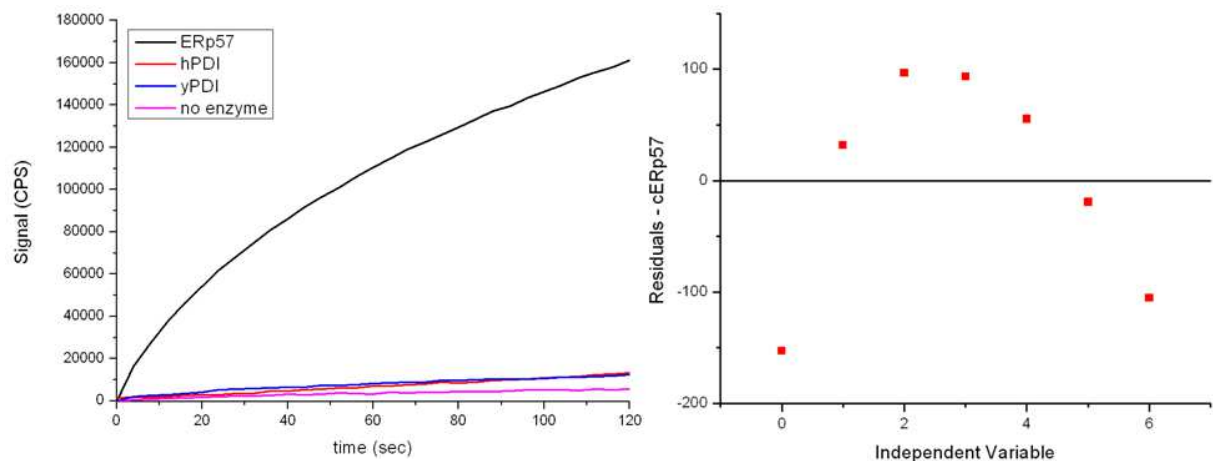
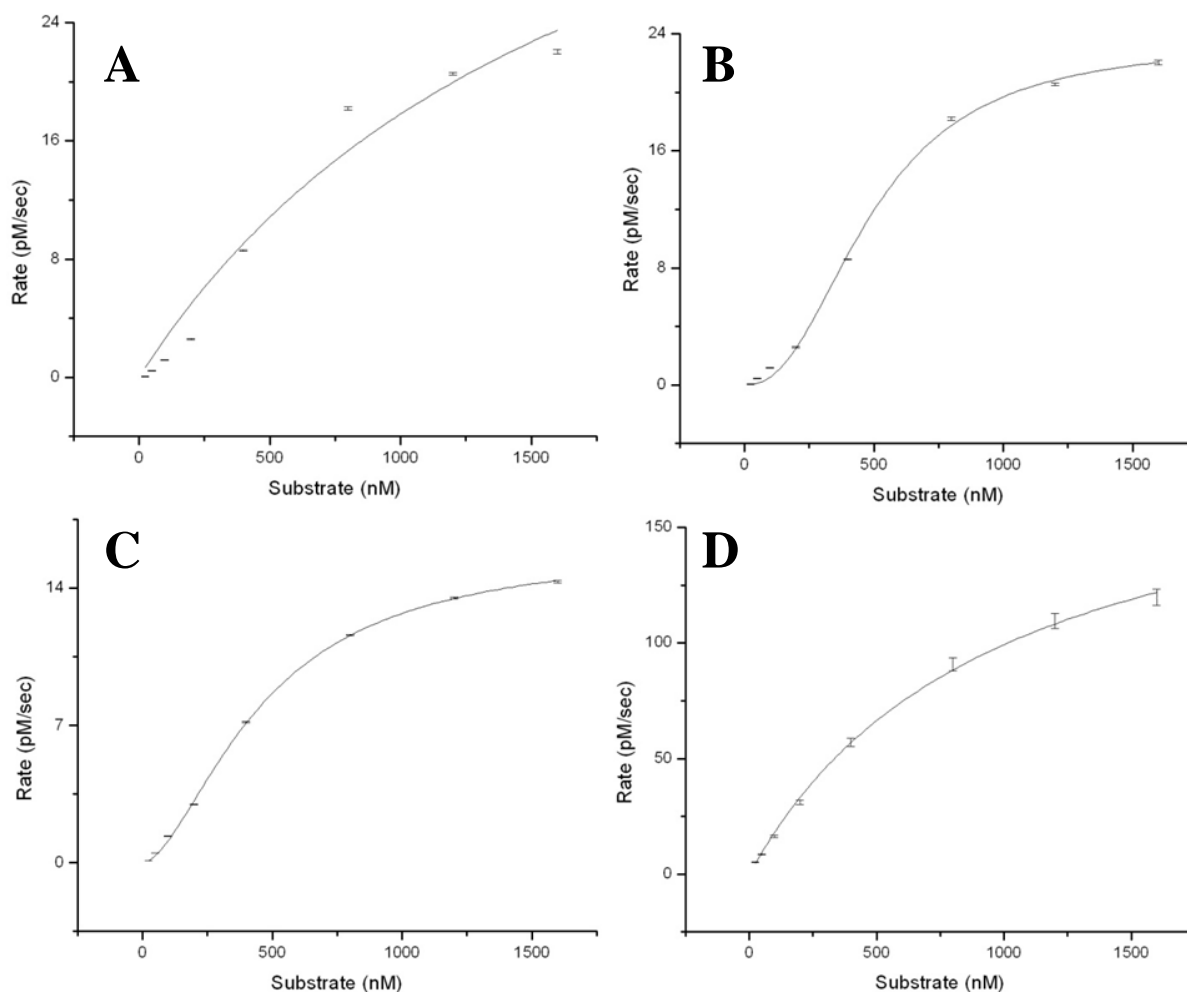


Figure 3.33 – Comparison of human and yeast PDI with chicken ERp57

The reduction of di-E-GSSG was observed either in the absence of enzyme or in the presence of human PDI (hPDI), yeast PDI (yPDI) or chicken ERp57 (cERp57) (left panel – 400 nM substrate). While the activity of the two PDI variants was very similar, the addition of chicken ERp57 caused a vastly accelerated reduction of the substrate. This led to the problem that the linear phase of the reduction is so short, that it cannot be properly resolved by the Fluoromax4 fluorimeter without a stopped-flow system. This is illustrated by the residuals of the linear fit for ERp57. Where the residuals of the fits for human and yeast PDI were randomly distributed around the base line, the residuals for ERp57 indicate a misfit (right panel). Yet a further reduction of the number of data points would greatly increase the influence of random noise.

As both Hill coefficients fitted to the data for human and yeast PDI are close to two, the number of active sites per PDI molecule, this might suggest a cooperative mode of substrate binding between both active sites. As both active sites face each other over the hydrophobic cleft (see Introduction for a model of PDI), they might indeed bind substrate together. On the other hand, cooperativity is usually moderated through conformational changes upon substrate binding of one active site, which are relayed through the protein to another active site. It has been shown, that redox changes in the a' domain result in conformational changes that are not limited to the a' domain, but are relayed through the flexible linker to at least the bb' core of PDI [89].

Also, the reduced reaction speed in the low concentration range might be explained by adsorption effects of the substrate, or the protein on the surface of the cuvette. These changes would lead to a more pronounced reduction of the reaction speed at low concentrations as the fixed amount of substrate that would be removed from solution would represent a greater fraction of the total.

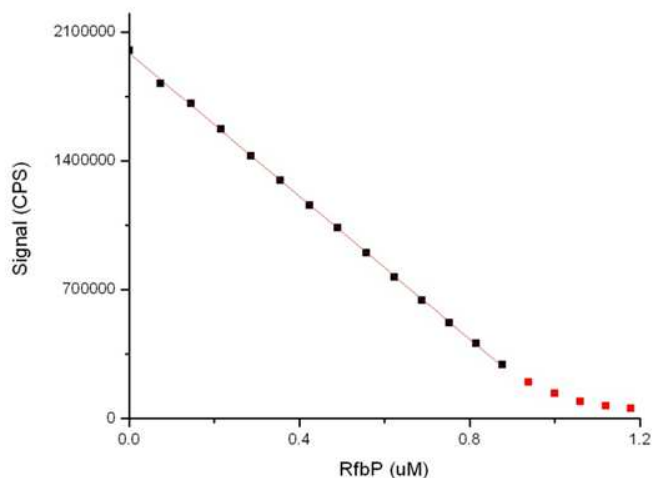


Michaelis-Menten fit				Hill fit				
Protein	K_m (nM)	v_{max} (pM/sec)	R^2 -factor	Protein	K_m (nM)	v_{max} (pM/sec)	Hill coefficient	R^2 -factor
yPDI	1810 ± 950	50.1 ± 16.2	0.9642	yPDI	492 ± 15.6	23.4 ± 0.70	2.34 ± 0.18	0.9980
hPDI	1001 ± 300	25.4 ± 3.58	0.9834	hPDI	289 ± 10.1	16.1 ± 0.34	1.68 ± 0.07	0.9994
cERp57	983 ± 76	196 ± 7.72	0.9983	cERp57	698 ± 61.7	167 ± 7.23	1.15 ± 0.05	0.9994

Figure 3.34 – Fits of the reduction of di-E-GSSG catalyzed by PDI or ERp57

The initial reaction rates were fitted to the di-E-GSSG reduction data (Fig 3.3.6) and plotted against the substrate concentration used. When the reaction rates of reactions catalyzed by yeast PDI were fitted using the Michaelis-Menten equation, the deviation of the data from the fit was rather large (R^2 value of 0.964, panel A). The Hill equation, which is an extension of the Michaelis-Menten kinetic used for cooperative enzymes, gave a much better fit of the same data (R^2 value of 0.998, panel B). The Hill equation also proved significantly better to fit the data from human PDI (panel C – Hill fit) whereas it made barely a difference for chicken ERp57 (panel D – Michaelis-Menten-equation). Although the Hill fit consistently gives a better R^2 value, this is mathematically to be expected due to the greater number of parameters.

Figure 3.35 – Standard curve of riboflavin fluorescence at 450/525 nm.



The fluorescence of a solution containing 1 μM of riboflavin was assayed for fluorescence with the settings as described in the Methods section (2.3.8). Apo-RfbP was then added, while keeping the total riboflavin concentration at 1 μM . The fluorescence decreases in a linear manner (red line indicates fit) up to a concentration of 900 nM RfbP. The data points marked in red have been excluded from the fit.

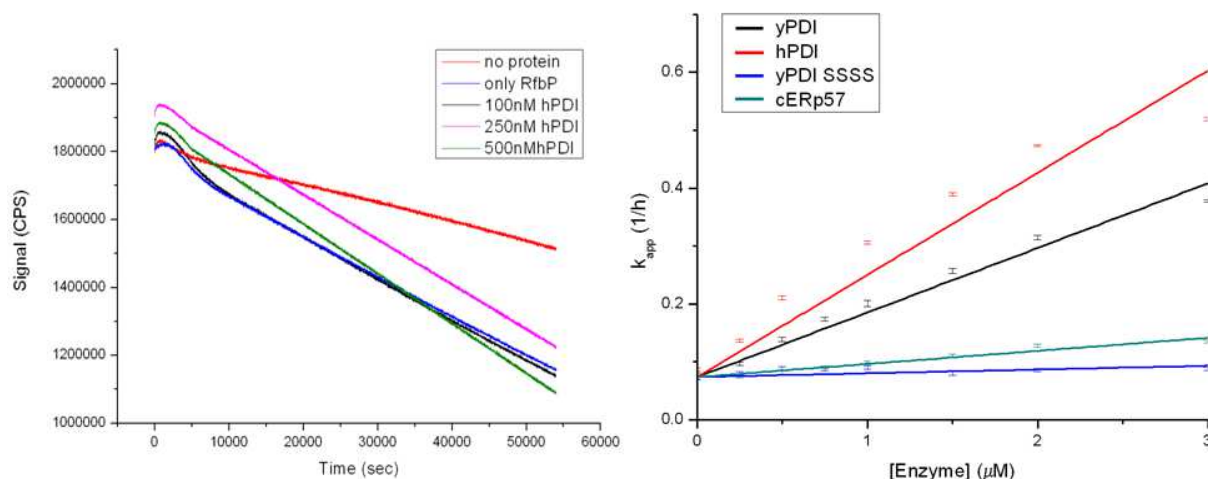


Figure 3.36 - Refolding of reduced RfbP

Reduced RfbP in 0.6 M guanidinium hydrochloride was diluted to a final concentration of 1 μM (6 mM guanidinium hydrochloride) into refolding buffer containing 1 μM riboflavin. The fluorescence of riboflavin was monitored at an excitation wavelength of 450 nm and an emission wavelength of 525 nm over 9 hours. After an initial surge, the fluorescence decreased in a linear fashion (left panel). As the fluorescence decreases significantly faster than the control without RfbP, the reduction can only in part be attributed to the bleaching of the riboflavin and can only be explained by the formation of native RfbP. The addition of PDI to the reaction increased the rate of folding in a concentration dependent manner. When the linear part was fitted and the rate was plotted against the concentration of the enzyme (after subtracting the 'no PDI rate'), a linear relationship could be discerned between the concentration of the folding catalyst and the apparent folding rate (right panel). This enables a comparison of the folding activity of each enzyme by comparing the slope of the linear plot.

At higher concentrations, filter effects could come into play as suggested by the reduction of di-E-GSSG by DTT alone. The reaction velocity values for DTT at substrate concentrations above 1 μM are lower than expected for the linear relationship between substrate concentration and reaction speed of a pseudo-first-order reaction. Such filter effects would further distort a hyperbolic into a sigmoidal curve, especially as the sigmoidal shape is more pronounced the lower the reaction speed as indicated by the Hill coefficients. Thus, while the reliability of the determined kinetic is questionable, the qualitative difference between ERp57 and PDI in their acceleration of the reduction is obvious. Yet this result is rather surprising, as ERp57 had been less active than PDI itself in classic assays such as RNase refolding [107] or insulin turbidity assays [108].

To put the surprising reductive capabilities of ERp57 in perspective, all three proteins were also tested in the refolding of riboflavin binding protein (RfbP). As described in the beginning of this chapter, this assay provides information about the combination of oxidase, isomerase and substrate binding capabilities of an enzyme. The very first step taken to establish a quantitative RfbP refolding assay is to establish a standard curve to relate the signal to the amount of native apo-RfbP formed. For this, the fluorescence of 1 μM riboflavin was determined with parameters designed to keep the fluorescence below two million counts per seconds (CPS), which is the upper limit for the linear detection range of the fluorimeter's detector. Then apo-RfbP, which had been prepared as described in the Methods section, was added while keeping the total riboflavin concentration at 1 μM . As expected, the fluorescence signal of the free riboflavin decreased as the added apo-RfbP bound the free riboflavin (Fig 3.3.8). The results could be fitted in a linear fashion yielding a factor of 1983 CPS/nM to convert the decrease in the fluorescence signal into the decrease in the concentration of free riboflavin in solution, which is equal to the increase in concentration of native RfbP. In contrast, addition of either reduced carboxymethylated RfbP which is permanently unfolded, or the addition of human PDI did not decrease the fluorescence signal. As the next step, denatured and reduced RfbP at a final concentration of 1 μM was added to 1 μM riboflavin in a redox refolding buffer (see Methods 2.3.8), and the fluorescence was monitored. Initial experimental setups with human PDI at a concentration of 100 nM revealed no difference in fluorescence within a time frame of 20 minutes, when compared to a setup containing no PDI or even no RfbP. When the time frame was significantly extended to 9 hours, a linear decrease in fluorescence was visible after an initial reaction phase. In the linear phase,

the fluorescence of assays containing RfbP decreased significantly stronger compared to just the free riboflavin. Thus, the additional decrease over the riboflavin only sample can only be attributed to the formation of native RfbP. When human PDI in varying concentrations was added to the refolding reaction, the rate increased, although it took near stoichiometric concentrations to achieve a significant increase in the refolding rate (Fig. 3.3.9).

$$\text{Eq.3.3.7} \quad v = k \times [\text{RfbP}] \times [\text{PDI}]; k_{app} = \frac{v}{[\text{RfbP}]}$$

$$\text{Eq.3.3.8} \quad y = y_0 + A \times e^{-x \times t} - b \times x$$

As the Michaelis-Menten kinetic is not applicable at those concentrations and, anyhow, the multi-step reaction of refolding cannot be described by this model, a simple model of the apparent rate comparison was used to fit the data. This model treats refolding as a pseudo-first-order reaction (eq. 3.3.7) with the concentration of RfbP staying constant. The model expects a

	RfbP refolding	Inserted flap reduction
Protein	in % (yPDI)	in % (yPDI)
yPDI	100.0 ± 3.8	100.0 ± 4.8
hPDI	158.5 ± 12	78.2 ± 1.6
cERp57	20.3 ± 1.2	6.89 ± 0.7
yPDI SSSS	5.8 ± 0.2	---

Table 3.2 – Comparison of the activities of selected folding catalysts

The k_{app} data was plotted against the enzyme concentration and fitted with a linear model. The correlation factor obtained in this way was normalized (yPDI = 100%) for a comparison of the activities of the compared enzymes in both the RfbP refolding assay and the IFS reduction assay.

linear relationship between the concentration of the folding catalyst and k_{app} , which represents the reaction rate divided by the RfbP concentration. As can be seen, for each tested folding catalyst k_{app} increases with the concentration of the folding catalyst and the increase can be reasonably fitted with a linear equation (Fig 3.3.9 right panel). While this means that the kinetic parameters K_M and v_{max} cannot be determined

using this assay, a general folding capability can be assayed using the rate constant – as determined by the slope of the fitted k_{app} – as a measure. This allows a qualitative comparison between the folding catalysts. As can be seen by comparison, the general refolding capabilities of both human and yeast PDI are comparable (Table 3.3.1). As expected, the PDI SSSS mutant without active site cysteines showed barely any activity. Chicken ERp57 shows only a very low activity in the refolding of RfbP. This is in stark contrast to the results of the FSQ assay where ERp57 outperformed both PDI variants in the reduction of the low molecular weight substrate di-E-GSSG.

The third assay was developed by my collaboration partner Stefan Lorenz from the group of Prof. Franz X. Schmid at the University of Bayreuth. The substrate used for this assay is a mutant version of the inserted flap domain (IFD), which is part of the phage protein BGX3. This domain relies on the spatial proximity of its N and C termini to remain folded. In the BGX3 protein, this is the case as the inserted flap domain is inserted into a loop of another domain and the fold of the parent domain ensures the spatial proximity of the IFD termini. When expressed and purified in isolation, the IFD remains unfolded. Only when a cysteine is introduced at each end of the polypeptide chain, the formation of a disulfide brings the termini together and triggers a spontaneous folding of the domain [105]. The introduced cysteines are the only cysteines present in this IFD mutant and thus there is only one intramolecular disulfide bond possible. Upon reduction of the disulfide bond the domain unfolds rapidly within 200 ms. The folding status of the domain can be determined by the fluorescence of the only tryptophan residue present in the mutant IFD. When excited, the emission spectrum of a tryptophan residue is influenced by the polarity of its microenvironment shifting to shorter wavelength in a hydrophobic environment.

As of the writing of this thesis, the assay based on this redox sensitive substrate has not been published and I was asked to refrain from using this assay until it was published. Therefore it is only possible to describe and use the data gathered by myself together with Stefan Lorenz in the lab of Prof. Schmid. As a first step, the reduction of inserted flap substrate (IFS) by DTT was tested. For this 1 μ M of IFS was added at time point zero into reduction buffer (100 mM potassium phosphate pH 7.0; 2 mM EDTA, 5 μ M DTT) and the fluorescence was monitored. The excitation wavelength was 295 nm at which the fluorescence of tryptophan residues

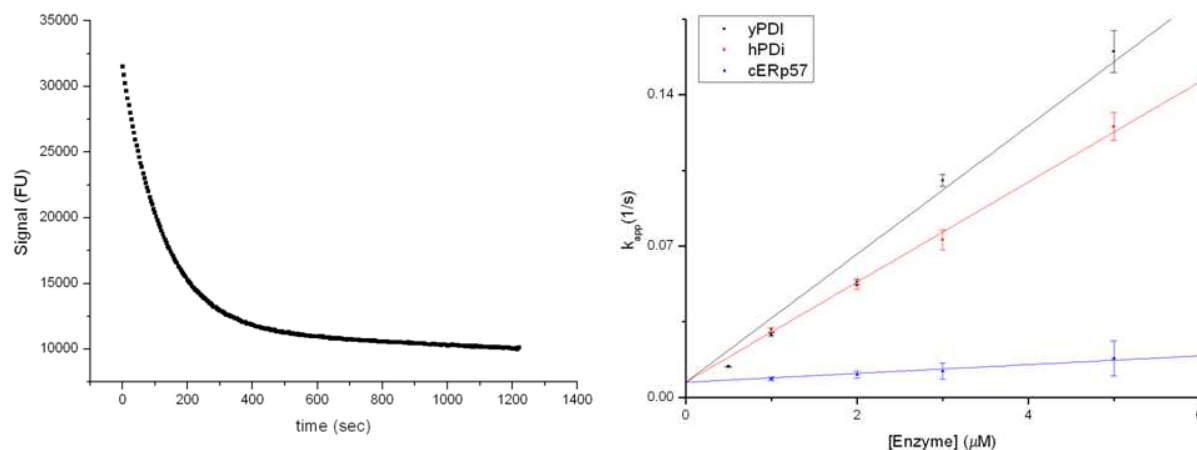


Figure 3.37 – Reduction of the inserted flap domain substrate

The inserted flap domain substrate, which is a mutant of the inserted flap domain of the phage protein BGX3 containing one cysteine at each the N and C terminus, was added to reduction buffer (100 mM potassium phosphate pH 7.0, 2 mM EDTA, 5 μ M DTT) to a final concentration of 1 μ M. The tryptophan fluorescence was monitored with an excitation wavelength of 295 nm and an emission wavelength of 330 nm. Upon reduction, the IFS unfolds placing the tryptophan in a polar environment, thus reducing the emission at 330 nm (left panel). The decrease in fluorescence was fitted with a mono-exponential decay function with the exponential factor being used as k_{app} . The experiment was repeated in the presence of the respective enzyme. The k values gained from each fit yielded a linear relation when plotted against the enzyme concentration (right panel). The correlation factor was then used to compare the catalytic abilities of the tested enzymes (Table 3.3.1).

dominates over the signal of tyrosine or phenylalanine residues. The emission was monitored at 330 nm which is the emission maximum for tryptophan residues buried inside a folded protein whereas the emission maximum of a tryptophan surrounded by water is at 352 nm [109]. Over time the monitored fluorescence signal dropped as the IFS was reduced and unfolded. During the unfolding the tryptophan used as fluorescent probe becomes surrounded by water and shifts its emission spectrum to higher wavelengths (Fig 3.3.10 left panel). These data were fitted with a mono-exponential decay function expanded by a linear term to account for the bleaching of the fluorescent probe (eq. 3.3.8). The time constant t was used as an apparent rate constant to characterize the rate of the reaction. Preliminary experiments by Stefan Lorenz showed no effect of folding catalysts at catalytic concentrations. At stoichiometric concentrations, however, the rate constants increased, showing a linear dependency on the concentration of the folding catalyst (Fig 3.3.10 right panel). From the plot of the apparent rate constants against the

concentration of folding catalyst the general catalytic ability for this reaction can be compared by a comparison of the slopes of the linear fit. As can be seen by a comparison of the slopes, human and yeast PDI both have comparable reductase activities for the inserted flap substrate (Table 3.3.1). In contrast, the activity of ERp57 is significantly lower. These results are similar to the results gained from the RfbP assay. In both cases, ERp57 displayed only an activity that was barely above the background while the two PDI variants were similar in their activity and significantly accelerated the reaction.

As outlined above, each of the three assays is ill-suited for the exact determination of the kinetic parameters K_M and v_{max} . As established the fluorescent self-quenching assay suffers from purity issues of the substrate and the data processing is difficult. For very active proteins such as chicken ERp57, neither the current setup of the Fluoromax4 fluorimeter nor the Fluostar plate reader can gather enough data points to allow a fit of the linear portion of the reaction. For this a fluorimeter equipped with a stopped-flow system is needed to obtain data during the short initial period of the reaction. For proteins with low activity such as human or yeast PDI, the filter effects at higher substrate concentrations make it very difficult to properly fit a Michaelis-Menten curve to the reaction velocities as such a fit is prone to error, if there are no data points at concentrations significantly higher than the K_M . Yet the method has also shown its merit to compare the reductase ability of the active sites in a qualitative fashion. Likewise the refolding assay using the riboflavin binding protein as well as the reduction assay with the inserted flap substrate suffers from the low activity of the folding catalyst necessitating in both cases stoichiometric concentrations of the enzyme. This means that those two assays too are not suitable for the determination of K_M and v_{max} , but they allow for an accurate comparison of activity. Nevertheless, the assays still represent very useful tools as each of the three methods probes a different set of functions and activities which can vary between different PDI family members. The activity of ERp57 greatly exceeds that of either human or yeast PDI when reducing a low molecular weight substrate. Yet it struggles to catalyze a refolding reaction. This indicates a problem with either substrate binding or isomerization as the FSQ assay has shown ERp57's great oxidoreductase activity. Yet ERp57 is also a very weak catalyst for the reduction of the inserted flap substrate, suggesting that the substrate binding and recognition ability of ERp57 is very poor. This is perfectly in line with the literature where ERp57 has been shown to work in collaboration with lectins such as calnexin or calreticulin which are thought to handle

the substrate recognition [99]. This further demonstrates that the combination of these three assays is still a very powerful tool to dissect the activity and role of various PDI family members.

3.4 Substrate recognition of the PDI family

With an estimated one third of all proteins undergoing oxidative folding [40], the PDI family has to interact with an extremely heterogeneous group of substrates. There have been studies reporting that PDI can attack native disulfide bonds in folded proteins [45, 110]. These results suggest that PDI might interact with folded and unfolded proteins alike, however, only members of the latter group would alter their conformation in the process. On the other hand those studies have been carried out under non-physiological redox conditions. Since chaperones in general are able to recognize hydrophobic residues and/or unstructured backbone regions associated with unfolded proteins [111], an alternative model would suggest that PDI selectively interacts with unfolded proteins relying on biochemical properties to discriminate the folding status of potential substrates. Using isothermal titration calorimetry (ITC), I investigated how PDI and selected family members interact with protein substrates. During an ITC experiment, PDI was added to potential substrate proteins like RNase A or riboflavin binding protein (RfbP). I compared the interaction between PDI and substrates with respect to the folding status of the substrate as well as the redox status and domain architecture of the respective PDI family member.

To investigate substrate recognition by members of the PDI family model substrates need to cover a range of molecular weights and, at the same time, contain a significant number of disulfide bonds. Selecting by those criteria, I chose bovine pancreatic trypsin inhibitor (BPTI), bovine pancreatic RNase A (RNase A) and riboflavin binding protein (RfbP) purified from chicken egg white that had already been employed in other studies as refolding substrates for PDI as well as bovine serum albumin (BSA) as a substrate with a higher molecular weight. All substrates were denatured and carboxymethylated as described in chapter 2.3.5 to generate modified and permanently unfolded substrates (denoted by 'RCM' for reduced carboxymethylated). Upon removal of guanidinium chloride through dialysis, RCM BPTI precipitated and could not be used in further experiments. RCM RNase A, RCM BSA and RCM RfbP, however, remained in solution up to concentrations of at least 150 μ M in the presence of 250 mM NaCl.

To confirm that the substrates are unfolded even in the absence of denaturant I performed CD spectroscopy with samples dialyzed against 50 mM potassium phosphate buffer and 250 mM

NaCl. The modified forms of both RNase A and RfbP generated similar spectra that were consistent with a random coil conformation, thus indicating that both proteins are in the unfolded state [112]. As expected the unmodified proteins exhibited CD spectra that were consistent with folded proteins and also their known three-dimensional structures (Fig. 3.37 A). In contrast RCM BSA displayed a less pronounced change in the CD spectrum compared to native BSA (Fig. 3.37 B), probably indicating that the protein retained at least some features of its native conformation. For initial ITC experiments we therefore employed only RfbP and RNaseA to investigate in each case the differential interactions of PDI with either a folded or an unfolded substrate.

To ensure a complete modification and uniformity of our modified substrate proteins, we tested samples of RfbP by native PAGE. We only observed a single band for RCM RfbP which migrated differently than unmodified RfbP (Fig. 3.38 A), suggesting that all thiol-groups had been modified. In order to account for possible aggregation, RCM RfbP was analyzed by dynamic light scattering (Fig. 3.38 B) and analytical size exclusion chromatography (Fig. 3.38 C). Neither method revealed significant amounts of aggregates and both methods indicated that RCM RfbP remains a monomeric protein. In general, our samples of the RCM-modified substrate proteins only contained a single species and these homogenous proteins were used in the ITC experiments. In order to ensure that the modification had been complete we also loaded a sample where the excess of the modifying agent over the free thiols had been increased from 1:2 to 1:20. We could observe additional weakly defined bands, most likely representing modifications of histidine residues.

In ITC control experiments I titrated native human PDI into a sample cell containing only buffer. Upon injection, there is a significant uptake of heat that is dependent on the concentration difference between the injectant and the current value in the cell (Figure S2). We have interpreted this signal to represent the dissociation of dimeric PDI upon dilution in the cell and could fit the heat from this experiment to a dissociation constant (K_D) of 1.15 mM. This behavior is consistent with the observation of a dimeric form of PDI during the purification as well as with reports of homologous PDIs forming homodimers [113-115]. This background heat has been subtracted from all ITCs using a separate buffer run for each tested protein as reference (see

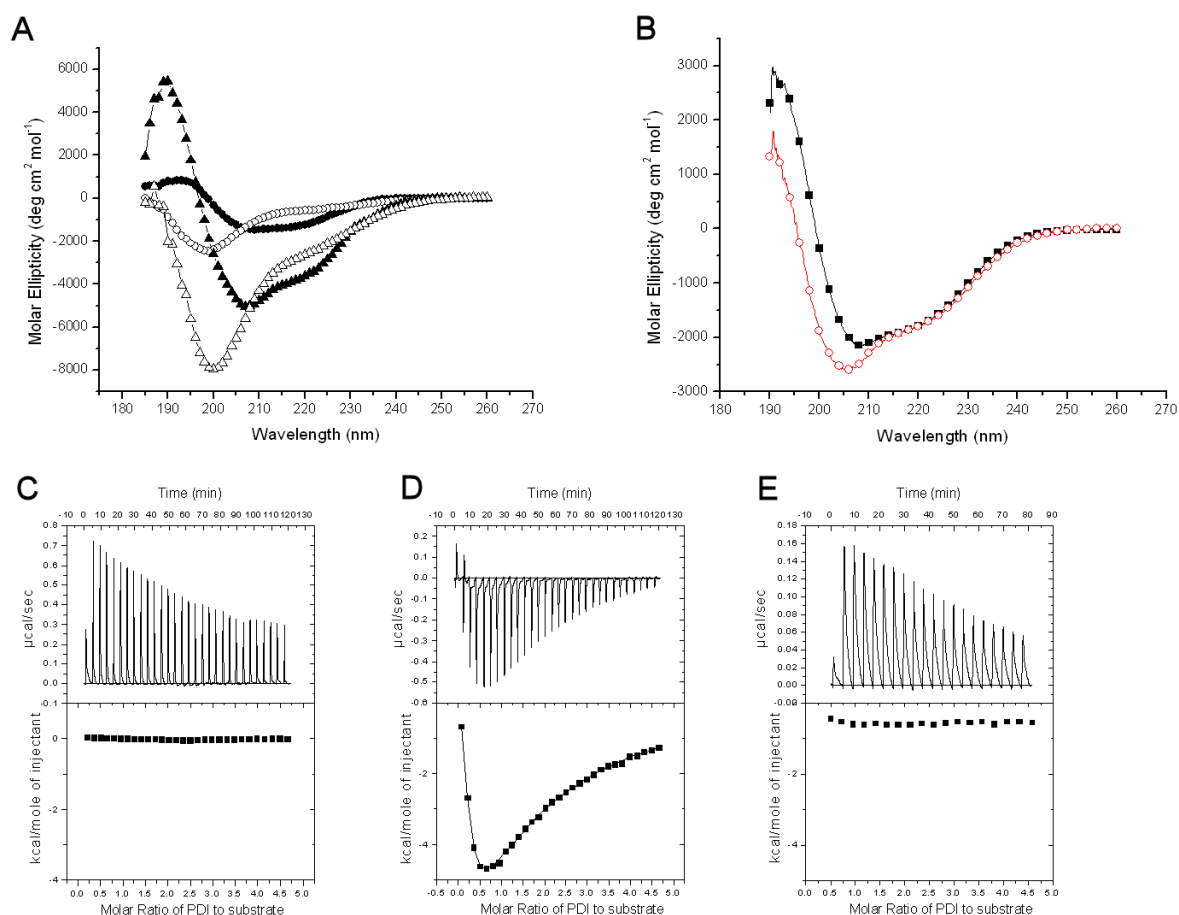


Figure 3.38 – CD spectroscopy of model substrates and initial ITC experiments

PDI only interacts with unfolded proteins. (A) CD spectra of RNase (■, □) and RfbP (●, ○). Scans using native proteins are displayed with filled symbols, whereas scans using denatured reduced carboxymethylated (RCM) proteins are displayed with empty symbols. (B), CD spectra of native (filled) or RCM BSA (empty). (C-E), ITC experiments injecting human PDI into native RfbP (C), RCM RfbP (D), or RCM BSA (E). For every ITC displayed in this article, the lower panel of the graph has been corrected by subtracting the buffer control (see Figure S2).

supplementary data Figure S2). When injecting PDI into cells containing native unmodified RfbP the heat release did not differ from the injections into buffer (Fig. 3.37 C), thus indicating that PDI does not interact with folded RfbP.

In contrast, RCM RfbP elicited a different response (Fig. 3.37 D). After subtraction of the PDI dilution control, a substantial and complex heat signature remained which could not be fitted reasonably using a one site binding model. However, the data could be fit applying a two site

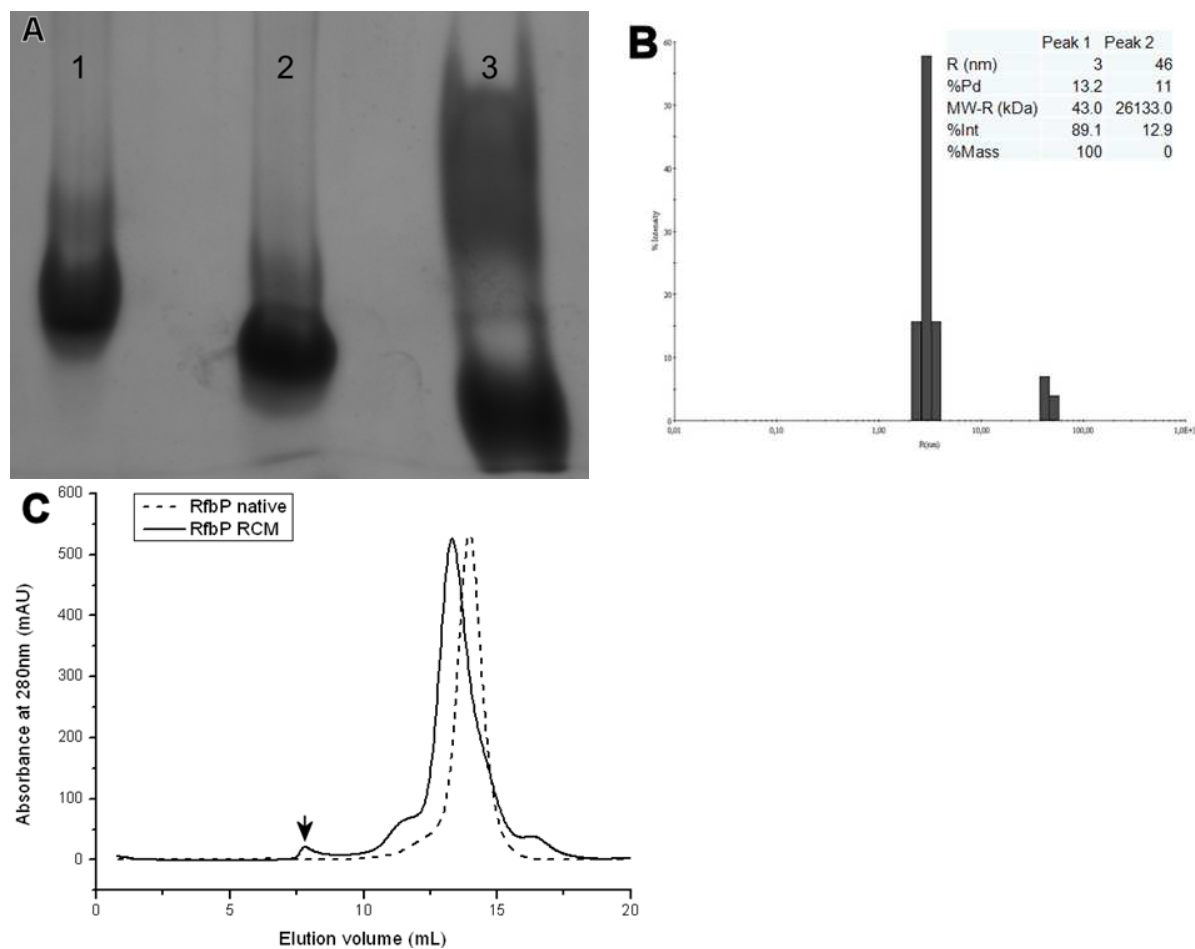


Figure 3.39 – Effects of carboxymethylation

(A) Native polyacrylamide gel electrophoresis of unmodified RfbP (lane 1), reduced and carboxymethylated (RCM) RfbP modified as described in chapter 2.3.5 (lane 2), and RfbP modified with a 10-fold greater excess of iodoacetic acid (lane 3). (B) Dynamic light scattering of a solution containing 10 μ M RCM RfbP. Two peaks are detected; peak one (100% of the mass and 89% of the intensity) corresponds to a RCM RfbP monomer (calculated molecular weight of 26 kDa compared to an observed mass of 43 kDa) with a low polydispersity of 13.2%. (C) Size exclusion chromatography analysis on a Superdex 200 10/100GL column. Only a very small amount (<5%) of the modified RCM RfbP forms an aggregate (arrow), eluting at the void volume of 8 mL. The modified protein has a higher apparent molecular weight, presumably due to an elongated form of the unfolded protein compared to the near-spherical shape of the native protein.

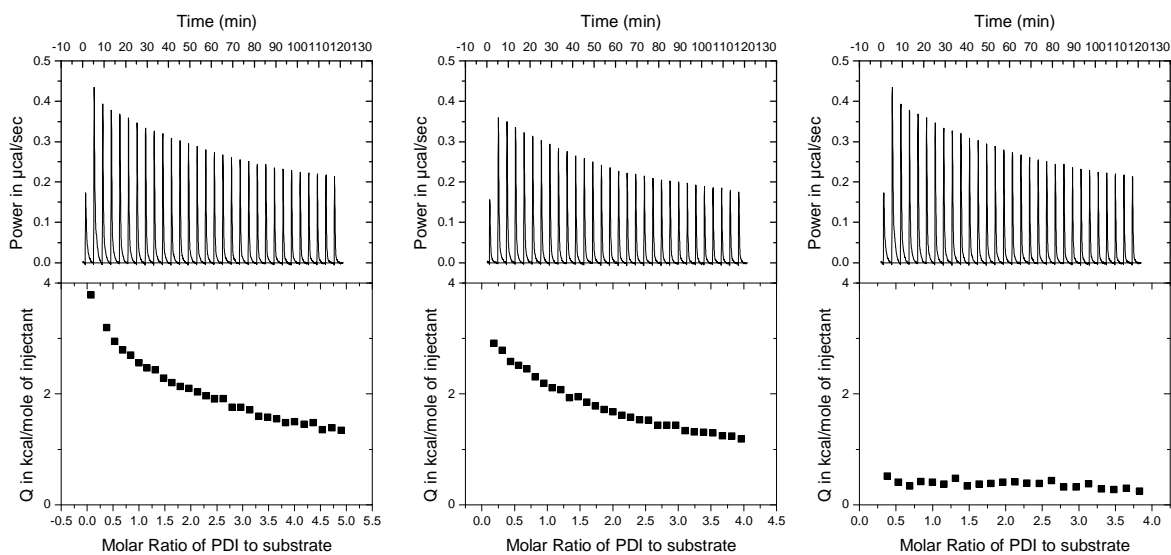


Figure 3.40 – Buffer correction of ITC experiments

ITC control titrations. (A) ITC experiment in which wild-type human PDI is titrated into native RfbP illustrating the heat of dissociation of PDI upon dilution. (B) Titration of wild-type human PDI into buffer only, illustrating that the resulting signal is identical to (A) and hence PDI does not interact with native RfbP. (C) For every ITC displayed in this thesis, the lower panel of the graph has been corrected by subtracting the buffer control. This can result in graphs where the raw data shows a heat release (C upper panel, which is identical to the upper panel of (A)) but the integration shows no signal relative to the background ((C) lower panel, resulting from the subtraction of the data shown in (B) from those shown in (A)).

binding model resulting in: (1) One binding event that is characterized by a stoichiometry of approximately seven RfbP molecules to one molecule of human PDI. This binding is driven by an increase in entropy and is characterized by an affinity in the low micromolar range ($K_d=1.6\pm 0.8 \mu\text{M}$) and at the same time is enthalpically opposed. (2) A second binding which is characterized by a stoichiometry of two PDI to one substrate molecule and is almost two orders of magnitude weaker ($K_d=117\pm 13\mu\text{M}$) than the entropy driven binding. This binding process is driven by an increase in enthalpy but is entropically opposed. With two simultaneous binding processes within one signal, it is difficult to properly derive the thermodynamic parameters of each binding as the fitting process can converge to multiple, slightly differing local minima. Thus, from these initial experiments I could only conclude that human PDI can distinguish between the folded and unfolded states of a protein and interacts with denatured RfbP in a complex manner.

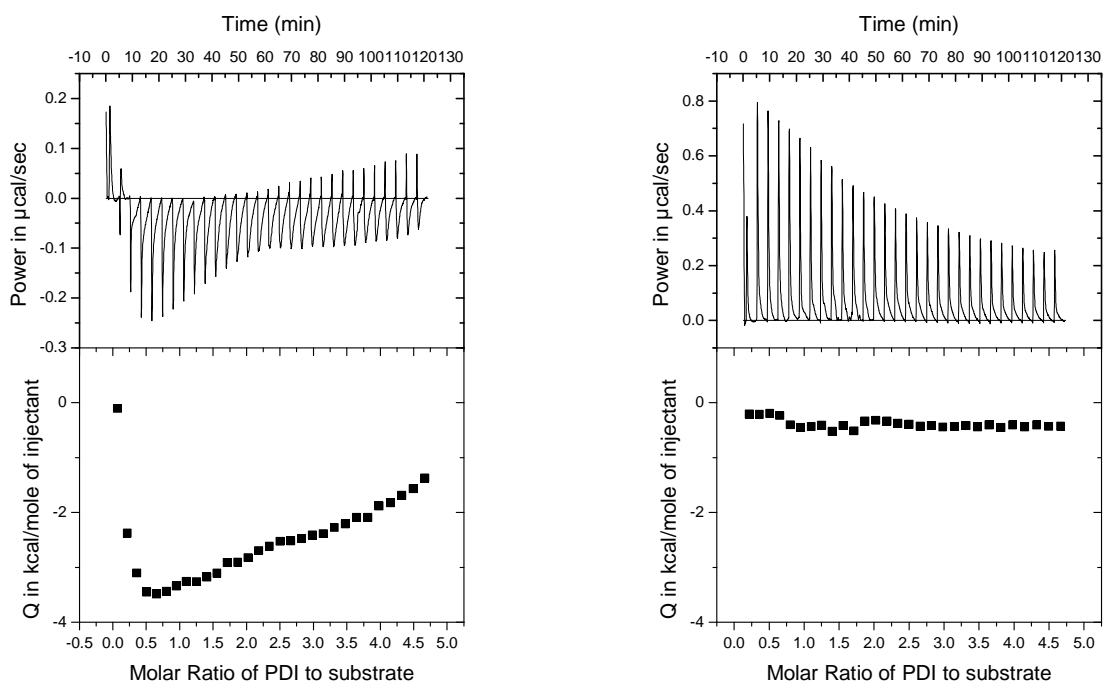


Figure 3.41 – RNase A as binding partner for human PDI

Binding of PDI to RNase A. Wild-type human PDI is titrated either into RCM RNase A (A), or native RNase A (B). This results in a biphasic binding or no interaction, respectively.

I subsequently carried out ITC experiments with PDI and native and modified RNase A. The results only differ marginally from the results obtained with RfbP (Fig. 3.40). When we titrated PDI into a cell containing RCM BSA, the signal did not differ from the background (Fig. 3.37 E) despite the modification. This is consistent with the CD spectra of BSA which displayed only minor differences between the native and modified forms of this protein, thus confirming that PDI only interacts with denatured substrates. More importantly this behavior excludes the possibility of PDI specifically recognizing carboxymethylated cysteine residues.

With PDI being able to recognize unfolded substrates, I investigated selected members of the PDI family for comparison. The first paralog under investigation was ERp57, and just like PDI, ERp57 did not show any binding to native RfbP. However, this enzyme also did not interact with RCM RfbP since the heat signature did not differ from that of the dilution run (Fig. 3.41 A). The observed absence of a binding signal is in line with the previously published results indicating that ERp57 interacts with its substrates only indirectly via the lectins calnexin and calreticulin [116].

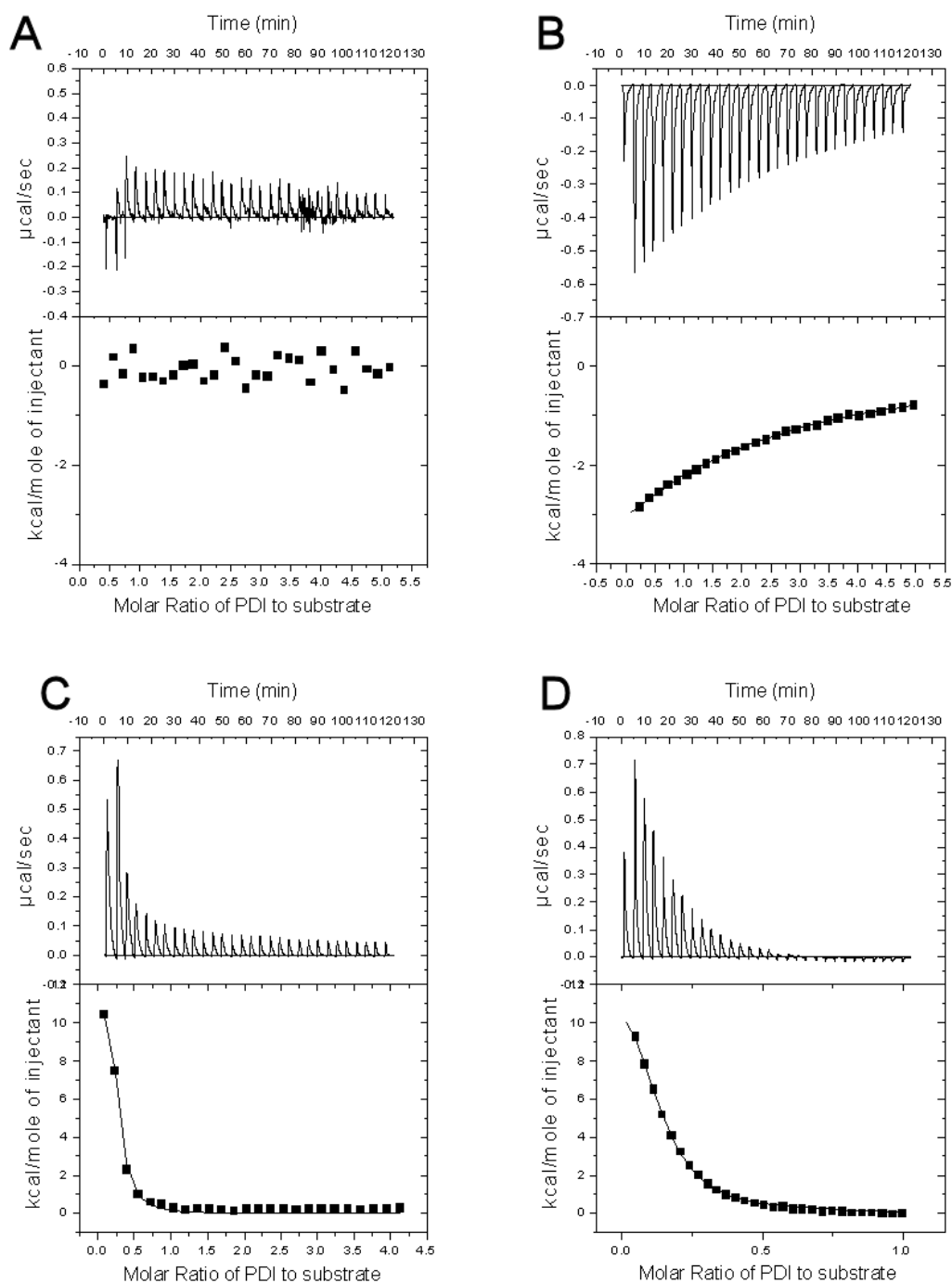


Figure 3.42 – Substrate binding of selected PDI family members

Some PDI family members display only one binding mode. (A-C) ITC experiments with human ERp57 (A), human ERp27 (B) and human P5 (C) as titrant and RCM RfbP as substrate. (D), ITC titration with P5 at a concentration of 500 μM and RCM RfbP at a concentration of 100 μM .

Human ERp27 is a family member lacking catalytic domains, yet it contains the bb' core. When ERp27 was titrated into native RfbP, there was no visible difference from the buffer control (data not shown). In contrast, ERp27 generated a distinct heat signature when titrated into RCM RfbP (Fig. 3.41 B) indicating a binding very similar to the enthalpy-driven, second binding mode of human PDI while lacking the entropy-driven binding. With only one binding mode being

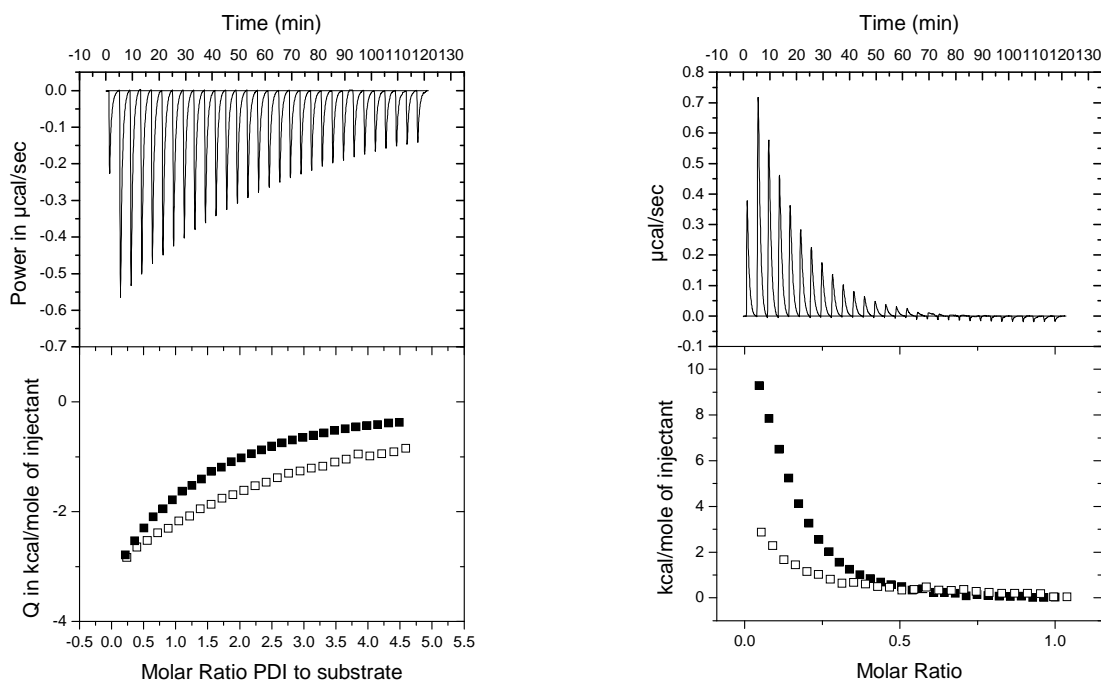


Figure 3.43 – Influence of temperature on ITC results

Comparison of ITC experiments at 10 °C and 25 °C. ERp27 (A) or P5 (B) were titrated into RfbP-RCM either at 25 °C (filled symbols) or 10 °C (empty symbols). Only the titrations at 25 °C are displayed in each top panel.

present, the thermodynamic parameters of the binding could be fitted with confidence (see Table 3.3), resulting in a K_d value of $168 \pm 17 \mu\text{M}$.

Human P5 (PDI_A6) also displayed the ability to distinguish between the folded and unfolded states of RfbP. As with all the other investigated proteins it showed no sign of binding to the native substrate (data not shown) but when titrated into RCM RfbP it displayed binding driven by entropy similar to the first binding mode of PDI (Fig. 3.41. C). To further investigate this binding, which under the initial conditions is only described by very few data points, the concentration of the titrant was reduced resulting in a run that ended with equimolar amounts of P5 and RCM RfbP. With more data points available for the fit, the thermodynamic parameters

Protein	$K_d(\mu\text{M})$	ΔH (kcal/mol)	$-T\Delta S$ (kcal/mol) [#]	N*
PDI (H)	78±10	-11.5±1.5	5.9	1.95±0.18
PDI (S)	1.6±0.8	3.1±1.1	-11.0	0.15±0.01
ERp57	n.d.	n.d.	n.d.	n.d.
PDIp	82±5	-15.7±0.7	10.2	1.84±0.05
P5	4.7±0.3	14.0±0.4	-21.1	0.14±0.01
ERp27	168±17	-7.4±0.9	2.2	1.90±0.16
PDI a domain	5.2±1.3	6.3±2.9	-11.7	0.18±0.02
PDI bb' domains	83±9	-11.9±2.1	6.6	1.93±0.21

Table 3.3 – Summary of thermodynamic parameters

Binding parameters as fitted to the integrated heat signatures. PDI (H) and (S) represent the enthalpy-driven and entropy-driving binding modes of human PDI, respectively. (#) Contribution of entropy to the binding is calculated from the affinity and the observed enthalpy at T=298 K. Since this quantity is not directly determined experimentally, no standard deviations can be given. (*) Stoichiometry of the binding as derived from the ITC analysis. One molecule of substrate interacts with n PDI molecules.

could be determined with confidence yielding a dissociation constant of 4.7±0.3 μM (Fig. 3.41 D).

To further test the different binding modes, additional experiments with ERp27 and P5 were performed at 10° C. I could observe similar signals at 10 °C compared to 25 °C when titrating ERp27 into RCM RfbP (Fig. 3.42 A). In contrast, P5 when titrated into RCM RfbP generated a less pronounced heat signature at the lower temperature (Fig. 3.42 B). This is in line with the contributions of the entropy term ‘-TΔS’. For ERp27 the contribution of the entropy is close to zero and therefore a change in temperature influences the observed enthalpy less than in the case of P5, where binding is driven by a change in entropy. Unfortunately, it is not possible to determine the binding enthalpies accurately enough to calculate the heat capacity of binding.

With three members of the PDI family already showing three different binding behaviors towards an unfolded protein substrate, I characterized the binding behavior of four additional

















<i>Protein</i>	<i>Domain architecture</i>	<i>ITC</i>	<i>Binding mode</i>
PDI			H+S
ERp57			n.d.
PDIp			H
P5			S
ERp27			H
PDIr			S
ERp18			S
ERp46			n.d.

Figure 3.44 – Domain architecture and binding modes of PDI family members

Summary of binding modes (H=enthalpy-driven and S=entropy-driven) together with the domain architectures of the respective PDI family member (n.d. means no binding could be detected). Under the heading ITC the raw ITC data are displayed

family members, namely PDIp, PDIr, ERp18 and ERp46, in order to gain a more complete picture of substrate interaction of the PDI family. As shown in Figure 3.43, PDI is the only protein simultaneously displaying both binding behaviors, while there are multiple family members which exhibit either the entropy-driven (P5, PDIr and ERp18), or the enthalpic binding mode (PDIp and ERp27), or do not interact with substrate at all (ERp57 and ERp46). In all cases where binding could be detected, PDI family members were able to distinguish between the folded and unfolded states of their offered substrates. The affinity for the observed entropy-driven binding modes was in a similar range with K_d values varying between 3 and 15 μM and a stoichiometry of approximately 6-8 substrates per PDI family member. In contrast, the enthalpy-driven binding is more than one order of magnitude weaker (K_d values in the range from 60 to 200 μM) with a stoichiometry of two enzyme molecules per substrate.

The survey of the PDI family suggested a relationship between the binding mode and the domain structure of the investigated protein. In all cases where an entropy-driven binding event was observed, the enzyme in question contained at least one a-type domain. In contrast, occurrences of enthalpy-driven binding correlated precisely with the presence of at least one b-type domain

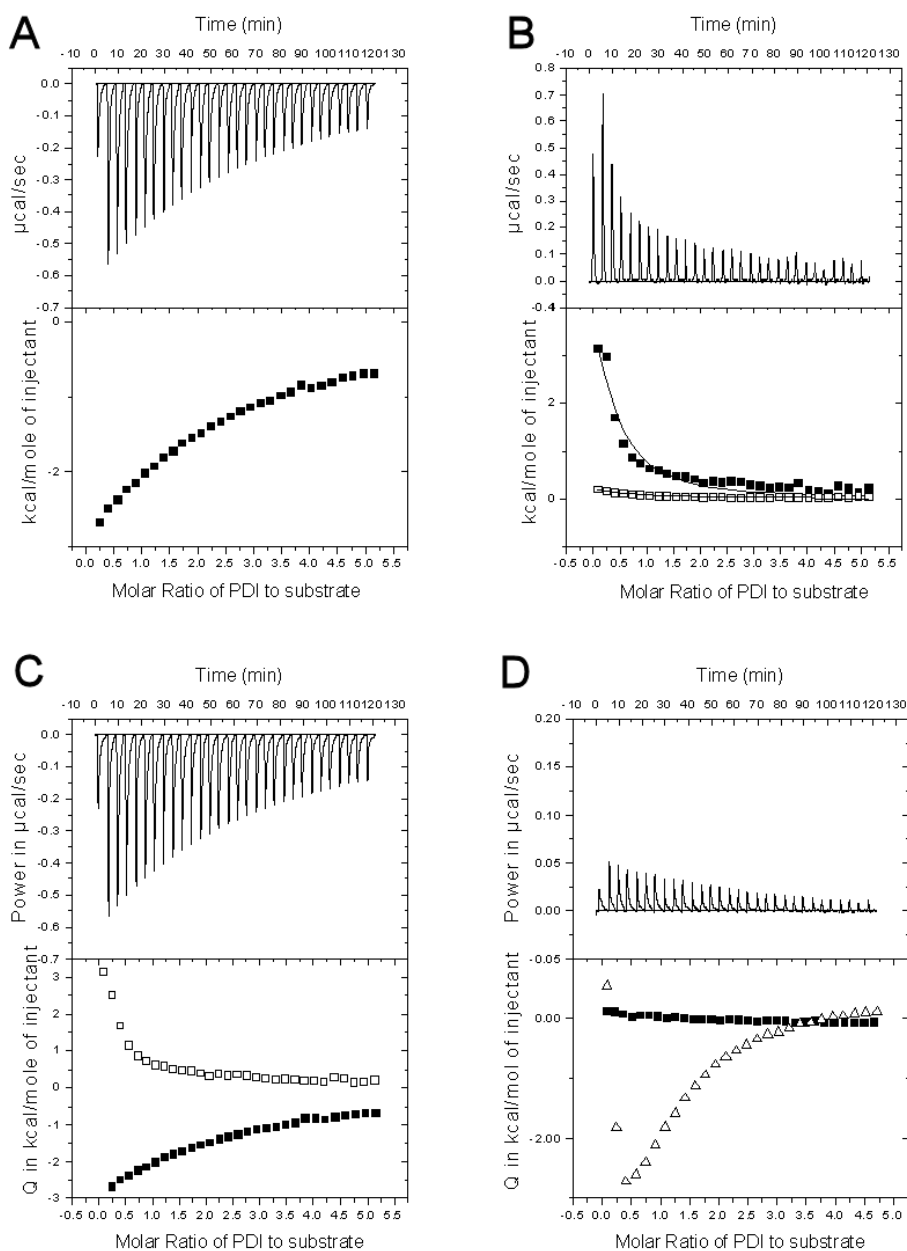


Figure 3.45 – Correlation between binding behavior and domain architecture

Mapping of binding modes to the domains of human PDI. (A) ITC analysis of the bb' fragment of human PDI (A136-K352) and its interaction with RCM RfbP. (B) ITC experiment injecting either the a domain (E1-P135; filled symbols) or the a' domain (K366-D476; open symbols) into RCM RfbP. The displayed raw data in the upper panel correspond to the filled symbols. (C) ITC experiment injecting either the N-terminal half (E1-L234; empty symbols) or the C-terminal half (P235-L508; filled symbols) into RCM RfbP. The raw data displayed in the upper panel correspond to the filled symbols. (D) ITC experiment with the mobility restricted mutant of yeast PDI and RCM RfbP. For comparison the data using the yeast PDI wild-type has been added to the lower panel (open symbols).

with a hydrophobic cleft, however, there were also cases (ERp57, ERp46, PDIr and PDIp) where no binding could be observed despite the presence of appropriate domains.

To better understand the contribution of a-type and b-type domains to substrate binding we therefore tested multiple truncations of human PDI for their interactions with the RfbP substrate. The first construct to be analyzed featured the bb' domain core of PDI. When titrated into RCM RfbP the bb' domains only displayed an enthalpy-driven binding with thermodynamic parameters similar to the enthalpic component of the full-length protein (Fig. 3.44 A). In contrast, the a domain interacts with RCM RfbP, and this binding event is driven by entropy. Surprisingly the a' domain does not bind RCM RfbP at all (Fig. 3.44 B). When I split PDI into its N- and C-terminal halves, the N-terminal part containing the a and b domains displayed only entropy-driven binding while the C-terminal part showed enthalpic binding (Fig. 3.44 C). Together these experiments mapped the entropy driven binding to the isolated a domain, whereas the b' domain is solely responsible for the enthalpy-driven binding.

In order to be fully functional, PDI requires its domains to be flexible with respect to each other, as has been previously shown for yeast PDI (12). As soon as the domain flexibility was restricted by the introduction of additional disulfide bonds, tethering the a and a' domain to the bb' core, *in vitro* and *in vivo* catalysis of disulfide bond formation was greatly diminished, or entirely abolished. When I titrated a mobility-restricted mutant containing the mutations E123C, S247C, N306C and D441C into RCM RfbP, no binding could be observed (Fig. 3.44 D). This is in line with a model suggesting that the opening and closing of the a-type domains allows access to the substrate-binding site (12).

Based on biochemical experiments and the structures of the bb' domains from human PDI as well as full-length PDI of *S. cerevisiae* a substrate binding site in the b' domain has been identified which is formed by a hydrophobic cleft in the b' domain [63, 89, 117, 118]. Sequence and structural comparisons of the b' domains of several PDI family members [97] revealed that the hydrophobic cleft is occluded by a network of hydrogen bonds in ERp57 and ERp72. Based on the structural alignment I introduced three point mutations (F249Q/F260Y/I301E) into the b' domain of human PDI. When I titrated this hydrophobic cleft mutant into RCM RfbP no binding was evident in the ITC (Fig. 3.45 A) even though the general fold of PDI was unperturbed as judged by CD spectroscopy (data not shown).

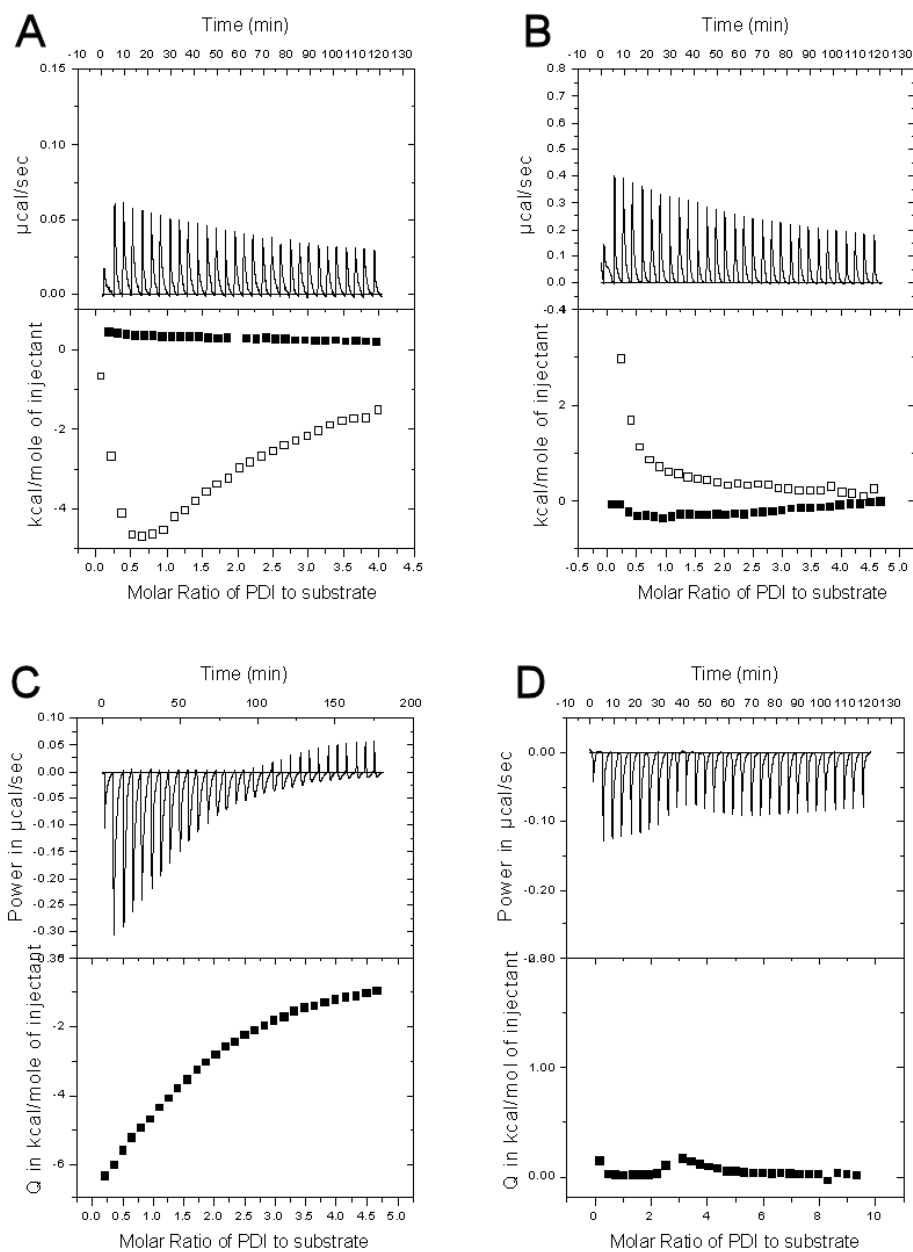


Figure 3.46 – Factors influencing substrate binding in the a-type and b-type domains

(A) ITC analysis with the hydrophobic cleft mutant of human PDI and RCM RfbP (raw data + filled symbols). For comparison, the data of wild-type human PDI has been added to the bottom panel (open symbols). (B) ITC experiment with the a domain of human PDI (E1-P135, raw data+filled symbols) pre-incubated with 5 mM reduced glutathione and then titrated into RCM RfbP. For comparison the data of the a domain under non-reducing conditions has been added (open symbols). (C) ITC experiment with human PDIp pre-incubated with 5 mM oxidized glutathione and RCM RfbP. (D) ITC experiment in which human P5 at a concentration of 1 mM was titrated into 200 μ M reduced glutathione.

For the a-type domains I focused on the Cys-X-X-Cys active site, as it is the defining feature of this domain. With redox active cysteines being the site of catalysis, I analyzed the impact of the redox status on the entropy-driven binding. I modified the ITC conditions to include 5 mM reduced glutathione in the buffer. This physiological reductant is also abundant in the cytoplasm as a redox buffer to keep proteins reduced. When I titrated the a domain of human PDI into unfolded substrate in the presence of reduced glutathione, I observed no binding (Fig. 3.45 B), whereas the same experiment carried out without reductant resulted in an entropy-driven binding as described above (Fig. 3.44 B). To further analyze the influence of the redox status we also tested reduced P5 for its ability to bind RCM RfbP and, in an analogous fashion, no binding could be observed (data not shown). When I tested PDIp for binding of RCM RfbP in the presence of 5 mM oxidized glutathione, no entropy-driven binding could be observed (Fig. 3.45 C) despite PDIp sharing 49.5% sequence identity. Thus, reduction of enzymes removes the entropy-driven binding, however, oxidation of enzymes that initially did not display entropy-driven binding such as PDIp does not result in an interaction with the substrate, indicating that these enzymes do not interact with substrates for reasons that are not connected to the redox status of their a-type domains and hence that there are two types of a-type domains.

With the entropy-driven binding being redox dependent, it has been suggested (Lloyd Ruddock, personal communication) that the heat uptake upon injection of an a-type domain into modified substrate could be caused by residual unmodified thiols that attack the active site disulfide bond of oxidized a-type domains. If only a small portion of the substrate population would carry a free thiol, it would explain the unusual stoichiometry observed. To test this hypothesis I titrated P5 into buffer containing reduced glutathione. Despite the presence of free thiols in the substrate, no heat release could be observed (Fig. 3.45 D). Thus, I can exclude that the observed heat signature is the result of a disulfide bond exchange reaction.

Members of the PDI family are key catalysts during the oxidative folding of newly synthesized proteins passing through the secretory pathway, and, whenever the PDI family network is severely compromised, cells suffer drastic effects[119-121]. Yet it is still unknown, how the functions of the different PDI family members are integrated to form the efficient processing network that is observed in higher eukaryotes. This study is the first to investigate substrate interactions of the PDI family using ITC, a method that yields more information about the

thermodynamic parameters describing the binding processes than the previously employed methods of pull-down and size exclusion chromatography experiments. In contrast to previous studies relying on peptides, I investigated substrate-binding using permanently unfolded full-length proteins. The chosen proteins are well-characterized model proteins for oxidative folding studies such as RNase A [44, 122, 123] or RfbP [76].

With those tools I tried to verify the hypothesis developed based on structural and biochemical data that substrates bind into the hydrophobic cleft of the b' domain [65, 66, 124]. Against my expectations I observed not one but two separate interaction modes between the surveyed PDI family members and potential substrates. These binding modes are driven either by entropy or by enthalpy and could be structurally localized to the a-type domains and b-type domains, respectively. With the aid of truncation mutants of human PDI I could further narrow down the interaction sites to the a domain for the interaction dictated by an increase in entropy and to the b' domain for the enthalpy-driven interaction.

A comparison of the behavior of the different PDI family members revealed that every protein displaying entropy-driven binding contained a thioredoxin domain with a redox sensitive active site (a-type domain), while every protein with an enthalpy-driven binding contained a b-type domain which is characterized by the absence of an active site. Yet the inverse does not hold true. Not every a-type domain can facilitate an entropy-driven binding and not every b-type domain can facilitate an enthalpy-driven binding. As the binding capabilities of b-type domains have already been known, these domains have been extensively studied [65, 66, 125]. Based on a structural comparison of yPDI, ERp57 and ERp72 [97] which suggested that the binding is facilitated by a hydrophobic cleft in the b' domain, I engineered an obstruction of this cleft by three hydrogen bonded side chains as it naturally occurs in ERp57 and ERp72. This triple mutant exhibited no binding at all, while its structure was not significantly altered as judged by CD spectroscopy (data not shown), thus demonstrating that the hydrophobic cleft is indeed responsible for this interaction.

As stated above, the entropy-driven binding is exhibited by domains containing a Cys-X-X-Cys active site. Here, substrate binds in a redox dependent fashion, where the active site needs to be oxidized for the interaction to take place. Yet an oxidized active site on its own is not sufficient for binding as shown by the oxidation experiment of PDI_p, which fails to bind substrates via the

entropy-driven mode despite being in the oxidized state. This suggests that there is an unknown structural feature that is responsible for binding, which is present in some a-type domains but not in others. This additional information will prove useful in analyzing the roles of single PDI family members.

Questions still remain concerning the measured stoichiometries, especially in the case of the entropy-driven binding. The two PDI to one substrate ratio of the enthalpic binding can be explained by the larger size of the substrates, which presumably offer multiple binding sites. As the target substrates for the hydrophobic cleft have been suggested to contain either two consecutive aromatic residues, or two aromatic residues separated by a single non-aromatic residue [126], it is interesting to note that the amino acid sequence of the RfbP offers three such sites with two being in close spatial proximity, so that only one or the other could be bound. On the other hand, RNase A does not offer even a single site, yet it is recognized by b' domains with a similar ratio, thus indicating an even broader recognition pattern.

The binding of the a domain on the other hand shows a ratio of one PDI protein to six or more substrates. Given the limited surface of the a domain and the fact that the binding of multiple reduced substrates for oxidation would promote the formation of disulfide-linked aggregates which are connected via intermolecular bridges such a stoichiometry seems unlikely. The stoichiometry could also be interpreted as PDI selectively binding to only a fraction of the available substrate molecules. This appears to be a more appropriate interpretation, however, the question for the distinguishing feature of the bound substrate arises.

When I tested substrate prepared with a more than tenfold excess of iodoacetic acid over free thiols, I could observe the appearance of additional molecular species in native PAGE experiments (Fig. 3.38 A). These most likely correspond to modifications of several histidine residues. This highly charged substrate was not recognized by either binding mode of PDI. Thus, it seems likely that the a domain actually recognizes unfolded proteins that are uncharged. When I prepared unfolded substrate using a twofold excess of IAA, most of the molecules are modified to an extent that renders them unrecognizable for the a domain. This results in the observed stoichiometry, where only a fraction of the offered substrate is permanently unfolded, yet not that strongly modified with charged groups as to abolish binding.

The redox sensitivity of substrate-binding by an a-type domain immediately suggests its involvement in the oxidase function of the protein. Since the binding only occurs with an oxidized active site, its only possible role is to bind unfolded substrate for the introduction of disulfide bonds. It has already been shown, that the a' domain of PDI undergoes significant conformational changes depending on the redox status [89]. Given the high degree of homology between both active site domains, a similar mechanism seems likely for the a domain. On the other hand the isomerase activity displayed by many members of the PDI family is only possible when the active sites are in the reduced state. This means that isomerization of incorrectly disulfide-linked proteins is mostly facilitated with the help of the b' domain binding site as the substrate-binding capability of an a-type domain is abolished upon reduction. Alternatively it is possible that within one protein the a-type domains capable of binding are predominantly in the oxidized form when present in their native ER redox environment. With the 'twisted U' shape adopted by PDI [63], ERp57 [68] and possibly other members of the PDI family, the a-type domains of one enzyme are in close spatial proximity. This would allow for one domain to bind substrate in need for isomerization and present it to a reduced active site of another domain within the same molecule.

Up to this point PDI is the only protein investigated that simultaneously shows both modes of binding. This is in line with the central role in oxidative protein folding that has been attributed to this protein. With both forms of binding being present, the binding of the a domain will take precedence over the binding of the b' domain due to the more than one order of magnitude higher binding affinity. Thus, the a domain binding is dominant when the protein is in the oxidized state, but as soon as PDI has oxidized its substrate, the b' domain binding will take over. Interestingly the structure of Ero1 α revealed a protruding β -hairpin [127] that has been suggested to bind to the hydrophobic cleft of the b' domain of PDI. As the hydrophobic cleft becomes the dominant binding site upon reduction, it ensures that Ero1 α can bind and reoxidize PDI for another cycle of catalysis. With this alternation of binding Ero1 α can selectively prefer reduced over oxidized PDI, which has been shown previously [127].

The exact nature as well as location of substrate recognition still requires further investigation. Yet the addition of another binding site allows to further classify members of the PDI family depending on the number and kind of substrate-binding domains. The results of this part of the

project further underlines the central role of PDI itself for the oxidative folding as it is the only protein within its family to contain both modes of substrate binding. With the role of many PDI family members hinted at through their mode of binding, their function in and contribution to the network responsible for oxidative folding can hopefully be elucidated.

4. Synopsis & Outlook

This project aimed to get a better understanding of the complex actions and interactions of the folding catalysts in the endoplasmic reticulum and the roles of specific PDI family member. This encompassed the development of new methods to characterize the activity and function of an investigated enzyme as well as a survey of the PDI family utilizing these methods. The different conformational or oligomeric states of PDI continue to be a topic in the literature since the first purifications of the protein. Despite multiple contradicting studies [64, 80-83], these potentially important mechanisms are still not fully recognized by the scientific community. The redox dependency gives a new angle to investigate the role of this mechanism, especially in regards to important topics like oxidative stress and the related fields of oxidizing damage and aging. While the results presented here suggest a functional role, it is very hard to determine the function of the different isoforms of PDI due to the necessity to maintain a strongly oxidative environment to prevent re-equilibration. This of course excludes the use of redox based assays. It might be possible to test the substrate binding capabilities using isothermal titration calorimetry (ITC), yet the strong oxidant hydrogen peroxide when tested, produced a very strong background heat that made it impossible to observe the actual binding reaction. While ITC based assays are by design not impossible to be executed under oxidative conditions, they would require extensive optimization of the technique. Thus the logical next step concerning the PDI isoforms is mass spectrometry to exclude either conformational or oligomeric changes.

While there was little success with the initial goals of crystallizing human PDI and PDI-substrate co-crystals the partial results achieved can hopefully be utilized for another more optimized approach. From the expression and purification of a wide variety of PDI family members it became obvious, that besides stability, an extreme solubility is the hallmark trait of this protein family. As crystallization is a controlled form of precipitation, the high solubility inhibits the formation of crystals. Therefore the next steps in the crystallization of human PDI could be a reduction of surface entropy as well as potential modification of surface residues with hydrophobic groups to reduce solubility. Another problem is the high degree of flexibility which leads to a heterogeneous population of PDI molecules again preventing the formation of uniform crystals. For this purification under oxidizing conditions might work to create a uniform population of PDI molecules as described in the first part of the Results section. Utilizing a

combination of these tools could be the step needed to create diffracting crystals of human PDI. In addition, we were able to solve the structure of human ERp27. While ERp27 itself contains no active site, it can recognize unfolded proteins and bind them as shown by the ITC results presented in this thesis. In the structure, the substrate binding site is indicated by molecules of polyethylene glycol which partially bind to a hydrophobic cleft in the surface of the C terminal domain. This cleft is homologous to the known substrate binding site in PDI. It is known, that ERp27 interacts with ERp57 [71]. While attempts at molecular docking *in silicio* were not successful so far, we could show in collaboration with the group of Heike Hermanns that the expression of ERp27 is increased as a response to oxidative stress as part of the unfolded protein response. ERp57 has been shown to suffer from poor substrate recognition by both poor results in refolding and reduction assays in the literature [107, 108] as well as from the results of the activity assays and ITC measurements presented in this thesis (Chapter 3.3 & 3.4). Therefore ERp27 might work as another adaptor protein in addition to the known adaptors calnexin and calreticulin [99] to enable ERp57 to perform as a folding catalyst.

Despite my failure to determine the kinetic parameters of PDI family members, the results from the established assays can be used for a better characterization. Especially in connection with the established ITC assay, different specializations of PDI family members can now be investigated further. This is an interesting prospect when looking at the ERp27-ERp57 interplay, as it might give a clue how ERp27 differs from the already known adaptor proteins calnexin and calreticulin. Especially the use of RfbP in both a refolding assay and the ITC assay can open up possibilities for a further characterization due to the glycosylation that RfbP carries when purified from chicken eggs. This is an important feature as both calnexin and calreticulin belong to the lectins – a class of proteins known to recognize proteoglycans. Classic refolding substrates mostly do not contain proteoglycans and standard recombinant substrates will not contain them due to the inability of *E. coli* to glycosylate. RfbP in contrast can be easily deglycosylated by PNGaseF to create substrates with and without glycosylations for comparison. As the ER is not only the location for oxidative protein folding but also the location for N-linked glycosylation, a specialization of specific PDI family members for glycosylated or non-glycosylated proteins is a likely possibility. With four different activity assays – three based on fluorescence and another one based on ITC – available now for a qualitative analysis of the different activities, an in-depth characterization of the PDI family would further our understanding of the specifics of each

protein. Such a general study would provide starting points for further investigations to unravel the relations in the ER folding network.

5. References

1. Weichsel, A., et al., *Crystal structures of reduced, oxidized, and mutated human thioredoxins: evidence for a regulatory homodimer*. Structure, 1996. **4**(6): p. 735-51.
2. Hall, G., et al., *Structure of Mycobacterium tuberculosis thioredoxin C*. Acta crystallographica. Section D, Biological crystallography, 2006. **62**(Pt 12): p. 1453-7.
3. Lancelin, J.M., et al., *NMR structures of thioredoxin m from the green alga Chlamydomonas reinhardtii*. Proteins, 2000. **41**(3): p. 334-49.
4. Hashimoto, M., et al., *Role of protein aggregation in mitochondrial dysfunction and neurodegeneration in Alzheimer's and Parkinson's diseases*. Neuromolecular medicine, 2003. **4**(1-2): p. 21-36.
5. Prusiner, S.B., *Prions and neurodegenerative diseases*. The New England journal of medicine, 1987. **317**(25): p. 1571-81.
6. Privalov, P.L., *Stability of proteins: small globular proteins*. Advances in protein chemistry, 1979. **33**: p. 167-241.
7. Levinthal, C., *Are There Pathways for Protein Folding*. Journal De Chimie Physique Et De Physico-Chimie Biologique, 1968. **65**(1): p. 44-&.
8. Levinthal, C., *Mossbauer Spectroscopy in Biological Systems*. Proceedings of a Meeting held at Allerton House, Monticello, IL, 1969. **1**(1): p. 2.
9. Matouschek, A., L. Serrano, and A.R. Fersht, *The folding of an enzyme. IV. Structure of an intermediate in the refolding of barnase analysed by a protein engineering procedure*. Journal of molecular biology, 1992. **224**(3): p. 819-35.
10. Anfinsen, C.B., et al., *The kinetics of formation of native ribonuclease during oxidation of the reduced polypeptide chain*. Proceedings of the National Academy of Sciences of the United States of America, 1961. **47**: p. 1309-14.
11. Anfinsen, C.B., *Principles that govern the folding of protein chains*. Science, 1973. **181**(96): p. 223-30.
12. Leopold, P.E., M. Montal, and J.N. Onuchic, *Protein folding funnels: a kinetic approach to the sequence-structure relationship*. Proceedings of the National Academy of Sciences of the United States of America, 1992. **89**(18): p. 8721-5.
13. Frauenfelder, H., F. Parak, and R.D. Young, *Conformational substates in proteins*. Annual review of biophysics and biophysical chemistry, 1988. **17**: p. 451-79.
14. Chaplin, M., *Do we underestimate the importance of water in cell biology?* Nature reviews. Molecular cell biology, 2006. **7**(11): p. 861-6.
15. LaBean, T.H., S.A. Kauffman, and T.R. Butt, *Libraries of random-sequence polypeptides produced with high yield as carboxy-terminal fusions with ubiquitin*. Molecular diversity, 1995. **1**(1): p. 29-38.
16. Davidson, A.R. and R.T. Sauer, *Folded proteins occur frequently in libraries of random amino acid sequences*. Proceedings of the National Academy of Sciences of the United States of America, 1994. **91**(6): p. 2146-50.
17. Presta, L.G. and G.D. Rose, *Helix signals in proteins*. Science, 1988. **240**(4859): p. 1632-41.
18. Goldenberg, D.P. and T.E. Creighton, *Circular and circularly permuted forms of bovine pancreatic trypsin inhibitor*. Journal of molecular biology, 1983. **165**(2): p. 407-13.

19. Bryngelson, J.D. and P.G. Wolynes, *Spin glasses and the statistical mechanics of protein folding*. Proceedings of the National Academy of Sciences of the United States of America, 1987. **84**(21): p. 7524-8.
20. Eisenberg, D. and A.D. McLachlan, *Solvation energy in protein folding and binding*. Nature, 1986. **319**(6050): p. 199-203.
21. Shoichet, B.K., et al., *A relationship between protein stability and protein function*. Proceedings of the National Academy of Sciences of the United States of America, 1995. **92**(2): p. 452-6.
22. Honeycutt, J.D. and D. Thirumalai, *The nature of folded states of globular proteins*. Biopolymers, 1992. **32**(6): p. 695-709.
23. Honeycutt, J.D. and D. Thirumalai, *Metastability of the folded states of globular proteins*. Proceedings of the National Academy of Sciences of the United States of America, 1990. **87**(9): p. 3526-9.
24. Bryngelson, J.D., et al., *Funnels, pathways, and the energy landscape of protein folding: a synthesis*. Proteins, 1995. **21**(3): p. 167-95.
25. Neufingerl, F.U., O.; Viehauser, M., *Chemie 1. Allgemeine und anorganische Chemie*2005, Wien: Jugend & Volk. 1.
26. Palade, G., *Intracellular aspects of the process of protein synthesis*. Science, 1975. **189**(4200): p. 347-58.
27. Gross, E., et al., *Structure of Ero1p, source of disulfide bonds for oxidative protein folding in the cell*. Cell, 2004. **117**(5): p. 601-10.
28. Vitale, A. and J. Denecke, *The ER folding helpers*, in *The Plant Endoplasmic Reticulum*, D.G. Robinson, Editor 2006. p. 64-66.
29. Dixon, D.P., et al., *Cloning and initial characterization of the Arabidopsis thaliana endoplasmic reticulum oxidoreductins*. Antioxidants & redox signaling, 2003. **5**(4): p. 389-96.
30. Cabibbo, A., et al., *ERO1-L, a human protein that favors disulfide bond formation in the endoplasmic reticulum*. The Journal of biological chemistry, 2000. **275**(7): p. 4827-33.
31. Bardwell, J.C., K. McGovern, and J. Beckwith, *Identification of a protein required for disulfide bond formation in vivo*. Cell, 1991. **67**(3): p. 581-9.
32. Bardwell, J.C., et al., *A pathway for disulfide bond formation in vivo*. Proceedings of the National Academy of Sciences of the United States of America, 1993. **90**(3): p. 1038-42.
33. Tavender, T.J. and N.J. Bulleid, *Peroxiredoxin IV protects cells from oxidative stress by removing H2O2 produced during disulphide formation*. Journal of cell science, 2010. **123**(Pt 15): p. 2672-9.
34. Karala, A.R., et al., *Efficient peroxide-mediated oxidative refolding of a protein at physiological pH and implications for oxidative folding in the endoplasmic reticulum*. Antioxidants & redox signaling, 2009. **11**(5): p. 963-70.
35. Nguyen, V.D., et al., *Two endoplasmic reticulum PDI peroxidases increase the efficiency of the use of peroxide during disulfide bond formation*. Journal of molecular biology, 2011. **406**(3): p. 503-15.
36. Aslund, F., K.D. Berndt, and A. Holmgren, *Redox potentials of glutaredoxins and other thiol-disulfide oxidoreductases of the thioredoxin superfamily determined by direct protein-protein redox equilibria*. The Journal of biological chemistry, 1997. **272**(49): p. 30780-6.

37. Joelson, T., B.M. Sjoberg, and H. Eklund, *Modifications of the active center of T4 thioredoxin by site-directed mutagenesis*. The Journal of biological chemistry, 1990. **265**(6): p. 3183-8.
38. Wunderlich, M. and R. Glockshuber, *Redox properties of protein disulfide isomerase (DsbA) from Escherichia coli*. Protein science : a publication of the Protein Society, 1993. **2**(5): p. 717-26.
39. Lundstrom-Ljung, J., et al., *Two resident ER-proteins, CaBP1 and CaBP2, with thioredoxin domains, are substrates for thioredoxin reductase: comparison with protein disulfide isomerase*. FEBS letters, 1995. **357**(3): p. 305-8.
40. Ghaemmaghami, S., et al., *Global analysis of protein expression in yeast*. Nature, 2003. **425**(6959): p. 737-41.
41. Dobson, C.M. and M. Karplus, *The fundamentals of protein folding: bringing together theory and experiment*. Current Opinion in Structural Biology, 1999. **9**(1): p. 92.
42. Saxena, V.P. and D.B. Wetlaufer, *Formation of three-dimensional structure in proteins. I. Rapid nonenzymic reactivation of reduced lysozyme*. Biochemistry, 1970. **9**(25): p. 5015-23.
43. Creighton, T.E., *Experimental studies of protein folding and unfolding*. Progress in biophysics and molecular biology, 1978. **33**(3): p. 231-97.
44. Low, L.K., et al., *Acceleration of oxidative folding of bovine pancreatic ribonuclease A by anion-induced stabilization and formation of structured native-like intermediates*. FEBS letters, 2000. **472**(1): p. 67-72.
45. Creighton, T.E., D.A. Hillson, and R.B. Freedman, *Catalysis by protein-disulphide isomerase of the unfolding and refolding of proteins with disulphide bonds*. Journal of molecular biology, 1980. **142**(1): p. 43-62.
46. Weissman, J.S. and P.S. Kim, *Efficient catalysis of disulphide bond rearrangements by protein disulphide isomerase*. Nature, 1993. **365**(6442): p. 185-8.
47. Goldberger, R.F., C.J. Epstein, and C.B. Anfinsen, *Acceleration of reactivation of reduced bovine pancreatic ribonuclease by a microsomal system from rat liver*. The Journal of biological chemistry, 1963. **238**: p. 628-35.
48. Tsai, B., et al., *Protein disulfide isomerase acts as a redox-dependent chaperone to unfold cholera toxin*. Cell, 2001. **104**(6): p. 937-48.
49. Wang, C.C. and C.L. Tsou, *Protein disulfide isomerase is both an enzyme and a chaperone*. The FASEB journal : official publication of the Federation of American Societies for Experimental Biology, 1993. **7**(15): p. 1515-7.
50. Puig, A. and H.F. Gilbert, *Protein disulfide isomerase exhibits chaperone and anti-chaperone activity in the oxidative refolding of lysozyme*. The Journal of biological chemistry, 1994. **269**(10): p. 7764-71.
51. Kivirikko, K.I., R. Myllyla, and T. Pihlajaniemi, *Protein hydroxylation: prolyl 4-hydroxylase, an enzyme with four cosubstrates and a multifunctional subunit*. The FASEB journal : official publication of the Federation of American Societies for Experimental Biology, 1989. **3**(5): p. 1609-17.
52. Wetterau, J.R., et al., *Protein disulfide isomerase appears necessary to maintain the catalytically active structure of the microsomal triglyceride transfer protein*. Biochemistry, 1991. **30**(40): p. 9728-35.
53. Turano, C., et al., *Proteins of the PDI family: unpredicted non-ER locations and functions*. Journal of cellular physiology, 2002. **193**(2): p. 154-63.

54. Lawrence, D.A., R. Song, and P. Weber, *Surface thiols of human lymphocytes and their changes after in vitro and in vivo activation*. Journal of leukocyte biology, 1996. **60**(5): p. 611-8.
55. Tager, M., et al., *Membrane-bound protein disulfide isomerase (PDI) is involved in regulation of surface expression of thiols and drug sensitivity of B-CLL cells*. Experimental hematology, 1997. **25**(7): p. 601-7.
56. Atkinson, H.J. and P.C. Babbitt, *An atlas of the thioredoxin fold class reveals the complexity of function-enabling adaptations*. PLoS computational biology, 2009. **5**(10): p. e1000541.
57. Kemmink, J., et al., *The folding catalyst protein disulfide isomerase is constructed of active and inactive thioredoxin modules*. Current biology : CB, 1997. **7**(4): p. 239-45.
58. Quan, S., et al., *The CXXC motif is more than a redox rheostat*. The Journal of biological chemistry, 2007. **282**(39): p. 28823-33.
59. Huber-Wunderlich, M. and R. Glockshuber, *A single dipeptide sequence modulates the redox properties of a whole enzyme family*. Folding & design, 1998. **3**(3): p. 161-71.
60. Gruber, C.W., et al., *Protein disulfide isomerase: the structure of oxidative folding*. Trends in biochemical sciences, 2006. **31**(8): p. 455-64.
61. Ren, G., et al., *Properties of the thioredoxin fold superfamily are modulated by a single amino acid residue*. The Journal of biological chemistry, 2009. **284**(15): p. 10150-9.
62. Klappa, P., et al., *The b' domain provides the principal peptide-binding site of protein disulfide isomerase but all domains contribute to binding of misfolded proteins*. The EMBO journal, 1998. **17**(4): p. 927-35.
63. Tian, G., et al., *The crystal structure of yeast protein disulfide isomerase suggests cooperativity between its active sites*. Cell, 2006. **124**(1): p. 61-73.
64. Tian, G., et al., *The catalytic activity of protein-disulfide isomerase requires a conformationally flexible molecule*. The Journal of biological chemistry, 2008. **283**(48): p. 33630-40.
65. Pirneskoski, A., et al., *Molecular characterization of the principal substrate binding site of the ubiquitous folding catalyst protein disulfide isomerase*. The Journal of biological chemistry, 2004. **279**(11): p. 10374-81.
66. Nguyen, V.D., et al., *Alternative conformations of the x region of human protein disulphide-isomerase modulate exposure of the substrate binding b' domain*. Journal of molecular biology, 2008. **383**(5): p. 1144-55.
67. Ellgaard, L. and L.W. Ruddock, *The human protein disulphide isomerase family: substrate interactions and functional properties*. EMBO reports, 2005. **6**(1): p. 28-32.
68. Dong, G., et al., *Insights into MHC class I peptide loading from the structure of the tapasin-ERp57 thiol oxidoreductase heterodimer*. Immunity, 2009. **30**(1): p. 21-32.
69. Kikuchi, M., et al., *Functional analysis of human P5, a protein disulfide isomerase homologue*. Journal of biochemistry, 2002. **132**(3): p. 451-5.
70. Alanen, H.I., et al., *pH dependence of the peptide thiol-disulfide oxidase activity of six members of the human protein disulfide isomerase family*. Antioxidants & redox signaling, 2006. **8**(3-4): p. 283-91.
71. Alanen, H.I., et al., *ERp27, a new non-catalytic endoplasmic reticulum-located human protein disulfide isomerase family member, interacts with ERp57*. The Journal of biological chemistry, 2006. **281**(44): p. 33727-38.

72. Sanger, F. and A.R. Coulson, *A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase*. Journal of molecular biology, 1975. **94**(3): p. 441-8.
73. Laemmli, U.K., *Cleavage of structural proteins during the assembly of the head of bacteriophage T4*. Nature, 1970. **227**(5259): p. 680-5.
74. Becvar, J. and G. Palmer, *The binding of flavin derivatives to the riboflavin-binding protein of egg white. A kinetic and thermodynamic study*. The Journal of biological chemistry, 1982. **257**(10): p. 5607-17.
75. Raturi, A. and B. Mutus, *Characterization of redox state and reductase activity of protein disulfide isomerase under different redox environments using a sensitive fluorescent assay*. Free radical biology & medicine, 2007. **43**(1): p. 62-70.
76. Rancy, P.C. and C. Thorpe, *Oxidative protein folding in vitro: a study of the cooperation between quiescin-sulphydryl oxidase and protein disulfide isomerase*. Biochemistry, 2008. **47**(46): p. 12047-56.
77. Brunger, A.T., *Free R value: a novel statistical quantity for assessing the accuracy of crystal structures*. Nature, 1992. **355**(6359): p. 472-5.
78. Engh, R.A. and R. Huber, *Accurate Bond and Angle Parameters for X-Ray Protein-Structure Refinement*. Acta Crystallographica Section A, 1991. **47**: p. 392-400.
79. Ramachandran, G.N., C. Ramakrishnan, and V. Sasisekharan, *Stereochemistry of Polypeptide Chain Configurations*. Journal of molecular biology, 1963. **7**(1): p. 95-&.
80. Hu, C.H. and C.L. Tsou, *C-terminal truncation of bovine protein disulfide isomerase increases its activity*. Biochemical and biophysical research communications, 1992. **183**(2): p. 714-8.
81. Morjana, N.A., B.J. McKeone, and H.F. Gilbert, *Guanidine hydrochloride stabilization of a partially unfolded intermediate during the reversible denaturation of protein disulfide isomerase*. Proceedings of the National Academy of Sciences of the United States of America, 1993. **90**(6): p. 2107-11.
82. Yu, X.C., C.C. Wang, and C.L. Tsou, *Association and dissociation of protein disulfide isomerase*. Biochimica et biophysica acta, 1994. **1207**(1): p. 109-13.
83. Solovyov, A. and H.F. Gilbert, *Zinc-dependent dimerization of the folding catalyst, protein disulfide isomerase*. Protein science : a publication of the Protein Society, 2004. **13**(7): p. 1902-7.
84. Solovyov, A., R. Xiao, and H.F. Gilbert, *Sulphydryl oxidation, not disulfide isomerization, is the principal function of protein disulfide isomerase in yeast Saccharomyces cerevisiae*. The Journal of biological chemistry, 2004. **279**(33): p. 34095-100.
85. Wilkinson, B., R. Xiao, and H.F. Gilbert, *A structural disulfide of yeast protein-disulfide isomerase destabilizes the active site disulfide of the N-terminal thioredoxin domain*. The Journal of biological chemistry, 2005. **280**(12): p. 11483-7.
86. Darby, N.J. and T.E. Creighton, *Characterization of the active site cysteine residues of the thioredoxin-like domains of protein disulfide isomerase*. Biochemistry, 1995. **34**(51): p. 16770-80.
87. Darby, N.J. and T.E. Creighton, *Catalytic mechanism of DsbA and its comparison with that of protein disulfide isomerase*. Biochemistry, 1995. **34**(11): p. 3576-87.

88. Karala, A.R., A.K. Lappi, and L.W. Ruddock, *Modulation of an active-site cysteine pKa allows PDI to act as a catalyst of both disulfide bond formation and isomerization*. Journal of molecular biology, 2010. **396**(4): p. 883-92.
89. Wang, C., et al., *Human protein disulfide isomerase is a redox-regulated chaperone activated by oxidation of domain a'*. The Journal of biological chemistry, 2011.
90. Grubisic, Z., P. Rempp, and H. Benoit, *A universal calibration for Gel Permeation Chromatography*. Journal of Polymer Science: Part B: Polymer Physics,, 1967. **34**: p. 1707-1713.
91. Mylonas, E. and D.I. Svergun, *Accuracy of molecular mass determination of proteins in solution by small-angle X-ray scattering*. Journal of Applied Crystallography, 2007. **40**: p. s245-s249.
92. Cai, H., C.C. Wang, and C.L. Tsou, *Chaperone-like activity of protein disulfide isomerase in the refolding of a protein with no disulfide bonds*. The Journal of biological chemistry, 1994. **269**(40): p. 24550-2.
93. Baud, F. and S. Karlin, *Measures of residue density in protein structures*. Proceedings of the National Academy of Sciences of the United States of America, 1999. **96**(22): p. 12494-9.
94. Lo Conte, L., C. Chothia, and J. Janin, *The atomic structure of protein-protein recognition sites*. Journal of molecular biology, 1999. **285**(5): p. 2177-98.
95. Derewenda, Z.S., *Rational protein crystallization by mutational surface engineering*. Structure, 2004. **12**(4): p. 529-35.
96. Cooper, D.R., et al., *Protein crystallization by surface entropy reduction: optimization of the SER strategy*. Acta crystallographica. Section D, Biological crystallography, 2007. **63**(Pt 5): p. 636-45.
97. Kozlov, G., et al., *A structural overview of the PDI family of proteins*. The FEBS journal, 2010. **277**(19): p. 3924-36.
98. Zapun, A., J.C. Bardwell, and T.E. Creighton, *The reactive and destabilizing disulfide bond of DsbA, a protein required for protein disulfide bond formation in vivo*. Biochemistry, 1993. **32**(19): p. 5083-92.
99. Oliver, J.D., et al., *ERp57 functions as a subunit of specific complexes formed with the ER lectins calreticulin and calnexin*. Molecular biology of the cell, 1999. **10**(8): p. 2573-82.
100. Schmid, F.X. and R.L. Baldwin, *Acid catalysis of the formation of the slow-folding species of RNase A: evidence that the reaction is proline isomerization*. Proceedings of the National Academy of Sciences of the United States of America, 1978. **75**(10): p. 4764-8.
101. Ruoppolo, M. and R.B. Freedman, *Refolding by disulfide isomerization: the mixed disulfide between ribonuclease T1 and glutathione as a model refolding substrate*. Biochemistry, 1995. **34**(29): p. 9380-8.
102. Hillson, D.A., N. Lambert, and R.B. Freedman, *Formation and isomerization of disulfide bonds in proteins: protein disulfide-isomerase*. Methods in enzymology, 1984. **107**: p. 281-94.
103. Luthman, M. and A. Holmgren, *Rat liver thioredoxin and thioredoxin reductase: purification and characterization*. Biochemistry, 1982. **21**(26): p. 6628-33.

104. Raturi, A., et al., *Platelet microparticle-associated protein disulfide isomerase promotes platelet aggregation and inactivates insulin*. *Biochimica et biophysica acta*, 2008. **1778**(12): p. 2790-6.
105. Lorenz, S.H. and F.X. Schmid, *The folding of the Inserted Flap domain is dependent on spatial proximity of its termini*, F.X. Kober and H. Schindelin, Editors. 2007: Wuerzburg. p. 4.
106. Barcroft, J. and A.V. Hill, *The nature of oxyhaemoglobin, with a note on its molecular weight*. *The Journal of physiology*, 1910. **39**(6): p. 411-28.
107. Zapun, A., et al., *Enhanced catalysis of ribonuclease B folding by the interaction of calnexin or calreticulin with ERp57*. *The Journal of biological chemistry*, 1998. **273**(11): p. 6009-12.
108. Frickel, E.M., et al., *ERp57 is a multifunctional thiol-disulfide oxidoreductase*. *The Journal of biological chemistry*, 2004. **279**(18): p. 18277-87.
109. Burstein, E.A., N.S. Vedenkina, and M.N. Ivkova, *Fluorescence and the location of tryptophan residues in protein molecules*. *Photochemistry and photobiology*, 1973. **18**(4): p. 263-79.
110. Zheng, J. and H.F. Gilbert, *Discrimination between native and non-native disulfides by protein-disulfide isomerase*. *The Journal of biological chemistry*, 2001. **276**(19): p. 15747-52.
111. Hartl, F.U. and M. Hayer-Hartl, *Molecular chaperones in the cytosol: from nascent chain to folded protein*. *Science*, 2002. **295**(5561): p. 1852-8.
112. Kelly, S.M., T.J. Jess, and N.C. Price, *How to study proteins by circular dichroism*. *Biochimica et biophysica acta*, 2005. **1751**(2): p. 119-39.
113. Lambert, N. and R.B. Freedman, *Structural properties of homogeneous protein disulphide-isomerase from bovine liver purified by a rapid high-yielding procedure*. *The Biochemical journal*, 1983. **213**(1): p. 225-34.
114. Kaska, D.D., K.I. Kivirikko, and R. Myllyla, *Purification and characterization of protein disulphide-isomerase from the unicellular green alga *Chlamydomonas reinhardtii*. A 120 kDa dimer antigenically distinct from the vertebrate enzyme*. *The Biochemical journal*, 1990. **268**(1): p. 63-8.
115. Sugiyama, H., et al., *Purification of protein disulfide isomerase from a thermophilic fungus*. *Bioscience, biotechnology, and biochemistry*, 1993. **57**(10): p. 1704-7.
116. Norgaard, P., et al., *Functional differences in yeast protein disulfide isomerases*. *The Journal of cell biology*, 2001. **152**(3): p. 553-62.
117. Denisov, A.Y., et al., *Solution structure of the bb' domains of human protein disulfide isomerase*. *The FEBS journal*, 2009. **276**(5): p. 1440-9.
118. Byrne, L.J., et al., *Mapping of the ligand-binding site on the b' domain of human PDI: interaction with peptide ligands and the x-linker region*. *The Biochemical journal*, 2009. **423**(2): p. 209-17.
119. LaMantia, M., et al., *Glycosylation site binding protein and protein disulfide isomerase are identical and essential for cell viability in yeast*. *Proceedings of the National Academy of Sciences of the United States of America*, 1991. **88**(10): p. 4453-7.
120. Winter, A.D., G. McCormack, and A.P. Page, *Protein disulfide isomerase activity is essential for viability and extracellular matrix formation in the nematode *Caenorhabditis elegans**. *Developmental biology*, 2007. **308**(2): p. 449-61.

121. Garbi, N., et al., *Impaired assembly of the major histocompatibility complex class I peptide-loading complex in mice deficient in the oxidoreductase ERp57*. *Nature immunology*, 2006. **7**(1): p. 93-102.
122. Geiger, R., et al., *Folding, quality control, and secretion of pancreatic ribonuclease in live cells*. *The Journal of biological chemistry*, 2011. **286**(7): p. 5813-22.
123. Shin, H.C. and H.A. Scheraga, *Catalysis of the oxidative folding of bovine pancreatic ribonuclease A by protein disulfide isomerase*. *Journal of molecular biology*, 2000. **300**(4): p. 995-1003.
124. Noiva, R., R.B. Freedman, and W.J. Lennarz, *Peptide binding to protein disulfide isomerase occurs at a site distinct from the active sites*. *The Journal of biological chemistry*, 1993. **268**(26): p. 19210-7.
125. Wallis, A.K., et al., *The ligand-binding b' domain of human protein disulphide-isomerase mediates homodimerization*. *Protein science : a publication of the Protein Society*, 2009. **18**(12): p. 2569-77.
126. Klappa, P., H.C. Hawkins, and R.B. Freedman, *Interactions between protein disulphide isomerase and peptides*. *European journal of biochemistry / FEBS*, 1997. **248**(1): p. 37-42.
127. Masui, S., et al., *Molecular bases of cyclic and specific disulfide interchange between human ERO1alpha protein and protein-disulfide isomerase (PDI)*. *The Journal of biological chemistry*, 2011. **286**(18): p. 16261-71.

E List of abbreviations

<u>Abbreviation</u>	<u>Definition</u>
BPTI	Bovine Pancreatic Trypsin Inhibitor
BSA	Bovine Serum Albumine
CD	Circular Dichroism
CPS	Counts Per Secound
CV	Column Volume
dATP	deoxy-AdenosineTriPhosphate
dCTP	deoxy-CytosineTriPhosphate
dGTP	deoxy-GuanineTriPhosphate
di-E-GSSG	di-Eosine-Gluthathione (oxidized)
DLS	Dynamic Light Scattering
DNA	DeoxyriboNucleic Acid
dNTP	deoxy-NucleotideTriPhosphate
DTT	DiThioThreitol
dTTP	deoxy-ThymineTriPhosphate
EDTA	EthyleneDiamineTetraAcetic acid
ER	Endoplasmic Reticulum
FPLC	Fast Protein Liquid Chromatography
FSQ	Fluorescence SelfQuenching
GAPDH	GlycerinAldehyde-3-Phosphate-DeHydrogenase
GSH	Gluthathione (reduced)
GSSG	Gluthathione (oxidzed)
IAA	IodoAcetic Acid
IFD	Inserted Flap Domain
IFS	Inserted Flap Substrate
ITC	Isothermal Titration Calorimetry
LB	Lysogeny Broth
MAD	Multiple wavelength Anomalous Dispersion
MPG	MethoxyPolyethylene Glycol 350
MTTP	Microsomal Triacylglycerol Transfer Protein
MW	Molecular Weight
MWCO	Molecular Weight CutOff
NCBI	National Center for Biotechnology Information

Abbreviation

NMR

NTB

OD

PCR

PDI

PEG

RCM

RfbP

RNA

SAD

SDS-PAGE

SeMet

SERp

TCEP

TMX

TrxA

Definition

Nuclear Magnetic Resonance

2-Nitro-5-ThioBenzoate

Optical Density

Polymerase Chain Reaction

Protein Disulfide Isomerase

PolyEthylene Glycol

Reduced CarboxyMethylated

Riboflavin binding Protein

RiboNucleic Acid

Single wavelength Anomalous Dispersion

Sodium DodecylSulfate PolyAcrylamide Gel Electrophoresis

Seleno Methionine

Surface Entrophy Reduction prediction

Tris(2-CarboxyEthyl)Phosphine

TransMembrane oXidoreductases

Thioredoxin A

F List of Figures

Figure 1.1 – Structure and function	Fehler! Textmarke nicht definiert.
Figure 1.2 – Schematic of a peptide backbone	- 12 -
Figure 1.3 – The ‘energy funnel’	- 13 -
Figure 1.4 – The role of PDI during folding	- 17 -
Figure 1.5 - Standard redox potentials.....	- 18 -
Figure 1.6 – Schematic of the thioredoxin fold	- 22 -
Figure 1.7 – Crystal structure of yeast PDI.....	- 23 -
Figure 1.8 – Surface hydrophobicity of PDI.....	- 24 -
Figure 2.1 – Di-Eosin-Glutathione	- 47 -
Figure 3.1 – Anion exchange chromatography during the purification of yeast PDI.....	- 66 -
Figure 3.2 – SDS gel electrophoresis of protein samples from the PDI purification	- 66 -
Figure 3.3 – Size exclusion chromatography of the pooled fractions 2-5 from Fig. 3.1	- 68 -
Figure 3.4 – Analytical Size Exclusion Chromatography	- 68 -
Figure 3.5 – Equilibrium formation at different concentrations	- 69 -
Figure 3.6 – Time course of the equilibrium formation.....	- 69 -
Figure 3.7 – Arresting the equilibration under oxidizing conditions.....	- 71 -
Figure 3.8 – Effects of reductant on the equilibrium	- 71 -
Figure 3.9 – Separation of PDI species by surface charge	- 72 -
Figure 3.10 – Size exclusion analysis of charge-separated samples.....	- 72 -
Figure 3.11 – Structure of a yeast PDI dimer	- 74 -
Figure 3.12 – Suggested model for the monomer-dimer-equilibrium.	- 76 -
Figure 3.13 – Protein crystals of PDI from <i>S. cerevisiae</i>	- 80 -
Figure 3.14 – Crystallization of human protein disulfide isomerase	- 82 -
Figure 3.15 – Proposed mutations for surface entropy reduction.....	- 83 -
Figure 3.16 – ERp27 crystallization	- 84 -
Figure 3.17 – Initial diffraction experiments with ERp27 crystals.....	- 86 -
Figure 3.18 – Limited proteolysis of ERp27	- 87 -
Figure 3.19 – Time course of ERp27 proteolysis by trypsin	- 88 -
Figure 3.20 – Purification of the tryptic fragment by size exclusion chromatography	- 89 -

Figure 3.21 – Crystallization of a tryptic fragment of ERp27	- 90 -
Figure 3.22 – Structure of ERp27- Crystal content	- 93 -
Figure 3.23 – Superposition of the five molecules in the asymmetric unit	- 94 -
Figure 3.24 – Sequence of ERp27	- 95 -
Figure 3.25 – Structure of ERp27 – Single molecule	- 96 -
Figure 3.26 – The hydrophobic cleft of ERp27	- 97 -
Figure 3.27 – Planar regression analysis of the domain orientation in the PDI family	- 99 -
Figure 3.28 – Fluorescence measurements during di-E-GSSG purification.....	- 104 -
Figure 3.29 – Fluorescence standard curve of di-E-GSSG.....	- 105 -
Figure 3.30 – Reduction of di-E-GSSG by DTT	- 106 -
Figure 3.31 – Reduction of di-E-GSSG catalyzed by yeast PDI.....	- 107 -
Figure 3.32 – Comparison of the wild-type, SSSS and 18E1 variants of yeast PDI.....	- 111 -
Figure 3.33 – Comparison of human and yeast PDI with chicken ERp57	- 113 -
Figure 3.34 – Fits of the reduction of di-E-GSSG catalyzed by PDI or ERp57.....	- 114 -
Figure 3.35 – Standard curve of riboflavin fluorescence at 450/525 nm.....	- 115 -
Figure 3.36 - Refolding of reduced RfbP.....	- 115 -
Figure 3.37 – Reduction of the inserted flap domain substrate	- 119 -
Figure 3.38 – CD spectroscopy of model substrates and initial ITC experiments	- 124 -
Figure 3.39 – Effects of carboxymethylation	- 125 -
Figure 3.40 – Buffer correction of ITC experiments	- 126 -
Figure 3.41 – RNase A as binding partner for human PDI.....	- 127 -
Figure 3.42 – Substrate binding of selected PDI family members	- 128 -
Figure 3.43 – Influence of temperature on ITC results.....	- 129 -
Figure 3.44 – Domain architecture and binding modes of PDI family members.....	- 131 -
Figure 3.45 – Correlation between binding behavior and domain architecture.....	- 132 -
Figure 3.46 – Factors influencing substrate binding in the a-type and b-type domains	- 134 -

G List of Publications and Congress contributions

Geng Tian, **Franz-Xaver Kober**, Urs Lewandrowski, Albert Sickmann, William J. Lennarz, and Hermann Schindelin

The catalytic activity of protein-disulfide isomerase requires a conformationally flexible molecule.

The Journal of biological chemistry, 2008. **283**(48): p. 33630-40.

Franz-Xaver Kober, Wolfgang Kölmel, Jochen Kuper, Johannes Drechsler and Hermann Schindelin

The structure of ERp27 reveals flexibility of the substrate binding site

Manuscript in preparation

Franz-Xaver Kober and Hermann Schindelin

Protein Disulfide Isomerase Family – Catalytic Characterization

Poster at the Evaluation of the Rudolf Virchow Center – February 2009

Franz-Xaver Kober and Hermann Schindelin

Protein Disulfide Isomerase: Substrate recognition and interaction

Poster at the Meeting of the European ER network – May 2009

Franz-Xaver Kober

Recognition of Substrates by Members of the Protein Disulfide Isomerase Family

Presentation at the Meeting of the European ER network – May 2011

Affidavit

I hereby confirm that my thesis entitled INSERT TITLE HERE is the result of my own work. I did not receive any help or support from commercial consultants. All sources and / or materials applied are listed and specified in the thesis.

Furthermore, I confirm that this thesis has not yet been submitted as part of another examination process neither in identical nor in similar form.

Place, Date

Signature

Eidesstattliche Erklärung

Hiermit erkläre ich an Eides statt, die Dissertation INSERT TITLE HERE eigenständig, d.h. insbesondere selbständig und ohne Hilfe kommerzieller Promotionsberater, angefertigt und keine anderen als die von mir angegebenen Quellen und Hilfsmittel verwendet zu haben.

Ich erkläre außerdem, dass die Dissertation weder in gleicher noch in ähnlicher Form bereits in einem anderen Prüfungsverfahren vorgelegen hat.

Ort, Datum

Unterschrift