

Studying macromolecular transitions by NMR and computer simulations

Lukas Sebastian Stelzl

Department of Biochemistry

University of Oxford

This dissertation is submitted for the degree of
Doctor of Philosophy

Acknowledgements

First, I would like to thank my supervisors Christina Redfield and Mark S.P. Sansom for their outstanding support. I am especially grateful to Christina, for providing superb mentoring, excellent supervision and for teaching me about NMR, enabling to study the microscopic mechanisms of complex biomolecules. Christina gave me the freedom to pursue my own ideas, and provided crucial insights when things did not quite work the first time I have tried them. Equally, important, Christina taught me a tremendous amount about the clear presentation of research and I am thankful for her help with fellowship applications. I am also enormously grateful for the mentoring provided by Mark, the many excellent suggestions, about research or about my plans for the future, and the interdisciplinary atmosphere in the Structural Bioinformatics and Computational Biochemistry Unit.

I want to thank Despoina A. I. Mavridou for her great mentoring and providing the samples I have used for the NMR measurements in this thesis. It was always great, discussing the work on DsbD with her. I am indebted to her for her good advice on fellowship and post-doc applications.

In addition I would like to thank Oliver Beckstein and Philip W. Fowler; I have learnt a lot, working on the LacY project. Oliver Beckstein (now at Arizona State University) has been a crucial mentor. I would like to particularly thank him for introducing me to statistical mechanics, which provided an important new perspective in my understanding of biology. Oli also introduced me to Python scripting which has been vital for my work. Phil's kind advice and many good suggestions, be it about the LacY manuscript or research in general, are deeply appreciated.

I am deeply grateful to Hedi Rochowanski for her help with data presentation, fellowship applications, encouragement and her tremendous support. Her good sense of judgement and humour, as well as many thoughtful suggestions have been instrumental in thesis writing.

Andrew J. Baldwin (in the Physical and Theoretical Chemistry Laboratory) has provided critical help with the analysis of the relaxation dispersion experiments. It has been great fun discussing the energy landscapes of polymers with him.

I want to thank Nick Soffe for his help and support with NMR experiments and for many

insightful conversations about NMR and computing. Moreover I want to thank all the other members of the Redfield and Sansom groups that have made the last four years such a great learning experience. Special thanks are due to the fellow thesis writers.

Finally I want to thank my parents and my sisters for their immense support and the many insightful as well as entertaining dinner table discussions over the years.

Für Hedi.

Abstract

Macromolecular transitions such as conformational changes and protein-protein association underlie many biological processes. Conformational changes in the N-terminal domain of the transmembrane protein DsbD (nDsbD) were studied by NMR and molecular dynamics (MD) simulations. nDsbD supplies reductant to biosynthetic pathways in the oxidising periplasm of Gram-negative bacteria after receiving reductant from the C-terminal domain of DsbD (cDsbD). Reductant transfer in the DsbD pathway happens *via* protein-protein association and subsequent thiol-disulphide exchange reactions. The cap loop shields the active-site cysteines in nDsbD from non-cognate oxidation, but needs to open when nDsbD bind its interaction partners. The loop was rigid in MD simulations of reduced nDsbD. More complicated dynamics were observed for oxidised nDsbD, as the disulphide bond introduces frustration which led to loop opening in some trajectories. The simulations of oxidised and reduced nDsbD agreed well with previous NMR spin-relaxation and residual dipolar coupling measurements as well as chemical shift-based torsion angle predictions. NMR relaxation dispersion experiments revealed that the cap loop of oxidised nDsbD exchanges between a major and a minor conformation. The differences in their conformational dynamics may explain why oxidised nDsbD binds its physiological partner cDsbD much tighter than reduced nDsbD. The redox-state dependent interaction between cDsbD and nDsbD is thought to enhance turnover. NMR relaxation dispersion experiments gave insight into the kinetics of the redox-state dependent interaction. MD simulations identified dynamic encounter complexes in the association of nDsbD with cDsbD. The mechanism of the conformational changes in the transport cycle of LacY were also investigated. LacY switches between periplasmic open and cytoplasmic open conformations to transport sugars across the cell membrane. Two mechanisms have been proposed for the conformational change, a rocker-switch mechanism based on rigid body motions and an “airlock” like mechanism in which the transporter would switch conformation *via* a fully occluded structure. In MD simulations using the novel dynamics importance sampling approach such a fully occluded structure was found. The simulations argued against a strict “rocker-switch” mechanism.

Publications

The work presented in Chapter 6 has resulted in the following publication

- Lukas S. Stelzl, Philip W. Fowler, Mark S.P. Sansom, Oliver Beckstein, Flexible gates generate occluded intermediates in the transport cycle of LacY. *J. Mol. Biol.*, **2014**, 426 (3) 735-751

Contents

Contents	vii
1 Introduction	1
1.1 Biology at the molecular level	1
1.2 Studying macromolecular transitions	2
1.3 NMR experiments	3
1.3.1 Studying fluctuations by NMR	3
1.3.2 Studying structural transitions by NMR	4
1.4 Molecular dynamics simulations	7
1.4.1 Atomistic molecular dynamics	7
1.4.2 Enhanced sampling methods	7
1.4.3 Comparison to experiment	8
1.5 The trans-membrane oxidoreductase DsbD	9
1.5.1 Isomerisation of disulphide bonds	9
1.5.2 DsbD supplies reductant to protect against oxidative damage	13
1.5.3 DsbD supplies reductant for cytochrome <i>c</i> maturation	14
1.5.4 Occurrence of DsbD related pathways	14
1.5.5 The reductant transfer reactions in the DsbD pathway	15
1.5.6 Control of the reductant transfer reactions	17
1.5.7 Studying the interaction of cDsbD and nDsbD	19
1.6 The mechanism of conformational changes in the transporter LacY	20
2 Theory and methodological background	23
2.1 Studying dynamics by NMR	23
2.1.1 Fast-timescale dynamics	23
2.1.2 Interpreting fast timescale NMR experiments	24
2.1.3 Chemical exchange	24

2.1.4	Exchange induced shifts in peak positions	26
2.2	Molecular dynamics simulations	27
2.2.1	The force field	27
2.2.2	Integration of the equation of motions	30
2.2.3	Temperature and pressure coupling	31
2.3	Enhanced sampling methods	32
2.3.1	Metadynamics free-energy calculations	32
2.3.2	Dynamics importance sampling calculations	33
2.3.3	Comparing molecular dynamics simulations to experiments	35
3	Molecular dynamics simulations of nDsbD and comparison to NMR data	37
3.1	Introduction	37
3.1.1	nDsbD is a "redox hub" in the periplasm of Gram-negative bacteria	37
3.1.2	Reductant transfer happens in a series of thiol-disulphide exchange reactions	40
3.1.3	The structure and dynamics of nDsbD	40
3.1.4	Studying nDsbD by molecular simulation and their comparison to NMR data	44
3.2	Methods	45
3.2.1	Atomistic molecular dynamics simulations	45
3.2.2	Comparing NMR spin-relaxation experiments with simulations	47
3.2.3	Calculating residual dipolar couplings from molecular dynamics simulations	49
3.2.4	TALOS-N prediction of the side-chain dihedral angles	50
3.3	Results	50
3.3.1	MD simulations of oxidised and reduced nDsbD are stable	50
3.3.2	Functional characterisation of the MD simulations of reduced nDsbD	53
3.3.3	Functional characterisation of the MD simulations of oxidised nDsbD	58
3.3.4	Comparing the simulations to the NMR spin-relaxation experiments	64
3.3.5	Comparing the simulations to experimental residual dipolar couplings	68
3.3.6	Comparing the MD simulations to the dihedral angle predicted from the chemical shifts	71
3.3.7	Comparing non-bonded interactions in the active sites of reduced and oxidised nDsbD	73
3.4	Discussion	75

3.4.1	The dynamics of the cap loop and association reactions of nDsbD	80
3.4.2	Further experiments	80
4	NMR relaxation dispersion experiments on the N-terminal domain of DsbD	81
4.1	Introduction	81
4.1.1	Relaxation dispersion experiments to study slow dynamics	82
4.1.2	The structure of sparsely populated states	82
4.1.3	Free-energy landscapes from relaxation dispersion experiments	83
4.2	Aims	84
4.3	Materials and methods	85
4.3.1	NMR spectrometers	85
4.3.2	Samples for NMR experiments	85
4.3.3	Implementation of the ^{15}N relaxation dispersion experiments	85
4.3.4	Relaxation dispersion experiments at multiple field strengths and temperatures	89
4.3.5	Analysis of the relaxation dispersion experiments	90
4.3.6	Determining the sign for chemical shift differences between ground and excited states	92
4.3.7	Metadynamics free energy calculations	94
4.4	Results	95
4.4.1	Implementation and testing of the relaxation dispersion pulse sequence	95
4.4.2	Relaxation dispersion experiments with reduced nDsbD	95
4.4.3	Relaxation dispersion experiments with oxidised nDsbD	98
4.4.4	Analysis of the relaxation dispersion experiments for oxidised nDsbD101	
4.4.5	Analysis of the chemical shift changes associated with the exchange process	104
4.4.6	The dynamics of the disulphide bond in computer simulations	108
4.4.7	Convergence of the metadynamics free-energy calculations	109
4.4.8	Characterisation of the "reduced-like" conformation accessed by the metadynamics simulations	113
4.5	Discussion	115
4.5.1	Oxidised nDsbD has an alternative sparsely populated state	116
4.5.2	The structure of the excited state of oxidised nDsbD	116
4.5.3	The free-energy landscape of nDsbD	117
4.5.4	Implications for biological activity	118

4.5.5	Future work	119
5	The association reaction of nDsbD with cDsbD	121
5.1	Introduction	121
5.1.1	Studying protein-protein association reactions	121
5.1.2	Encounter complexes	122
5.1.3	The association of cDsbD and nDsbD	124
5.1.4	Aims	125
5.2	Materials and methods	125
5.2.1	Sample for NMR measurements	125
5.2.2	Spectrometers	126
5.2.3	NMR relaxation dispersion experiments	126
5.2.4	Analysis of the relaxation dispersion experiments	126
5.2.5	Analysis of 1D NMR experiments	126
5.2.6	Electrostatics calculations	127
5.2.7	Molecular dynamics simulations of the association of reduced cDsbD and oxidised nDsbD	127
5.3	Results: NMR experiments	128
5.3.1	Association of C461A-cDsbD with oxidised nDsbD	128
5.3.2	Association oxidised cDsbD with oxidised nDsbD	134
5.4	Results: Computational study of the association reaction	138
5.4.1	A simplified model of the association reaction	138
5.4.2	Electrostatics calculations	142
5.4.3	Molecular dynamics simulations of the association of cDsbD and nDsbD	142
5.5	Discussion	148
5.5.1	The association of cDsbD with nDsbD	148
5.5.2	The encounter complex	150
5.5.3	Further experimental work	151
5.5.4	Further computational work	151
6	Studying conformational changes in the transport cycle of LacY	153
6.1	Introduction	153
6.1.1	LacY is a paradigm for membrane transport	153
6.1.2	The transport cycle of LacY	154
6.1.3	The mechanism of the conformational transition in LacY	156

6.1.4	Evidence for an occluded intermediate state	158
6.1.5	Studying conformational changes in the transport cycle of LacY . .	158
6.1.6	Aims	159
6.2	Methods	159
6.2.1	Membrane protein insertion	159
6.2.2	Equilibrium MD simulations	161
6.2.3	Analysis of simulation stability	164
6.2.4	Cluster analysis	164
6.2.5	DIMS simulations	164
6.2.6	Pore profile analysis	165
6.2.7	Water density analysis	166
6.2.8	DEER distance analysis	166
6.2.9	Poisson-Boltzmann electrostatics calculations	166
6.2.10	Order parameter analysis	168
6.3	Results: Molecular dynamics of cytoplasmic and periplasmic open LacY . .	169
6.3.1	Stability of the simulations	169
6.3.2	Functional analysis of the simulations of cytoplasmic and periplas- mic open LacY	170
6.3.3	Pore profile analysis	173
6.3.4	Water density analysis	173
6.3.5	Electrostatic free-energy for proton permeation	175
6.4	Results: DIMS simulations of the conformational changes in LacY	177
6.4.1	Stability of the model for the occluded state	177
6.4.2	Function characterisation of the occluded conformation	178
6.4.3	The R302-E325 salt-bridge	179
6.4.4	Comparison to experimental Double Electron Electron resonance data	180
6.4.5	Comparison to Major Facilitator Superfamily crystal structures . . .	183
6.4.6	DIMS simulations of the transition between cytoplasmic and periplas- mic open LacY	188
6.4.7	Structural comparison of cytoplasmic, occluded and periplasmic open LacY	190
6.5	Discussion	191
6.5.1	Conformational changes in LacY	191
6.5.2	The occluded state of LacY	192
6.5.3	The mechanism of the conformational transition	193

6.5.4	Conclusions	194
Appendix A	Relaxation dispersion experiments on nDsbD	195
A.1	Relaxation dispersion in the secondary binding site of nDsbD	195
Appendix B	Additional analysis of the simulations of LacY	197
B.1	Additional analysis of the simulations of LacY	197
B.1.1	Additional pore profile and water density analysis	197
B.1.2	Evaluation of the occluded model from DIMS and MD simulation .	200
B.1.3	Additional analysis of DEER distances	200
B.1.4	Order parameter analysis of the MD simulations	202
References		205

Chapter 1

Introduction

1.1 Biology at the molecular level

In this work the mechanism of macromolecular transitions is investigated by nuclear magnetic resonance (NMR) experiments and computer simulations. Transitions between different structural states of macromolecules underlie much of biology at the molecular level. For instance, proteins exchange between bound and unbound conformations. Cascades of binding events transmit a signal from a cell surface receptor to the nucleus where gene expression can then adapt to the signal. Moreover in many biological processes, proteins change their structure. Perhaps most strikingly are transporters that bind a substrate on one side of a biological membrane to release it on the other side after a conformational change. Conformational changes are vital in the enzymes that catalyse the myriad of chemical processes occurring in a cell at a given time. Structural changes in enzymes can orchestrate substrate binding and product release and the progression through a complex catalytic cycle (Boehr et al., 2006). Transitions such as these enable proteins to function as "molecular machines". Here the conformational changes in the N-terminal domain of the transmembrane oxidoreductase DsbD as well as the association of nDsbD with the C-terminal domain of DsbD (cDsbD) are studied by NMR and molecular simulations. In addition the mechanism of the conformational changes of the transporter Lactose permease (LacY) is probed by computer simulations.

At first, the set of transitions may look disparate. After all a conformational change and the association of two proteins are different processes. Yet the stable states of proteins in the set of examples and the transitions between them are shaped by the same interactions, electrostatics, hydrogen bonding and the hydrophobic effect in addition to the conformational preferences of the polymer shaped by its electronic structure. Lastly, changes in entropy

play a critical role in the transitions. The structured states of macromolecules are stabilised by many different such interactions and the interactions re-enforce each other. Such cooperativity has been identified as key feature of structures at the molecular level (Dill and Bromberg, 2011) and defines the energy landscapes of macromolecules. There is therefore another similarity between these processes. The product state of a folding reaction or of a protein-protein association reaction are both minima on a free-energy landscape stabilised by cooperative interactions.

1.2 Studying macromolecular transitions

How does one then study these ubiquitous transitions? Before discussing specific methodologies it is instructive to consider what is known about structural changes of proteins, based on simple yet exact statistical mechanical models (Dill, 1999). These models were devised to study protein folding but provide guidance for the study of conformational changes and protein binding which are the focus of this study. Modelling has emphasised that the structures of proteins are defined by many yet relatively weak interactions such as hydrogen bonds, salt-bridges, and while more complicated, also hydrophobic contacts (Chandler, 2005). Such interactions can thus be disrupted by thermal fluctuations, a hydrogen bond can be broken and a hydrophobic residue might become transiently exposed to the solvent. The importance of the weak nature of these interactions becomes clear when considering individual transition events. As each interaction contributes only marginally to the overall free energy of the system, the exact order in which they are formed and broken during a transition may not matter and their order may vary between individual transition paths. This does not mean that there are no preferences (Dill, 1999; Noé et al., 2009), *e.g.* on average a given helix might be formed before other elements in a folding reaction.

The probabilistic nature of the transition might then suggest that the tools of structural biology, high resolution NMR and X-ray experiments, can have only limited bearing on the study of the transition mechanisms. NMR and X-ray structures capture the dominant structure in solution or in a crystal and do not report on the fleeting interactions during the transitions. Indeed, approaches that give only limited (atomic) structural information have given important microscopic insights. For instance, mutations can be designed to probe the structures of the transition state ensembles for macromolecular transitions (Fersht, 1999). Such experiments report on whether a given residue is in a native-like conformation in the transition-state ensemble.

Single-molecule experiments only provide relatively limited structural information, such

as change in the distance between two fluorophores. However by following the trajectories of individual molecules critical mechanistic information can be gained (Bustamante et al., 2011), the sequence of events that convert one state to another can be followed directly. Other experiments report on ensemble averages and sparsely populated states such as transition intermediates might be difficult to detect. A good example are single-molecule experiments on the regulation of transcription by chromatin (Hodges et al., 2009). The experiments showed that nucleosomes block the progress of RNA polymerase on a strand of DNA. Occasionally, DNA dissociates from a nucleosome and the RNA polymerase can transcribe the gene past the obstacle.

Yet by trapping proteins in multiple different states and resolving their structures by X-ray crystallography, the molecular mechanism of the sodium coupled transporters has become increasingly well understood (Krishnamurthy et al., 2009). The recent work on Mhp1 is a good example (Shimamura et al., 2010). Crystal structures for structures open to either the periplasmic or the cytoplasmic sides of the membrane as well as an occluded structure were determined. This led to the conclusion that the alternating access to either side of the membrane is provided by the rigid body motion of a four helix motif. For many proteins such intermediate structures may exist, which can be studied by X-ray crystallography and NMR, and can provide crucial information about their transition mechanisms.

1.3 NMR experiments

NMR experiments that report on dynamics are a vital tool for studying the conformational landscapes of biological macromolecules (Henzler-Wildman and Kern, 2007; Palmer, 2004). NMR dynamics measurements give residue-specific insight and often provide atomic detail, revealing both fluctuations around the dominant structures and transitions to alternative conformational states.

1.3.1 Studying fluctuations by NMR

Studying fast-timescale dynamics by NMR

Spin-relaxation experiments report on very fast (ps-ns) fluctuations (Jarymowycz and Stone, 2006). Typically T_1 , T_2 and $^1\text{H} - ^{15}\text{N}$ Nuclear Overhauser Enhancements (NOEs) are recorded for backbone amides and are interpreted using the "model-free" approach. S^2 order parameters are obtained that describe the flexibility of individual bond vectors. However such experiments are insensitive to fluctuations slower than the overall rotational diffusion of

the molecule. As such spin-relaxation measurements can provide only indirect information about the mechanisms of macromolecular transitions. Nonetheless, such information can provide important clues about transitions. For instance, spin-relaxation experiments might reveal that a segment is disordered and becomes folded upon the binding of a ligand. While spin-relaxation experiments might only give limited insight into how a transition happens, they have explained, in some cases, why transitions occur. The fluctuations measured in the fast-timescale dynamics experiments are related to conformational entropy (Akke et al., 1993). An increase in the conformational entropy of protein side chains drives the binding of some sugar ligands to galectin-3 (Diehl et al., 2010).

Studying dynamics by residual dipolar coupling experiments

Residual dipolar coupling (RDC) measurements have also been used to study the fluctuations of proteins (Lakomek et al., 2006). RDCs are sensitive to a broader timescale of motions than spin-relaxation experiments and can probe dynamics up to the ms timescale. RDC experiments have highlighted correlated slow motions in the β -sheet of protein G (Bouvignies et al., 2005). Such correlated motions may be involved in allostery as they could transmit changes in one site of the protein, due *e.g.* ligand binding, to another, distal site.

1.3.2 Studying structural transitions by NMR

NMR relaxation dispersion experiments probe chemical exchange

NMR experiments cannot only reveal the fluctuations of proteins but also follow their transitions between alternative structures. Transitions between structural states exchange spins between different electronic environments, which is referred to as chemical exchange. Chemical exchange is detected in NMR experiments (Palmer et al., 2001). By detecting chemical exchange, NMR experiments such as relaxation dispersion measurements, monitor structural changes under equilibrium conditions. The experiments provide information about the populations of exchanging states, exchange rates as well as chemical shift differences between the states. The chemical shift differences report on the structural differences between the states. Relaxation dispersion experiments detect "invisible, excited" states, *i.e.* states that are somewhat higher in energy than the most populated state and which are consequently lowly populated. Such sparsely populated states are not detected by most experimental methods. "Excited" states are frequently important intermediates. An interesting example is the sparsely populated state of protein tyrosine phosphatases (Whittier et al.,

2013). In the absence of substrate the WPB loop that is part of the active site switches between a major open conformation and a minor closed conformation. Loop closure above the substrates positions the side chains for catalysis. Similar rates were obtained for the loop closure and the cleavage of the phosphate group. The conformational change is directly coupled to enzyme catalysis.

NMR relaxation dispersion has also been used successfully (Li et al., 2013; Sugase et al., 2007a) to study the association mechanisms of proteins. The work on the association between CREB and CBP is especially interesting (Sugase et al., 2007a). The NMR experiments revealed a three-state mechanism. The pKID domain of CREB is unfolded on its own. It binds the KIX domains of CBP in an encounter complex. One of the two helices of pKID is folded in the encounter complex. The other helix only folds after the formation of the encounter complex, to form the native complex. In addition, the study demonstrated that the encounter complex is significantly stabilised by the hydrophobic contacts. Previous work on encounter complexes in the association reactions of protein had suggested that electrostatics would primarily stabilise encounter complexes (Ubbink, 2009).

The chemical shift differences from relaxation dispersion enable the determination of high-resolution structures for excited states. Chemical shifts can be used to infer the likely local geometry of a residue (Shen and Bax, 2013) and thus restraints to determine a structural model. By collecting relaxation dispersion experiments for multiple nuclei enough restraints can be obtained for structure calculations (Vallurupalli et al., 2008). Residual dipolar couplings (RDCs) and residual chemical shift anisotropies for sparsely populated states can also be measured by relaxation dispersion experiments (Igumenova et al., 2007; Vallurupalli et al., 2007). RDCs and residual chemical shift anisotropies report on the relative orientation of a bond vector with respect to a reference frame and therefore provide important information on the structure of biomolecules. Through this approach the structure of a folding intermediate of the Fyn SH3 domain was elucidated (Neudecker et al., 2012). The backbone ^{15}N , ^1HN , ^{13}C , $^{13}\text{C}_\alpha$, $^1\text{H}_\alpha$ shifts and $^{15}\text{N} - ^1\text{H}$ RDCs as well as ^{13}CO residual chemical shift anisotropies of the excited state were determined for the structure calculation. The intermediate lacks one of the native β -strands which exposes an aggregation prone β -strand, such a folding intermediate may contribute to amyloid formation. A truncated mutant protein that mimics the folding intermediate was prone to aggregation, confirming the conclusions from the structural model for the intermediate state. This example and other such as the structure of the minor state of T4L99A lysozyme (Bouvignies et al., 2011) underline the potential of NMR experiments to characterise important intermediate structures in atomic detail.

Furthermore, the motions of methyl groups in a sparsely populated state have been investigated (Hansen et al., 2009). More recently hydrogen exchange measurements on excited states have been introduced (Long et al., 2014). Following the exchange of hydrogens with the solvent can reveal local folding and unfolding. NMR experiments that are sensitive to μs -ms dynamics can give highly detailed views of intermediate structures on the conformational landscapes of biomolecules, revealing both their structures and their dynamics, and therefore a powerful tool to study macromolecular transitions.

NMR experiments have been developed to study processes that are too fast or too slow to be studied by relaxation dispersion measurements. Relaxation dispersion experiments enable the investigation of processes with exchange rates between approximately 200 s^{-1} and 2000 s^{-1} . NMR chemical exchange saturation transfer (CEST) experiments capture dynamics slower than that, with exchange rates $k_{ex} \leq 200\text{ s}^{-1}$ (Vallurupalli et al., 2012). Slow chemical exchange processes can also be studied by magnetisation exchange experiments (Palmer et al., 2001). Transitions with $k_{ex} \gg 2000\text{ s}^{-1}$ can be investigated in spin-lock experiments.

Further NMR approaches to study macromolecular transitions

NMR paramagnetic relaxation enhancement (PRE) measurements also provide valuable insights into the conformational landscapes of biomolecules (Tang et al., 2006). PRE experiments have revealed the encounter complexes that are formed in protein-protein association reactions. Unspecific encounter complexes were found and it is thought that unspecific interactions enhance many protein-protein binding reactions by bringing the binding sites into close proximity (Kim et al., 2008; Kozakov et al., 2014).

Hydrogen exchange experiments are highly sensitive even to very sparsely populated states. Such experiments have suggested a step-wise folding mechanism for cytochrome *c* (Maity et al., 2005). The melting of RNAs can also be followed by such experiments (Rinnenthal et al., 2010, 2011). This revealed the base pairs that play a crucial role in the temperature-dependent folding and unfolding mechanism of the fourU thermometers in mRNAs. Temperature dependent unfolding of these RNA hairpins enables ribosome binding to the mRNA, enhancing the production of key proteins.

Experiments have been developed to study the folding of biomolecules as it happens, in real time (Balbach et al., 1995), which provides direct insight into transitions. It has become possible to record multi-dimensional NMR experiments during folding reactions, which provide detailed residue-specific information. The folding of the adenine riboswitch was followed by real-time NMR experiments (Lee et al., 2010). The measurements showed

how a key helix was only stabilised after the formation of tertiary contacts. The prospect of directly following reactions in atomic resolution is tantalising, yet the approach is still limited by the speed at which two- or three dimensional spectra can be acquired.

1.4 Molecular dynamics simulations

1.4.1 Atomistic molecular dynamics

Atomistic molecular dynamics (MD) calculations provide direct insight into the transition mechanism of biological macromolecules. MD simulations yield trajectories that give a highly detailed description of how atomic coordinates change with time. Typically, trajectories on the order of several μs can be generated. Many biochemical processes, such as folding, protein-protein association and conformational changes occur infrequently and may not be observed on the timescale accessible to atomistic MD simulations. Even if the process of interest is observed, conclusions are best drawn carefully, keeping in mind limitations of the simulation parameters and sampling. Ideally one would observe processes multiple times in simulations, further underscoring the importance of good sampling. Many schemes have been proposed to alleviate the insufficient sampling in MD simulations. The Shaw group has used specialised hardware to achieve extremely long simulations (Lindorff-Larsen et al., 2011), from hundreds of μs to ms, to probe the mechanism of protein folding. These simulations have demonstrated that the folding mechanism of proteins is largely determined by the interactions in the native state (Best et al., 2013).

1.4.2 Enhanced sampling methods

Sampling cannot not only be enhanced by brute force but also by the employing methods designed to enhance the exploration of the free-energy landscapes. A common strategy is the introduction of a bias potential into the simulation system. In the well-established umbrella sampling approach (Leach, 2001), bias potentials are added along a coordinate of interest that describes the state of the system. The free-energy along this coordinate is then obtained in the calculations which can give critical insight into intermediates (Fowler and Sansom, 2013) and transition states of macromolecular processes. A related technique is metadynamics (Barducci et al., 2011, 2008) which has been used in this study. In metadynamics, bias potentials are progressively added during the simulations as the space defined by the variables of interest are explored. The history dependent bias built up during the simulations yields the free energy surface of interest. Metadynamics has been used for example

to investigate the folding of the FBP28 WW domain (Juraszek and Bolhuis, 2010). In this study, the free-energy profiles calculated by metadynamics agreed with one of the proposed folding mechanism and ruled out a second mechanism. Metadynamics calculations have also been used to explore "excited states" from NMR relaxation dispersion experiments (Palazzesi et al., 2013).

Rare events can also be studied using a different approach. Free-energy calculations based on umbrella sampling and metadynamics require substantial knowledge about the system of interest. In the absence of this, the choice of coordinates in which to enhance sampling is highly non-trivial (Barducci et al., 2011). Using methods such as Transition Path Sampling (Juraszek and Bolhuis, 2010) and Dynamics Importance Sampling (DIMS) (Perilla et al., 2011) ensembles of transition events can be generated. For DIMS only a target structure is required to sample transition events. DIMS MD calculations have shown that a set of salt bridges in adenylate kinase functions as "zipper", breaking one salt-bridge facilitates the disruption of the next one (Beckstein et al., 2009). The cooperative "zipper" is thought to favour the open and closed conformation of adenylate kinase at the expense of intermediate structures. In this work DIMS MD was used to study rare transition events between the cytoplasmic and periplasmic open structures of LacY.

1.4.3 Comparison to experiment

Computer simulations are theoretical models (Frenkel and Smit, 2002), and not an *in silico* representation of reality. Such models, and the conclusions based on them, require careful testing in experiments. For instance, molecular simulations suggested that the conformational change that activates the signalling protein NtrC is facilitated by transient non-native hydrogen bonds that stabilise transition intermediates (Gardino et al., 2009). The mechanism was then confirmed by mutations and NMR relaxation dispersion experiments. NMR is in general well-suited to probe computational models. Many parameters commonly measured by NMR such as inter-atomic distances, dihedral angles and spin-relaxation parameters can be calculated straightforwardly from MD trajectories (Smith et al., 1995). Comparison of MD simulations of ribonuclease H to spin-relaxation experiments identified the conformational states most representative of the experimental solution ensemble (Robustelli et al., 2012). In one segment of a MD trajectory of ribonuclease H, a serine side chain formed a hydrogen bond with the backbone amide of a leucine to cap a helix. In another segment of the trajectory the hydrogen bond was not yet formed. The computed S^2 order parameter for the backbone amide of the leucine did not agree with experiment

in absence of the hydrogen bond, while the experimental S^2 is well reproduced in the segment featuring the hydrogen bond. Calculating chemical shifts from the trajectory using empirical models further strengthened the conclusion that the hydrogen bond caps the helix.

1.5 The trans-membrane oxidoreductase DsbD

In this work, transition events in two different biological systems, (1) in the periplasmic domains of the transmembrane oxidoreductase DsbD and (2) the transporter LacY, were investigated. Conformational changes in nDsbD and its association with cDsbD were probed by NMR and molecular simulations. The mechanism of the conformational changes in the transport cycle of LacY were studied by computer simulations alone.

In Gram-negative bacteria, the transmembrane oxidoreductase DsbD moves reductant from the cytoplasm into the periplasm. The periplasm of Gram-negative bacteria is an oxidising environment, but three important biosynthetic processes in the periplasm are dependent on the reductant provided by DsbD:

1. The isomerisation of incorrectly formed disulphide bonds requires reductant.
2. Reductant is needed in the protection against oxidative damage.
3. Pathways that lead to assembly of cytochrome *c* in the periplasm utilise reductant.

1.5.1 Isomerisation of disulphide bonds

The DsbD pathway channels reductant to disulphide bond isomerases that ensure that periplasmic proteins attain their cognate disulphide bonds (Cho and Collet, 2012). Disulphide bonds are present in many proteins in the periplasm of bacteria. For instance, in *Escherichia coli* (Dutton et al., 2008) hundreds of periplasmic proteins may feature disulphide bonds. Disulphide bonds are also common in extracytoplasmic proteins in eukaryotes, *e.g.* in secreted proteins. The covalent link between cysteines confers additional stability to these proteins, which is the main role of protein disulphide bonds. Disulphide bonds stabilise folded conformations by disavouring unfolded states by reducing their conformational entropy (Fersht, 1999). A disulphide bond covalently locks together parts of polymers chains which become less disordered as a consequence. Incorrectly formed disulphide bonds by contrast, may lock proteins in non-native, non-functional conformations. Disulphide bond isomerase resect non-cognate disulphide bond, making it possible for such proteins to attain their native structure.

To understand why bacteria need disulphide bond isomerases it is instructive to consider the pathway that introduces the disulphide bonds. Disulphide bonds are introduced into proteins by the powerful oxidase DsbA (Ito and Inaba, 2008). DsbA acts on largely unfolded proteins as they are translocated from their cytoplasmic origin into the periplasm, as illustrated in Fig. 1.1, by the Sec machinery. Two electrons are acquired by DsbA as it links two cysteines to form a disulphide bridge. DsbA is re-oxidised and readied for introducing the next disulphide bond by transferring the two electrons to DsbB. DsbB then feeds the electrons into the respiratory chain. The oxidation of two cysteines to form a disulphide bond by DsbA is relatively unspecific, as there are few structural constraints on disulphide bond formation in the absence of a stable, folded structure. Many proteins have only two cysteines that might form a disulphide bond and therefore unspecific oxidation does not pose a problem. Yet some proteins have multiple native disulphide bonds that are between non-consecutive cysteines. These cysteines may be incorrectly linked by DsbA, trapping such proteins in non-native, non-functional conformations.

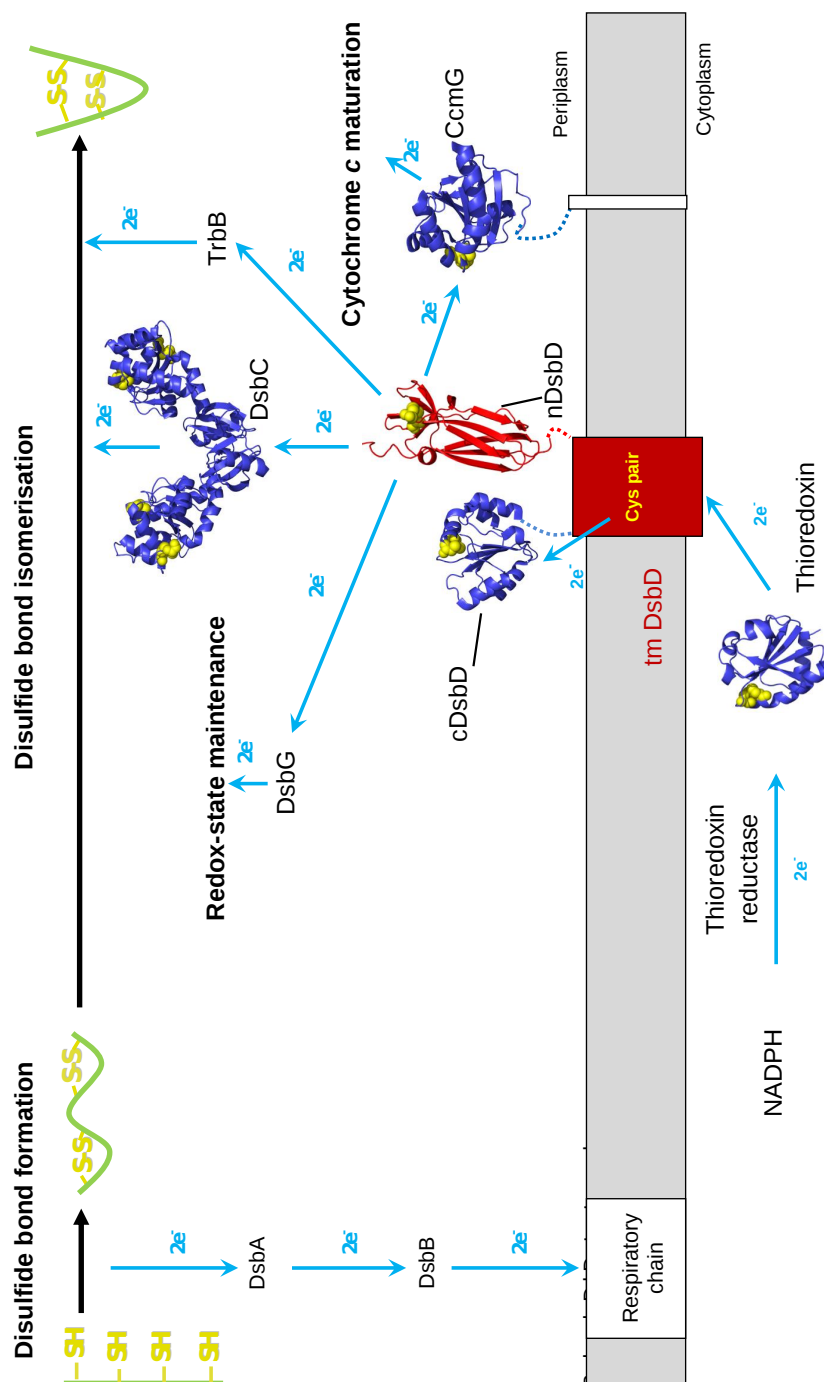


Fig. 1.1 The DsbD pathway provides reductant in the oxidising periplasm of Gram-negative bacteria. Reductant is used for disulphide bond isomerisation by DsbC or TrbB, protection of redox-sensitive groups by proteins such as DsbG as well as for the maturation of cytochrome c. Disulphide bonds are formed by DsbA in largely unfolded proteins.

In *E. coli* and many other Gram-negative bacteria DsbC rescues such misfolded proteins. DsbC has two thioredoxin fold domains and they form a relatively hydrophobic V-shaped binding site (Haebel et al., 2002) for its substrates. For its activity DsbC requires a supply of reductant from DsbD as indicated on Fig. 1.1. Two mechanisms have been proposed for the function of DsbC (Cho and Collet, 2012). In the first mechanism, DsbC reduces the non-cognate disulphide bond of its substrate. Subsequently, the substrate protein forms the native disulphide bond and returns the electrons to DsbC. In the second mechanism, DsbC would reduce the disulphide bond. The substrate would then be re-oxidised by DsbA. Such a cycle, or even cycles, of reduction and oxidation would then lead to the correct folding of the substrates and the formation of the cognate disulphide bridges. In the first mechanism DsbD would maintain DsbC in its reduced isoform in the oxidising environment of the periplasm, while in the second mechanism reductant would have to be resupplied by DsbD for each disulphide bond that is resected.

Only a handful of substrates have been identified for DsbC. Among these are RcsF (Cho and Collet, 2012) and possibly LptD (Depuydt et al., 2011), which both have very interesting biological functions. RcsF is part of a signalling cascade that monitors the integrity of the periplasm, the cell wall and the outer membrane (Cho and Collet, 2012). LptD introduces lipopolysaccharide (LPS) into the outer membrane of Gram-negative bacteria. The sugar moieties of LPS extend far beyond the core of the membrane and protect bacteria. LptD is the target of a new class of antibiotics and it is therefore the focus of much recent work. One group has found that LptD is a substrate of DsbC (Denoncin et al., 2010) while another group concluded that DsbC is not required for its assembly (Ruiz et al., 2010).

The "house keeping" role of DsbD and its downstream effectors

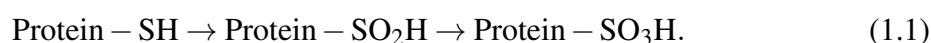
Collet and co-workers have proposed an explanation for this discrepancy (Depuydt et al., 2011). DsbA may correctly oxidise the cysteines of LptD frequently enough so that a reductive pathway is not absolutely required for LptD assembly, but it might make the process more effective. Many aspects of the DsbD pathways might follow such as paradigm, they are not essential in many bacteria, but they have important "housekeeping" functions that can make periplasmic pathways work more efficiently (Mavridou et al., 2012b). Nonetheless DsbD appears to be absolutely required in *Neisseria meningitidis* (Kumar et al., 2011).

The disulphide bond isomerase TrbB is involved in bacterial conjugation

Recently, the disulphide isomerase TrbB was discovered as another downstream effector of DsbD (Hemmis et al., 2011). TrbB is encoded by a plasmid and is required for the assembly of the bacterial conjugation apparatus, but its exact function is unclear. Plasmids and bacterial conjugation are of course important in the spread of antibiotic resistance genes among bacterial populations. A large number of plasmid carry proteins with a thioredoxin fold (Hemmis and Schildbach, 2013) and may catalyse disulphide bond isomerisation. A disulphide isomerase was also identified on a plasmid that carries drug-resistance genes (Premkumar et al., 2014).

1.5.2 DsbD supplies reductant to protect against oxidative damage

The second important periplasmic process powered by the DsbD system is the protection against oxidative damage (Cho and Collet, 2012). Due to the oxidising nature of the compartment non-specific oxidation can occur. For instance, catalytic cysteine residues can become oxidised (Depuydt et al., 2009) to sulfenic acid and subsequently to sulfinic and sulfonic acid:



Conversion to sulfonic acid can lead to irreversible loss of enzyme activity. Before this happens DsbG of *E. coli* can rescue the activity of proteins by reducing sulfenic acid back into cysteine thiols. As shown in Fig. 1.1, DsbG receives reductant from the N-terminal domain of DsbD (nDsbD). Trans-peptidases have been identified as DsbG substrates. These enzymes cross-link peptidoglycans of the cell wall.

In *Neisseria gonorrhoeae* (Brot et al., 2006) and *N. meningitidis* (Quintern et al., 2009) DsbD provides electrons to protect methionine from oxidation. In these Gram-negative bacteria methionine of periplasmic proteins can be oxidised to methionine sulfoxide. The modification reduces the hydrophobicity of methionine by introducing a polar oxygen to the sulphur moiety. Methionine oxidation can affect both protein conformation and function. PilB is thought to protect *N. gonorrhoeae* (Brot et al., 2006) from radical oxygen species generated by the host immune response to the infection with this pathogen. Protection against radical oxygen species may be a critical function of DsbD. For example, *Caulobacter crescentus* features a protein related to DsbD, ScsB, that may help to scavenge peroxide species.

1.5.3 DsbD supplies reductant for cytochrome *c* maturation

The third important function of reductive pathways in the bacterial periplasm is in the assembly of cytochrome *c*. Cytochrome *c* is of course a critical component of respiratory chains. In *E. coli*, DsbD also supplies CcmG of the cytochrome *c* biogenesis system with reductant, as depicted in Fig. 1.1 (Ito and Inaba, 2008). In *E. coli*, DsbD might play a "house keeping" role in this biogenesis pathway (Mavridou et al., 2012b). Apo cytochrome *c* is imported into the periplasm by the Sec translocase. The apo cytochrome *c* can be used straight away by the cytochrome *c* biogenesis machinery. Two cysteines of apo cytochrome *c* are covalently linked to the haeme, by thioether bonds, to form the native protein. Alternatively as reported by Mavridou *et. al.*, the apo protein can become oxidised by DsbA. CcmG can then cleave the disulphide bond introduced by DsbA and cytochrome *c* can be assembled correctly.

1.5.4 Occurrence of DsbD related pathways

Before discussing the function of DsbD in more detail, it is interesting to consider how prevalent this protein is and the occurrence of reductive pathways in oxidising environments. DsbD is present in many Gram-negative bacteria such as *E. coli*, *N. meningitidis* and *N. gonorrhoeae*. A simpler, related protein, CcdA, is present in the chloroplasts of plants, archae and most bacteria (Cho and Collet, 2012). CcdA has a transmembrane domain similar to DsbD but lacks the two periplasmic domains of DsbD, the C-terminal thioredoxin fold domain, cDsbD, and the N-terminal immunoglobulin fold domain, nDsbD. The domain structure of DsbD is highlighted in Fig. 1.2. nDsbD interacts with proteins of the disulphide isomerase, oxidative damage prevention and cytochrome *c* maturation pathways. The transmembrane domain of CcdA directly interacts with proteins like CcmG to supply reductant to the cytochrome *c* maturation pathway. CcdA lacks a domain such as nDsbD that could acts as "redox hub" and is thought to have fewer substrates than DsbD (Katzen et al., 2002). It is thought that DsbD evolved from an ancestral CcdA-like protein acquiring the periplasmic domains to serve a wider array of pathways than just cytochrome *c* biogenesis.

Eukaryotes can also isomerise incorrectly formed disulphide bonds (Depuydt et al., 2011). However, the proteins do not appear to be related to DsbD and DsbC. In the endoplasmic reticulum, disulphide bonds are introduced to proteins. PDI, protein disulphide-isomerase, introduces disulphide bonds into proteins. This enzymes also reduces incorrectly formed disulphide bonds so that substrates can attain the cognate disulphide bonds. The

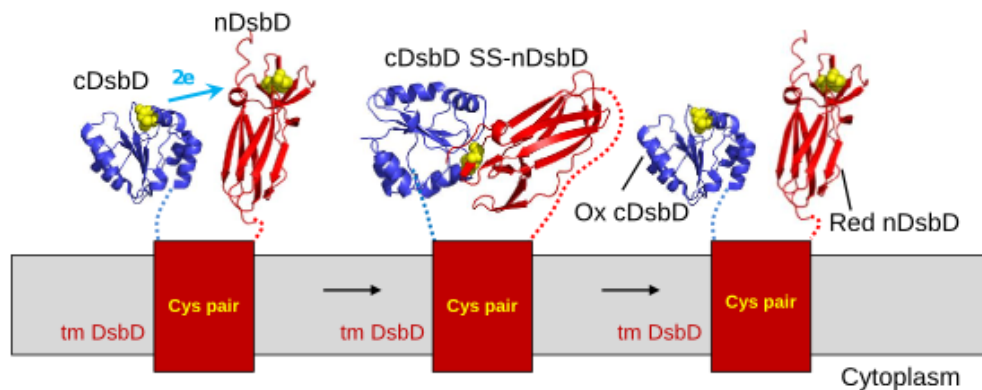


Fig. 1.2 The C-terminal domain of DsbD, cDsbD, receives reductant from the transmembrane domain of DsbD. cDsbD passes on reductant to nDsbD.

endoplasmic reticulum is an oxidising environment. Similar to DsbC, PDI engages in a reductive process in an oxidising environment to isomerise disulphide bonds.

1.5.5 The reductant transfer reactions in the DsbD pathway

Having discussed the main functions of the reductive DsbD pathway, it is important to consider the reductant transfer reactions themselves. As indicated in Fig. 1.3, reductant transfer happens in a series of thiol-disulphide exchange reactions (Ito and Inaba, 2008). Such reactions are a general feature of the Dsb proteins, both in the oxidising DsbA/DsbB pathway and in the reductive DsbD pathway (Fig. 1.1). No metals are involved in contrast to many other biological redox reactions. To transfer reductant *via* thiol-disulphide exchange reactions proteins first have to associate to form a complex. The chemical reaction commences with one of the thiols, of the reduced protein, losing a proton to form a thiolate. The thiolate nucleophilically attacks the disulphide bond of the oxidised protein, which is broken as a consequence. A new inter-molecular disulphide bond is formed. Crystal structures of such intermediates have been obtained. Structures of nDsbD in covalent complexes with cDsbD (Rozhkova et al., 2004), DsbC (Haebel et al., 2002) and CcmG (Stirnemann et al., 2005) have been solved. The intermolecular disulphide bond is severed in the next step. In this step the second cysteine thiol on the reductant providing protein attacks the intermolecular disulphide bond, to form an intramolecular disulphide bond. Overall, the reductant is transferred from one protein to the other, by swapping two thiols for a disulphide bond.

The reductant supplied by nDsbD to the periplasmic pathways involved in disulphide bond isomerisation, protection against oxidative damage and maturation of cytochrome *c* originates in the cytoplasm (Cho and Collet, 2012; Ito and Inaba, 2008). Reductant in the

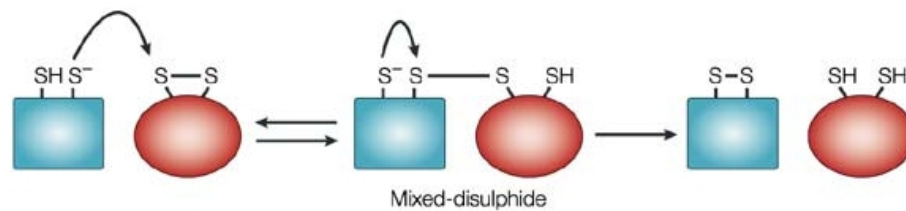


Fig. 1.3 Thiol-disulphide bond exchange reactions transfer reductant between the Dsb proteins. Proteins have to physically associate for such reactions. One of the thiols (one the blue protein) loses its proton and forms a negatively charged thiolate. The thiolate can then attack the disulphide bond. An intermolecular disulphide bond is formed. The other thiol (on the blue protein) will then attack the intermolecular bond, forming a new, intramolecular disulphide bond (on the blue protein). The two proteins exchange thiols for a disulphide bond in this reaction. The figure was reproduced with permissions from Sevier and Kaiser, 2002.

form of nicotinamide adenine dinucleotide phosphate (NADPH) is the starting point of the cascade of thiol-disulphide exchange reactions. Two electrons of NADPH are transferred to thioredoxin. Thioredoxin channels the reductant to the transmembrane domain of DsbD. The C-terminal domain of DsbD, cDsbD, then is in turn reduced by the transmembrane domain of DsbD. cDsbD then passes on the electrons to the N-terminal domain of DsbD, nDsbD. nDsbD then distributes the reductant in the periplasm as discussed above and functions as a "redox hub" (Stirnemann et al., 2006).

Reductant transfer through the transmembrane domain of DsbD

Much is already known about this pathway. Experiments have even given a description of reductant transfer across the membrane although no high-resolution structure is available for the transmembrane domain of DsbD. Site-directed alkylation was used to infer the water accessibility of the transmembrane domain of DsbD. The transmembrane domain was found to have an hour-glass shape (Cho et al., 2007) as shown in Fig. 1.4. This pore is water accessible. Site-directed alkylation experiments, did not detect major conformational changes depending on the redox state of the catalytic cysteine residues (Cho and Beckwith, 2009). The protein then functions according to the following mechanism. Thioredoxin, which is in the cytoplasm would interact with the transmembrane domain, which may induce conformational changes to accommodate it in the pore. Reductant is transferred from thioredoxin to the two cysteines in the centre of the membrane. After dissociation of thioredoxin, cDsbD can bind from the periplasmic side of the membrane. Again, the binding might be facilitated by conformational changes in the transmembrane domain. cDsbD then

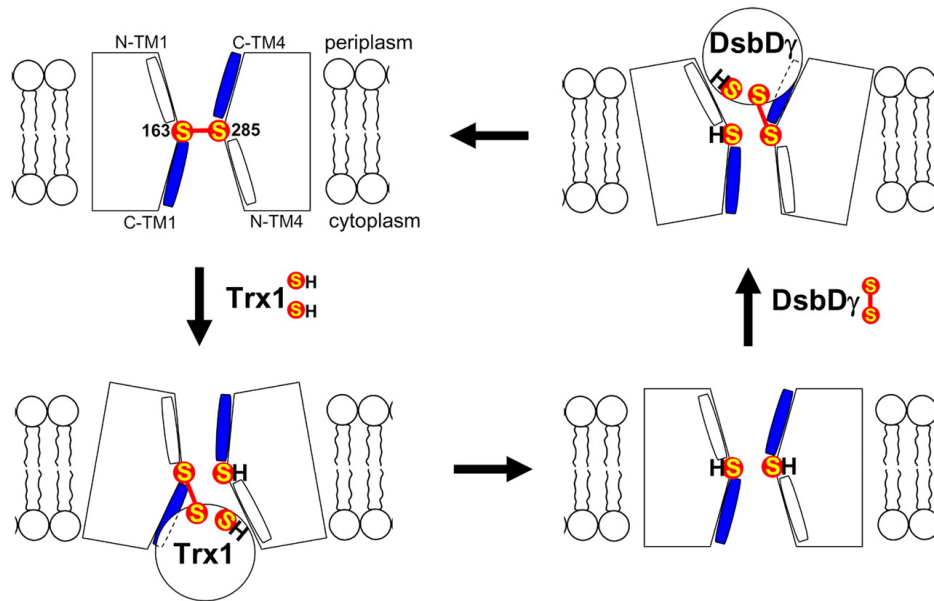


Fig. 1.4 A model for transmembrane reductant transfer has been proposed by Beckwith and co-workers. This figure was originally published in the *Journal of Biological Chemistry*. Cho, S.H. and Beckwith, J. Two snapshots of electron transport across the membrane: insights into the structure and function of DsbD. *J. Biol. Chem.* 2009; 284:11416-11424. the American Society for Biochemistry and Molecular Biology.

receives reductant from the two cysteines at the centre of the pore, before it dissociates from the transmembrane domain.

1.5.6 Control of the reductant transfer reactions

Moving electrons from cytoplasmic thioredoxin to downstream effectors such as DsbC, DsbG or CcmG is thermodynamically favourable (Rozhkova and Glockshuber, 2008). This can be seen from the standard reduction potentials in Table 1.1, which are all positive. The reductant transfer from cDsbD to nDsbD is only barely favourable, by about 7 mV according to Table 1.1, and this step is best considered thermodynamically neutral. The reactions in this pathway are sequentially coupled, *i.e.* reduced nDsbD is converted back into oxidised nDsbD as it provides reductant for DsbC, DsbG or CcmG. This should shift the equilibrium for the thermodynamically neutral step towards products of the reaction, oxidised cDsbD and reduced nDsbD.

The periplasm is an oxidising environment 1.1 and reactivity in the reductive DsbD pathway has to be controlled. As discussed above and depicted in Fig. 1.3, de-protonation of a cysteine thiol to form a thiolate is the first step in the reductant transfer by a thiol-

Table 1.1 Standard redox potentials E^0 for the Dsb pathways. Values were reproduced from Fig. 3 in Rozhkova and Glockshuber, 2008 and Table 1 in Ito and Inaba, 2008. The ΔG^0 , the free-energy differences for the species in standard state with concentrations of 1 M at 25 °C are favourable when E^0 for is positive considering, $\Delta G^0 = -nFE^0$, where F is Faraday's constant.

Reductive DsbD pathway	Standard redox potential E^0
NADPH	−320 mV
Thioredoxin	−284 mV
Transmembrane DsbD	−246 mV
cDsbD	−235 mV
nDsbD	−232 mV
CcmG	−209 mV
DsbC	−140 mV
DsbG	−129 mV
Oxidative pathway	
DsbA	−122 mV
DsbB	−224 mV
Periplasm	−165 mV

disulphide exchange reaction. Therefore reactivity can be controlled *via* the acidity of the cysteine thiols. cDsbD is particularly noteworthy in this respect. The cysteine that initiates the thiol-disulphide exchange reaction has a pK_a of 10.5 (Mavridou et al., 2007); the thiol will not lose its proton to form a thiolate. The formation of the negatively charged thiolate is disfavoured by close proximity of the negatively charged side chain of D455. D455 is shielded from the solvent in the complex of cDsbD with nDsbD and becomes protonated (Mavridou et al., 2009). C461 can then lose its proton and engage in the thiol-disulphide reaction that transfers reductant to nDsbD.

The redox potentials also show that a futile cycle could be formed in the periplasm. Futile reactions are prevented by kinetic barriers that separate the reductive DsbD pathway and the oxidative DsbA/DsbB system. For instance, nDsbD could be oxidised by DsbA based on the standard potentials. The rate of oxidation by of nDsbD by DsbA is slow (Rozhkova et al., 2004) and in particular much slower than the rate for the cognate reaction with DsbC. Crystal structures of nDsbD have shown that its active site cysteines are covered by a loop (Goulding et al., 2002; Mavridou et al., 2011), the so-called cap loop, as indicated in Fig. 1.5. I have previously demonstrated by NMR spin-relaxation measurements that the loop is not disordered (Stelzl, 2010). The cap loop might prevent DsbA and other non-cognate reaction partners from accessing the binding site and oxidising the active site cysteines.

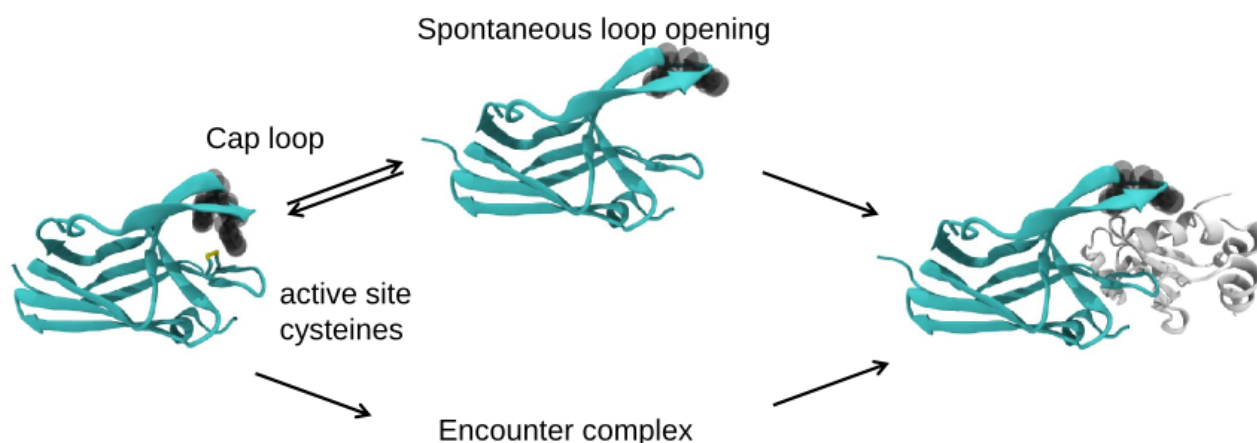


Fig. 1.5 nDsbD and cDsbD could interact by conformational selection or induced fit mechanisms.

1.5.7 Studying the interaction of cDsbD and nDsbD

The cap loop however needs to open for nDsbD to form complexes with its physiological interaction partners. The loop is open, albeit to different extents in the crystal structures of the covalent complexes of nDsbD with cDsbD, DsbC and CcmG. The mechanism of loop opening is not understood. The loop opening could be spontaneous and partners such as cDsbD would then bind in a conformational selection mechanism. Alternatively, the loop opening in nDsbD might only occur upon encountering its interaction partners. The loop would then open in an induced-fit mechanism. The two mechanisms are illustrated in Fig. 1.5. Understanding the mechanism of cap loop opening in the association reaction between nDsbD and cDsbD was a key aim of this project.

In addition to the loop opening mechanism, there is another interesting aspect of the association reaction of cDsbD and nDsbD that is poorly understood. Namely, the interaction affinity between the two domains depends on their redox-states (Mavridou et al., 2011). Reduced cDsbD associates with oxidised nDsbD. Reductant is then transferred from reduced cDsbD to oxidised nDsbD. The oxidised cDsbD and reduced nDsbD form a much weaker complex which subsequently falls apart. The rate at which the complex would fall apart is not known. Rapid dissociation of the complex after the reaction might prevent reverse reactions which would be expected to occur given that the reductant transfer is just about thermodynamically favourable. The decrease in the interaction affinities after the reductant transfer also frees up the two periplasmic domains for further reactions. cDsbD can receive reductant again from the transmembrane domain of DsbD while nDsbD can distribute the reductant to the pathways that require it. The redox-state dependent formation of the cDsbD

and nDsbD complex is thought to increase turnover in the pathway. It has been proposed that differences in packaging constraints could explain the redox-state dependence of the interaction affinities. Two thiols fill more space than a disulphide bond. Both the redox-state of nDsbD and cDsbD contribute to the observed redox-state dependence of the interaction affinities. Understanding the redox-state dependent association and dissociation of cDsbD and nDsbD was another important aim of this project.

The loop opening mechanism was investigated by atomistic MD simulations of nDsbD and their comparison to NMR data. Relaxation dispersion experiments were used to explore the conformational landscape nDsbD. MD trajectories for oxidised and reduced nDsbD were compared to understand the physical basis of the redox-state dependent interaction affinities.

To understand the association reaction between cDsbD and nDsbD NMR relaxation dispersion experiments were employed. One aim was to determine the kinetic basis of the redox-state dependent interaction between the two domains. The association reaction was also probed in MD simulations.

1.6 The mechanism of conformational changes in the transporter LacY

The Major Facilitator Superfamily (MFS) of transporters is the largest group of transporters (Hunter et al., 2012). MFS transporters are involved in many crucial functions, such as glucose permeation into human cells (Sun et al., 2012). LacY is the best studied member of this group (Forrest et al., 2011) and serves as paradigm for transport across biological membranes according to the alternating access mechanism. LacY switches between a periplasmic open and a cytoplasmic open conformation in its transport cycle. The transporter binds protons and sugars on the periplasmic side and releases them in the cytoplasm. LacY hence acts as a symporter using the proton electrochemical gradient to transport sugars into cells.

The mechanism of the conformational change between the cytoplasmic and periplasmic open states is however not understood. Rigid body motions in a so-called "rocker-switch" mechanism have been proposed to interconvert the two structural states (Dang et al., 2010). Such a rocker-switch mechanism may however allow the leakage of vital transmembrane gradients. An "airlock" mechanism (Forrest, 2013) has been proposed to explain how transporters prevent uncontrolled permeation. The mechanism of LacY was explored by DIMS and equilibrium MD calculations.

Before discussing the results from the investigation of DsbD and LacY, a brief overview

about NMR relaxation experiments and MD simulations is provided in the next chapter.

Chapter 2

Theory and methodological background

I will briefly discuss some of the methodological background for the study of fast (ps-ns) and slow (μ s-ms) dynamics by NMR. The second part of the chapter explains some of the computer simulation methods employed in this project.

2.1 Studying dynamics by NMR

2.1.1 Fast-timescale dynamics

To study ps-ns dynamics by NMR the relaxation rate for longitudinal magnetisation T_1 , and transverse magnetisation T_2 and $^{15}\text{N}-^1\text{H}$ Nuclear Overhauser Enhancements (NOEs) are typically measured for backbone amide ^{15}N (Jarymowycz and Stone, 2006). Two mechanisms dominate the relaxation of amide spins, dipole-dipole interactions and the chemical shift anisotropy (CSA). The amide proton will interact with the attached nitrogen nucleus *via* a dipole-dipole interaction. The strength of the dipole-dipole interaction is sensitive to the relative orientation of the amide bond vector with respect to the static magnetic field. The relative orientation of the amide is determined not only by the rotational diffusion of the protein molecule but also by its internal dynamics. The motions modulate the dipole-dipole interaction resulting in an oscillating field that can trigger spin relaxation.

A similar mechanism underpins relaxation of the amide ^{15}N nitrogen by CSA. The ^{15}N chemical shift tensor is anisotropic. The shielding of a spin by the nearby electrons thus depends on the relative orientation of the molecule towards the static magnetic field (Cavanagh et al., 2007). The shielding that the ^{15}N nitrogen experiences changes as the molecule tumbles in solution and the position of the amide changes due to the internal dynamics of the protein. The fluctuations in the shielding can result in the relaxation of the spin.

2.1.2 Interpreting fast timescale NMR experiments

NMR spin relaxation experiments are usually interpreted using the "model-free" formalism (Lipari and Szabo, 1982a,b). In this approach, contributions to the relaxation of the NMR signal from the overall tumbling of the molecule in solution and the internal motions of bond vectors can be distinguished. The rotational correlation function for a bond vector in a static molecule is given by eqn. 2.1.

$$C(t) = S^2 + (1 - S^2)e^{-t/\tau_e} \quad (2.1)$$

The rotational correlation function depends on τ_e , the internal correlation time for the motion of the bond vector, and order parameter S^2 . S^2 describes the flexibility of the bond vector. Its value can vary between 1, for a perfectly rigid bond vector, and 0 for unrestricted motions of the amide bond. Both simpler and more complex models than eqn. 2.1 are used to describe bond vector motions by the "model-free" approach (Mandel et al., 1995). In this project previously determined S^2 backbone amide order parameters (Stelzl, 2010) were compared to MD simulations (Chapter 3).

2.1.3 Chemical exchange

The exchange between different electronic environments leads to decay of the transverse magnetisation in a NMR experiment (Palmer et al., 2001). Here the simple case of a single spin exchanging between two states, $A \rightleftharpoons B$, is considered. The decay of the transverse magnetisation is described by eqn. 2.2.

$$\frac{d}{dt} \begin{bmatrix} M_A^+(t) \\ M_B^+(t) \end{bmatrix} = \begin{bmatrix} -i\omega_A - R_{2A}^\circ - p_B k_{ex} & p_A k_{ex} \\ p_B k_{ex} & -i\omega_B - R_{2B}^\circ - p_A k_{ex} \end{bmatrix} \begin{bmatrix} M_A^+(t) \\ M_B^+(t) \end{bmatrix} \quad (2.2)$$

Here ω_A and ω_B are the resonance frequencies of states A and B . The populations of the two states are denoted p_A and p_B . Note that the exchange rate k_{ex} and the populations then give the forward and backward rate constants for the reaction $k_f = p_B k_{ex}$ and $k_b = p_A k_{ex}$. R_{2B}° and R_{2A}° are the intrinsic relaxation rates for transverse magnetisation for the two states.

The magnetisation at a given time is then calculated by:

$$M(t) = e^{[i\hat{\omega} - \hat{R} + \hat{k}]t} M(0), \quad (2.3)$$

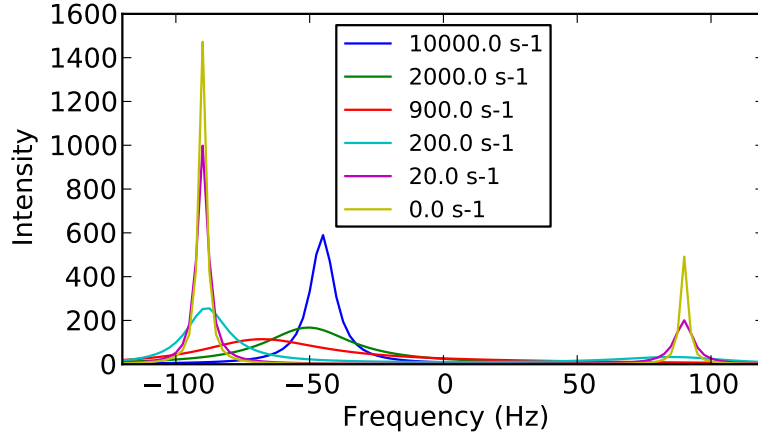


Fig. 2.1 The effect of chemical exchange on NMR spectra is illustrated using the Bloch-McConnell equation. The exchange was simulated using eqn. 2.3 and spectra were obtained by Fourier transformation. The resonance frequency difference between the two exchanging states was 180 rad/s. The exchange processes shown are strongly skewed with $p_A = 0.75$ and $p_B = 0.25$. The parameters used were chosen based on the example in Fig.1 in Palmer et al., 2001.

$$\text{where } M = \begin{bmatrix} M_A^+(t) \\ M_B^+(t) \end{bmatrix}, \hat{\omega} = \begin{bmatrix} \Omega_A & \\ & \Omega_B \end{bmatrix}, \hat{R} = \begin{bmatrix} R_{2A}^o & \\ & R_{2B}^o \end{bmatrix} \text{ and } \hat{k} = \begin{bmatrix} -k_f & k_b \\ k_f & -k_b \end{bmatrix}$$

Relaxation dispersion experiments make it possible to recover the parameters in eqn. 2.2 or eqn. 2.3 that determine the decay of transverse magnetisation. Probing the resonance frequency differences, exchange rates and the population of the exchanging states enables detailed investigations of protein mechanisms. Relaxation dispersion experiments are used in Chapter 4 and Chapter 5 to study a conformational change and protein-protein interactions.

Fast, intermediate and slow exchange

Before proceeding it can be noted that Fig. 2.1 shows the three prototypical cases for chemical exchange (Palmer et al., 2001). First, small exchange rates lead to two well resolved peaks in the simulated spectrum. The system is in slow exchange on the NMR chemical shift timescale when $k_{ex} < \Delta\omega$. The purple line in Fig. 2.1 also shows that the minor peak can be strongly broadened even for a slowly exchanging system and might not be observable in experiments (Palmer et al., 2001). Secondly, increasing the exchange rate, so that the system is in intermediate exchange, $k_{ex} \approx \Delta\omega$, results in a broad peak (light blue line,

$k_{ex} = 200 \text{ s}^{-1}$, Fig. 2.1) for the major state. The resonance for the minor state is very broad and might be too broad to be detected in experiments. The peak position of the major state is appreciably shifted towards the minor state. Thirdly, increasing the exchange much further leads to fast exchange on the chemical shift timescale, $k_{ex} > \Delta\omega$. A single peak with a population weighted resonance frequency, $p_A\omega_A + p_B\omega_B$, is obtained.

2.1.4 Exchange induced shifts in peak positions

Chemical exchange does not only contribute to the decay of the NMR signal, but also affects the position of the peaks as noted above. The positions of the peaks is shifted compared to the case where there is no exchange between the two states (Fig. 2.1). The exchange induced change in peak position δ_{ex} can be calculated from eqn. 2.4 (Skrynnikov et al., 2002).

$$\delta_{ex} = k_f \frac{\xi}{(1 + \frac{\Delta R}{k_b})^2 + \xi^2} \quad (2.4)$$

In eqn. 2.4 $\xi = \frac{\Delta\omega}{k_b}$. The rates k_f , k_b are the forward and reverse rate constants for the exchange process, $\Delta\omega$ and ΔR is the resonance frequency difference and the difference in the relaxation rates for the two states. One can see from eqn. 2.4 that the peak position is sensitive to the resonance frequency difference $\Delta\omega$ between the two exchanging states. Analysis of relaxation dispersion experiments gives the absolute resonance frequency differences. Exchange induced shifts in peak positions can be used to recover the missing sign information, which greatly facilitates the interpretation of chemical shift differences from relaxation dispersion experiments.

Based on 2.5 it becomes clear that missing shift information can be recover using a pair of HSQC and HMQC experiments (Skrynnikov et al., 2002). The first term within the curled bracket describes the exchange induced shift in an HSQC experiment. The second term denotes the exchange induced shift for multi-quantum coherence in an HMQC experiment.

$$\Omega_N = k_f \left\{ \frac{\xi_N}{(1 + \rho)^2 + \xi_N^2} - \frac{1}{2} \left[\frac{\xi_N + \xi_H}{(1 + \rho)^2 + \xi_N + \xi_H} + \frac{\xi_N - \xi_H}{(1 + \rho)^2 + \xi_N - \xi_H} \right] \right\} \quad (2.5)$$

In eqn. 2.5 $\rho = \frac{\Delta R}{k_b}$ and $\xi = \frac{\Delta\omega}{k_b}$. The rates k_f and k_b are the forward and the reverse rates for the exchange process, $\Delta\omega$ and ΔR is the resonance frequency difference and the difference in the relaxation rates for the two states.

Microscopic basis of chemical exchange

It is interesting to briefly consider the microscopic basis for the signal decay described by eqn. 2.2. The following discussion is based on Monte-Carlo calculations presented by Lippens et al., 2004. Spins exchange stochastically between different structural states due to exchange processes such as the binding of one protein to another. The states will differ in their electronic environments and in their Larmor frequencies. Spins will acquire a phase difference that depends on the time that they have spent in the different states. As the spins jump stochastically between the states, the spins will become progressively out of phase. Each spin will nonetheless contribute to the observed spectrum. Yet the different phase the spins have acquired means that their signals in the spectrum will all be slightly offset. Instead of a narrow line one will observe a broadened peak, such as the broad peaks shown in Fig. 2.1.

2.2 Molecular dynamics simulations

So far I have considered some important aspects of the NMR experiments that were recorded and analysed for this work. In the following I will give a brief overview of the theoretical and methodological basis of some of the computer simulation techniques I have used. In molecular dynamics (MD) simulations, the time-evolution of a system is followed using classical mechanics. Empirical force fields such as the CHARMM and GROMOS force fields used in this work (MacKerell et al., 1998; van Gunsteren et al., 1996), enable the calculation of potential energies for molecules and thus the calculation of forces. By calculating the forces acting on the particles in the simulation system, their dynamics can be followed using Newton's equations of motion.

2.2.1 The force field

The interactions of the atoms or particles in a MD calculation are described by potential energy functions that are also referred to as force fields. Force fields describe both the interactions due to the stereochemistry of covalent bonds in a molecule, the bonded interactions, and non-bonded interactions, through space, between atoms that are not directly linked by chemical bonds (Dill and Bromberg, 2011; Levitt, 2001). In force fields for biochemical applications bonded terms for bond stretching (U_{bond}), angle bending (U_{angle}) and dihedral angle fluctuations ($U_{dihedral}$) are included. Non-bonded interactions terms for electrostatic ($U_{Coulomb}$) and van der Waals (U_{vdW}) interactions are generally used. This is illustrated in

eqn. 2.6

$$U = U_{bond} + U_{angle} + U_{dihedral} + U_{Coulomb} + U_{vdW} \quad (2.6)$$

Before discussing the functional forms of these terms it is important to point out that their calculation critically depends on many parameters that have to be carefully chosen and optimised. Force field parameters may be derived from spectroscopic experiments, quantum mechanical calculations or parameters may be chosen to reproduce thermodynamical data.

Bond stretching

Bond stretching is described using a simple harmonic potential (eqn. 2.7). The energy depends on the deviation of a covalent bond from its reference length b_0 and the force constant k_{bond} .

$$U_{bond} = \frac{k_{bond}}{2} (b - b_0)^2 \quad (2.7)$$

Angle bending

Angle bending is also described by a harmonic potential, with a force constant k_{angle} and reference angle θ_0 .

$$U_{angle} = \frac{k_{angle}}{2} (\theta - \theta_0)^2 \quad (2.8)$$

In general, the force constants for angle bending are somewhat lower than for bond stretching (Leach, 2001). Bond angles can be deformed more easily than bond lengths. Both bond stretching and angle bending are often referred to as "hard degrees of freedom". Torsion angles can vary more easily. Torsion angles and the non-bonded interactions, which are also more malleable, are therefore called "soft degrees of freedom".

Dihedrals

The dependence of the potential energy on changes in torsion angles are captured by eqn. 2.9 (Leach, 2001; van der Spoel et al., 2005),

$$U_{dihedral}(\phi_{ijkl}) = k_\phi (1 + \cos(n\phi - \phi_s)), \quad (2.9)$$

where k_ϕ is a force constant, n is the multiplicity and ϕ_s is the phase factor. The multiplicity determines the number of minima within a 360° rotation of the dihedral angle. The phase shift locates the position of the minimum in $V_d(\phi_{ijkl})$. In addition, improper dihedral terms are used in force fields to lock some groups that are known to be planar, *i.e.* aromatic rings, in

such configurations. In the CHARMM force field an additional dihedral term, the CMAP correction (Mackerell et al., 2004), improves the representation of the protein backbone (Buck et al., 2006).

Electrostatic interactions

Atoms that are not linked by covalent bonds interact *via* electrostatics and van der Waals interactions (Leach, 2001). Both interactions are evaluated for pairs of atoms. The electrostatic interaction between two charged moieties is described by Coulomb's law:

$$U_{Coulomb} = \frac{q_1 q_2}{4\pi\epsilon_0 r_{ij}}, \quad (2.10)$$

In eqn. 2.10 q_1 and q_2 are the partial charges of the interacting atoms, ϵ_0 is the permittivity of the vacuum, while r_{ij} is the distance between the interacting groups. The assignment of partial charges to atoms is an important part of the parametrisation of force fields. From eqn. 2.10 it can be seen that $U_{Coulomb}$ depends on $\frac{1}{r_{ij}}$. The strength of electrostatic interactions decays only slowly as the separation between two particles i and j increases. Ignoring the long-range nature of electrostatics, can introduce errors into MD calculations. For example, electrostatic interactions beyond a certain cut-off distance are sometimes ignored. As a consequence a charged atom might not "see" another charged moiety in one time step. In the next time-step the distance might be below the cut-off and the two atoms might interact strongly. Such cut-off artifacts can also make the calculation unstable.

The most accurate method for describing long-range electrostatics is based on Ewald summation. Electrostatic interactions are not limited to the simulation cell. The charged groups in the central box also interact with an infinite array of periodic boxes. The summation of the interactions with this array of periodic boxes converges very poorly and is computationally very expensive. The calculation can be split into a summation in real-space and in reciprocal space, which greatly increases its computational efficiency, as the two terms converge more quickly than the direct summation of the contributions from the periodic boxes. The particle mesh Ewald approach was used in the MD calculations presented in this thesis (Darden et al., 1993; Essmann et al., 1995).

Van der Waals interactions

In addition to electrostatic interactions atoms also experience van der Waals interactions (Dill and Bromberg, 2011; Leach, 2001). Even uncharged atoms are subject to attractive van

der Waals interactions. Fluctuations in the electron clouds of atoms generate instantaneous dipoles. The dipole on one atom can induce a dipole on a second atom. The dipoles will interact and cause a weak, relatively short-range attraction. These interactions are also referred to as London dispersion forces. At very close range the van der Waals interaction is repulsive. This reflects that atoms repel each other at close separation due to the Pauli exclusion principle. The van der Waals interactions are typically described using a Lennard-Jones potential,

$$U_{vdW}(r) = 4\varepsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 \right] \quad (2.11)$$

where r is the separation between two atoms, ε is the well depth for the interaction potential and σ is the collision diameter, which defines the separation at which the interaction energy is zero.

2.2.2 Integration of the equation of motions

Unlike for simple systems, Newton's equations of motion cannot be solved analytically for molecular systems with tens of thousands of particles. The equations of motion have to be integrated numerically (Leach, 2001). The calculation is broken down in many steps with a small constant time step Δt . In each step the forces acting on the particles have to be calculated. The forces, the current positions of the particles and their velocities can then be used to calculate update their positions. Repeating this calculation many times then yields a MD trajectory. The Verlet algorithm is a simple and robust algorithm for generating a MD trajectory and illustrates how the time evolution of particles is followed in MD calculations. Small changes in *e.g.* the position of a particle can be expressed as a Taylor series expansion as shown in eqn. 2.12. The first and second derivatives of r with respect to the time are the velocity, $v(t)$, and the acceleration, $a(t)$, of the particle.

$$r(t + \Delta t) = r(t) + \Delta t v(t) + \frac{1}{2} \Delta t^2 a(t) + \dots \quad (2.12)$$

The particle position can not only be expanded for $r(t + \Delta t)$ but also for $r(t - \Delta t)$:

$$r(t - \Delta t) = r(t) - \Delta t v(t) + \frac{1}{2} \Delta t^2 a(t) - \dots \quad (2.13)$$

The two equations can be combined by addition:

$$r(t + \Delta t) = 2r(t) - r(t - \Delta t) + \Delta t^2 a(t) \quad (2.14)$$

Using eqn. 2.14 the position at time $t + \Delta t$ can be calculated based on the current position $r(t)$ and the position in the previous time step $t - \Delta t$ as well as the acceleration $a(t)$. The acceleration is of course obtained from calculating the forces acting on the particles. It is important to note that multiple forces can act on a single particle at the same time. The forces are then combined as a vector sum which is then used in the MD algorithm. In the GROMACS simulations the related leap-frog algorithm was used.

2.2.3 Temperature and pressure coupling

In MD calculations one samples from statistical-mechanical ensembles such as the micro-canonical ensemble (NVE), the canonical ensemble (NVT) or the isothermal-isobaric ensemble (NPT) (Frenkel and Smit, 2002). In the canonical ensemble a system is specified by its fixed number number of particles N , its volume V and the temperature T . The isothermal-isobaric ensemble (NPT) is described by the fixed number of particles N , the pressure P and the temperature T . These two ensembles describe the conditions under which most biochemical studies are conducted and therefore are of special interest when studying biological macromolecules.

To study a system at a given temperature in the NVT or NPT ensembles temperature coupling algorithms are employed. The temperature is related to the time average of the kinetic energy of the system $\langle K \rangle$ (Leach, 2001):

$$\langle K \rangle = \frac{N_f k_B T}{2}, \quad (2.15)$$

where N_f is the number of degrees of freedom and k_B is Boltzmann's constant. The kinetic energy is also defined *via* the velocities of the particles v_i ,

$$K = \frac{1}{2} \sum_{i=1}^N m_i v_i^2, \quad (2.16)$$

where m_i is the mass of particle i . By rescaling the particle velocities, the desired average system temperature can be achieved. In the velocity-rescaling thermostat by Bussi et al., 2007 that is used in this work, the velocities are rescaled by α which is defined as:

$$\alpha = \sqrt{\frac{K_t}{K}} \quad (2.17)$$

The value of K_t in eqn. 2.17 is chosen from a distribution of kinetic energies. By drawing

K_i from a distribution of kinetic energies canonical sampling is achieved.

To study a system in the isothermal-isobaric ensemble (NPT), the pressure of the system has to be controlled. The box volume and dimensions are scaled by pressure coupling algorithms to fix the average pressure in the simulation systems. The approach by Parrinello and Rahman, 1981 was used for pressure coupling.

2.3 Enhanced sampling methods

Macromolecular transitions are often rare events and may not be observed in the standard equilibrium MD simulations. Typically a system can be simulated for a couple of μs . I have employed two methods that enhance sampling by introducing a bias: (1) metadynamics free-energy calculations and (2) dynamics importance sampling MD.

2.3.1 Metadynamics free-energy calculations

In metadynamics simulations, a history-dependent bias potential is used to enhance sampling along one of more collective variables (Barducci et al., 2011, 2008). A collective variable is a metric that describes the process of interest and can be loosely thought of as a reaction coordinate. In metadynamics Gaussian bias potentials are added at regular intervals in the space that is defined by the collective variables. More bias potentials are added in the favourable regions of the collective variable space that are frequently visited by the simulation. A large number of bias potentials will be added in such regions. Conversely, only few bias potentials will be added to regions that are poorly sampled. The accumulation of the bias potentials has two consequences. First, the atoms that define the collective variables will experience a repulsive force in the well-sampled low free-energy regions of the landscape. Metadynamics simulations can thus escape from favourable "valleys" on the free-energy landscapes and can cross free-energy barriers, which enhances the exploration of the conformational landscape. New, energetically favourable conformational states can be discovered in the simulations. Secondly, over time the bias potentials will start to correspond to the inverse of the free-energy landscape for the collective variables.

Here well-tempered metadynamics was used. In the original implementations of metadynamics (Laio and Parrinello, 2002), bias potentials of uniform height were added to the simulations. As bias potentials are constantly accumulated the simulation system can start to explore unphysical high free-energy regions. In well-tempered metadynamics the height

of the Gaussian bias-potentials added w is scaled as shown in eqn. 2.18.

$$w = \omega e^{-V(s,t)/\Delta T} \tau_G \quad (2.18)$$

Here ω is the initial bias deposition rate (Barducci et al., 2008) and τ_G is the time interval at which bias potentials are added. ΔT is a temperature. Increased values of ΔT enables the crossing of larger free-energy barriers. It is important to point out the temperature of the remainder of the simulation system is not affected by this. $V(s,t)$ represents the sum of the bias potentials that the simulation experiences. Its value depends on s , the collective variable, and changes with time. The eqn. 2.18 suggests that as the simulations visits already explored regions in the collective variable space the height of the hills should decrease. The bias is defined by eqn. 2.19.

$$V(s,t) = \Delta T \ln(1 + \omega \frac{N(s,t)}{\Delta T}) \quad (2.19)$$

$N(s,t)$ is the histogram for the collective variable s in the metadynamics simulation. From $V(s,t)$ the free-energy surface for s can be estimated by,

$$FES(s) = -\frac{T + \Delta T}{\Delta T} V(s,t) \quad (2.20)$$

2.3.2 Dynamics importance sampling calculations

In dynamics importance sampling (DIMS) MD transitions between conformational states are generated (Perilla et al., 2011; Woolf, 1998). To enhance the sampling, a bias is introduced. Consider the average $\langle g \rangle$ that depends on the coordinate x , that is distributed according to Q . The coordinate x may report on the transition between two conformational states. Q is a Boltzmann factor that describes the canonical distribution of x . In the case of conformational changes, the average would report on many transition events.

$$\langle g \rangle = \int dx(x) Q(x) \quad (2.21)$$

A bias D is then introduced to speed up the calculation of $\langle g \rangle$. A judicious choice of the bias can drastically speed up the calculation of transition events.

$$\langle g \rangle = \int [D(x) dx] \frac{g(x) Q(x)}{D(x)} \quad (2.22)$$

Here the transitions were biased using a soft-ratcheting algorithm that resembles Maxwell's demon (Beckstein et al., 2009). Trial steps are generated by molecular dynamics. Progress in the trial steps towards the target state is assessed using a progress variable, ϕ . Steps that move the system closer to the target will be accepted. Backwards steps will only be accepted with a certain probability. The probability for accepting backwards steps depends on ϕ_0 which determines the softness of the ratchet. A stiffer ratchet $\phi_0 < 10^{-6}$ typically results in more successful transitions for a given amount of compute time. Allowing more backwards steps by setting $\phi_0 > 10^{-6}$ can lead to a more diverse ensemble of transition pathways which may give a better representation for the process.

$$p_{acc}(\Delta\phi) = \begin{cases} 1 & \text{if } \Delta\phi \leq 0 \\ \exp[-\Delta|\frac{\Delta\phi}{\phi_0}|^2] & \text{if } \Delta\phi > 0 \end{cases} \quad (2.23)$$

The backbone root mean square deviation to the target state is a robust choice of progress variable and was employed in this work. The bias introduced in the generation of a trajectory is tracked by the DIMS score. The score for given step i for which N MD steps were generated by the soft-ratcheting algorithm is determined by eqn. 2.24.

$$R_i = \frac{N}{p_{acc}(\Delta\phi)} \quad (2.24)$$

To illustrate eqn. 2.24, a score of 1 would be obtained if the first trail step moves the simulation towards its target. The total DIMS score is then defined as the product of the scores for the individual steps. The bias is known, so in principle a quantitative description of the transitions in terms of reaction rates is possible. In practice, many transitions, with a diversity of starting states would be required (Perilla et al., 2011). Here the DIMS was used to sample reasonable transition paths that can give insight into transition mechanisms. Previous work has established that DIMS transition paths follow the contours of free-energy landscapes (Beckstein et al., 2009).

DIMS simulations were conducted with an implicit representation of the membrane system, following the work of Denning and Woolf, 2010. Langevin dynamics (eqn. 2.25) was used to follow the time evolution of the simulation system (Leach, 2001). In Langevin dynamics the influence of the solvent on the solute is represented by a friction term and a random force. The motions of the solute particles are slowed down by solvent friction that depends on the frictional coefficient γ and the velocity of the particle $\frac{dx_i(t)}{dt}$. The random force $R_i(t)$ models the influence of collisions with solvent particles on the dynamics of the

solute.

$$m_i \frac{d^2 x_i(t)}{dt^2} = F_i(x_i(t)) - \gamma \frac{dx_i(t)}{dt} + R_i(t) \quad (2.25)$$

The force $F_i(x_i(t))$ originates from the non-bonded and bonded interactions of the solute particle.

2.3.3 Comparing molecular dynamics simulations to experiments

Experiments typically probe large ensembles of molecules. Hence experimental investigations of molecular structure and dynamics yield ensemble averaged measurements (Frenkel and Smit, 2002). Computer simulations on the other hand produce trajectories that describe the evolution of single molecules. A long enough trajectory would visit all states that constitute the ensemble that was probed in experiment. The time and ensemble averages of a particular property of the molecule would be equivalent and system would be said to be ergodic. While the simulations conducted in this thesis are certainly not long enough, time averages from MD simulations were nonetheless used as estimates for ensemble averages and compared with experiments.

Chapter 3

Molecular dynamics simulations of nDsbD and comparison to NMR data

3.1 Introduction

The conformational equilibrium of nDsbD was studied by MD simulations. MD simulations may reveal fluctuations in the cap loop that covers the active site. This loop needs to open if nDsbD is to bind its cognate interaction partners. For this the sub-millisecond dynamics of nDsbD were investigated by molecular dynamics (MD) simulations and by comparing the simulations to NMR data. The simulations may reveal how nDsbD functions as a "redox hub" in the oxidising periplasm of Gram negative bacteria (Stirnemann et al., 2006).

3.1.1 nDsbD is a "redox hub" in the periplasm of Gram-negative bacteria

The periplasm of Gram negative bacteria is very oxidising. Yet reductant is still required in this compartment for vital biosynthetic tasks (Ito and Inaba, 2008). There are three main groups of periplasmic processes that require the reductant supplied by DsbD as discussed in the Chapter 1.5 and shown in Fig. 3.1:

1. Reductant is required to isomerise incorrectly formed disulphide bonds. Disulphide bonds are non-specifically introduced by DsbA as shown in Fig. 1.1. Non-native disulphide bonds are then resected by disulphide bond such as DsbC which enables proteins to access their cognate fold. Among the substrates of DsbC (Cho and Collet, 2012) is RcsF and possibly LptD which are critical for the integrity of the cell enve-

lopes of Gram-negative bacteria. Disulphide bond isomerisation is also critical in the assembly of the bacterial conjugation apparatus (Hemmis et al., 2011).

2. Reductant is required to maintain the redox-state of some proteins. DsbG has been shown to maintain the catalytic cysteine residues in periplasmic enzymes (Depuydt et al., 2009). DsbD and related proteins may (Cho and Collet, 2012) supply further pathways that protect against non-specific oxidation by radical oxygen species.
3. CcmG in the cytochrome c maturation pathway also requires reductant.

The trans-membrane oxidoreductase DsbD supply reductant for such processes. The trans-membrane protein DsbD channels reductant from the cytoplasm in to the periplasm. First reductant from thioredoxin is transferred via the trans-membrane domain of DsbD to the C-terminal domain of DsbD (cDsbD). cDsbD then passes on the reductant to nDsbD, which acts as a "redox-hub" (Stirnemann et al., 2006). nDsbD then distributes reductant to three sets of partners. (1) Disulphide isomerases that resect incorrectly formed disulphide bonds, such as DsbC and TrbB, which is important for the assembly of the bacterial conjugation apparatus. (2) DsbG, which maintains the redox state of sulphur moieties in the periplasm (Depuydt et al., 2009). (3) CcmG, part of the cytochrome c maturation pathway, crucial for bacterial respiration.

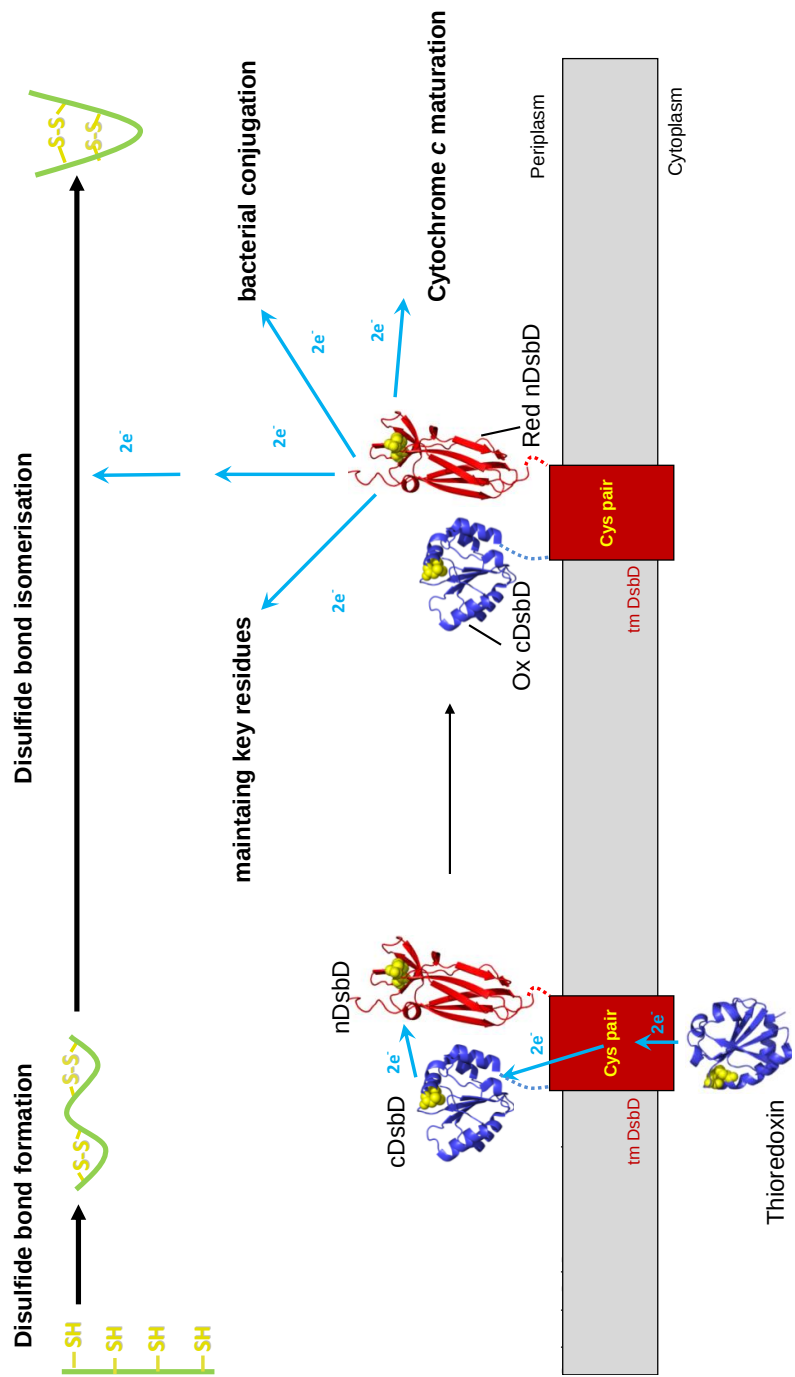


Fig. 3.1 The DsbD pathway provides reductant in the oxidising periplasm of Gram-negative bacteria.

3.1.2 Reductant transfer happens in a series of thiol-disulphide exchange reactions

Reductant transfer in the *Dsb* pathway is accomplished in a series of thiol-disulphide exchange reactions. Unlike in other redox transferring pathways, no metal centres are involved. The reaction partners associate, an intermolecular disulphide bond is formed and subsequently broken, before the complex falls apart after transferring the reductant. Given the oxidising nature of the periplasm, the reactivity of the thiols in the DsbD has to be regulated. Thiols should only be oxidated when transferring reductant to the cognate interaction partners. The cap loop is thought to protect the thiols in the active site of nDsbD (Goulding et al., 2002; Mavridou et al., 2009). The phenyl group of F70 shields the sulphur atom of C109, the more reactive of the two cysteines, from the periplasm (Fig. 3.2). C103 is even more buried.

3.1.3 The structure and dynamics of nDsbD

The conformational equilibrium of the cap loop

Previously, I have used to $^1\text{H} - ^{15}\text{N}$ residual dipolar coupling experiments (Stelzl, 2010) to test whether the cap loop would really cover the active site of nDsbD. Overall, the X-ray structures for reduced and oxidised nDsbD were consistent with NMR residual dipolar coupling (RDC) measurements. The RDCs were also consistent with the cap loop being closed on average in both oxidised and reduced nDsbD, but this was not definitely proven. The cap loop may be able to shield the active thiols of reduced nDsbD from non-cognate side reactions. However, to draw firm conclusions on the cap loop structure from the RDC data, additional measurements would be required.

However, the loop has to open of nDsbD to form complexes with its partners. The cap loop is in an open conformation in the crystal structures of the covalent complexes between nDsbD with CcmG (Edeling et al., 2002), cDsbD (Rozhkova et al., 2004), and DsbC (Haebel et al., 2002). The opening mechanism is not understood. Firstly, it has been suggested that the loop is very flexible and would thus readily open when nDsbD associates with its partners. Analysis of NMR spin relaxation experiments by the "model-free" approach (Lipari and Szabo, 1982a,b) suggested that the cap loop is rigid in both oxidised and reduced nDsbD (Stelzl, 2010). The need to protect the active of nDsbD also argues against the loop being very flexible. Secondly, the loop might occasionally open spontaneously. nDsbD would then bind its interaction partners in a conformational selection

mechanism. Thirdly, the cap loop might only ever open on encountering its cognate binding partners such as cDsbD. Here the flexibility of the loop and the conformations it can access were investigated using MD simulations.

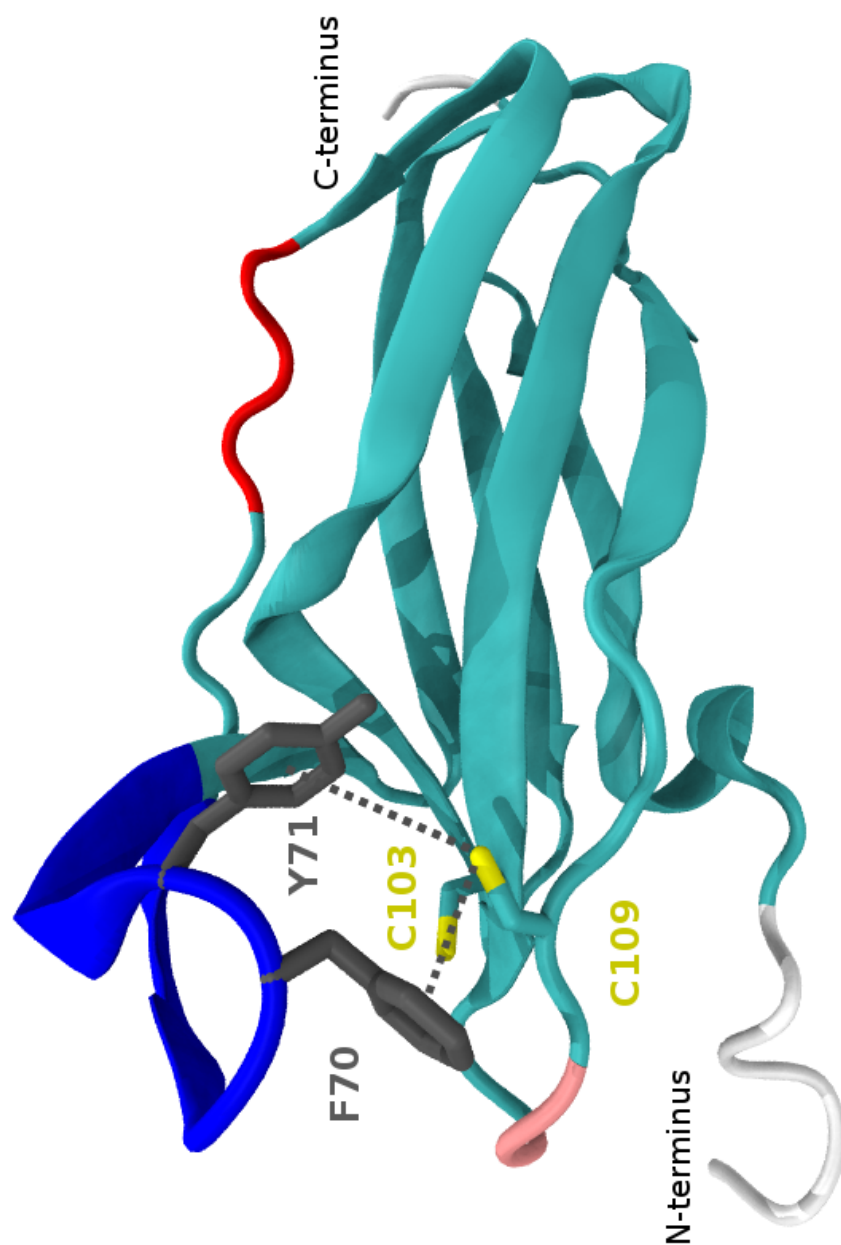


Fig. 3.2 The structure of nDsbD is illustrated based on the crystal structure of reduced nDsbD (PDB: 3PFU). The cap loop covers the active site. The cap loop region (residues 65-75) is highlighted in blue. The side chains of F70 and Y71 (shown in grey) may shield the thiols of the active site cysteines C103 and C109 from non-cognate side-reactions. The conformational state of the cap loop was followed in MD simulations by tracking the distances between the centres of the aromatic rings of F70 and Y71 and the sulphur atom of C109 (as indicated by the grey-dashed lines). In addition to structural data, previous NMR spin-relaxation experiments that reported on the ps-ns dynamics of nDsbD (Stelzl, 2010) were the starting points for this study. Most residues of nDsbD do not show extensive fluctuations, with the exception of the disordered N- and C-terminal regions (residues 1-12 and 124-137 respectively), and have $S^2 > 0.8$. Residues 57 to 60 which lack regular secondary structure and were found to be mobile in NMR spin-relaxation experiments (S^2 all < 0.80) are shown in red. Only a few residues, such as G107 and F108 in the active site, showed redox-state dependent differences in their flexibility. Residues G107 and F108 are more flexible in oxidised nDsbD and are highlighted in pink.

The association reaction between cDsbD and nDsbD is redox-state dependent

The reactivity of nDsbD is not only controlled by the cap loop but also modulated by its redox state-dependent association with cDsbD (Mavridou et al., 2011). Reduced cDsbD associates with oxidised nDsbD. Reductant is subsequently transferred from reduced cDsbD to oxidised nDsbD. Upon reductant transfer, the affinity between the two domains drops by 30-40 fold and the two domains will dissociate. The thiol-disulphide exchange reaction is only just favourable ($E^0 = 3 \text{ mV}$ see Table. 1.1 and Rozhkova and Glockshuber, 2008). The redox-state dependence of the interaction affinity between cDsbD and nDsbD will limit the extent to which the thermodynamically feasible reverse reaction, reductant transfer from reduced nDsbD to oxidised cDsbD, can happen. Moreover, the drop in the interaction between cDsbD and nDsbD frees up the two domains for subsequent reactions. Oxidised cDsbD will interact with trans-membrane domain of DsbD and will cycle back to its reduced isoform. Reduced nDsbD will supply reductant to its partners in the periplasm. The redox-state dependent interaction between cDsbD and nDsbD may therefore enhance turnover in the pathway.

Neither the crystal structures nor the NMR data have provided a clear structural explanation. The crystal structures of oxidised and reduced nDsbD are remarkably similar. There are subtle differences in active site and the cap loop region between the two redox isoforms (Mavridou et al., 2011). Perhaps the most notable differences are small changes in the orientations of the aromatic rings in the cap loop and that the backbone of the loop is moved towards the active site cysteines in reduced nDsbD. The functional relevance of the differences is not obvious. RDC data and the "model-free" analysis of the ps-ns dynamics of nDsbD have not shown clear differences between oxidised and reduced nDsbD that would explain the difference in the affinity for cDsbD. Previously determined chemical shifts however suggest that the solution ensembles of oxidised and reduced nDsbD differ, not just in the active site (Mavridou et al., 2012a, 2011) where the switch between two cysteine thiol to a disulphide bond is expected to affect at least the local chemical shifts, but also in the cap loop region. Furthermore exchanging a disulphide bond with a two thiols could change the packaging constraints in the active site which may contribute to the redox-state dependent interaction between cDsbD and nDsbD (Mavridou et al., 2011). Such packaging effects will be explored in the simulations presented here.

3.1.4 Studying nDsbD by molecular simulation and their comparison to NMR data

The simulations were compared to experiments not only for validation, but also to refine the interpretation of crystal structures and the NMR measurements. Computer simulations could show whether the subtle differences in the crystal structures persist over time and consequently might be functionally relevant. The structures accessed in the simulations will be assessed by comparison to the RDC measurements which report on time averaged structure. The simulations may explain some of the differences between the RDC measurements for oxidised and reduced nDsbD which are hard to understand from the crystal structures alone. Comparing the simulations to the NMR spin-relaxation experiments will indicate whether the fluctuations are represented well in the molecular dynamics calculations. Chemical shifts are also sensitive to the time-averaged structure of proteins. Directly comparing chemical shifts to molecular dynamics simulations is however still challenging and therefore a different approach was used. The local geometry expected on the basis of chemical shifts was compared with the average structures from the simulations (Shen and Bax, 2013).

Aims

By running extensive equilibrium MD simulations I investigated whether the cap loop is flexible in oxidised and reduced nDsbD. In the simulations I tested whether the active site cysteines of reduced nDsbD would be protected from non-cognate oxidation. Simulations were started from both open and closed conformations which allowed us to test whether structures with the cap loop open are ever accessed for significant periods of time. Hence it might be possible to discern if nDsbD binds its partners via induced fit or conformational selection mechanisms. Differences in the dynamics of oxidised and reduced nDsbD could point towards explanations for the redox-state dependence of the interaction affinities of nDsbD with cDsbD. The simulations were compared to NMR experiments to assess the quality of the simulation ensembles and to refine the interpretation of the experimental data.

Table 3.1 pK_a s for titratable side-chains of oxidised and reduced nDsbD

Side-chain	NMR reduced nDsbD	NMR oxidised nDsbD
Glu 67	not fully analysed	not fully analysed
Glu 113	5.3 and 3, peak shifts with two pK_a s	5.3 and 3, peak shifts with two pK_a s
His 27	~6.9 (unreliable)	~6.9 (unreliable)
His 52	7	not fully analysed
His 66	6.6	6.7
Cys 103	$pK_a \gg 7$	-
Cys 109	$pK_a \gg 7$	-

3.2 Methods

3.2.1 Atomistic molecular dynamics simulations

Setup of the simulations

Simulations were prepared and run with GROMACS 4.5 (www.gromacs.org, Hess et al., 2008). Simulation systems were setup for crystal structures of reduced (PDB 3PFU, Table 3.2) and oxidised nDsbD (PDB 1L6P, Table 3.4). The cap loop is closed in these structures. In addition simulations were also started with the cap loop open, for both reduced (Table 3.3) and oxidised nDsbD (Table 3.5). Such simulations were based on chain B in the 1VRS crystal structure of the covalent intermediate of the cDsbD and nDsbD redox transfer reaction or on open structures accessed in other simulations (Rozhkova et al., 2004). Missing atoms in the experimental structures were added using the WHAT-IF web-server (Hekkelman et al., 2010; Vriend, 1990).

The protonation state of titratable side chains was chosen with care to model nDsbD at pH 6.5, matching the conditions used in the NMR experiments. For some side chains the pK_a was determined in NMR titration experiments by Prof Christina Redfield (Table 3.1). In all simulations, the glutamate and aspartate side chains were de-protonated, as expected. For H27 and H66 in both redox isoforms the $pK_a \geq \text{pH}$ and the side-chains were represented as positively charged. For H52 in reduced nDsbD, it was also found that the side-chain should positively charged. For oxidised nDsbD, the pK_a was not available and the side chain was also assumed to be positively charged. The NMR titration experiments also showed that the side chains of C103 and C109 have $pK_a \gg 7$ and should be represented as thiols in the simulations.

The proteins were simulated using periodic boundary conditions (PBC) in rhombic dodecahedral boxes that are filled with water and ions. To avoid periodicity artifacts, the min-

imum distance from the protein to the box edges was either 12Å or 15Å, resulting in simulations systems containing around 40,000 to 50,000 atoms respectively. Ions were added to ensure that the simulation system did not have an overall charge. In addition, Na⁺ and Cl⁻ were added to system, up to an ion concentration of 0.1 M to prevent electrostatic PBC artifacts. Given the long-range nature of electrostatics, residues on one edge of the protein could interact with residues on the opposite side of the biomolecule across the periodic boundary, possibly seriously distorting the sampling of structures in the simulations.

Simulation parameters

The CHARMM 22 energy function, i.e. the force field, was used to calculate potential energies in molecular dynamics and energy minimisation calculations (MacKerell et al., 1998). The CMAP correction (Mackerell et al., 2004) was used for description of the backbone dihedral angles by CHARMM. The correction has been shown to more accurately represent the conformational equilibria of biomolecules (Buck et al., 2006). The solvent was described with the CHARMM TIP3P water model. While computationally more demanding than the standard SPC or TIP3P models, the CHARMM TIP3P water model should be used when using a CHARMM force field, ensuring consistency in the simulation parameters.

Electrostatic interactions were calculated with the Particle Mesh Ewald Method (PME) (Darden et al., 1993; Essmann et al., 1995). The cut-off for treating electrostatic interactions as short range was 12Å. It is important to point out that this cut-off does not imply that long-range electrostatic interactions are neglected. On the contrary PME enables a relatively accurate treatment of long-range electrostatics. The Lennard-Jones potential was switched to zero, between 10Å and 12Å, following the setup introduced when the CHARMM force field was added to GROMACS (Bjellmar et al., 2010). The neighbour list was updated every 5 integration steps with cut-off for the neighbour list update of 12Å.

The length of chemical bonds involving hydrogen atoms was constrained, as it is typically done in simulations with CHARMM potential energy functions, using the PLINCS algorithm (Hess, 2008). The equations of motions were integrated with a 2 fs time step.

Equilibration and production simulations

To remove any close contacts between atoms the simulation systems containing water, protein and ions were energy minimised. Production runs were preceded by 4 ns of equilibration in the NVT ensemble using position restraints. Equilibrium simulations were run in the NPT ensemble for 10 ns, 20 ns or 200 ns. Separate temperature groups were used for

protein and non-protein atoms. Simulation temperature was kept at 298.15 K by rescaling the velocities, with a τ_T of 0.1 ps (Bussi et al., 2007). The pressure in the simulation box was maintained at 1 bar by the method proposed by Parrinello and Rahman with a τ_P of 0.5 ps and a compressibility of $4.5 \times 10^{-5} \text{bar}^{-1}$ (Parrinello and Rahman, 1981).

3.2.2 Comparing NMR spin-relaxation experiments with simulations

Fast time-scale dynamics measured by NMR spin relaxation experiments can be readily compared to computer simulations. Typically the spin relaxation of nuclei in the protein backbone, in particular the ^1H and ^{15}N nuclei, are characterised in such NMR experiments. NMR spin-relaxation experiments are usually interpreted by the "model-free" formalism that allows to deconvolve contributions to spin relaxation from the overall diffusion of the macromolecule and internal motions of individual bond vectors (Jarymowycz and Stone, 2006). Both the rotational diffusion of the macromolecule and the internal dynamics can be calculated from molecular dynamics trajectories. Here the focus is on the comparison of internal dynamics in experiment and the simulations. The overall tumbling is removed from the simulations by aligning the simulation trajectory to a reference structure.

The S^2 order parameter that describes the amplitude of bond vector motions can be calculated from molecular dynamics simulations. S^2 is defined as the long term limit of the rotational auto-correlation function for a bond vector, $C_{\hat{\mu}}$:

$$C_{\hat{\mu}}(t) = S^2 + (1 - S^2)e^{-t/\tau_e}. \quad (2.1 \text{ revisited})$$

Then as $\lim_{t \rightarrow \infty} e^{-t/\tau_e} = 0$:

$$S^2 = \lim_{t \rightarrow \infty} C_{\hat{\mu}}(t). \quad (3.1)$$

The angular rotational auto-correlation function for a bond vector μ compares orientation of the vector at a given time with its orientation at the time origin by taking the dot or scalar product of the two vectors:

$$C_{\hat{\mu}}(t) = P_2[\hat{\mu}(0) \hat{\mu}(t)], \quad (3.2)$$

where $P_2 = \frac{1}{2}(3x^2 - 1)$ is the second Legendre polynomial and $\hat{\mu}$ is the unit vector along the bond of interest. Since this function is highly noisy, averaging over time origins is performed as indicated by the angular brackets,

$$C_{\hat{\mu}}(t) = \langle P_2[\hat{\mu}(0) \hat{\mu}(t)] \rangle. \quad (3.3)$$

For a simulation trajectory with N frames that are spaced at regular time intervals Δt , the rotational auto-correlation function can then be calculated as

$$C_{\hat{\mu}}(j\Delta t) = \frac{1}{N-j} \sum_{i=0}^{N-1-j} P_2[\hat{\mu}(i\Delta t) \hat{\mu}((i+j)\Delta t)], \quad (3.4)$$

where j denotes the frame for which the function is evaluated and i specifies the time origins used for the time averaging. Using eqn. 3.4 the value of the rotational auto correlation cannot be calculated with the same accuracy for all frames in a trajectory. The first value for this function is calculated by averaging over $N-1$ frames, whereas for $j=N-1$ only one frame can be used as a time origin. The later time points of correlation functions calculated by eqn. 3.4 will be more severely affected by noise which may be problematic for calculating S^2 from a trajectory, especially given its definition as the long term limit of the correlation function. By restricting the calculation to $M \leq N/2$ frames each value for $C_{\hat{\mu}}(j\Delta t)$ is obtained with the same accuracy, overcoming this problem.

$$C_{\hat{\mu}}(j\Delta t) = \frac{1}{M} \sum_{i=0}^{N-1-M} P_2[\hat{\mu}(i\Delta t) \hat{\mu}((i+j)\Delta t)] \quad (3.5)$$

The long-term limit of this function can now be compared to experimental S^2 values. Usually the average value of $C_{\hat{\mu}}(j\Delta t)$ is compared to S^2 from experiment (Chandrasekhar et al., 1992):

$$S^2 = \frac{1}{M\Delta t} \sum_{i=0}^{N-1-M} C_{\hat{\mu}}(j\Delta t) \quad (3.6)$$

The average in eqn. 3.6 also includes the initial values of $C_{\hat{\mu}}(j\Delta t)$ which will be higher than the values the function plateaus at. However, the initial drop in $C_{\hat{\mu}}(j\Delta t)$ is usually so fast that including the whole auto correlation function in the averaging in eqn. 3.6 does not affect the result greatly. Using the long-time limit from eqn. 3.5 or the average defined by eqn. 3.6 will give virtually identical results for bond vectors that are subjected to fast motions only. The correlation function for these inter-atomic vectors will converge onto a single value. As pointed out in an early study (Chandrasekhar et al., 1992), the rotational correlation functions for bond vectors that are subjected to more complicated dynamics converge more poorly. In this case taking the average using eqn. 3.6 will give a better estimate of S^2 and hence this equation is generally employed.

Eqn. 3.6 is equivalent to the following equation often seen in the literature (Smith et al.,

1995; Zeiske et al., 2013)

$$S^2 = \frac{1}{2} [3 \sum_{\alpha=1}^3 \sum_{\beta=1}^3 \langle \hat{\mu}_\alpha \hat{\mu}_\beta \rangle], \quad (3.7)$$

where α and β denote the x,y,z components of the bond vector $\hat{\mu}$. In eqn. 3.7 the average dot product for bond vector $\hat{\mu}$ is calculated first before applying the second Legendre polynomial to get a value for S^2 for the simulation analysed, while in eqn. 3.6 (using eqn. 3.5) the rotational correlation function is calculated explicitly before obtaining S^2 .

Finally, the time scale of the motions that S^2 values from NMR spin relaxation experiments are sensitive to has to be taken into account when comparing experiment to MD simulations. In the "model-free" approach, it is assumed that the overall tumbling of the macromolecule and the local dynamics of the individual bond vectors are independent. Thus, S^2 order parameters from experiments report on the amplitude of motions that are faster than τ_M , the correlation time for the diffusion of the molecule. Hence, S^2 should be calculated for time windows (Zeiske et al., 2013) that are on the order of the rotational correlation time, τ_M , to compare like with like, fast time scale dynamics in simulations and NMR experiments.

3.2.3 Calculating residual dipolar couplings from molecular dynamics simulations

To calculate RDCs from the MD simulations, the structures in the simulations were aligned to a reference structure. The use of a single reference structure is justified given the stability of the overall fold of nDsbD detected in the simulations (*c.f.* the C_α RMSD calculations presented in Figure 3.3). In addition, the radius of gyration was calculated for representative simulations and was found to be stable (*not shown*).

The reference structure has its principle axes aligned with the alignment tensor. The residual dipolar coupling for a given residue in given structure is then calculated using

$$\overline{D} = \frac{\kappa}{R^3} (A_x \cos^2 \theta_1 + A_y \cos^2 \theta_2 + A_z \cos^2 \theta_3) \quad (3.8)$$

where κ depends on physical constants as defined earlier and R is the bond length and is uniformly set to 1.04 Å. The angles θ_1 , θ_2 and θ_3 describe the orientation of the individual bonds in respect to the three principal axes of the alignment tensor. A_x , A_y and A_z are the principal components of the alignment tensor. RDCs from all trajectories started with the x-ray structures of oxidised and reduced nDsbD respectively were averaged. The calculation

has been implemented in Python, based on the MDAnalysis library (Michaud-Agrawal et al., 2011). The RDC calculation will be made publicly available in a future release of the MDAnalysis library.

3.2.4 TALOS-N prediction of the side-chain dihedral angles

TALOS-N uses chemical shifts to find protein fragments that may resemble the local geometry of protein residues (Shen and Bax, 2013). The side-chain dihedral angles for reduced and oxidised nDsbD were predicted from their chemical shifts using TALOS-N. The 1L6P crystal structure of oxidised nDsbD was deleted manually from the database of the local installation of TALOS-N to ensure unbiased predictions for nDsbD.

3.3 Results

3.3.1 MD simulations of oxidised and reduced nDsbD are stable

Multiple simulations of oxidised and reduced nDsbD were started from closed as well as from open conformations. Crystal structures for the closed conformations of both redox states are available. There is no experimental structure of nDsbD in its apo-state with an open cap loop for either redox state. However, nDsbD is open in the (1VRS) X-ray crystal structure of the covalent intermediate of the reductant transfer between cDsbD and nDsbD. This structure, featuring either a disulphide bond or two thiols in the active site, was used to simulate the open conformations of oxidised and reduced nDsbD respectively. For reduced nDsbD, starting with the closed 3PFU crystal structure 14 simulations with a total simulation time of 1 μ s were run (Table 3.2). In addition, 18 short simulations, a total of 190 ns, were initiated with reduced nDsbD in an open conformation based on chain B in the 1VRS crystal structure (Table 3.3). For oxidised nDsbD, 14 simulations with the loop closed were started from the 1L6P crystal structure, yielding 1 μ s of simulation time (Table 3.4). Open conformations of the oxidised isoform were investigated in 13 simulations with a combined simulation time of 160 ns (Table 3.5). Three of these simulations were started from the final structure of *oxI* where the cap loop had opened spontaneously.

The structural stability of the proteins in the simulations was investigated for representative trajectories by calculating C_{α} root mean square deviations (RMSDs) from the starting structures. In simulations based on high-quality experimental protein structures only limited deviations (C_{α} RMSD 1.0 – 1.5 Å for the structured regions) from the starting structures are expected in MD trajectories on the order of tens to hundreds of ns. Larger deviations can

Table 3.2 MD simulations with the reduced, closed crystal structure of nDsbD

Name	Starting structure	length	Δ protein-box
red1	3PFU	200 ns	12 Å
red2	3PFU	200 ns	12 Å
red3	3PFU	200 ns	12 Å
red4	3PFU	200 ns	12 Å
red5	3PFU	20 ns	12 Å
red6	3PFU	20 ns	12 Å
red7	3PFU	20 ns	12 Å
red8	3PFU	20 ns	12 Å
red9	3PFU	20 ns	12 Å
red10	3PFU	20 ns	12 Å
red11	3PFU	20 ns	12 Å
red12	3PFU	20 ns	12 Å
red13	3PFU	20 ns	12 Å
red14	3PFU	20 ns	12 Å

Table 3.3 MD simulations with an open reduced structure of nDsbD. *redO1S1* is a continuation of *redO1*, no new velocities were generated. The other simulations spawned from the end structure of *redO1* were initiated with new starting velocities.

Name	Starting structure	length	Δ protein-box
redO1	red 1VRS B	10 ns	12 Å
redO2	red 1VRS B	10 ns	12 Å
redO3	red 1VRS B	10 ns	12 Å
redO4	red 1VRS B	10 ns	12 Å
redO5	red 1VRS B	10 ns	12 Å
redO6	red 1VRS B	10 ns	12 Å
redO7	red 1VRS B	10 ns	12 Å
redO8	red 1VRS B	10 ns	12 Å
redO1S1	final frame redO1	20 ns	12 Å
redO1S2	final frame redO1	10 ns	12 Å
redO1S3	final frame redO1	10 ns	12 Å
redO1S4	final frame redO1	10 ns	12 Å
redO1S5	final frame redO1	10 ns	12 Å
redO1S6	final frame redO1	10 ns	12 Å
redO1S7	final frame redO1	10 ns	12 Å
redO1S8	final frame redO1	10 ns	12 Å
redO1S9	final frame redO1	10 ns	12 Å
redO1S10	final frame redO1	10 ns	12 Å

Table 3.4 MD simulations with the oxidised, closed crystal structure of nDsbD

Name	Starting structure	length	Δ protein-box
ox1	1L6P	200 ns	12 Å
ox2	1L6P	200 ns	12 Å
ox3	1L6P	20 ns	12 Å
ox4	1L6P	20 ns	12 Å
ox5	1L6P	20 ns	12 Å
ox6	1L6P	20 ns	12 Å
ox7	1L6P	20 ns	12 Å
ox8	1L6P	200 ns	15 Å
ox9	1L6P	200 ns	15 Å
ox10	1L6P	20 ns	15 Å
ox11	1L6P	20 ns	15 Å
ox12	1L6P	20 ns	15 Å
ox13	1L6P	20 ns	15 Å
ox14	1L6P	20 ns	15 Å

Table 3.5 MD simulations with oxidised open structures of nDsbD

Name	Starting structure	length	Δ protein-box
oxO1	ox 1VRS B	10 ns	12 Å
oxO2	ox 1VRS B	10 ns	12 Å
oxO3	ox 1VRS B	10 ns	12 Å
oxO4	ox 1VRS B	10 ns	12 Å
oxO5	ox 1VRS B	10 ns	12 Å
oxO6	ox 1VRS B	10 ns	12 Å
oxO7	ox 1VRS B	10 ns	12 Å
oxO8	ox 1VRS B	10 ns	12 Å
oxO9	ox 1VRS B	10 ns	12 Å
oxO10	ox 1VRS B	10 ns	12 Å
ox1s1	final frame ox1	20 ns	12 Å
ox1s2	final frame ox1	20 ns	12 Å
ox1s3	final frame ox1	20 ns	12 Å

point to deficiencies in the simulation force field or in the starting structures. For trajectories started with reduced nDsbD only very little deviation from the crystal structure is detected in the structured segments of the protein (residues 14-121). Specifically, for trajectory *red1* the average C_{α} RMSD was 0.9 Å (Figure 3.3 D) and for simulation *red4* the average C_{α} RMSD was even less, 0.8 Å (Figure 3.3 E). Simulation *redO2* started with an open conformation of reduced nDsbD shows somewhat larger fluctuations in the backbone with an average C_{α} RMSD of 1.1 Å.

Simulations of oxidised nDsbD were stable as well judged by the C_{α} RMSD calculated for typical trajectories. For simulation *ox1* the C_{α} RMSD to the equilibrated oxidised closed crystal structure averaged 0.9 Å. Similarly the mean C_{α} RMSD for trajectory *ox9* was 0.9 Å. For simulation *oxO7* started from an oxidised nDsbD structure with an open cap loop, a slightly larger deviation from the starting structure was found, with average C_{α} RMSD of 1.2 Å. Given the overall stability of our simulations, one can proceed to analyse them in more detail.

RMSD analysis of the simulations

The RMSD calculations can give of an initial view of the structural preferences in the simulation ensemble. The open simulation started with the 1VRS B structures appear to be closing. Simulations *oxO7* and *redO2* move away from their open starting structures and move closer to the respective crystal structure for oxidised and reduced nDsbD. Open structures appear to be unstable and seem to rapidly switch to closed conformations. Given that RMSD differences between open and closed structures are quite small as seen for example in Fig. 3.3 C and D, it is important to use more sensitive metrics to track the cap loop in the simulations.

3.3.2 Functional characterisation of the MD simulations of reduced nDsbD

Cap loop conformation in the simulations of reduced nDsbD

Through the extensive set of MD simulations (Table. 3.2) it was investigated whether the cap loop protects the active site thiols of reduced nDsbD. Such protective function has been proposed based on the crystal structures of nDsbD (Goulding et al., 2002; Mavridou et al., 2011). For this the conformation of the cap loop in the simulations was tracked using simple distances. The conclusions from the analysis were subsequently tested by calculating the

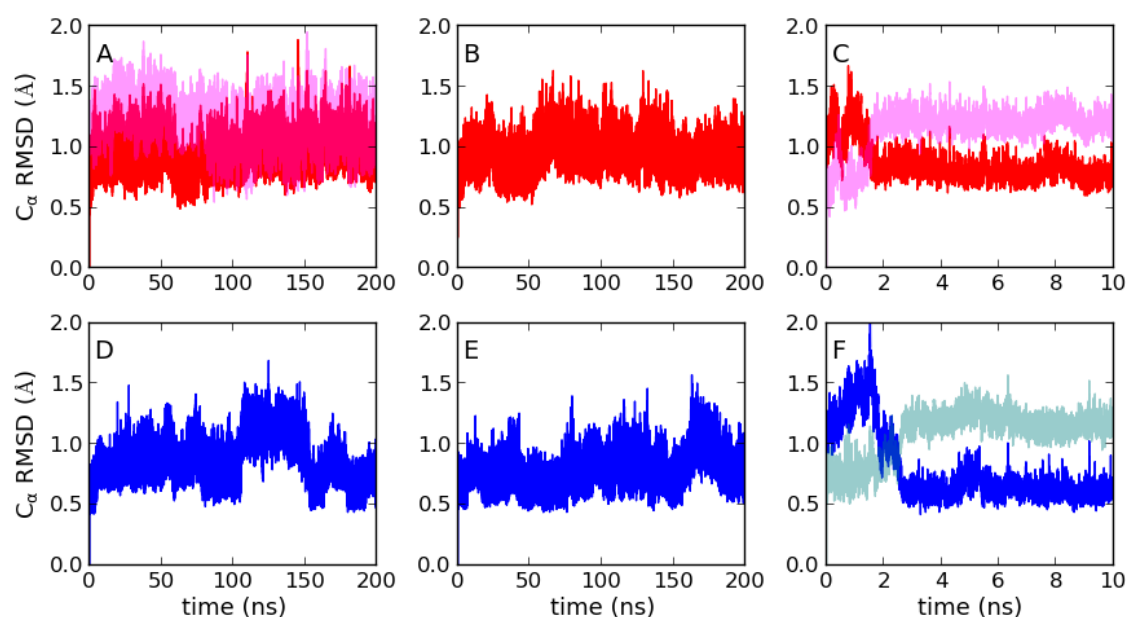


Fig. 3.3 C_{α} RMSD for representative simulations with open and closed structures of both oxidised and reduced nDsbD. In A and B the C_{α} RMSD to the starting structure for simulations *ox1* and *ox9*, started with the closed 1L6P crystal structure of oxidised nDsbD is shown in red. For simulation *ox1* the C_{α} RMSD to the open structure of nDsbD is also shown in magenta. In panel C, the time course for simulation *ox07*, compared to the open structure and the closed 1L6P structure are shown, in magenta and red. The C_{α} RMSD traces for simulation *red1* and *red4* are shown D and E. In plot F, the C_{α} RMSD, was calculated in respect to the open starting structure and the closed 3PFU structure in blue and cyan. The C_{α} RMSD was calculated after superimposing all the frames in a given simulation with the respective open or closed reference structure with MDAnalysis. Residues 14-121, which are well structured according to the NMR dynamics experiments, were included in the calculations.

surface accessibility of the active-site cysteines in the simulations.

The distance between the sulphur of C109 in the active site and the centre of the aromatic ring of F70 in the cap loop was used to track the state of the loop in the simulations. As indicated in Fig. 3.4, the F70-C109 distance clearly distinguishes between open and closed structures of nDsbD seen in X-ray crystallography. Parameters that distinguish biologically important states are often referred to as order parameters, and should not be confused with the S^2 backbone order parameter that describes the flexibility of bond vectors. Figs. 3.4a (*red1*) and 3.4b (*red4*) show typical trajectories from the set of 14 simulations. The distance between F70 and C109 stays close to the value obtained for the closed starting structure and open structures are not accessed in either simulation.

Monitoring an additional distance, the separation between the centre of the aromatic ring of Y71 in the cap loop and C109 in the active site, does not change the picture. This can be seen by representing the 1 μ s ensemble of simulations started with the closed crystal structure of reduced nDsbD, as a 2D histogram, using the F70-C109 as well as the Y71-C109 distance (Fig. 3.6a). Contour lines, indicating the most populated conformations, make it clear that structures with the cap loop closed dominate. Some structures with more open conformations are occasionally accessed as evidenced by the grey line that indicates the extent of sampling on the plot. Nonetheless a fully open conformation such as the one in the 1VRS crystal structure is not sampled. The 1VRS B structure is clearly outside the region of the 2D order parameter space explored by the 1 μ s of simulations.

In addition, the conformation of the cap loop was followed in the simulations started with open structures of reduced nDsbD (Table. 3.3), using the same distances. The loop closed in five out of eight short (10 ns) MD simulations started with a reduced open structure based on the 1VRS B crystal structure. A typical trajectory is shown in Fig. 3.4c. The distance between F70-C109 quickly decreases to resemble values which are typically seen in the simulations started with closed crystal structure, *c.f.* Figs. 3.4a and 3.4b. The closed conformation is the most common structure observed in simulations started with the open structure when averaging the 2D order parameter histograms, as evidenced by the contour lines in Fig. 3.6b. More simulations were launched from the final open structure of one of the three simulations, *red01*, that did not show closure within 10 ns. While none of this set of ten 10 ns simulations entered a fully closed state, all of them showed some loop closure, as evidenced by the reduction in the distances between C109 SG and the rings of F70 and Y71. The closure is apparent in Fig. 3.6c where the final structures of this set of simulations are indicated and compared to their starting structure, the final snapshot of *red01*. The analysis of the simulations using the F70-C109 and Y71-C109 distances thus suggest that

open conformations of the cap loop are unfavourable and that loop closure can happen on the timescale of tens of ns.

Surface accessibility in the simulations of reduced nDsbD

The solvent accessible surface areas (SASAs) for the sulphur atoms of the active site cysteines were calculated for representative simulations of reduced nDsbD. This gives a direct indication of whether the sulphur atoms would be exposed to the periplasm and thus to side reactions, complementing the distance based analysis of the functional state of nDsbD. The calculation of SASA from atomistic trajectories is relatively slow and thus SASAs were only calculated for two typical trajectories of reduced nDsbD.

In the (energy minimised) crystal structure of reduced nDsbD, both sulphur atoms are largely buried, with a combined SASA of $\sim 0.04\text{nm}^2$. Only a fraction of the total surface area for the two sulphur atoms, *i.e.* $2 \times 4\pi r_{\text{vdW}}^2 \text{Sulphur} \sim 0.89\text{nm}^2$, using $r_{\text{vdW}} = 0.188\text{nm}$ is exposed. The van der Waals radius is the same as the one used for sulphur in the CHARMM potential energy function. In simulation *red1* the SASA for the sulphurs increased to an average of $\sim 0.09\text{nm}^2$. The active-site sulphurs remain largely shielded from the solvent, which is consistent with the distance analysis for this trajectory which showed that the loop remained closed and visual inspection.

The calculations also showed that the sulphur atoms are indeed exposed when the cap loop is open, as judged by the distance analysis. In the initial stages of simulation *redO2*, started with the open 1VRS B structure, the two sulphur atoms are clearly exposed, with SASA as high as 0.8nm^2 . The cap loop is also clearly open as judged by the F70-C109 distance (Fig. 3.4c), which is mostly above $10\text{\AA}$. After $\sim 2 - 3\text{ ns}$ of simulation the sulphur atoms become much less solvent exposed with an average SASA of 0.1nm^2 for the last 6 ns of *redO2*, which is very similar to the average from *red1*.

Considering solvent exposed surface area, also helps to rationalise why the open cap loop is unstable (Fig. 3.8). The hydrophobic residues F70 and Y71 are much more solvent exposed in open structures, such as in the initial part of trajectory *redO2*, with a SASA above 3.5nm^2 , than in closed structures. The average SASA for F70 and Y71 is 1.9nm^2 for closed nDsbD in *red1* and for the last 6 ns of *redO2*. Following the argument of Chandler, burying the side chains of F70 and Y71 maximises the hydrogen bonding between the water molecules solvating the cap loop (Chandler, 2005). The closure of the cap loop is therefore enthalpically favourable. This assumes that loop opening reduces the number of hydrogen bonds the water molecules surrounding the loop can form. Unlike the situation for small molecules, there is no way of maintaining the maximum number of hydrogen bonds possible

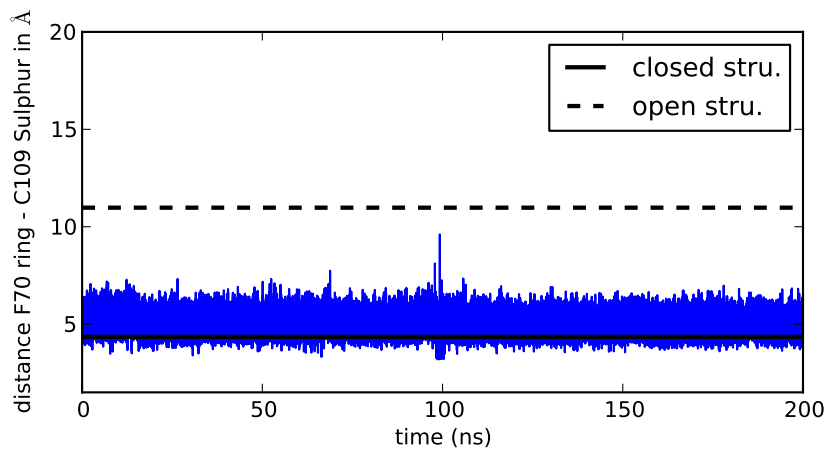
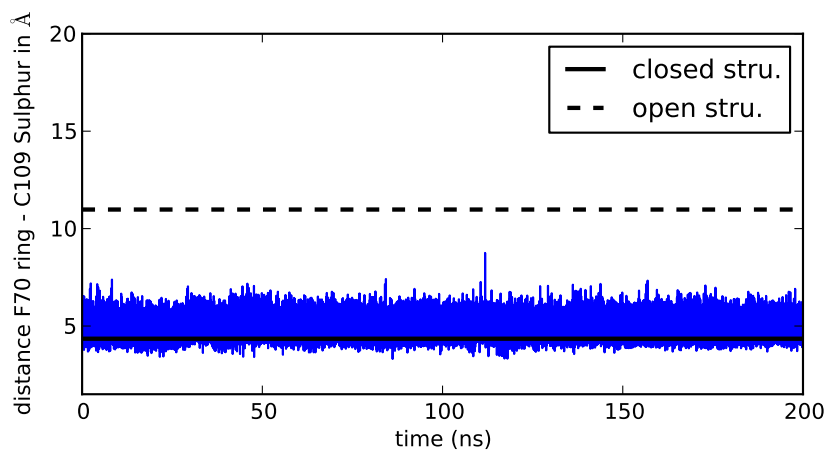
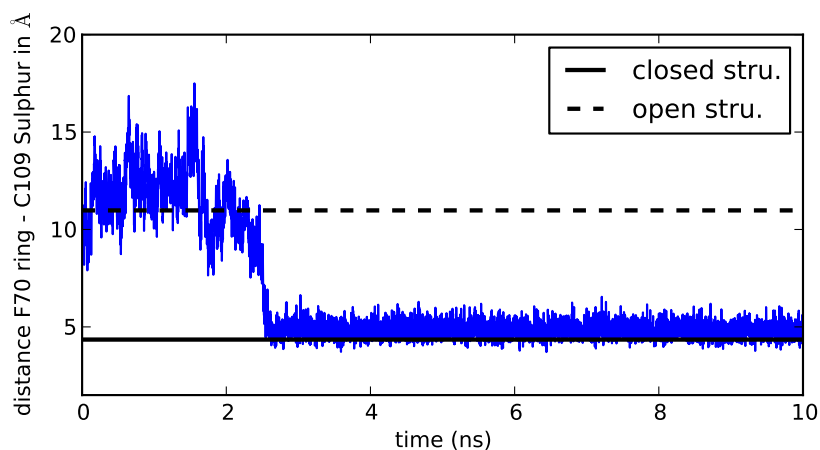
(a) Simulation *red1* started with the closed 3PFU structure(b) Simulation *red4* started with the closed 3PFU structure(c) Simulation *redO2* started with an open structure based on the 1VRS crystal structure

Fig. 3.4 The dynamics of the cap loop in selected simulations of reduced nDsbD. Cap loop opening and closing is followed by tracking the distance between the centre of the aromatic ring of F70 and the sulphur of C109.

by restricting the interfacial water molecules. The interpretation seems reasonable in the case of loop opening in the case of nDsbD, and could be tested by analysis of the hydrogen bonding patterns in the simulations.

3.3.3 Functional characterisation of the MD simulations of oxidised nDsbD

Cap loop conformation assessed in the simulations of oxidised nDsbD

As for reduced nDsbD, both distances and solvent accessible surface areas were used to characterise the functional state of the cap loop in the MD simulations of oxidised nDsbD. MD simulations started with the closed 1L6P crystal structure of oxidised nDsbD show more complicated dynamics in the cap loop than the simulations of the reduced protein. In some simulations the cap loop remained closed such as in trajectory *ox8* (Fig. 3.5a), only showing occasional fluctuations, as assessed by the distance between the aromatic ring of F70 and the sulphur of C109. In simulation *ox1* on the other hand, the fluctuations in the cap loop are much more pronounced (Fig. 3.5b). After 80 ns the cap loop opens as judged by the F70-C109 distance and stays open for the remainder of the 200 ns equilibrium MD trajectory.

To confirm the analysis the distance between the aromatic ring of Y71 and sulphur atom of C109 was also tracked. This distance serves as an additional order parameter that can distinguish between open and closed structures. In Fig. 3.7a the two distances are used to represent the 14 simulations with a total simulation time of 1 μ s as a two dimensional histogram. A grey line is drawn around the bins accessed by simulations started from the closed 1L6P crystal structure of oxidised nDsbD. The open structure of nDsbD, from the 1VRS crystal structure, is clearly within the area accessed by the simulations started from the closed crystal structure of oxidised nDsbD. Nonetheless, closed structures predominate, as highlighted by the contour lines in Fig. 3.7a that map right next to the 1L6P structure on the two dimensional distance space for the cap loop.

The F70-C109 and the Y71-C109 distances were also used to track the cap loop in the simulations of oxidised nDsbD started with the cap loop open. Simulations started from the open 1VRS B structure had a tendency to close the cap loop. Closure can be seen for example for simulation *ox07* in Fig 3.5c. The conformational preferences of the 10 simulations started with the open 1VRS B structure are represented in Fig. 3.7b using both the F70-C109 and Y71-C109 distances. Contour lines show that the most populated conformations for this set of simulations map next to the closed crystal structure of oxidised

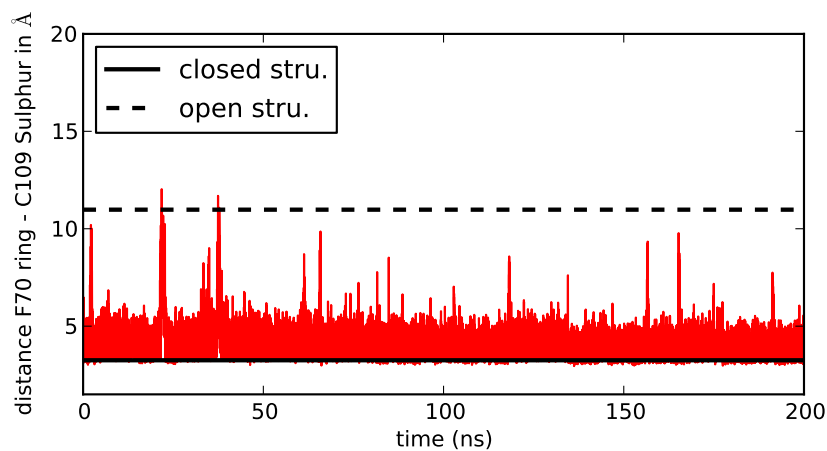
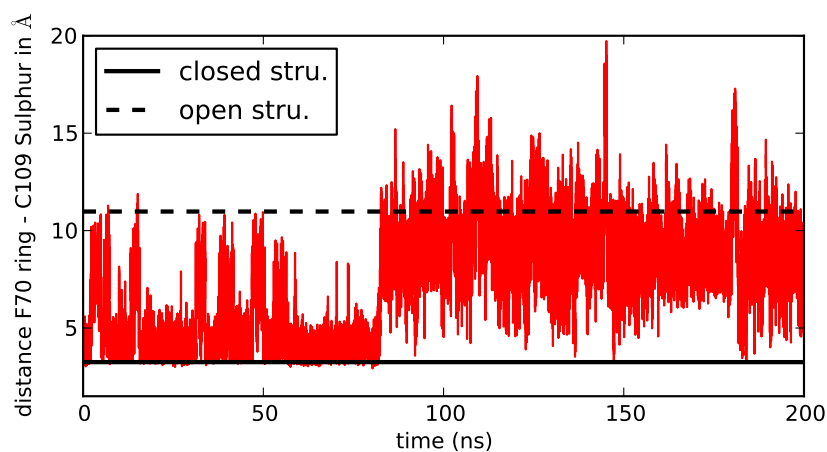
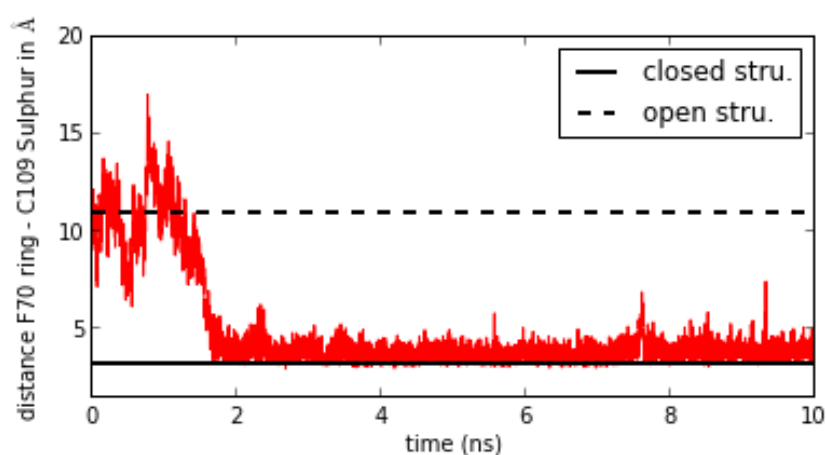
(a) Simulation *ox8* started with the closed 1L6P structure(b) Simulation *ox1* started with the closed 1L6P structure(c) Simulation *ox07* started with an open structure based on the 1VRS crystal structure

Fig. 3.5 The dynamics of the cap loop in selected simulations of oxidised nDsbD. Cap loop opening and closing is followed by tracking the distance between the centre of the aromatic ring of F70 and the sulphur of C109.

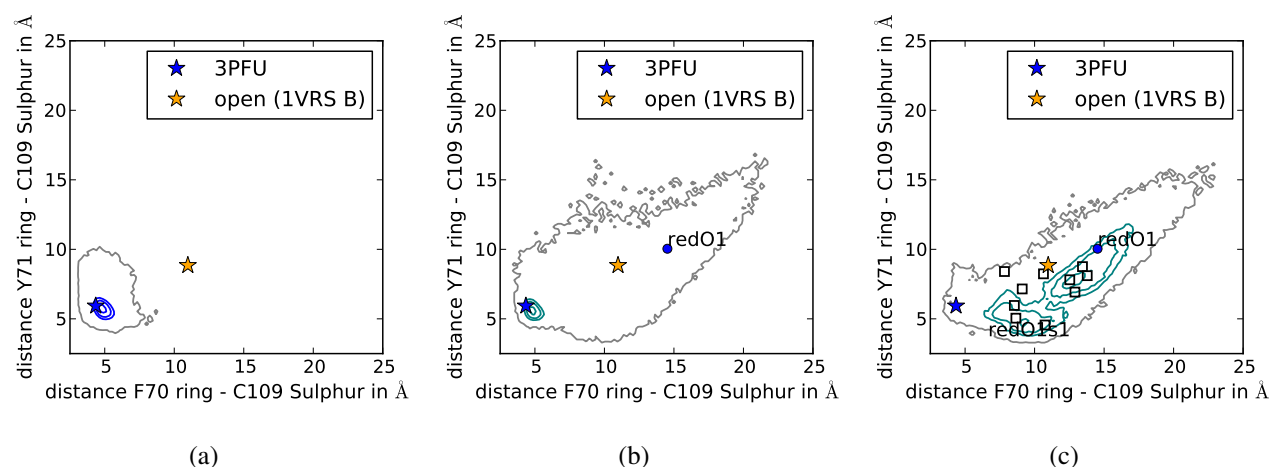


Fig. 3.6 Analysis of the cap loop structure in simulations of reduced nDsbD. Simulations started with the closed structure of reduced nDsbD are shown in (a). Multiple simulations were also started with the 1VRS B open structure with the active site reduced (b). *redO1* did not close in the simulations. The simulation was continued and further simulations were initiated from its final frame (c). Contour lines are drawn at 20, 40 and 80 % of the maximum bin density in blue (a) and cyan (b and c) to illustrate the most common structures. The gray lines surround all bins accessed by the simulation ensemble and indicate the extent of sampling.

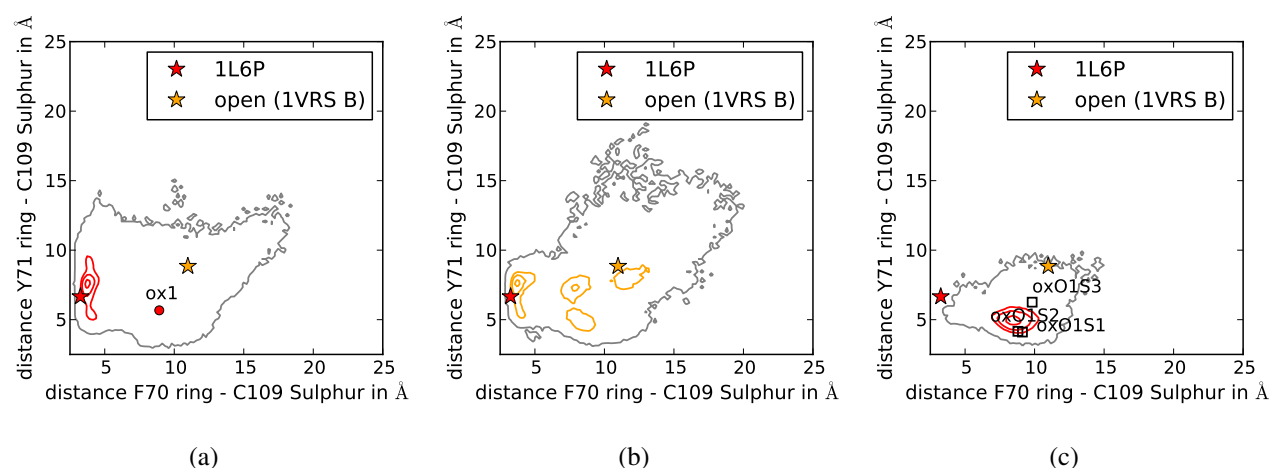


Fig. 3.7 Analysis of the cap loop structure in simulations of oxidised nDsbD. The simulations analysed were started with the closed 1L6P structure (a), the 1VRS B structure with a disulphide bond in the active site (b), and the final structure of *ox1* (c). Contour lines are drawn at 20, 40 and 80 % of the maximum bin density. Contour lines are in red (a and c) and orange (b) respectively. The gray lines surround all bins populated by the simulation ensemble.

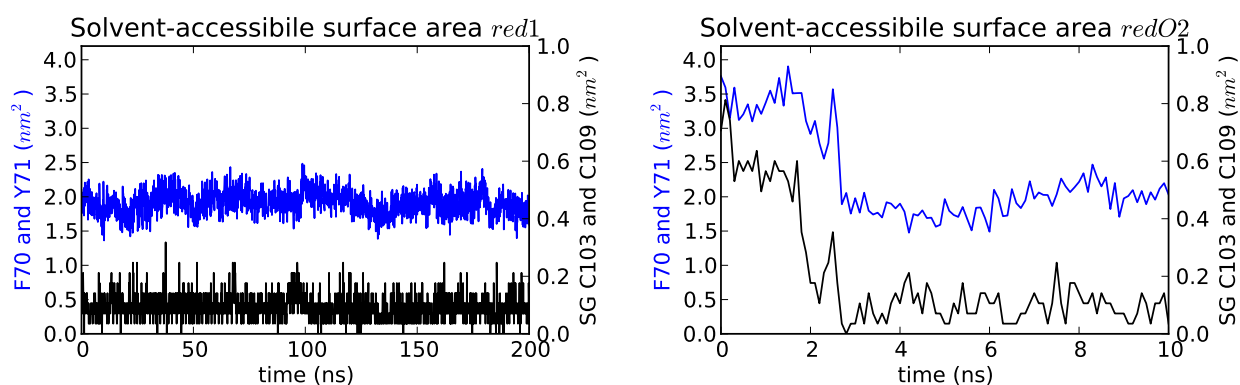


Fig. 3.8 Solvent accessible surface area (SASA) for the aromatic cap loop residues F70 and Y71 as well as for S_γ of C103 and C109 in simulations *red1* (Fig 3.8a) and *redO2* (Fig 3.8b).

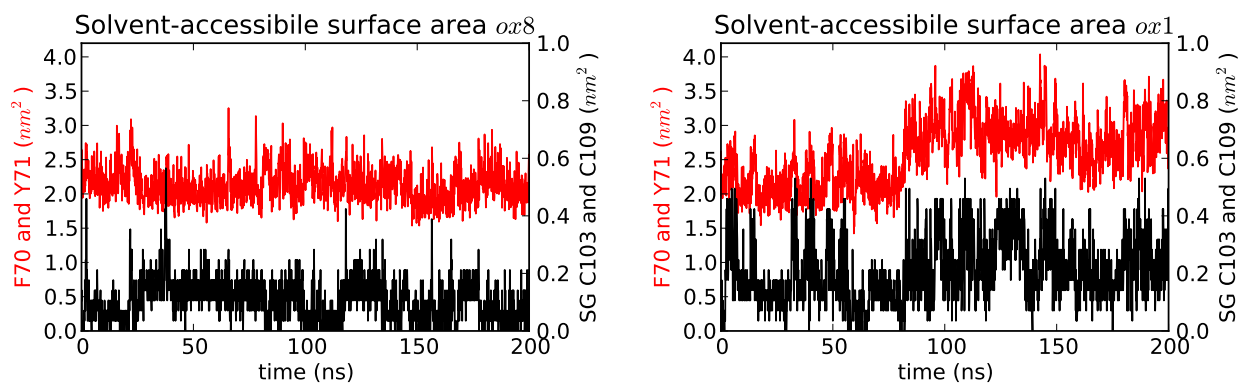


Fig. 3.9 Solvent accessible surface area (SASA) for the aromatic cap loop residues F70 and Y71 as well as for S_γ of C103 and C109 in simulations *ox8* (Fig 3.9a) and *ox1* (Fig 3.9b).

nDsbD. The result is not as clear cut as for reduced nDsbD (Fig. 3.6b). The contour lines show a significant population of oxidised nDsbD with an open loop, for the set of ten 10 ns trajectories.

Three simulations, *ox1s1*, *ox1s2* and *ox1s3*, were started from the open final structure of *ox1*. On average, the cap loop remained open in the three 20 ns simulations, as evidenced by the contour lines in Fig. 3.7c on the two dimensional cap loop distance plot. The line indicating the region of the conformational space accessed by *ox1s1*, *ox1s2* and *ox1s3* comes close to the 1L6P structure, indicating a degree of transient loop closure in the simulations. The final structures of the three simulations are however open, as shown in Fig. 3.7c. The two sets of simulations started with oxidised nDsbD with an open cap loop appear to behave somewhat differently. The cap loop seemed more prone to closure in simulations started with the 1VRS B structure than in simulations *ox1s1*, *ox1s2* and *ox1s3*. This may reflect true differences in the stability of the starting structures. In the 1VRS crystal, nDsbD is in a covalent complex with cDsbD. The structure of open nDsbD from the 1VRS crystal structure may hence be considerably affected by interactions with cDsbD and the intermolecular disulphide bond. Therefore it might not be a good representation of an open structure of nDsbD in the absence of cDsbD and less stable than the open structure spontaneously accessed in *ox1*. However differences between the two sets of simulations may be due to insufficient sampling for either set. The combined simulation lengths of 100 ns and 60 ns are clearly too short for the two sets to converge.

Surface accessibility in the simulations of oxidised nDsbD

Calculating the solvent accessible surface areas (SASAs) for selected simulations of oxidised nDsbD confirmed the analysis based on the cap loop distances. In simulation *ox8*, the average SASA for the sulphur atoms of C103 and C109 in the active site is 0.1nm^2 . Only a small part of the total surface area, $\sim 0.89\text{nm}^2$, of the two sulphurs, is exposed to the solvent. SASA for the sulphur atoms show transient spikes for trajectory *ox8*, similar to the pattern seen in the cap loop order parameter analysis in Fig. 3.5a. The correspondence between the sulphur becoming more exposed according to the SASA calculation and cap loop opening detected by the F70-C109 distance is particularly evident when showing the time course for the two parameters on the same plot (Fig. 3.10). The SASA for F70 and Y71 also shows transient increases. The average SASA for F70 and Y71 is 2.13nm^2 , and remained nonetheless similar to the average calculated for the simulation *red1* $\sim 1.9\text{nm}^2$ where the cap loop stayed fully closed. Overall, the SASA calculations for the aromatic residues in the ring and the sulphur atoms in the active site strongly confirm the conclusion

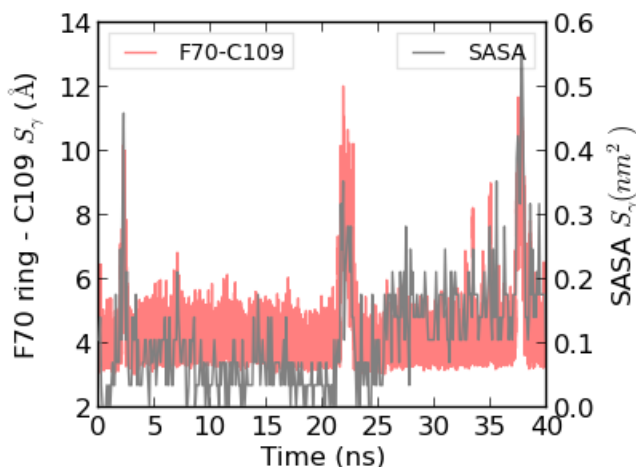


Fig. 3.10 The F70-C109 order parameter and the SASA calculated for S_γ of C109 are compared. Both measures are sensitive to loop opening. Here the initial segment of the 200 ns trajectory for *ox8* is presented.

from the cap loop distance analysis for *ox8*. The cap loop remains closed on average, but shows signs of transient opening.

Given the sustained loop opening in simulation *ox1*, it is instructive to also calculate surface accessibilities for the trajectory. In the initial stages (< 80 ns) of *ox1* before the loop opened, the average SASA for the sulphurs is already $\sim 0.17\text{nm}^2$, which is clearly larger than for *ox8*. The cap loop shows significant fluctuations in this segment, that lead to the transient spikes in accessibility for the catalytic sulphur atoms. The fluctuations are also larger than for *ox8* as judged by the F70-C109 distance even before the loop fully opens. The difference is evident from comparing the time course of the F70-C109 distance up to 80 ns in Fig. 3.5b and Fig. 3.5a. The accessibility of the active-site sulphur atoms increases further after cap loop opening (> 85 ns) to an average value of 0.24nm^2 .

Loop opening is clearly evident in the time course for the exposed surface area of F70 and Y71 in Fig. 3.9b. The average SASA for the two aromatic residues in the cap loop increases from $\sim 2.2\text{nm}^2$ before loop opening (< 80 ns) to $\sim 2.9\text{nm}^2$ after the loop open. The average for the open state was calculated from 85 ns to 200 ns. The increase in SASA of the two hydrophobic side chains upon loop opening may explain why closed structures are the most common structures not only for reduced but also oxidised nDsbD, which is apparent from the 2D histograms of the cap loop order parameters. The analysis does not explain yet why oxidised but not reduced nDsbD showed spontaneous loop opening in the extensive set of simulations.

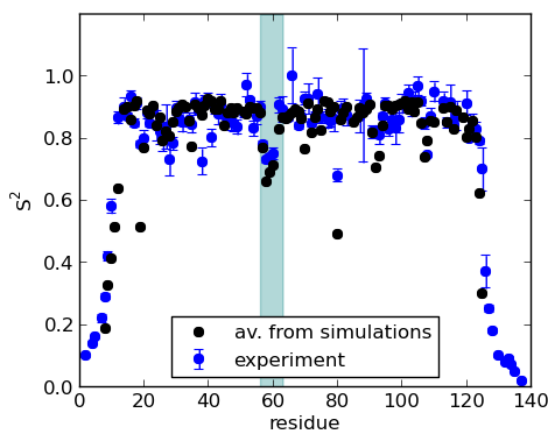
3.3.4 Comparing the simulations to the NMR spin-relaxation experiments

MD simulations, unlike the analysis of the NMR spin-relaxation experiments, showed clear differences between oxidised and reduced nDsbD. Thus it is important to ask whether the simulations are consistent with the experiments, *i.e.* to compare the S^2 order parameters determined from the model-free analysis of the experiments and the values back-calculated from the MD simulations. NMR spin relaxation experiments analysed by the "model-free" approach are mostly sensitive to motions faster than the overall tumbling of the protein. To compare fast timescale fluctuations in experiment and simulation it is therefore important to calculate S^2 for time intervals similar to the rotational tumbling time of the protein. The rotational correlation time, τ_c , of both oxidised and reduced nDsbD is close to 7.5 ns. We have calculated S^2 for 5 ns segments of the 20 ns and 200 ns trajectories that constitute the 1 μ s ensembles for both oxidised and reduced nDsbD. The choice of 5 ns is close to τ_c ensures all segments are of equal length, which facilitates the analysis.

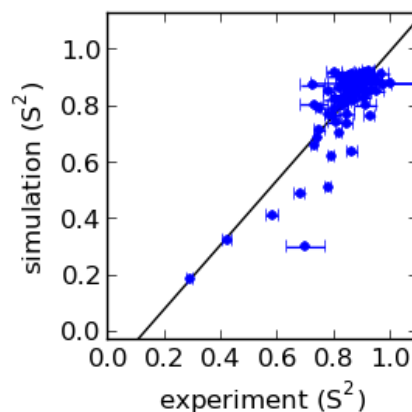
Comparing experimental and calculated S^2 for reduced nDsbD

Overall there is a good agreement between the order parameters calculated from the 1 μ s MD ensemble for reduced nDsbD and experiment. The root mean-square deviation (RMSD) between experiment and S^2 from the simulations is 0.08, *i.e.* the S^2 from simulations and experiment disagree on average by about 8 %. The good correlation, with a correlation coefficient of 0.82, is evident from Fig 3.11b. While there are some outliers, the extent of fast-time scale dynamics in the experiment is on the whole correctly reproduced by the simulations. The N- and C-termini are very flexible, but most residues in reduced nDsbD are well ordered with $S^2 > 0.80$.

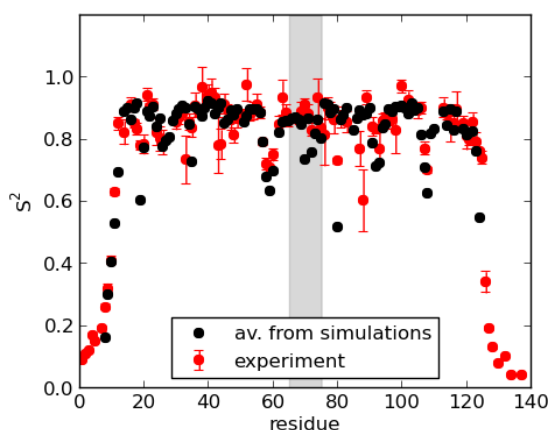
Encouragingly the simulations also managed to reproduce more subtle features of the residue-specific variations in S^2 . For example, residues 57-60 have S^2 order parameters below 0.80, in both experiments and simulations. In particular, for V58 and Q59 S^2 order parameters below 0.75 were found in both experiments and simulations. Residues A56-G63 are not part of the regular regular secondary structure of the domain which explains the increased flexibility in this region. The cap loop is not very flexible in the simulations of reduced nDsbD. Residues in the cap loop region, *i.e.* between W65 and S74, have backbone S^2 above 0.75 which is consistent with the "model-free" analysis of the spin relaxation experiments.



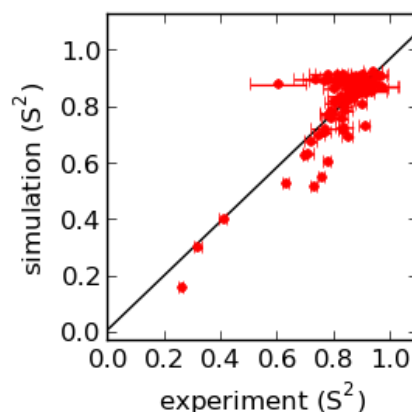
(a) S^2 reduced nDsbD from MD simulations and experiments



(b) correlation for reduced nDsbD



(c) S^2 oxidised nDsbD from MD simulations and experiment.



(d) correlation for oxidised nDsbD.

Fig. 3.11 S^2 calculated from simulations of oxidised and reduced nDsbD is compared to experiment. In (a) and (c) the experimental and calculated S^2 for reduced and oxidised nDsbD are compared on a per-residue basis. Experimental S^2 order parameters are shown in blue and red in (a) and (c) for oxidised nDsbD, while the S^2 parameters from MD are depicted in black. Experiments and MD are also compared correlation plots in (b) and (d) for both reduced and oxidised nDsbD.

Comparing experimental and calculated S^2 for oxidised nDsbD

For oxidised nDsbD, good agreement is found between the experimental backbone amide order parameters and the 1 μ s of MD simulations started with the crystal structure of oxidised nDsbD. In general, the agreement between experiment and simulation is similar to the reduced protein. The RMSD between experiment and S^2 from the simulations is again 0.08 and a good correlation, with a correlation coefficient of 0.83, is obtained, as shown in Fig 3.11d. Most residues do not show extensive fluctuations and have $S^2 > 0.8$, while the N- and C-terminal residues are very dynamic.

It was investigated whether the simulations reproduced the trends in S^2 in the folded segments of oxidised nDsbD detected in experiments. Similar to reduced nDsbD, S^2 is lowered for residues 57-60 in both experiment and simulation as this region does not have regular secondary structure. The amide order parameters for G107 (experiment 0.77, MD 0.71) and F108 (experiment 0.70, MD 0.63) are lower than for the preceding residues in both simulation and experiment of oxidised nDsbD. These two residues, close to the C109, are some of the few residues that show significant differences in their fast timescale dynamics between the two redox-states of nDsbD according to the NMR experiments and the simulations recapitulate this. Residues 57-60 and 107-108 are highlighted on the structure of nDsbD in Fig. 3.2.

Opening of the cap loop was detected in the simulations of oxidised nDsbD but not of reduced nDsbD, which clearly distinguishes the two redox states *in silico*. Thus it is particularly important to assess whether loop opening is consistent with the NMR experiments. Dynamics in the cap loop, such as opening of the loop, is expected to decrease the amide S^2 order parameters calculated from the simulations. Indeed loop opening, reduces the S^2 order parameter for the backbone amide of F70 calculated from the simulations as shown in Fig 3.12b. A large drop in the rotational auto-correlation function calculated for the time block between 80-85 ns coincides with a substantial increase of the F70-C109 distance that reports on loop opening. The 1 μ s MD ensemble for oxidised nDsbD somewhat overestimates the flexibility in the cap loop region (residue 65-75) as can be seen in Fig. 3.12a. Loop opening may be a relatively rare event and equilibrium simulations that can routinely access only 1 μ s will consequently give incorrect estimates of the opening frequency.

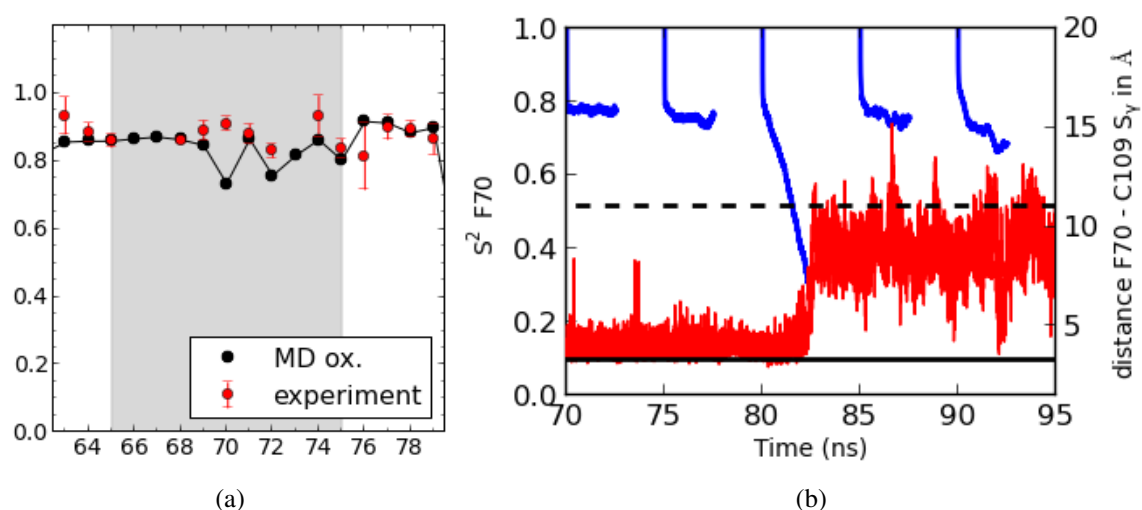
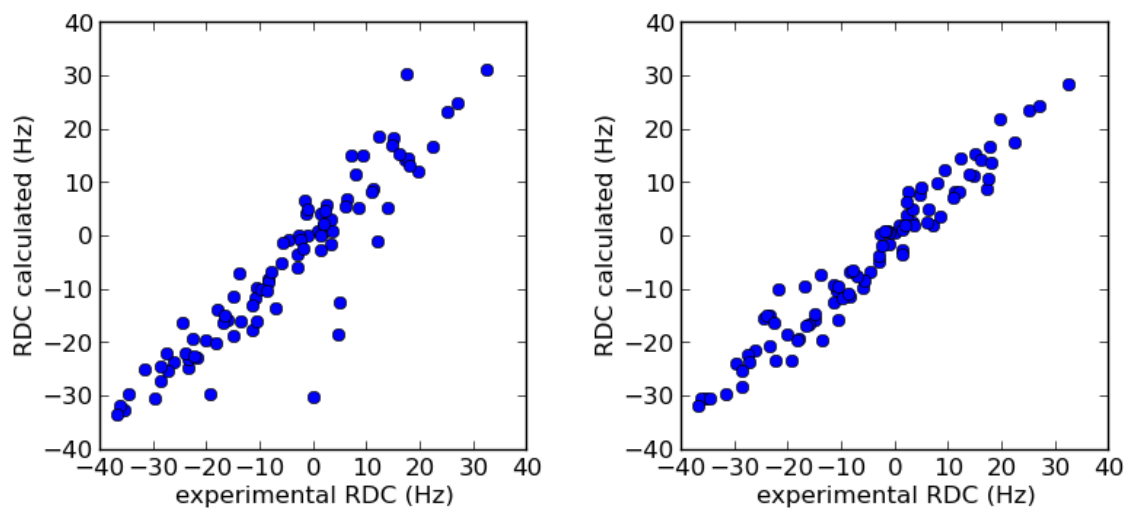


Fig. 3.12 The influence of cap loop opening in oxidised nDsbD on the calculated S^2 order parameter was assessed. In (a) the S^2 values calculated for 1 μ s of MD of oxidised nDsbD is shown and compared to the experimental values. The comparison shows that the simulations somewhat overestimate the flexibility in the cap loop. In (b) the rotational auto-correlation functions for the amide bond vector of F70 are shown in blue. In the distance between the aromatic ring of F70 and S_γ of C109 is shown in red. The distance expected for a closed and open conformation of the cap loop are shown in black as continuous and dashed lines.

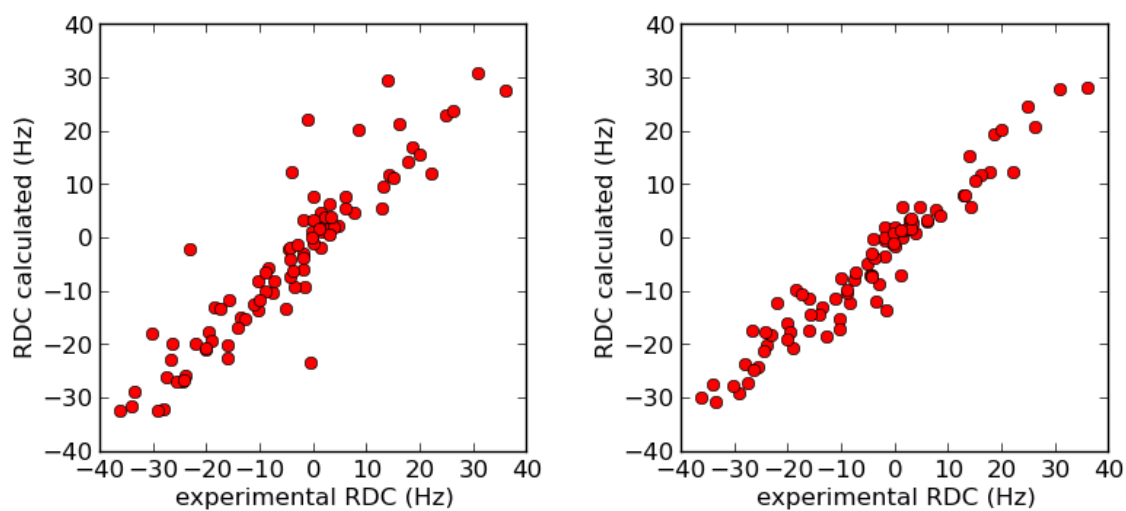
3.3.5 Comparing the simulations to experimental residual dipolar couplings

Residual dipolar coupling (RDC) measurements provide highly accurate information on the average structure of molecules in solution. RDCs report on the relative orientations of bond vectors with respect to the direction of the static magnetic field. RDCs for the NH bond vectors were calculated for the simulation ensembles of both oxidised and reduced nDsbD. For comparison, the agreement between experiment and the RDCs calculated from crystal structures is shown in Fig. 3.13. For reduced nDsbD, the agreement between calculated and measured $^{15}\text{N} - ^1\text{H}$ RDCs improves (RMSD 3.9 Hz) compared to the crystal structure (RMSD 6.1 Hz). For oxidised nDsbD, the difference between calculated and measured RDCs is also smaller (RMSD 4.2 Hz) for the 1 μs MD ensemble than for the 1L6P x-ray structure (RMSD 6.2 Hz). All residues for which RDCs could be measured and calculated were included in the comparisons. Simulations of both oxidised and reduced nDsbD are a better representation of the average backbone conformation in solution than the crystal structures.

It was tested whether the differences in cap loop conformations detected in the simulations agree with the RDC measurements. In Fig. 3.14) the predicted and measured RDCs for the residues in the cap loop of oxidised and reduced nDsbD are compared. For reduced nDsbD, the simulations (Fig. 3.14b) do not provide a better picture of the time averaged conformation of the cap loop than the crystal structure (Fig. 3.14a). The divergence between the calculated RDCs and experiment is higher for the MD ensemble for reduced nDsbD (RMSD 5.69 Hz) than for the 3PFU X-ray structure (RMSD 2.65 Hz). By contrast for oxidised nDsbD, a somewhat lower root mean square deviation is found for the cap loop region in the MD ensemble (RMSD 5.58 Hz) compared to the 1L6P crystal structure (RMSD 7.93 Hz). The MD ensemble greatly improves the agreement between the experimental and calculated RDC for K73. However, the MD ensemble gives a poorer match for F70 than the static oxidised crystal structure. Overall, the RDC measurements do not provide direct evidence for the differences in the cap loop seen in the simulations of oxidised and reduced nDsbD. Nonetheless the simulations appear to be good representation of the average solution structures of oxidised and reduced nDsbD, given the overall good agreement with the RDC experiments.

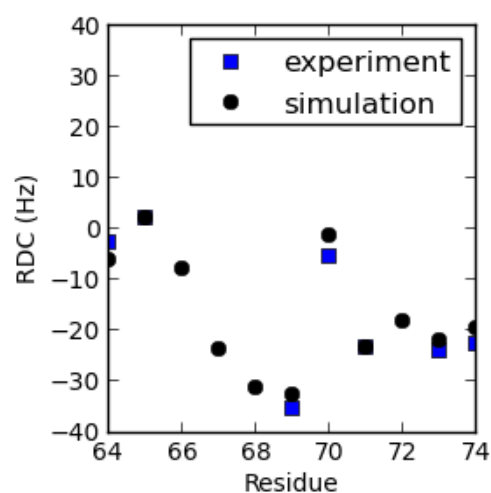


(a) Reduced RDCs: experiment compared to 3PFU. (b) Reduced RDCs: experiment compared to 1 μ s MD.

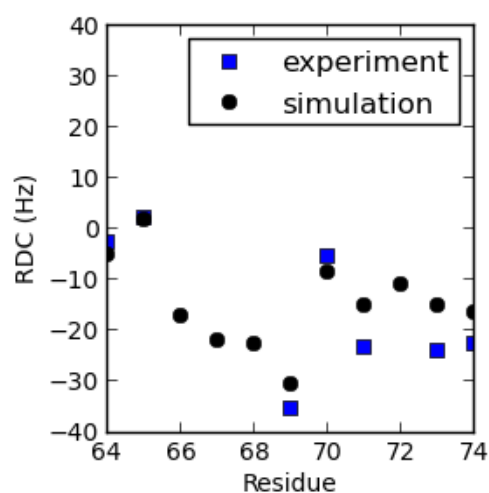
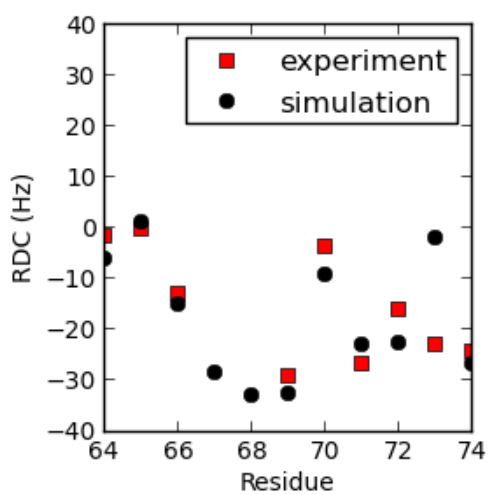


(c) Oxidised nDsbD: experiment compared to 1L6P (d) Oxidised nDsbD: experiment compared to 1 μ s MD

Fig. 3.13 $^{15}\text{N} - ^1\text{H}$ RDCs calculated from crystal structures and simulations of oxidised and reduced nDsbD compared to experiment.



(a) 3PFU crystal structure

(b) reduced nDsbD: 1 μ s MD ensemble

(c) 1L6P crystal structure

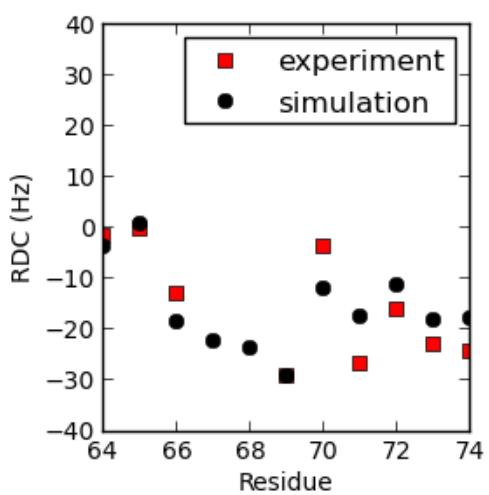
(d) oxidised nDsbD: 1 μ s MD ensemble

Fig. 3.14 Comparing calculated and measured RDCs for the cap loop

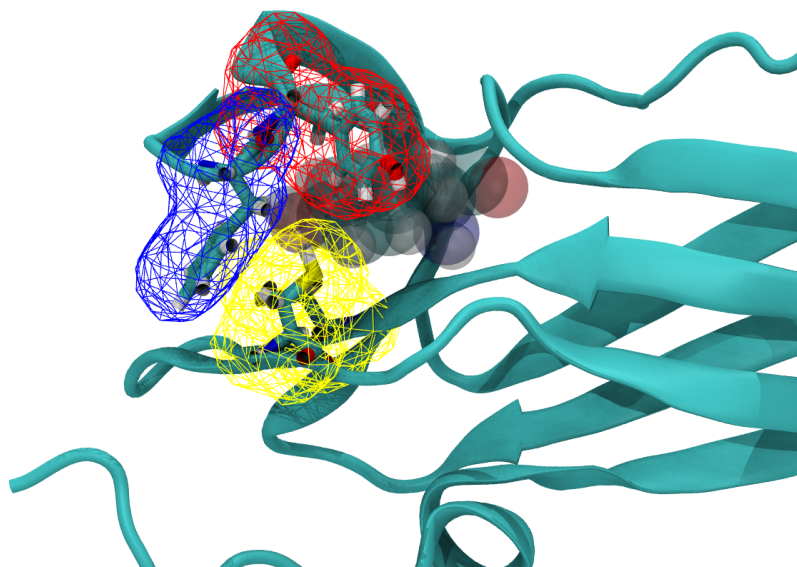
3.3.6 Comparing the MD simulations to the dihedral angle predicted from the chemical shifts

Empirical methods have been making great strides in predicting the likely local geometry of biomolecules from NMR chemical shifts. The simulations should provide a very good picture of local geometry (Robustelli et al., 2012) and thus it is instructive to compare them to the predictions based on experimental chemical shifts. Here, TALOS-N (Shen and Bax, 2013) was used to predict the χ_1 side-chain dihedral angles for both oxidised and reduced nDsbD from the previously determined chemical shifts of both redox states (Mavridou et al., 2011). The focus of the analysis was on the side-chain dihedral angles as visual inspection of the simulations suggested that some of the key differences between oxidised and reduced nDsbD may be due to changes in side-chain geometry. The average conformations of F70 and Y71 in *red1*, a typical trajectory for reduced nDsbD and *ox1*, which showed spontaneous and sustained cap loop opening, are visualised as atomic densities in Fig. 3.15. F70 in particular is much more flexible for oxidised than for reduced nDsbD, which is evident when comparing the well-defined density for *red1* and the much wider volume sampled in *ox1*.

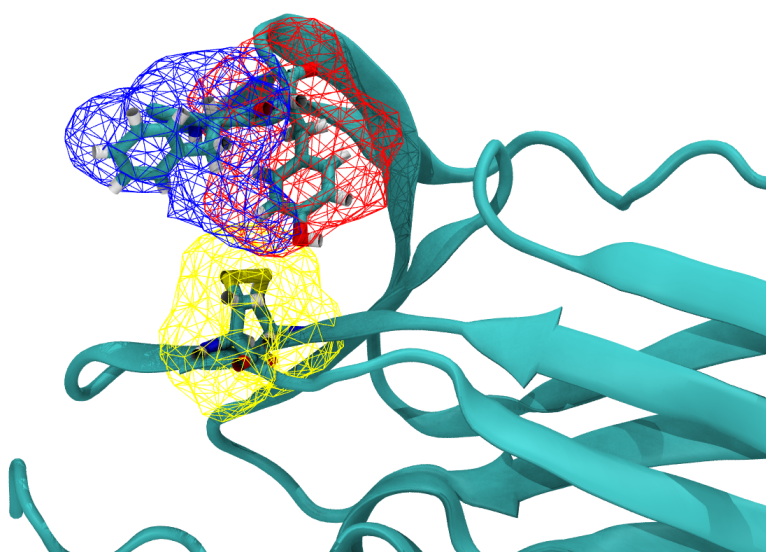
TALOS-N uses experimental chemical shifts to search a database of protein fragments to assign dihedral angles and thus whether the side-chains are in gauche-, gauche+ or trans conformations. For a number of residues, TALOS-N did not find a single dominant (populated over 60% of the time) side-chain dihedral and no prediction was made, which may be due to limitations in this empirical approach. However, many residues will show averaging between different side chain conformations in solution and the inability of TALOS-N to find a single dominant rotamer correctly reflects this.

For most residues for which TALOS-N can assign a gauche-, gauche+ or trans orientation the results agree between the two redox states (Fig. 3.16a). For example, for Y100 TALOS-N predicts, the side-chain to be in a gauche- conformation for both oxidised and reduced nDsbD. As shown in Fig. 3.17 this pattern is obtained from the 1 μ s simulation ensembles for oxidised and reduced nDsbD. Y100 is on one of the β -strands in the Ig-fold of nDsbD and its side chain is closely packed against the rest of the protein, restricting its flexibility.

Interestingly, TALOS-N finds some differences between the two redox states in the cap loop region. Unlike for Y100, the side-chain conformation for F70, the key residue that covers C109 in the active site, is predicted to differ between the two redox states. For reduced nDsbD, a gauche- conformation is predicted by TALOS-N (Fig. 3.16b) which is



(a) The active site in simulations of reduced nDsbD.



(b) The active site in simulations of oxidised nDsbD

Fig. 3.15 The conformations sampled by F70,Y71, C103 and C109 in simulations of oxidised and reduced nDsbD. Mass densities were calculated with VMD (Humphrey et al., 1996) to indicate the time-averaged atomic structures of these residues. The wire-frames surfaces are at 0.3 times the maximal densities.

also seen in the simulations (Fig. 3.18). For F70 in oxidised nDsbD, no dominant rotamer is found by TALOS-N. Fragments with both trans and gauche- conformations match the chemical shifts and the populations of these two rotamers are estimated to be approximately equal. TALOS-N uses the information from 18 chemical shift to estimate the χ_1 dihedral angle. All 18 chemical shifts were available for F70 and included in the calculation. In 1 μ s of simulation frequent transitions between the trans (Fig. 3.18d) and gauche- (Fig. 3.18c) conformations were observed. The transition resulted in nearly equal populations for the two conformers (Fig. 3.18). Predictions based on chemical shifts alone directly confirm that the cap loop is more flexible in oxidised than reduced nDsbD.

3.3.7 Comparing non-bonded interactions in the active sites of reduced and oxidised nDsbD

To rationalise the difference in flexibility between oxidised and reduced nDsbD the van der Waals and electrostatic interactions in their active sites were compared. First, representative conformations of F70, Y71, C103 and C109 were selected from the trajectories *red1* and *ox1*. For this five and six frames respectively spaced at a constant interval were chosen to represent a total of 200,000 frames (for 200ns simulation time). Thus a first look at the detailed interactions that distinguish the two redox states in the simulations was made possible. For oxidised nDsbD, the ring of F70 interacts with the S_γ of C109. The interaction will be dominated by the van der Waals interactions given that sulphurs forming a disulphide bond only feature a very small partial charge (-0.08) in the CHARMM22 potential energy function. F70 can pack closely onto the the sulphur, that forms a relatively flat interaction surface, both in its trans and its gauche- conformation, which may explain why it can readily switch between the conformations.

For reduced nDsbD we see a very different interaction between the ring of F70 and C109. The ring now packs next to the C_β of C109. The ring of F70 interacts closely with one of the H_β atoms of the C109 (Fig. 3.19). The CD1, CE1, CZ, CE2 and CD atoms of the ring carry a small negative partial charge (-0.115) in the CHARMM22 potential energy function. The negative charges on the carbon atoms are balanced by equally sized positive partial charges on the phenyl protons. In this way the π system of the aromatic group can be represented in MD simulations. The ring carbon atoms thus interact electrostatically with the very small positive partial charge (+0.09) on $H_{\beta 1}$. This interaction can be thought of as a C-H $\cdots \pi$ hydrogen bond (Salonen et al., 2011). Such hydrogen bonds are considered to be weak but it might help to restrict the F70 side chain. Additionally there will of course be

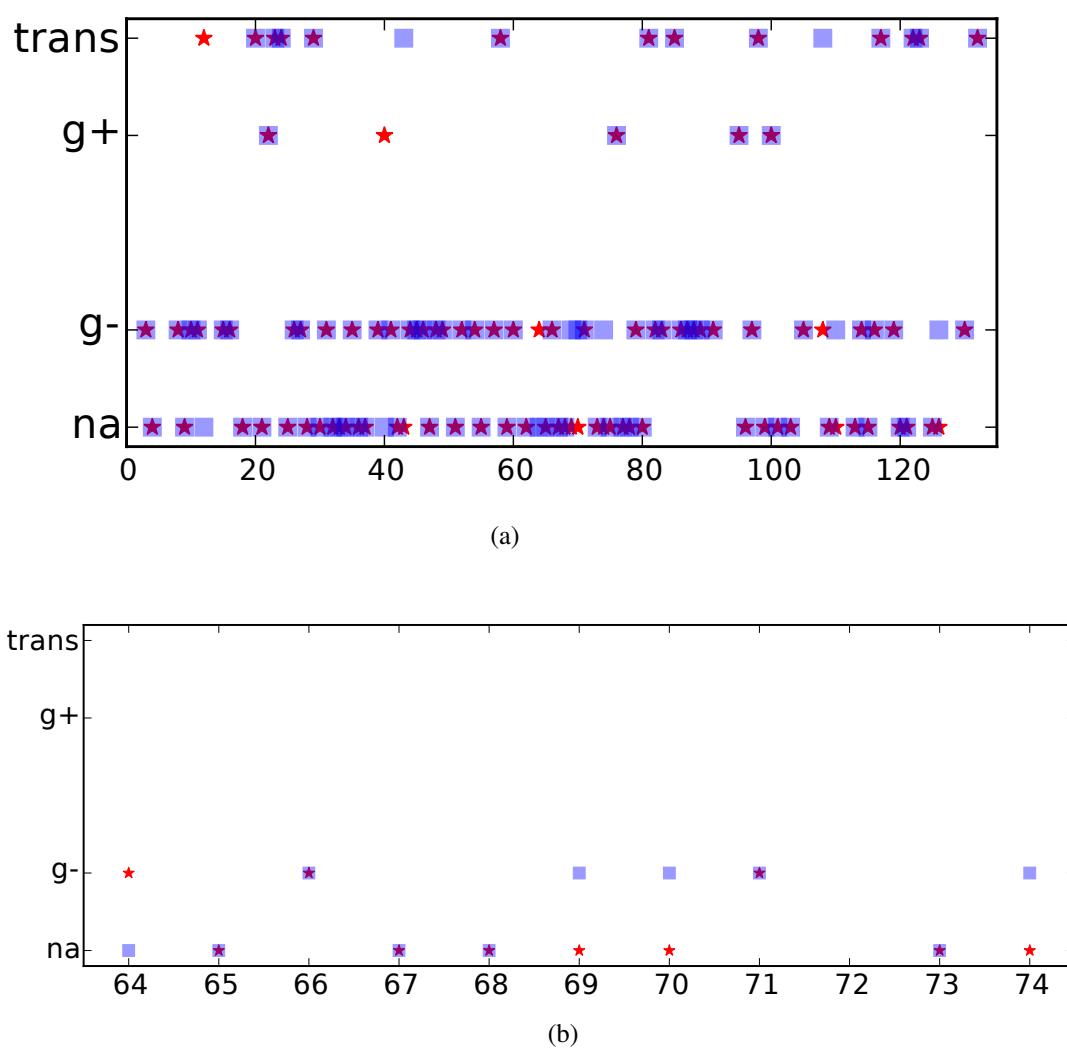


Fig. 3.16 The χ_1 side chain dihedral angles for reduced and oxidised nDsbD were predicted from the chemical shifts using TALOS-N. Predictions for reduced and oxidised nDsbD are shown as blue squares and red stars. Overall the side-chain dihedral angles for the redox isoforms are very similar (a), but there are differences in the cap loop region (b).

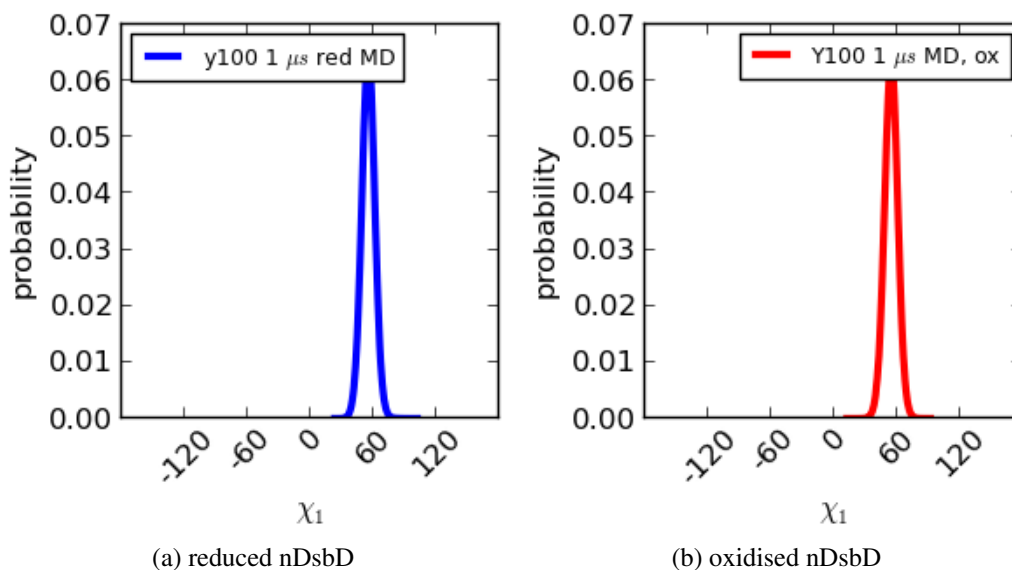


Fig. 3.17 The χ_1 side-chain dihedral distributions for Y100 in 1 μ s of simulations with oxidised and reduced nDsbD are shown.

van der Waals interactions between F70 and C109.

3.4 Discussion

Oxidised and reduced nDsbD were both simulated for 1 μ s in atomistic equilibrium MD simulations. There are clear differences in the simulations of the two redox isoforms, which were particularly apparent when analysing the cap loop conformation with distance-based order parameters. In reduced nDsbD, the cap loop showed only limited fluctuations and remained closed, as shown in the 1D histogram for the F70-C109 distance in Fig. 3.21a. The cap loop and the active site pack together tightly for reduced nDsbD, as visualised in Fig. 3.15a, using densities to represent simulation structures. Experiments by Prof. Redfield have found that cysteine C103 and C109 have very high pK_a s and the shielding of the active site by the aromatic residues on the cap loop likely contributes to this.

For oxidised nDsbD, on the other hand, cap loop opening was observed in our simulations. Closed conformations are however still dominant (Fig. 3.21b) but the cap loop is more flexible in oxidised than in reduced nDsbD. This can be seen both by comparing densities calculated for key residues in Fig. 3.15 (for *ox1*) or by looking at a series of representative snapshots in Fig. 3.20b (for *ox8*).

The simulations are in good agreement with spin relaxation and RDC measurements.

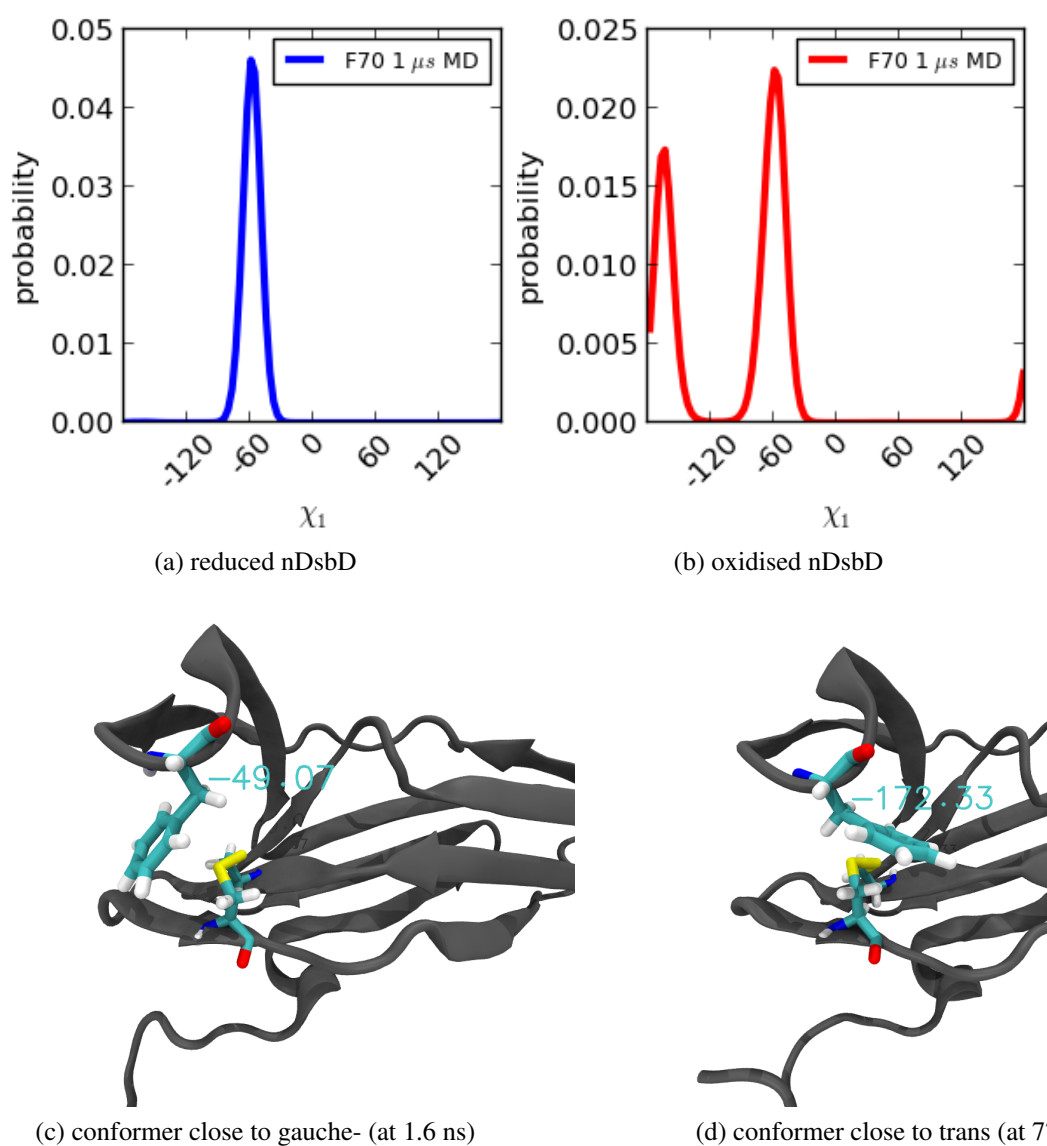


Fig. 3.18 The χ_1 side chain dihedral distributions for F70 in 1 μ s of simulations with oxidised and reduced nDsbD are shown. The averaging seen for F70 in simulations of oxidised nDsbD is visualised with two representative snapshots for *ox8*. A conformer close to gauche- (snap short at 1.6 ns) and one closer to trans (snapshot at 77 ns) are shown.

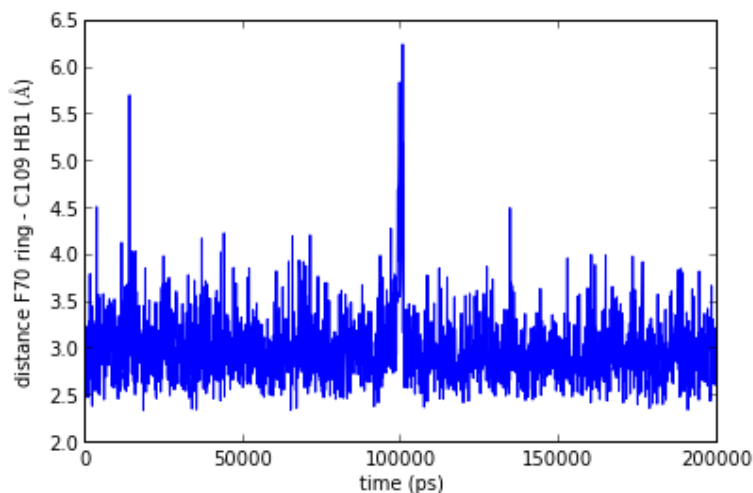
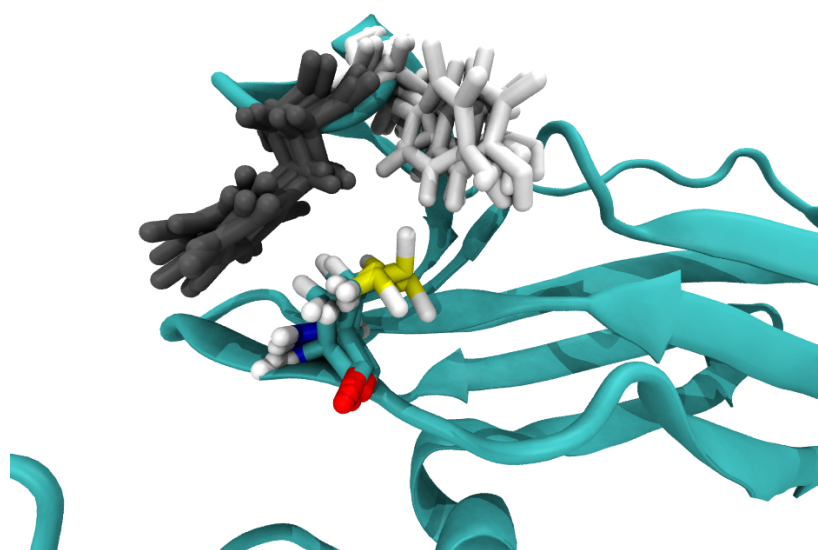


Fig. 3.19 The aromatic ring of F70 interacts closely with $H_{\beta 1}$ of C109 in simulation *red1*. The interaction can be considered a $C-H \cdots \pi$ hydrogen bond.

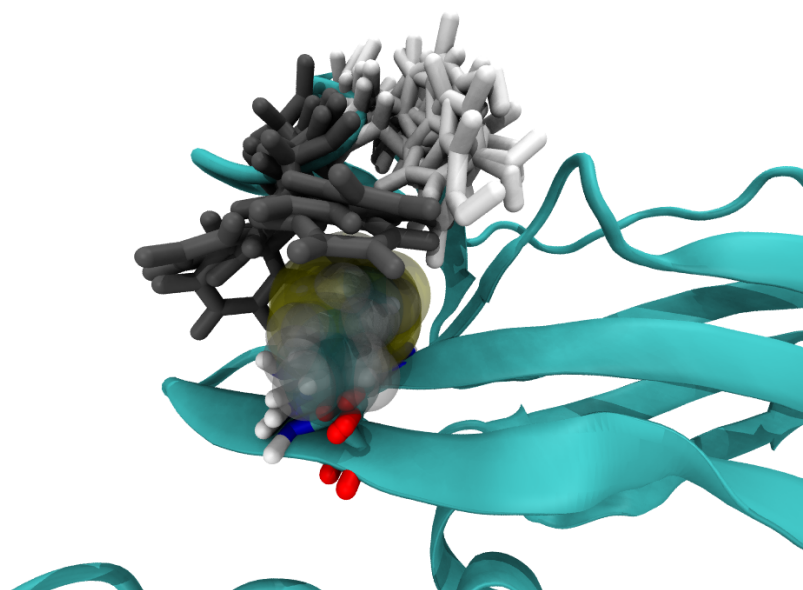
The model-free analysis and the measured RDCs do not directly support the cap loop opening in oxidised nDsbD, which is perhaps unsurprising considering that open conformations constitute only a small fraction of the ensemble (Fig. 3.21b). Encouragingly, TALOS-N predictions agree with a key differences between the simulation ensembles for oxidised and reduced nDsbD. The side chain of F70 is flexible in oxidised nDsbD which is likely important in cap loop opening. The simulations showed that the side chain of F70 switches between gauche- and trans in oxidised nDsbD while it is locked in the gauche- conformation in reduced nDsbD. TALOS-N predicts this difference for F70 of oxidised and reduced nDsbD from the chemical shifts alone (Shen and Bax, 2013).

Based on the simulations one can try to explain the differences in the cap loop and the active site of the two redox states. In oxidised nDsbD, the disulphide bond and specifically the S_{γ} of C109 provides a relatively flat surface for binding the phenyl ring of F70, as shown in Fig. 3.20b. F70 can thus interact with C109 both in its gauche- and trans orientations. Attractive interactions between the sulphur and the ring have been highlighted before (Goulding et al., 2002). The interaction will however be dominated by weak van der Waals forces. It is therefore not surprising that F70 can disengage from the sulphur, opening the cap loop. Open conformations are nonetheless sparsely populated (*c.f.* Fig. 3.21b) as two hydrophobic side chains become exposed to the solvent (Fig. 3.9b).

In reduced nDsbD, the interaction between F70 and C109 is different. The side chain of C109 forms an inclined plane for binding the phenyl ring (Fig. 3.20a). The thiols also



(a) Active site thiols and cap loop for *redI*



(b) Active site disulphide bond and cap loop for *ox8*

Fig. 3.20 To illustrate the differences in the active sites of oxidised and reduced nDsbD in two representative simulations, F70, Y71, C103 and C109 are shown for five and six equally space frames from the 200,000 frames trajectories for *redI* and *ox8* respectively. For oxidised nDsbD the van der Waals surface of the cysteine atoms is shown.

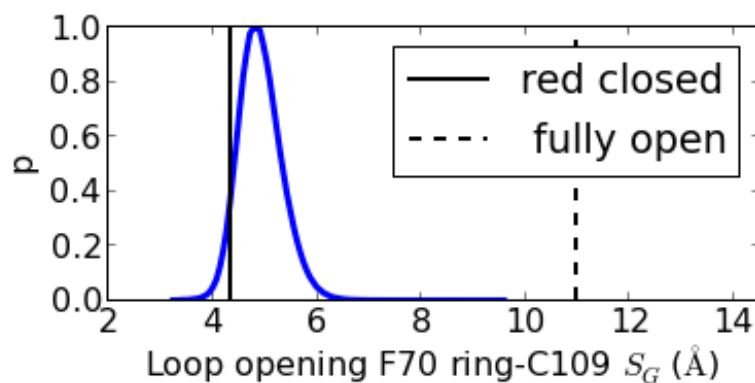
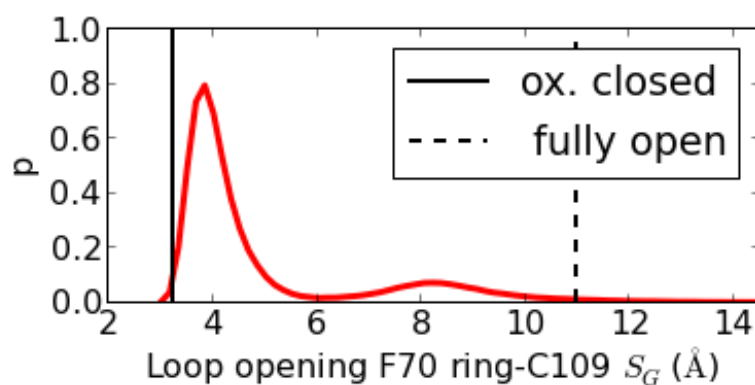
(a) Active site thiols and cap loop for *redI*(b) Active site thiols and cap loop for *redI*

Fig. 3.21 The conformational equilibrium of the cap loop as captured by F70-C109 distance. The simulation ensembles for reduced (a) and oxidised (b) nDsbD are compared.

occupy a larger volume than a disulphide bond (Mavridou et al., 2011). These structural features may contribute to locking F70 in a gauche- conformation close to the C_β of C109. In addition to van der Waals forces, there is a $C-H \cdots \pi$ hydrogen bond between H_β of C109 and the aromatic ring of F70. $C-H \cdots \pi$ hydrogen bonds are not considered to be strong but the bond might still contribute to stabilise the packing of the cap loop onto the active site.

Small differences in the positioning of F70 and Y71 were apparent from the crystal structure (Mavridou et al., 2011). Here I was able to go beyond static structures to show how details of the side chain packing of oxidised and reduced nDsbD might be important for its function as a redox hub (Stirnemann et al., 2006) by controlling loop opening.

The disulphide bond itself may also affect the dynamics of oxidised nDsbD. Indeed, the S^2 order parameters for G107 and F108, which are right next to the disulphide bond, are lower for oxidised than reduced nDsbD. The difference was recapitulated in our simulations. Disulphide bonds often have complicated μ s-ms dynamics and understanding them might be important given the interplay between the cap loop and the active site.

3.4.1 The dynamics of the cap loop and association reactions of nDsbD

The differences in the flexibility of the cap loop may explain why oxidised nDsbD has a higher affinity for cDsbD than the reduced protein. Cap loop opening is required for nDsbD to form its complexes. Loop opening in oxidised nDsbD, may therefore facilitate the binding of cDsbD which would happen in an induced-fit mechanism. Reduced nDsbD which provides reductant to thioredoxin fold proteins in *e.g.* the cytochrome *c* maturation and disulphide bond isomerisation pathways on the other hand does not show loop opening, shielding the thiols in the active site from oxidation. The cap loop of reduced nDsbD would only ever open upon binding its partners such as DsbC of the disulphide bond isomerisation and CcmG of the cytochrome *c* maturation pathway. Reduced nDsbD would bind its partners in an induced-fit mechanism.

3.4.2 Further experiments

The simulations emphasised the importance of rare events for understanding the function of nDsbD and I went onto study them with NMR relaxation dispersion experiments. Such experiments are especially interesting as they may provide a way to study the μ s-ms dynamics of the cap loop and the disulphide bond in oxidised nDsbD. The predictions from TALOS-N could be further verified by experimentally determining the F70 χ_1 angle by measuring three bond J-couplings.

Chapter 4

NMR relaxation dispersion experiments on the N-terminal domain of DsbD

4.1 Introduction

Computer simulations and analysis of chemical shift differences have detected differences between oxidised and reduced nDsbD, as discussed in the previous chapter. The differences between the redox states may explain why oxidised nDsbD forms a stronger interaction with cDsbD than reduced nDsbD. The redox-state dependence of the interaction affinity between cDsbD and nDsbD speeds up turnover in the *Dsb* pathway (Mavridou et al., 2011). In particular, the side-chain of F70 in the cap loop is flexible in oxidised but not reduced nDsbD. The cap loop opened spontaneously in MD simulations of oxidised nDsbD. In reduced nDsbD the cap stays closed at all times, protecting the active-site thiols from the oxidising periplasm.

Slower motions than the ones studied in the previous chapters may also play a role in the conformational equilibria of reduced and oxidised nDsbD. Cap loop opening in oxidised nDsbD is likely a rare event and might occur on a μ s-ms time-scale. In addition, the disulphide bond of oxidised nDsbD may undergo dynamics in this time window. Slow dynamics of the disulphide bond would explain the missing amide peaks for C103 and C109, which form the bond, and Y110 as well as the very weak peak for A104 as observed by HSQC spectra for oxidised nDsbD (Mavridou et al., 2012a). Such motions may play an important role in differentiating the two redox states and can be characterised in detail by NMR relaxation dispersion experiments.

4.1.1 Relaxation dispersion experiments to study slow dynamics

Motions on the μs -ms time-scale often exchange spins between distinct electronic environments with different resonance frequencies (Palmer et al., 2001). This is referred to as chemical exchange. Chemical exchange leads to decay of transverse magnetisation that for a two-state exchange process is described by,

$$\frac{d}{dt} \begin{bmatrix} M_A^+(t) \\ M_B^+(t) \end{bmatrix} = \begin{bmatrix} -i\Omega_A - R_{2A}^\circ - k_f & k_b \\ k_f & -i\Omega_B - R_{2B}^\circ - k_b \end{bmatrix} \begin{bmatrix} M_A^+(t) \\ M_B^+(t) \end{bmatrix}, \quad (2.2 \text{ revisited})$$

where Ω_A and Ω_B are the resonance frequencies, *i.e.* the chemical shifts, of the ground and excited states; R_{2A}° and R_{2B}° are the intrinsic relaxation rates of transverse magnetisation in the two states; p_A and p_B are their populations; while k_f and k_b are the forward and backwards rate constants for the reaction.

The decay of transverse magnetisation described by eqn. 2.2 can be reversed by spin-echos. However, the refocusing does not work when the spin jumps from one electronic environment to the other during the spin-echo (Baldwin and Kay, 2009). By applying many spin-echo pulses in short succession, exchange during the echo becomes less likely. The refocusing is enhanced and the observed relaxation rate $R_{2,eff}$ decreases. By measuring the relaxation rate $R_{2,eff}$ at varying rates of applying the spin echoes, ν_{CPMG} , the parameters in eqn. 2.2 are recovered. The experiments thus yield the kinetics and thermodynamics of the exchange process as well as the chemical shift differences between the exchanging states.

4.1.2 The structure of sparsely populated states

Relaxation dispersion experiments enable the structural characterisation of sparsely populated states. Structures that constitute only a very small fraction, as low of 0.5 %, of the ensemble can be accurately determined. Such states are "invisible" to most other experimental approaches and often referred to as "excited states", as they are somewhat higher in energy than the most-populated state (Baldwin and Kay, 2009).

The most important reporters on the structures of excited states are the chemical shifts from the relaxation dispersion experiments. In some cases, chemical shift differences between ground and excited states resemble the differences between the protein in distinct conditions, such as in absence or presence of a ligand. In this case the interpretation can be straightforward. Dihedral angles changes have been inferred from chemical shift differences between ground and excited states (Massi et al., 2005). In principle, if enough chemical

shift differences between the exchanging species can be measured, a structural model of the excited state can be calculated (Bouvignies et al., 2011). Here only amide nitrogen chemical shift differences are measured and no such model has been calculated.

Instead, computer simulations are used to gain some microscopic insight into measured chemical exchange processes. The chemical exchange processes probed by the experiments are very slow compared to the time-scales that are routinely accessible in atomistic MD simulations. Consequently, this approach has only been explored in a handful of cases, using either specialised hardware (Xue et al., 2012) or enhanced sampling methods (Palazzesi et al., 2013). The sampling of the process of interest can be enhanced by adding a bias to simulations. Here the metadynamics algorithm, that builds up a history-dependent bias is used to sample sparsely populated states detected by NMR. While the algorithm is well-established, its application to the interpretation of NMR relaxation dispersion experiments is novel (Palazzesi et al., 2013).

4.1.3 Free-energy landscapes from relaxation dispersion experiments

Experimentally one way of exploring the exchange reaction in more detail is by perturbing it, *e.g.* by changing the solvent conditions. A straightforward perturbation is changing the temperature. Repeating the experiments at multiple temperatures can be used to tease apart multi-step exchange processes (Grey et al., 2003). Moreover, the usual assumption that there is one single, global exchange process can be tested. Different motions are expected to have a divergent response to temperature. Overall, collecting relaxation dispersion experiments under different conditions improves the quality of the fits (Kovrigin et al., 2006). Finally, experiments at multiple temperatures give a detailed picture of the free-energy landscape for an exchange process (Mulder et al., 2001a).

What can be gained from an experimental characterisation of the free-energy landscape? First, the free energy difference between the ground and excited state, $\frac{p_B}{p_A} = e^{-\Delta G/RT}$, can be decomposed into an enthalpy and entropy difference. In principle, experimental ΔH measurements provide a stringent test for statistical mechanical models (Dill and Bromberg, 2011) such as MD simulations. Very long (>ms) MD simulations of protein folding have been assessed by comparison of the experimental folding enthalpies (Piana et al., 2014). Note however that accurate calculation of ΔH and ΔS from simulations requires even more sampling than calculating free-energies differences. While challenging (Chodera and Mobley, 2013), the enthalpy and entropy changes can sometimes be interpreted in the absence of expensive calculations. For example for BPTI, the thermodynamic changes between differ-

ent structures detected by relaxation dispersion experiments were rationalised based on the strain energy of the disulphide bond that underlies the conformational changes (Grey et al., 2003).

Secondly, experiments at multiple temperatures will also report on the temperature dependence of the rates and thus on the barrier between the reactant and product states. Again, experimentally determined barriers should in principle be reproduced by models of biomolecular conformational changes. However, the comparison to MD simulation is not straightforward; poor sampling in atomistic MD simulation is of course a severe limitation. Moreover, the height of experimentally determined free-energy barriers $\Delta G^\ddagger$ depends on the choice of the transmission coefficient κ which makes a direct comparison to computed barriers more difficult.

A qualitative interpretation of the experimental barriers can be nonetheless insightful and there is some interest in characterising the nature of the barriers between different protein conformations (Henzler-Wildman and Kern, 2007). An enthalpic barrier suggests that a specific interaction has to be broken during the reaction. An entropic barrier might mean that diffusion between the states itself and not the disruption of an interaction such as a hydrogen-bond or a salt-bridge limits the reaction rate (Dill and Bromberg, 2011). It has been hypothesised that barriers between the different polymer conformations are largely entropic. In contrast, many covalent reactions of small molecules are dominated by enthalpic barriers.

4.2 Aims

Here ^{15}N relaxation dispersion experiments are used to explore the μs -ms dynamics of both oxidised and reduced nDsbD. It will be interesting to see whether any sparsely-populated states are detected by relaxation dispersion experiments. Cap loop opening and dynamics of the disulphide bond might conceivably give rise to excited states. Analysing the chemical shift changes between ground and excited states will give an initial view as to the nature of alternative structures. The interpretation of the relaxation dispersion experiments will be guided by metadynamics free-energy calculations that can capture slow dynamics.

4.3 Materials and methods

4.3.1 NMR spectrometers

NMR experiments discussed in this chapter were recorded on home-built 500 and 750 MHz spectrometers. The spectrometers are equipped with Oxford Instruments Co magnets and home-built triple-resonance probes with triple axis gradients. The spectrometers are controlled by GE/Omega data acquisition computers and software.

4.3.2 Samples for NMR experiments

Dr Despoina Mavridou kindly provided two samples of oxidised nDsbD for NMR experiments. Sample concentrations were 0.3 mM and 0.6 mM, the pH was adjusted to 6.5 and 5 % D₂O was added. Reduced nDsbD was obtained by reduction with DTT, as previously established. Reduction of oxidised nDsbD by DTT is typically very slow, requiring around 12 h (Mavridou et al., 2011). The sample was typically incubated overnight to achieve full reduction, which was then confirmed by NMR.

4.3.3 Implementation of the ¹⁵N relaxation dispersion experiments

In relaxation dispersion experiments the relaxation rate is measured at different spin-echo repetition rates. The strength of this re-focusing field is determined by τ_{CPMG} , the spacing between the 180° re-focusing pulses:

$$v_{CPMG} = \frac{1}{4\tau_{CPMG}} \quad (4.1)$$

To measure relaxation dispersion for amide nitrogens a pulse sequence had to be implemented. The pulse sequence was implemented in the SEQUEL language on the GE Omega spectrometers. To aid de-bugging of the sequence, I have started out with a working pulse sequence for measuring the transverse relaxation of NH spin systems, similar to the one shown in Fig 4.1b (Millet et al., 2000). In Fig 4.1, the relaxation dispersion pulse sequence from the Palmer group is shown (Loria et al., 1999).

To convert the conventional sequence, two sets of changes were required. First, the J-coupled spin-echo after the transfer from the ¹H to ¹⁵N was removed. In the conventional T₂ sequence anti-phase magnetisation is converted to in-phase before the repeated Carr-Purcell-Meiboom-Gill (CPMG) spin-echos. In contrast, magnetisation is anti-phase $I_z S_y$ at the start of the CPMG train in the relaxation dispersion sequence. Secondly, a J-coupled spin

echo is introduced in the centre of the spin-echo elements, during the period U, highlighted in Fig. 4.1a. Here, $I_z S_y$ evolves into $-S_x$ in-phase magnetisation.

Averaging between in-phase and anti-phase magnetisation during the spin-echoes

The interconversion between in-phase and anti-phase magnetisation during U in the centre of the CPMG refocusing pulses is crucial (Loria et al., 1999). To emphasise the importance of the interconversion, it is instructive to consider a conventional T_2 experiment (Fig. 4.1b). During τ_{CPMG} in the spin-echo elements, magnetisation evolves due to scalar-coupling, interconverting in-phase and anti-phase magnetisation.

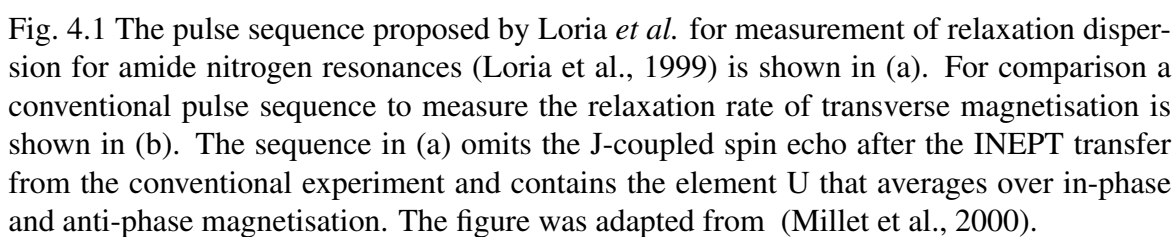
$$S_x \xrightarrow{2\pi J_{IS}\tau_{CPMG}I_z S_z} \cos(\pi J_{IS}\tau_{CPMG})S_x + \sin(\pi J_{IS}\tau_{CPMG})2I_z S_y \quad (4.2)$$

This evolution, is of course, reversed by the 180° pulse and the subsequent free-precession in the delay τ_{CPMG} . In-phase magnetisation is recovered.

$$\cos(\pi J_{IS}\tau_{CPMG})S_x + \sin(\pi J_{IS}\tau_{CPMG})2I_z S_y \xrightarrow{\pi S_x} \xrightarrow{2\pi J_{IS}\tau_{CPMG}I_z S_z} S_x \quad (4.3)$$

The relaxation rate of anti-phase magnetisation is however larger than for in-phase magnetisation. Increasing the value of τ_{CPMG} will increase the relaxation rate as the spin spends more time in the faster relaxing anti-phase state. Thus the effective relaxation measured for a particular value of τ_{CPMG} does not only depend on the extent of re-focusing of the signal broadened by chemical exchange.

The pulse sequence (Fig. 4.1a) breaks the correlation between τ_{CPMG} and the relaxation rate caused by scalar coupling (Loria et al., 1999). Before U anti-phase magnetisation $I_z S_y$ evolves to in-phase magnetisation $-S_x$ during τ_{CPMG} . The 180° pulse reverses the direction of precession and the spin returns to anti-phase at the end of the spin-echo at $2\tau_{CPMG}$. The J-coupled spin-echo, U, in the middle of the spin-echos converts anti-phase into in-phase magnetisation. After U, in-phase magnetisation evolves into anti-phase magnetisation, but of course the in-phase magnetisation is recovered at the end of each spin-echo. Increasing τ_{CPMG} then means that there is more in-phase magnetisation before U, but more anti-phase magnetisation after U. Overall, the spins spend one half of the time as anti-phase ($I_z S_y$) and the other half as in-phase magnetisation ($-S_x$) during the CPMG spin-echo elements in the sequence by Loria *et al.*. The change of $R_{2,eff}$ with ν_{CPMG} can now be interpreted confidently.



Constant length of the spin-echo element

The $R_{2,eff}$ for a given ν_{CPMG} frequency can be measured in two ways. First, a series of experiments can be recorded for a given τ_{CPMG} , varying the number of refocusing loops, n , in Fig. 4.1a). The relaxation rate is then obtained by following the decay in intensities as the number of refocusing loops increases as in a conventional R_2 measurement. This approach leads to very long measurement times and makes it difficult to measure $R_{2,eff}$ at many ν_{CPMG} frequencies. Good sampling of $R_{2,eff}$ ν_{CPMG} curves is important to accurately describe the exchange process.

Alternatively, only a single experiment is recorded for a given re-focusing frequency ν_{CPMG} (Mulder et al., 2001b). For a given ν_{CPMG} the intensity measured for a peak is converted into the relaxation rate $R_{2,eff}$ by eqn. 4.4,

$$R_{2,eff} = \frac{-1}{T_{CPMG}} \ln \frac{I_{\nu_{CPMG}}}{I_{ref}}. \quad (4.4)$$

Using this approach, T_{CPMG} is kept constant throughout the experimental series, which ensures that the $R_{2,eff}$ values measured throughout the series are comparable. I_{ref} in eqn. 4.4, is the intensity from a reference experiment. In the reference experiment the spin-echo period in Fig. 4.1a is omitted. The reduced accuracy in each $R_{2,eff}$ compared to collecting extensive data points for each ν_{CPMG} is compensated by the improved sampling of the ν_{CPMG} re-focusing frequencies and hence the approach by Mulder *et al.* was used here.

Temperature control in the relaxation dispersion experiments

In the relaxation dispersion experiments the rate of applying ^{15}N refocusing pulses is varied between experiments. Radio-frequency pulses heat up the NMR sample. The sample temperature can therefore differ between a spectrum recorded with a high and a low refocusing frequency. To ensure that the overall sample temperature does not change with ν_{CPMG} , compensatory spin-echoes were added to the pulse sequence, preceding the recycle delay. In experiments with low ν_{CPMG} frequencies, more heating pulse are applied than in experiments with larger ν_{CPMG} frequency. In Fig. 4.2a, spectra with refocusing field strengths ν_{CPMG} of 66.7 Hz and 800.0 Hz are overlayed. The peak positions are virtually identical which shows that the sample temperature was not significantly affected by the spin-echo pulses.

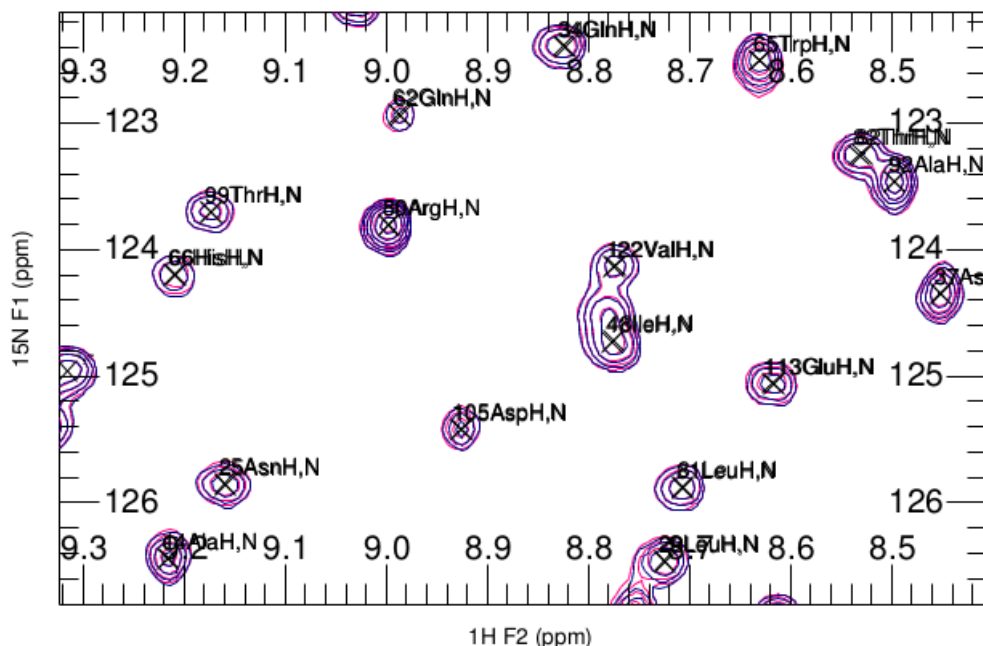


Fig. 4.2 The relaxation dispersion experiments recorded with refocusing frequencies of 66.67 Hz and 800 Hz are shown. The peak positions do not depend on the refocusing field strengths. Therefore, the temperature was kept constant during the experiments

4.3.4 Relaxation dispersion experiments at multiple field strengths and temperatures

A list of NMR relaxation dispersion experiments collected for nDsbD is given in Table. 4.1. Experiments were collected at five temperatures between 15 °C and 35 °C. Experiments at 15 °C and 25 °C were repeated at two magnetic field strengths, corresponding to ^1H frequencies of 500.1 MHz and 750.0 MHz.

Refocusing field strengths

Experiments with different refocusing frequencies were queued in a semi-random order. It typically took around three days to record an experimental series at a given combination of static magnetic field strength and temperature. It was found that refocusing frequencies up to 850 Hz could be routinely used on the Omega spectrometers. At higher refocusing frequencies the experimental series was occasionally stopped by the fail-safe mechanism in the probe and routine experiments were not possible. The fail-safe mechanism aborts experiments using amounts of radio-frequency power that may be damaging to the probe of the spectrometer.

Table 4.1 Relaxation dispersion experiments with oxidised nDsbD

T	^1H frequency	^{15}N frequency	T_{CPMG}
10 °C	500.1 MHz	50.7 MHz	40 ms
15 °C	500.1 MHz	50.7 MHz	40 ms
15 °C	750.0 MHz	76.0 MHz	40 ms
20 °C	500.1 MHz	50.7 MHz	40 ms
25 °C	500.1 MHz	50.7 MHz	30 ms
25 °C	750.0 MHz	76.0 MHz	30 ms
35 °C	500.1 MHz	50.7 MHz	40 ms

The length of the spin echoes was either 40 ms and 30 ms as detailed in Table 4.1. Shortening the spin-echo period improves signal to noise but increases the smallest possible ν_{CPMG} that can be accessed, which is a problem when studying slow processes. This is not a problem in the experiments at 25 °C with T_{CPMG} 30 ms as the smallest possible ν_{CPMG} was only slightly increased from 50 Hz to 66.7 Hz. Initial experimental curves already suggested a relatively fast process with curves levelling off around 500 Hz, which made it less important to sample low ν_{CPMG} values.

Details of data acquisition

The direct dimension F_2 was typically recorded at ^1H frequencies of 500.068 745 0 MHz and 749.954 165 0 MHz. 1024 complex data points were collected with a sweep width of 6289.31 Hz and 9433.96 Hz in the direct dimension. The indirect dimension F_1 was recorded at ^{15}N frequencies of 50.677 233 0 MHz and 76.000 755 0 MHz.

The indirect dimension was recorded with a sweep width of 1538.46 Hz (500.1 MHz) and 2304.15 Hz (750.0 MHz). The number of complex increments included in the final processing varied and was between 64 and 128. As the resonances of nDsbD are well resolved, spectral resolution was sacrificed to improve signal to noise, by reducing the number of increments processed. The recycle delay was 1.2 s and 1.5 s at 500.1 MHz and 750.0 MHz. Experiments were typically preceded by 64 dummy scans. The experiments were typically preceded by a dummy experiment that was not analysed. The dummy experiment enabled temperature equilibration.

4.3.5 Analysis of the relaxation dispersion experiments

The relaxation dispersion experiments were processed by standard methods using NMRPipe (Delaglio et al., 1995). The intensities of the amide resonances were followed in CCPNMR

Table 4.2 Analysis of the relaxation dispersion experiments with oxidised nDsbD

T	¹ H frequency	Residues included in multi-temperature fit
10 °C	500.1 MHz	G63, A64, W65, Y71, K73, S74
15 °C	500.1 MHz	G63, A64, W65, H66, D68, F70, Y71, G72, K73, S74, E75
15 °C	750.0 MHz	
20 °C	500.1 MHz	G63, A64, W65, H66, D68, E69, F70, Y71, G72, K73, S74, E75
25 °C	500.1 MHz	G63, A64, W65, H66, D68, E69, F70, Y71, G72, K73, S74, E75
25 °C	750.0 MHz	
35 °C	500.1 MHz	G63, A64, W65, E69, F70, Y71, K73, S74

Analysis (Vranken et al., 2005). The relaxation dispersion curves were fitted using CATIA (Hansen, 2008). In all cases it was assumed that a single exchange process gave rise to the relaxation dispersion curves. Furthermore the intrinsic relaxation rates R_{2A} and R_{2B} were assumed to be the same. This assumption is likely unjustified as the ground and excited state structures may be significantly different. However only a relatively small error is introduced by setting $R_{2A} = R_{2B}$ (Ishima and Torchia, 2006).

In addition the temperature dependence of p_B for a two-state reaction was taken to follow:

$$p_B = \frac{e^{-G_B/(RT)}}{Q} = \frac{e^{-G_B/(RT)}}{e^{-G_A/(RT)} + e^{-G_B/(RT)}} = \frac{e^{-(\Delta H - T\Delta S)/(RT)}}{1 + e^{-(\Delta H - T\Delta S)/(RT)}}, \quad (4.5)$$

where Q is the partition function, which is specified *via* the free energies for state A and B, G_A and G_B . ΔH and ΔS are the enthalpy and entropy differences for going from state A to state B. This neglects changes in heat capacity (Choy et al., 2005). To introduce heat capacity in the fitting, further experiments would be needed to prevent over-fitting.

It was assumed that the temperature dependence of the reaction rates can be described by Transition State Theory and the Eyring equation (Dill and Bromberg, 2011).

$$k = \kappa \left(\frac{k_B T}{h} \right) e^{-\Delta H^\ddagger / (k_B T)} e^{\Delta S^\ddagger / k_B} \quad (4.6)$$

$\Delta H^\ddagger$ and $\Delta S^\ddagger$ are the activation enthalpy and the activation entropy for the reaction. The transmission coefficient κ was set to $1.4 \times 10^{-7} \text{ s}^{-1} \text{ K}^{-1}$, the default value in CATIA and similar to $\kappa 1.6 \times 10^{-7} \text{ s}^{-1} \text{ K}^{-1}$ used in a previous relaxation dispersion study (Choy et al., 2005).

The chemical shift difference between ground and excited state for a given residue was allowed to change linearly with temperature in fitting with CATIA. An initial comparison of the chemical shifts obtained from fitting at 15 °C and 25 °C suggested that this is justified (*not shown*).

$$|\Delta ppm|(T) = |\Delta ppm| \times T + |\Delta ppm_{offset}| \quad (4.7)$$

4.3.6 Determining the sign for chemical shift differences between ground and excited states

The direction, that is the sign, of the chemical shift changes between ground and excited states can be extracted from comparing HSQC and HMQC spectra (Skrynnikov *et al.*, 2002). The peak positions in the nitrogen dimension in the HMQC experiment is subtracted from the peak position in the HSQC experiment:

$$\tilde{\omega}_N = HSQC - HMQC \quad (4.8)$$

Differences in ppm ($\tilde{\omega}_N$) are then converted into Hz (Ω_N). Note that this changes the sign of the differences as ^{15}N has a negative gyromagnetic ratio. A positive value of Ω_N in Hz indicates that the peak of the excited state is upfield of the ground state. The excited state peak is at a smaller ppm value than the ground state.

To accurately determine the sign of the chemical shift differences between ground and excited states, the HSQC and HMQC pulse sequences proposed by Skrynnikov *et al.* were employed. Using a standard HMQC sequence distorted peaks were detected in a trial experiment and thus the sequences from Skrynnikov *et al.* were implemented in SEQUEL on the GE/Omega consoles. In the modified HMQC pulse sequence by Skrynnikov *et al.* a selective REBURP-1 ^1H 180° pulse on the amide protons ensures absorptive line shapes. In both the HMQC and HSQC experiments, an e-snob soft pulse was used to suppress the water in the initial segment of the pulse sequence. In the HSQC experiment rectangular water gate pulses at the end of the pulse sequence in the reverse INEPT element are required for water suppression. A phase offset was required for the water gate pulses.

Sign determination experiments were recorded with a ^1H frequency of 749.954 1620 MHz and a ^{15}N frequency of 76.000 7550 MHz. In the direct dimension, a sweep width of 9433.96 Hz was used. The sweep width in the indirect dimension was 2304.15 Hz. 64 scans and 128 increments in the indirect dimension were collected. Each experiment was preceded by 16 dummy scans. Individual experiments took about 7.7 h to finish.

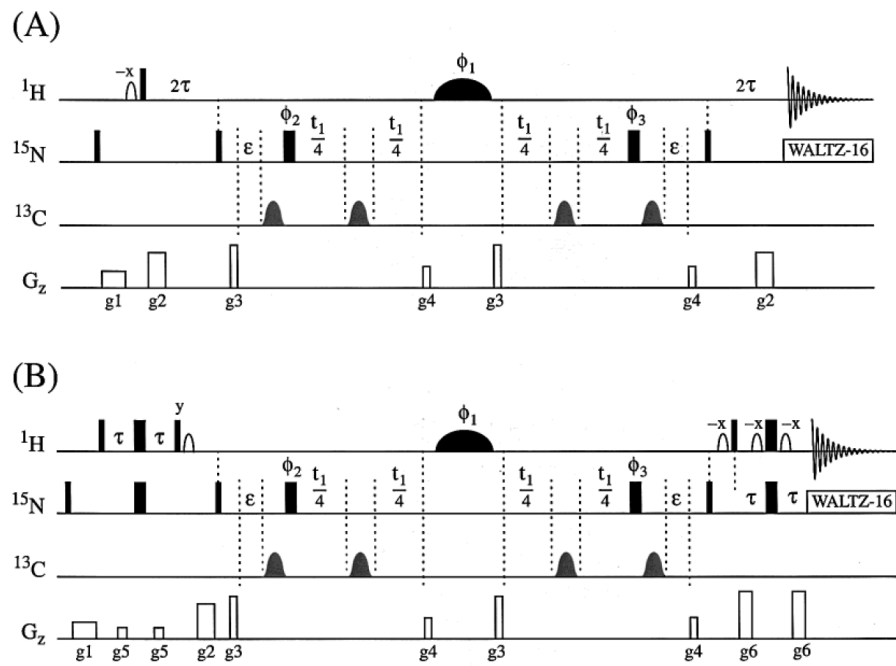


Fig. 4.3 The HMQC (a) and HSQC (b) pulse sequences for measuring the sign differences between ground and excited states. Note that the evolution periods are practically identical (Skrynnikov *et al.*, 2002) to facilitate the analysis. In a deviation from the implementation of Skrynnikov *et al.* no ^{13}C decoupling pulses were added to the sequence.

4.3.7 Metadynamics free energy calculations

Computer simulations enable calculations of free-energy surfaces. Metadynamics is a very flexible method to calculate a free-energy profile along a collective variable (Barducci et al., 2011). In metadynamics, a history-dependent potential speeds up the sampling and enables the simulation system to escape local minima. Gaussian potentials are added in the space defined by the collective variables. When the simulation returns to previously accessed regions of conformational space, the system experiences a repulsive force from the added bias potentials in addition to the forces calculated from the potential energy function. At long times, the sum of the Gaussian potentials will recover the inverse of the free-energy surface for the collective variables.

Here well-tempered metadynamics is used. In the well-tempered implementation of metadynamics (Barducci et al., 2008), the Gaussian height is rescaled as the simulation progresses. This reduces problems with the continuous build-up of the bias potentials which can lead to the sampling of unphysical high energy regions of phase space. Well-tempered metadynamics can converge to an accurate representation of a free-energy surface in a single simulation.

Here, the χ_1 side-chain dihedral angles for C103 and C109 were used as a collective variable. Gaussians were deposited every 2500 integration steps. The initial height of the Gaussian was 0.1 kJ mol^{-1} . A bias factor of 20 was used in the simulations. The width of the Gaussian was 0.2 rad. The metadynamics simulations were run using the PLUMED plugin (Bonomi et al., 2009) for Gromacs (Hess et al., 2008).

The simulations were setup as described in the preceding chapter. Briefly, the simulation system contained the 1L6P protein structure, water and Na^+ as well as Cl^- up to a salt concentration of 0.1 M. The box dimensions were chosen so that the minimum distance between the protein and the box edges was 1.5 nm. The CHARMM 22 potential energy function with the CMAP correction was used as in the preceding chapter (Bjelkmar et al., 2010). The production simulation at 298.15 K in the NPT ensemble was preceded by energy minimisation and equilibration using position restraints.

4.4 Results

4.4.1 Implementation and testing of the relaxation dispersion pulse sequence

A ^{15}N relaxation dispersion pulse sequence was implemented to characterise μs - ms dynamics of nDsbD. The association of cDsbD with nDsbD was used to test the implemented pulse sequence. cDsbD is thought to be quite rigid in absence of nDsbD. Thus one expects flat profiles of $R_{2,\text{eff}}$ plotted against the refocusing frequency ν_{CPMG} , *i.e.* no relaxation dispersion. Several peaks in the HSQC spectrum of oxidised cDsbD, such as G509, are broadened when adding a small amount of oxidised nDsbD (Mavridou et al., 2011). Broadening of G509 in the nitrogen dimension upon addition of cDsbD is shown in Fig. 4.4a. The peaks are broad since cDsbD exchanges between a small nDsbD-bound conformation and the dominant unbound state. In this instance a relaxation dispersion effect is expected. The predicted pattern was seen (Fig. 4.4b), in absence of nDsbD a flat profile was recorded for G509, but in presence of nDsbD, a strong relaxation dispersion effect was detected. From this it can be concluded that the pulse sequence is working as expected. Relaxation dispersion experiment to probe the association reaction cDsbD and nDsbD are described in more detail in Chapter 5.

4.4.2 Relaxation dispersion experiments with reduced nDsbD

For reduced nDsbD, no evidence for extensive μs - ms dynamics was detected. In particular no significant chemical exchange process was found for the cap loop region. Typical curves are shown in Fig. 4.5 for residues W65, E69 and F70 in the cap loop region of nDsbD. The curves are flat which suggests that there are no significant dynamics on the timescale that the relaxation dispersion experiment probes. For F22 which is in the DsbC binding site there is evidence for chemical exchange which is discussed in more detail in Appendix A.

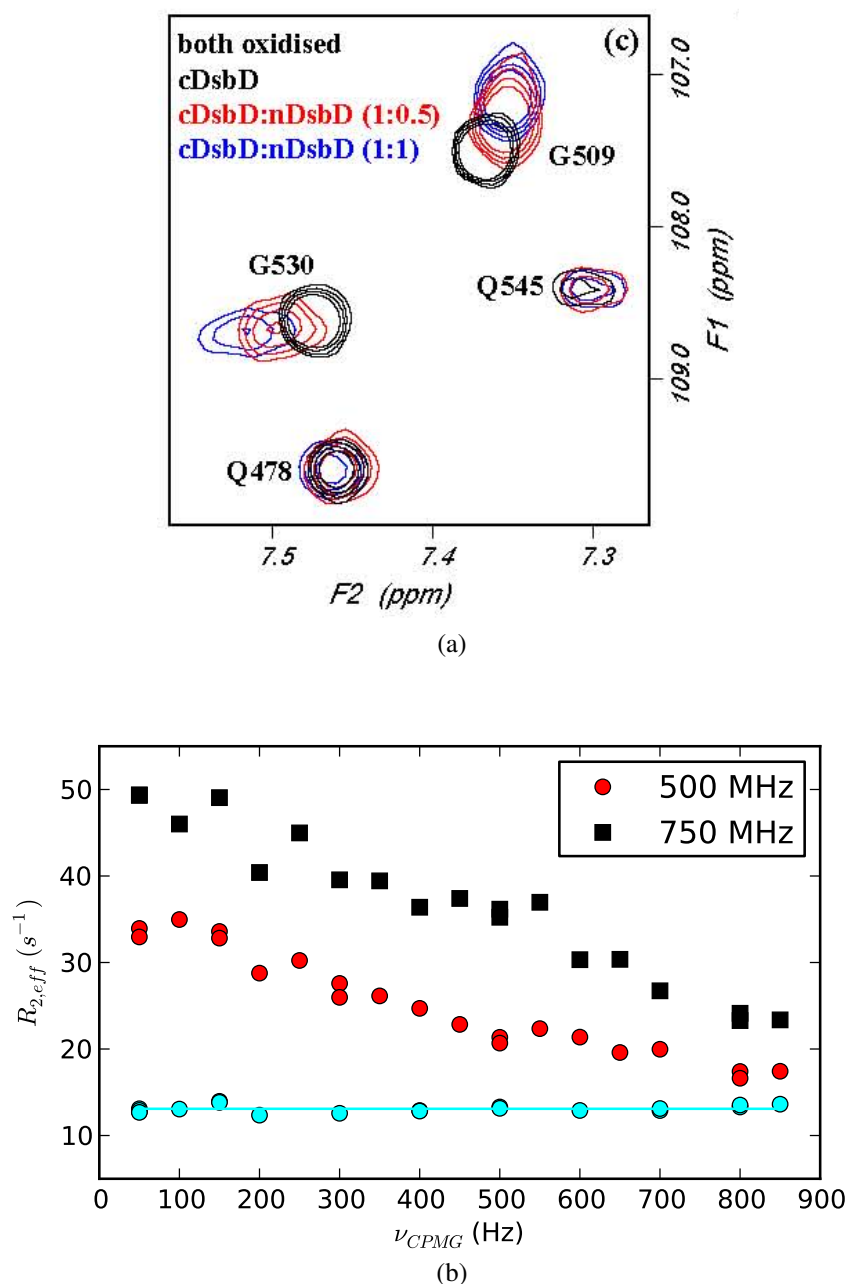


Fig. 4.4 G509 of cDsbD becomes broadened as nDsbD is added (a). The implemented pulse sequence was tested for ¹⁵N cDsbD in absence and presence of unlabelled nDsbD. G509 did not show a relaxation dispersion effect in the absence of nDsbD. Measurements in absence of nDsbD, at 500 MHz, are shown as cyan circles. The average relaxation rate in absence of nDsbD is plotted as well (also in cyan) to illustrate that the profile is flat. Relaxation rates recorded after the addition of nDsbD are shown as red circles (at 500 MHz) and black squares (at 750 MHz). Subfigure (a) was kindly provided by Prof. Christina Redfield.

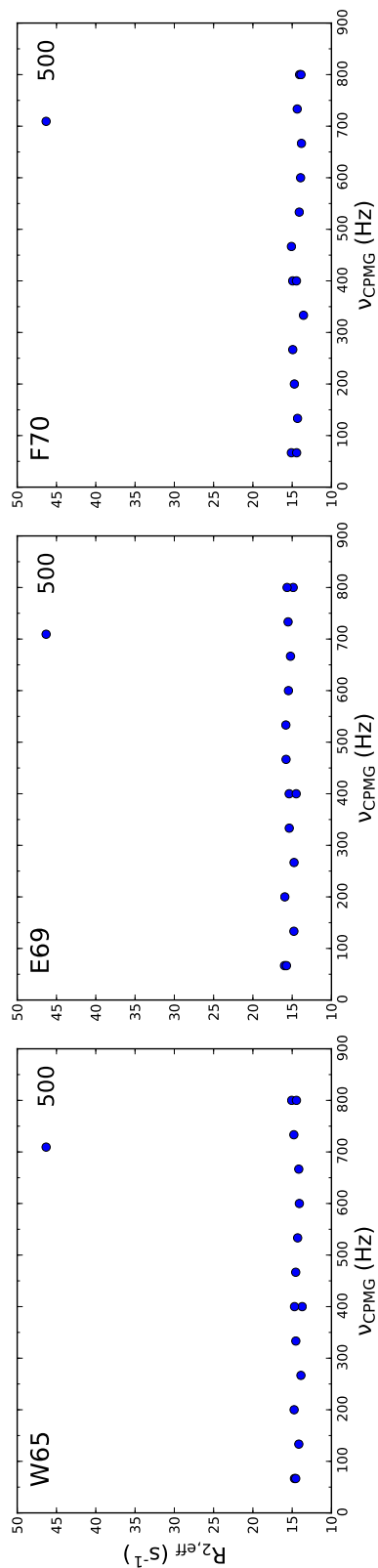


Fig. 4.5 Relaxation dispersion curves for cap loop residues in reduced nDsbD recorded at 25 °C.

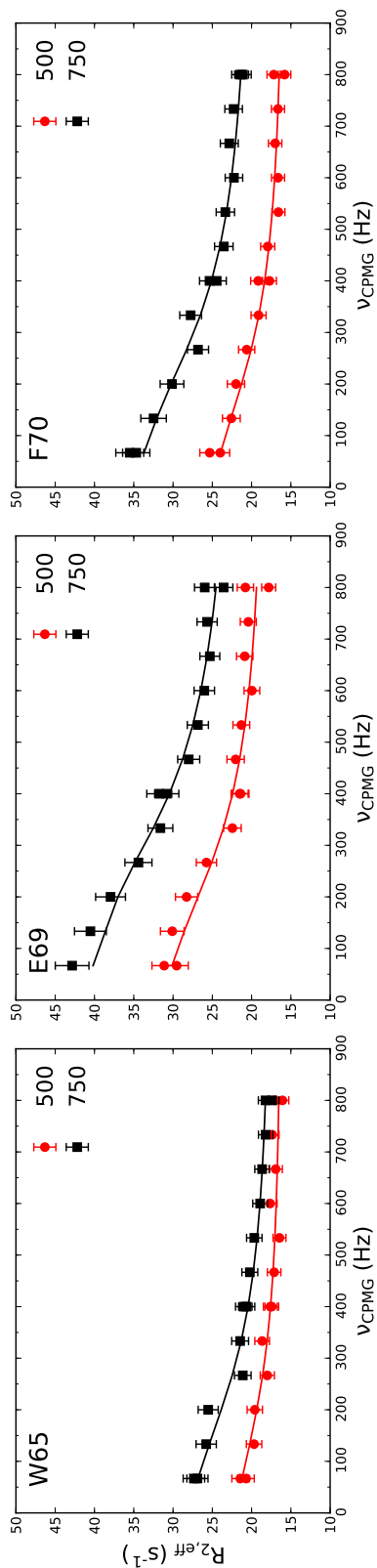


Fig. 4.6 Relaxation dispersion curves for cap loop residues in oxidised nDsbD recorded at 25 °C. *NB* the lines are from global fitting across multiple temperatures. Equally good fits were obtained fitting at just 25 °C (*not shown*).

4.4.3 Relaxation dispersion experiments with oxidised nDsbD

For oxidised nDsbD the relaxation dispersion experiments detected conformational exchange in the cap loop region. Eight residues A64, W65, E69, F70, Y71, K73 and S74 showed strong relaxation dispersion effects and are highlighted in Fig. 4.7. For these residues, as shown in Fig. 4.6 for E69, Y71 and K73, the effective relaxation rate drops as more spin-echo pulses are applied in the relaxation period T_{CPMG} , the signal is refocused. There is also a relaxation dispersion effect for F22 in the secondary DsbC binding site as described in more detail in Appendix A.

The experiment was repeated at higher magnetic field (^1H frequency of 750 MHz). Recording relaxation dispersion data for at least two magnetic field strengths is a prerequisite (Kovrigin et al., 2006) for accurate determination of the rates and populations for the exchange process as well as the chemical shift differences between the exchanging species. As expected the size of the effect increased at higher magnetic field strength, which can be seen by comparing the black (^1H frequency of 750.0 MHz) and red data (^1H frequency of 500.1 MHz) points on Fig. 4.6.

Relaxation dispersion experiments at four additional temperatures were collected for oxidised nDsbD. Data at multiple temperatures can further improve the confidence in the fit parameters. Different exchange process will differ in their response to changing the temperature (Grey et al., 2003). The additional experiments will thus provide a good test for the usual assumption that all exchanging residues are part of the same process. Experiments at multiple temperatures also help to distinguish between a two and a three-state exchange process.

Typical relaxation dispersion curves for the five temperatures are shown in Fig. 4.8 for W65, F70 and Y71. For W65 and Y71 good data were obtained at all temperatures. For some residues amide peaks were not resolved at all of the five temperatures. For F70, for instance, the amide peak could not be identified with confidence at 10 °C (Fig. 4.8b) due spectral overlap and dispersion curves were obtained at four temperatures for this residue.

The temperature dependence of the equilibrium constant (eqn. 4.5) and the reaction rates (eqn. 4.6) is also useful for characterising the energy landscape of the conformational exchange. Even before fitting the data, comparing the curves at 15 °C and 25 °C in Fig. 4.8 suggests that the exchange effect is more pronounced at 15 °C than at 25 °C. The process may be exothermic ($\Delta H < 0$).

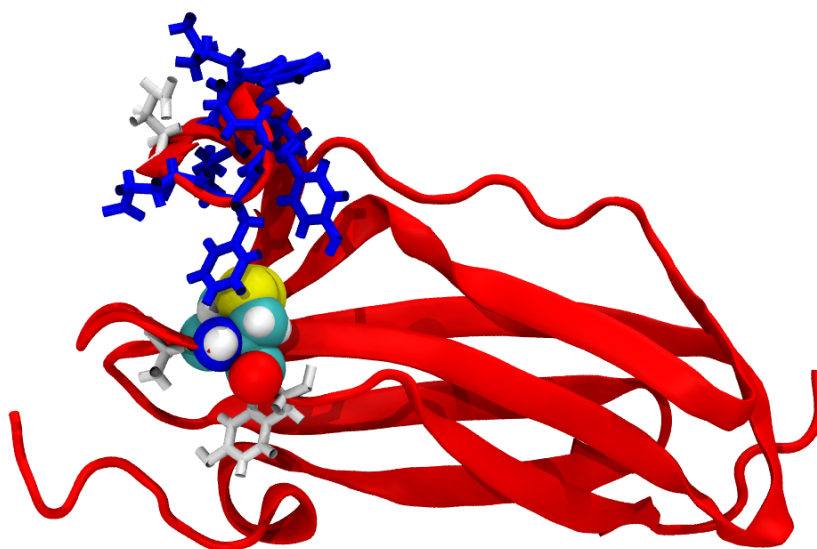
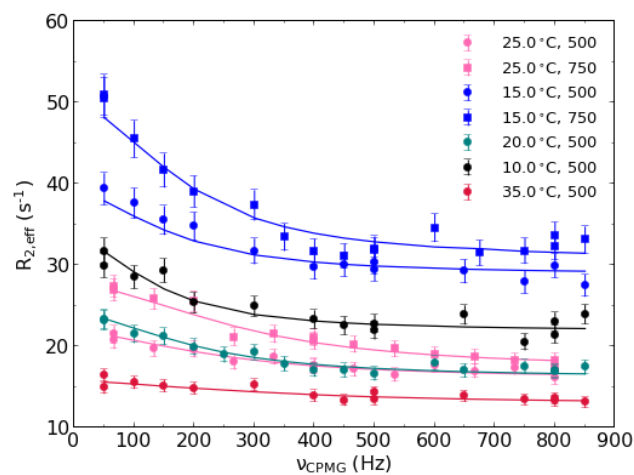
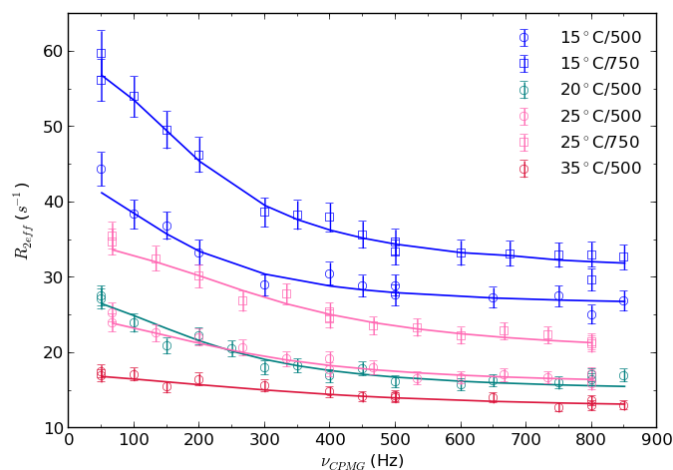


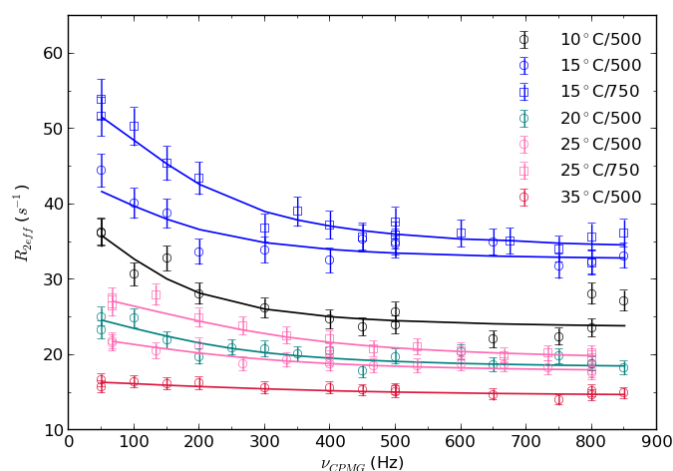
Fig. 4.7 Relaxation dispersion and invisible amide peaks in the active site and the cap loop of oxidised nDsbD. Residues A64, W65, E69, F70, Y71, K73 and S74 for which strong relaxation dispersion was detected for their amide nitrogens are shown in blue. Amide peaks that are likely broadened beyond detection (E67 and Y110) or are extremely weak (A104) are shown in white. The disulphide bond is shown in van der Waals representation.



(a) W65



(b) F70



(c) Y71

Fig. 4.8 Example curves to illustrate the fitting across experimental conditions are shown. Relaxation dispersion curves collected at 10 °C, 15 °C, 20 °C, 25 °C and 35 °C were fitted simultaneously using CATIA.

Table 4.3 Fit parameters for relaxation dispersion experiments of nDsbD

	15 °C	15 °C (five temperatures)	25 C	25 C (five temperatures)
k_A	40 s ⁻¹	60 s ⁻¹	30 s ⁻¹	30 s ⁻¹
k_B	820 s ⁻¹	950 s ⁻¹	1360 s ⁻¹	1430 s ⁻¹
k_{ex}	860 s ⁻¹	1010 s ⁻¹	1390 s ⁻¹	1460 s ⁻¹
p_B	4.9 %	5.7 %	2 %	2 %

4.4.4 Analysis of the relaxation dispersion experiments for oxidised nDsbD

Fits for the relaxation dispersion curves

Relaxation dispersion data were analysed using CATIA (Hansen, 2008). It was assumed that all eight residues in the cap loop that showed relaxation dispersion are part of the same, global exchange process. F22 which appeared to have a relaxation dispersion effect in both oxidised and reduced nDsbD and is far removed from the cap loop region was not included. Likely, it undergoes different dynamics than the cap loop region. The experiments recorded at 25 °C at two magnetic field strengths gave good fits, suggesting that a single, global exchange process largely accounts for the experimental curves. Transitions to the excited state are rare with (forward) rate constant $k_A = 30 \text{ s}^{-1}$ (Table 4.3). The rate for the return to the ground state is much larger, $k_B = 1360 \text{ s}^{-1}$. Only about 2 % of oxidised nDsbD molecules are in the alternative, excited conformation.

Good fits were also obtained for the experiments collected at 15 °C. Fitting the data recorded at 15 °C confirms the conclusion from comparing the curves recorded at different temperatures that the reaction is exothermic. The population of the excited state increased to about 5 %.

Multi-temperature fits of the relaxation dispersion experiments with oxidised nDsbD

Data collected at five different temperatures could be fitted together. The populations of the excited and the ground state were assumed to vary with temperature according to eqn. 4.5. The temperature dependence of the reaction rate was described using the Eyring equation (eqn. 4.6). The chemical shift difference between ground and excited states for each residue was assumed to vary linearly (eqn. 4.7). All residues included in the fitting (Table 4.2), 12 in total, were assumed to be part of the same exchange process. The residues included in the fits were all in the cap loop region. F22 which appears to have a relaxation dispersion effect but is far removed from the cap loop region was not included.

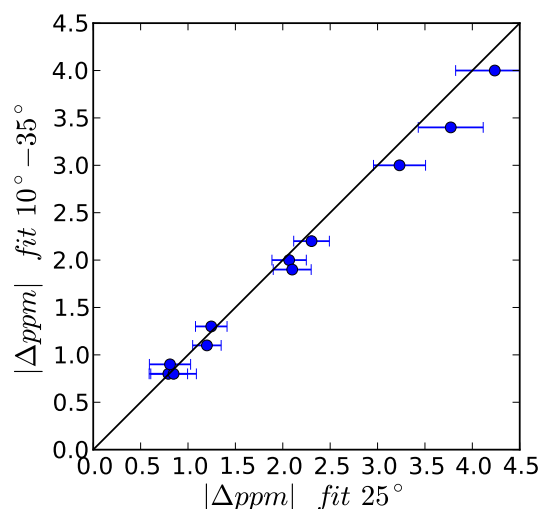


Fig. 4.9 Very similar absolute chemical shift differences were obtained when fitting data just at 25 °C and at five different temperatures from 10 °C to 35 °C.

Overall, good fits were obtained for all residues in the cap loop region as evidenced by Figs. 4.6 and 4.8. Thus there is no need to evoke more than a single global exchange process. Residues that are part of different exchange process are expected to show poorer fits especially when fitting data collected at multiple temperatures. Moreover, a two-state mechanism appears to describe the data well, there is no evidence for a more complicated three-state exchange process which has previously been detected using relaxation dispersion experiments at multiple temperatures (Grey et al., 2003).

To validate the fits across five temperatures, the fit parameters were compared with ones extracted at 15 °C or 25 °C. The k_{ex} value at 15C differs by around 200 s^{-1} between the two approaches. The difference is somewhat higher than expected. Both approaches gave a relatively large population of the excited state at 15C (4.9 and 5.9). The results from fits at five temperatures were also compared with the fits at 25 °C. At 25 °C very similar exchange rates, k_{ex} , of 1390 s^{-1} and 1460 s^{-1} were obtained in the two approaches. Similarly, p_B was 2% at 25 °C in both cases. The chemical shift differences at 25 °C also agree extremely well when comparing the fits, as shown in Fig. 4.9. Overall the agreement is good which shows that the fitting of experiments collected at five temperatures worked as expected.

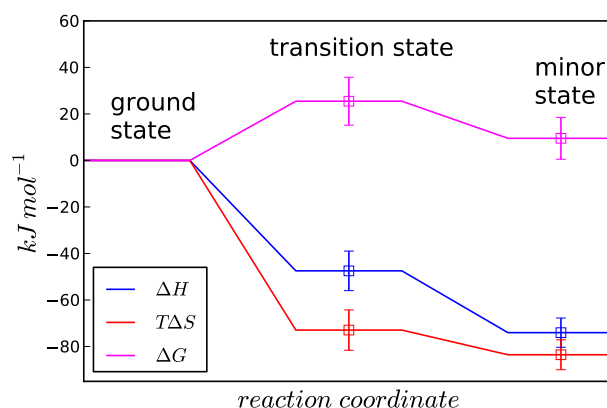


Fig. 4.10 A 1D energy landscape is obtained from global fitting of the relaxation dispersion experiments recorded at temperatures between 10 °C and 35 °C. The exact value of $\Delta S^\ddagger$ and thus $\Delta G^\ddagger$ depend on the choice of transmission coefficient κ . Here κ was set to 1.6×10^{-7} which is close previously used values (Choy et al., 2005).

The experimental free-energy landscape

The fits at five different temperatures yield a description of the energy landscape of the exchange process. The excited state of oxidised nDsbD is enthalpically more favourable than the ground state according to the fits of the relaxation dispersion curves at five temperatures. This agrees with visual inspection of the curves in Fig. 4.8 and the conclusion from comparing the fits obtained at 15 °C and 25 °C. The excited state is entropically less favourable than the ground state and overall it is about 10 kJ mol^{-1} less favourable (Fig. 4.10).

The fitting of the experimental curves at the five temperatures also reports on the barrier between the two states. The transition state appears to be more enthalpically favourable than the ground state. There is however a sizable free energy barrier as the transition state is lower in entropy than the ground state. Note that $\Delta S^\ddagger$ and hence $\Delta G^\ddagger$ depend on the transmission coefficient κ that is used in eqn. 4.6. However, the absence of a significant enthalpic barrier was already suggested by comparing the reaction rates obtained at 15 °C and 25 °C (Table 4.3) when fitting data at only a single temperature. There is no obvious rate increase in going from 15 °C, with $k_A = 40 \text{ s}^{-1}$, and 25 °C, with $k_A = 30 \text{ s}^{-1}$.

Relaxation dispersion experiments at different sample concentrations

Relaxation dispersion experiments for oxidised nDsbD were repeated with a second sample at a lower sample concentration to test whether the observed chemical exchange is concentration dependent. Establishing whether the conformational exchange is concentration

dependent will provide important clues for identifying the nature of the process. For instance, one would expect a strong concentration dependence if aggregation of nDsbD gives rise to the measured relaxation dispersion profiles.

Typical relaxation dispersion profiles at the two sample concentrations are compared in Fig. 4.11. For E69 (Fig. 4.11d), which shows a strong dispersion effect, and A64 (Fig. 4.11b) very similar curves are obtained at the two concentrations. The curves also show clearly that experiments at lower concentrations give noisier $R_{2,eff}$ curves. A somewhat different trend is seen for W65 (Fig. 4.11c), F70 (Fig. 4.11e) and S74 (Fig. 4.11e). It looks as if lowering the concentration shifts the curves to lower $R_{2,eff}$ values. At lower sample concentrations, viscosity will be reduced and this could explain the shift towards lower $R_{2,eff}$ values. On the whole, there was not a strong concentration dependence of the relaxation dispersion curves for oxidised nDsbD, suggesting aggregation is not involved.

4.4.5 Analysis of the chemical shift changes associated with the exchange process

Absolute chemical shift changes

The nitrogen chemical shift differences obtained from the fitting of the relaxation dispersion experiments provided an initial view of the sparsely populated alternative state, that is the excited state, of oxidised nDsbD. It is important to note that the fits only yield the magnitude of the chemical shift changes between the ground state and the excited state. The absolute chemical shift differences can nevertheless be interpreted.

In principle one could use the chemical shift differences to test whether the excited state could correspond to an open conformation. The errors in predicting ^{15}N chemical shifts are however similar in magnitude to the shift differences between ground and excited states. Thus the open conformation from simulations (in the previous chapter) and the 1VRS crystal structure were not compared to the excited state chemical shifts.

First the possibility that the excited state corresponds to an unfolded conformation was tested. The chemical shift changes expected for converting the cap loop region into a random-coil were calculated using a program previously written by Prof Redfield. The calculation took the sequence dependence of the random coil shifts into account. The lack of correlation in Fig. 4.12a demonstrates that the cap loop does not lose its structure in the lowly populated alternative state.

The chemical shift changes between the ground and excited state structures of oxidised nDsbD resemble the chemical shift changes between oxidised and reduced nDsbD (Fig.

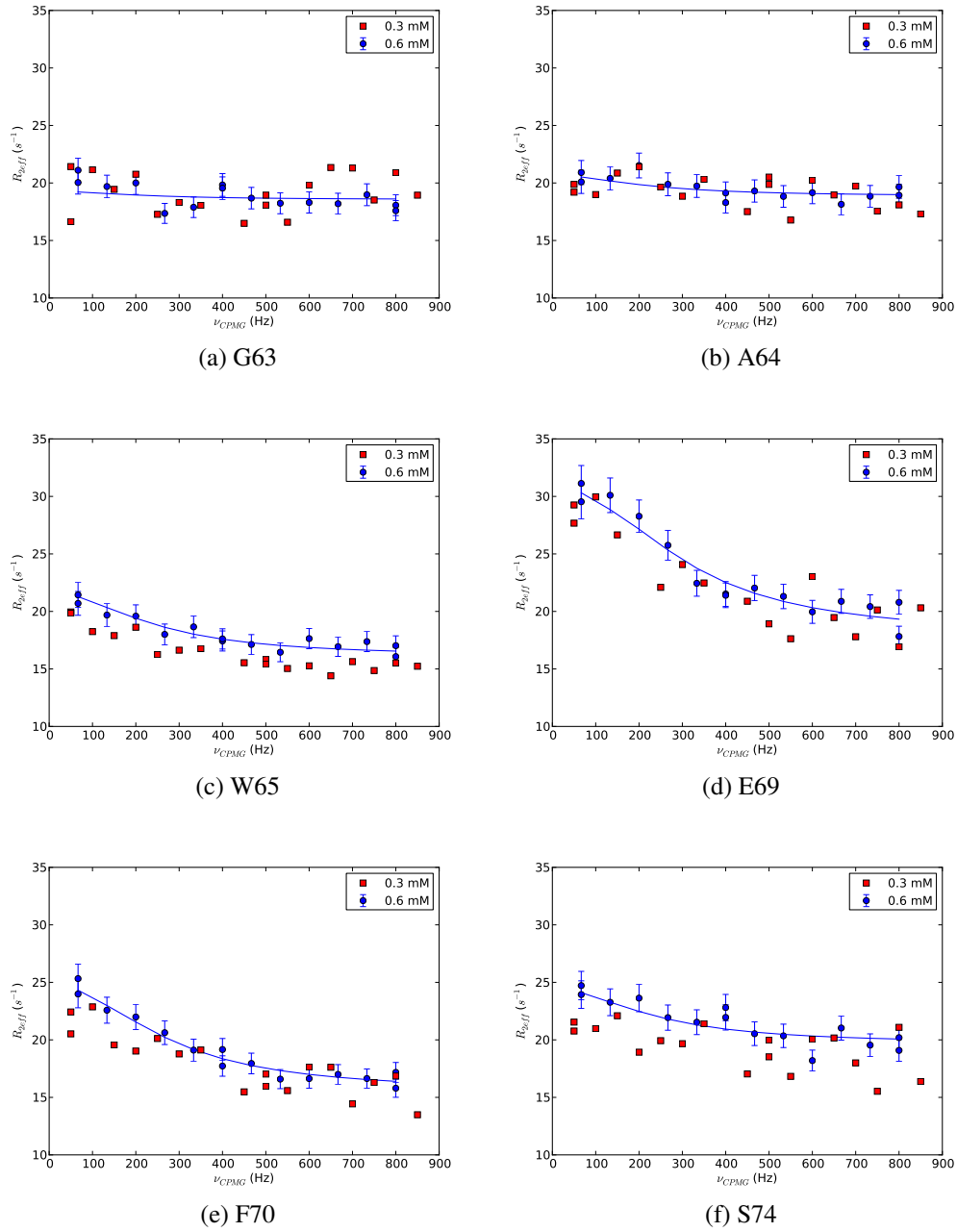


Fig. 4.11 Relaxation dispersion experiments of oxidised nDsbD with sample concentrations of 0.3 mM and 0.6 mM are compared.

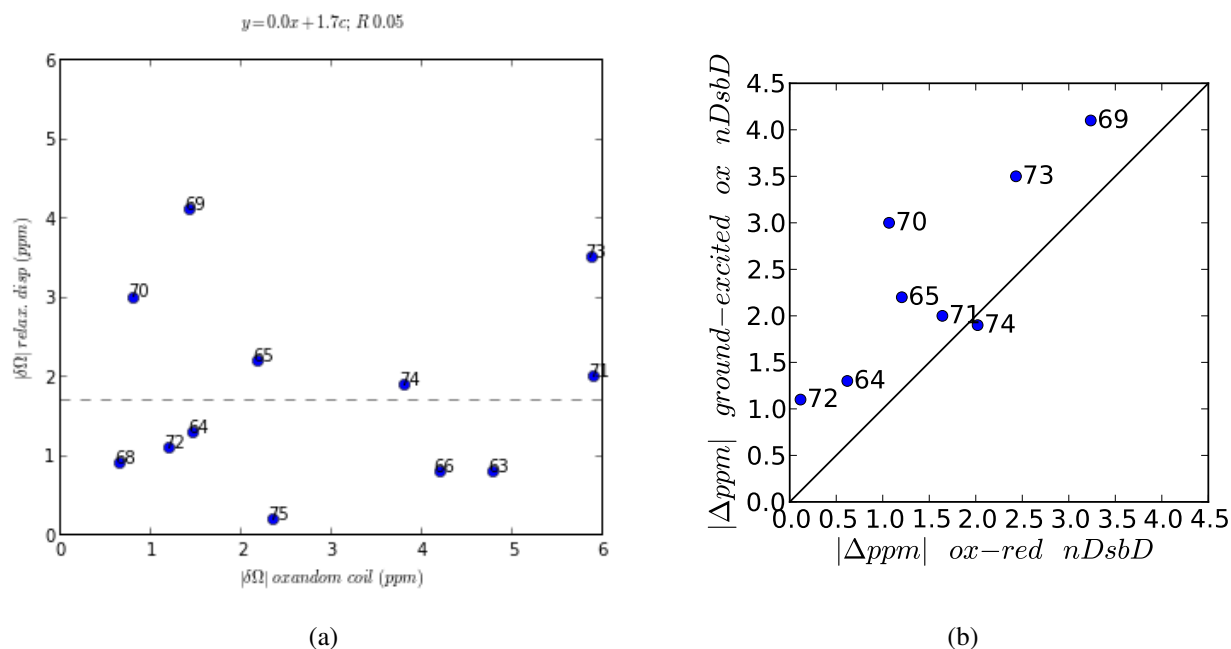


Fig. 4.12 The absolute chemical shift changes between the ground and excited state of oxidised nDsbD are compared with the expected changes for local unfolding (a) and differences between reduced and oxidised nDsbD (b).

4.12b). The correlation is not perfect, for example the chemical shift of F70 changes by 3 ppm between the two states of oxidised nDsbD. The nitrogen shift for this residue differs by about 1.1 ppm between the two redox isoforms of nDsbD. The changes between ground and excited state are general larger than the differences between the two redox states. Nonetheless, the correlation shown in Fig. 4.12b suggests that the structure of the excited state of oxidised nDsbD would be somewhat similar to reduced nDsbD. Before drawing any further conclusions, an attempt was made to determine the sign of chemical shift differences, *i.e.* the direction, of the chemical shift changes between the ground and excited states of oxidised nDsbD.

Determining the sign of the chemical shift changes

The absolute chemical shift differences between the ground and excited states of oxidised nDsbD can be complemented by determining the missing sign information in HSQC-HMQC experiments (Skrynnikov et al., 2002). Chemical exchange affects the peak positions. The peaks in HMQC and HSQC spectra are differentially affected by chemical exchange which allows recovering the sign for the chemical shift difference. The sign can be determined with confidence when the difference between the HSQC and the HMQC experiment, $|\Omega_N|$,

Table 4.4 ^{15}N chemical shifts in HSQC and HMQC experiments. The sign information could be extracted with confidence for E69, Y71 and K73.

	HSQC	HMQC	$\tilde{\Omega}_{Nexp.}(\text{ppm})$	sign $\Omega_N(\text{Hz})$
A64	117.491	117.495	-0.004	[+ve]
W65	122.508	122.506	+0.002	[+ve]
E69	118.900	118.931	-0.031	+ve
F70	120.811	120.810	+0.001	[-ve]
Y71	116.268	116.273	-0.005	+ve
K73	126.896	126.873	+0.023	-ve

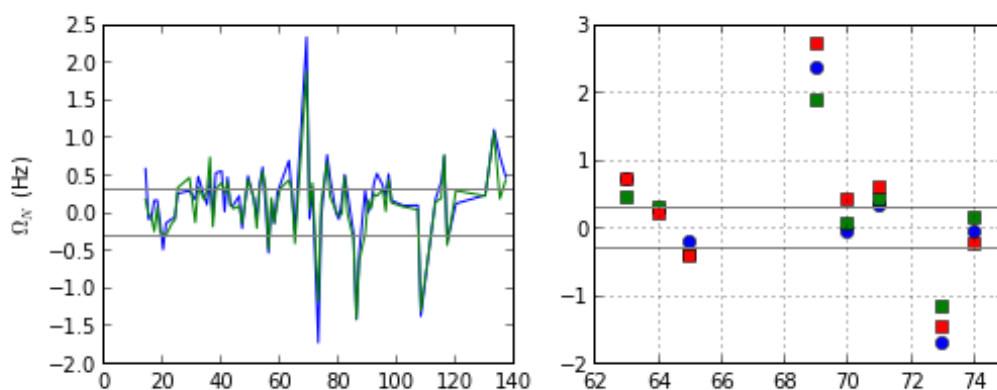


Fig. 4.13 The left panel shows the differences in the ^{15}N resonance frequencies between the ground and excited states of oxidised nDsbD as determined by a pair of HSQC-HMQC. The results across the sequence of nDsbD are shown for two repeat sets of experiments in blue and green. The panel on the right shows the results for three pairs of HSQC-HMQCs experiments for the cap loop region.

is $> 0.3\text{Hz}$. As shown in Fig. 4.13a the two largest differences in peak positions between the HSQC and HMQC spectra, Ω_N , are seen for E69 and K73 in the cap loop. For E69 and K73 we find $|\Omega_N|$, is $> 0.3\text{Hz}$, while for Y71 the $|\Omega_N|$ is close to the cut-off value of 0.3Hz . The positive $|\Omega_N|$ for E69 and Y71 (Table 4.4) shows that their excited state peaks are upfield of the ground state peaks. The excited state peaks are at lower ppm value than the ground state resonances (Fig. 4.14a). Similarly, the peaks for E69 and Y71 in reduced nDsbD have smaller ^{15}N chemical shift than in the oxidised protein. For K73, a negative $|\Omega_N|$ was found. The excited state peak for K73 has larger ppm value than in the ground state. The ^{15}N resonance frequency is at a larger ppm value in reduced than oxidised nDsbD. The results for these three residues together with absolute chemical shift differences suggest that the excited state structure of oxidised nDsbD indeed somewhat resembles the reduced isoform.

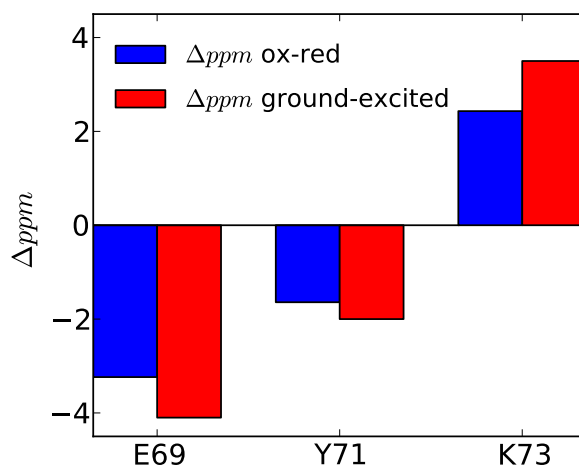


Fig. 4.14 The chemical shift differences for E69, Y71 and K73 between the ground and excited states of oxidised nDsbD are compared to the difference in shift between oxidised and reduced nDsbD

To test the conclusions, two additional pairs of HSQC and HMQC spectra were recorded. For E69, Y71 and K73, all three pairs of HSQC and HMQC experiments agree with each other (Fig. 4.13). The sign determination confirmed the conclusion that the excited state of oxidised nDsbD somewhat resembles reduced nDsbD. In reduced nDsbD the cap loop packs stably onto the active site, as demonstrated by the MD simulations. The "reduced-like" excited state of oxidised nDsbD is unlikely to feature an open cap loop.

4.4.6 The dynamics of the disulphide bond in computer simulations

The dynamics of the disulphide bond in molecular dynamics simulations

The nature of the excited state of oxidised nDsbD is not clear. Computer simulations were used to test whether slow dynamics of the disulphide bond could underlie the equilibrium detected by the relaxation dispersion experiments. In oxidised nDsbD peaks for C103, C109, Y110 are missing and the peak for A104 is very weak. The peaks are missing or weak most likely due to extreme broadening which points towards μ s-ms dynamics of the disulphide bond.

A conformational change involving the disulphide bond would be consistent with the excited state of oxidised nDsbD resembling the reduced protein. The cysteine side-chains adopt different structures in reduced and oxidised nDsbD and this can be seen by analysing the MD simulations presented in the preceding chapter. For reduced nDsbD the most com-

mon χ_1 side-chain dihedrals for both C103 and C109 are close to $-\pi$ (Fig. 4.15a). For oxidised nDsbD, the C103 and C109 χ_1 side chain dihedral angles both cluster around $-\frac{2\pi}{5}$ (Fig. 4.15b). The cysteine side-chains might adopt a "reduced-like" conformation in the excited state of oxidised nDsbD, that would favour close packing of the cap loop onto the active site. This would explain the "reduced-like" excited state shifts of the cap loop.

Exploring the free-energy landscape of the disulphide bond by metadynamics

The existence of "reduced-like" ways of arranging the cysteine side-chains in the active site of oxidised nDsbD was tested using metadynamics. In metadynamics, free-energy surfaces along a variable of interest, referred to as a collective variable, can be calculated. Collective variables are conceptually similar to the geometric order parameters that were used to distinguish different conformational states.

A free energy surface for oxidised nDsbD was calculated using the cysteine C103 and C109 χ_1 side chain dihedral angles as collective variables (Fig. 4.16a). To initially assess the calculated free energy surface it is instructive to compare it to the equilibrium MD simulations. Highly populated states in the equilibrium MD simulations of oxidised nDsbD (Fig. 4.15b) are expected to correspond to minima on the free energy surface calculated by metadynamics (Fig. 4.16a). Two of the additional minima of the metadynamics free-energy surface were within the region sampled in the equilibrium simulations, one with C103 $\chi_1 > -\frac{\pi}{2}$ and C109 $\chi_1 > -\pi$, the other with C103 $\chi_1 > -\pi$ and C109 $\chi_1 > -\frac{\pi}{2}$.

The qualitative agreement between equilibrium MD and the free-energy surface from metadynamics was encouraging and the sampled C103 and C109 χ_1 dihedral angles were compared to reduced nDsbD (Fig. 4.15a). Indeed the free energy surface for oxidised nDsbD contains "reduced-like" states. For example, there is a state with the χ_1 of C103 close to $-\pi$ and C109 $0 < \chi_1 < \frac{\pi}{2}$ in the metadynamics simulations and the equilibrium MD of reduced nDsbD. Importantly, a state similar to the most populated conformation in the MD simulations of reduced nDsbD, with χ_1 of both C103 and C109 close to $-\pi$, was found. The crystal structure of reduced nDsbD maps onto this region on dihedral angle surface, as indicated by the purple marker on Fig. 4.16a. This might be the "reduced-like" structure discovered in the relaxation dispersion experiments.

4.4.7 Convergence of the metadynamics free-energy calculations

Before analysing the reduced-like conformation accessed in the metadynamics simulations of oxidised nDsbD in more detail it is important to critically assess the calculation. The

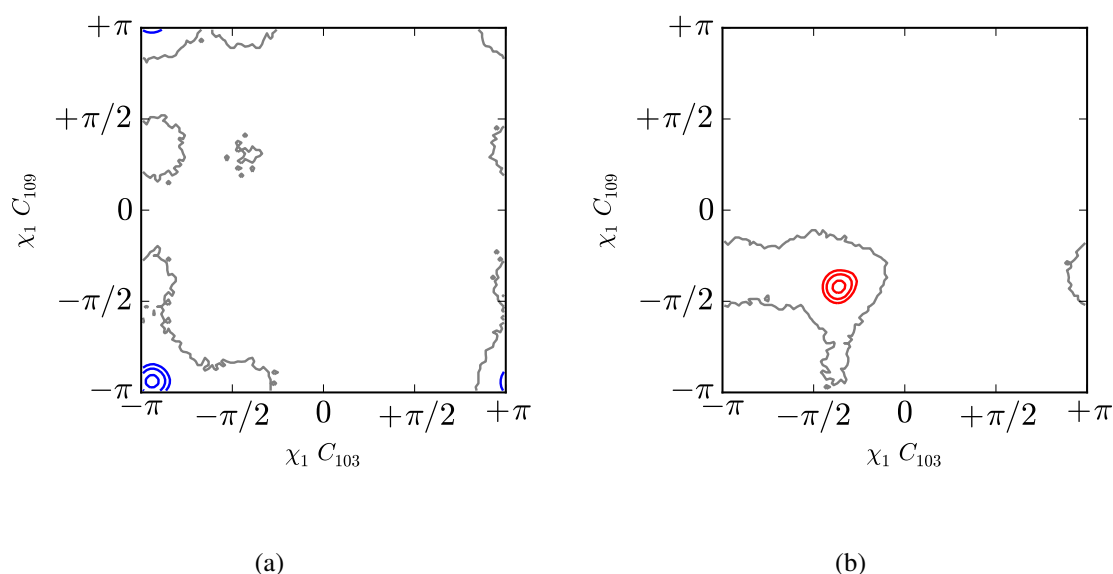


Fig. 4.15 The side-chains of C103 and C109 in the active site of nDsbD adopt different conformations in the reduced (a) and oxidised (b) isoforms as shown by contour plots of their χ_1 dihedral angles. Here 1 μ s of MD simulations are analysed for each redox-state.

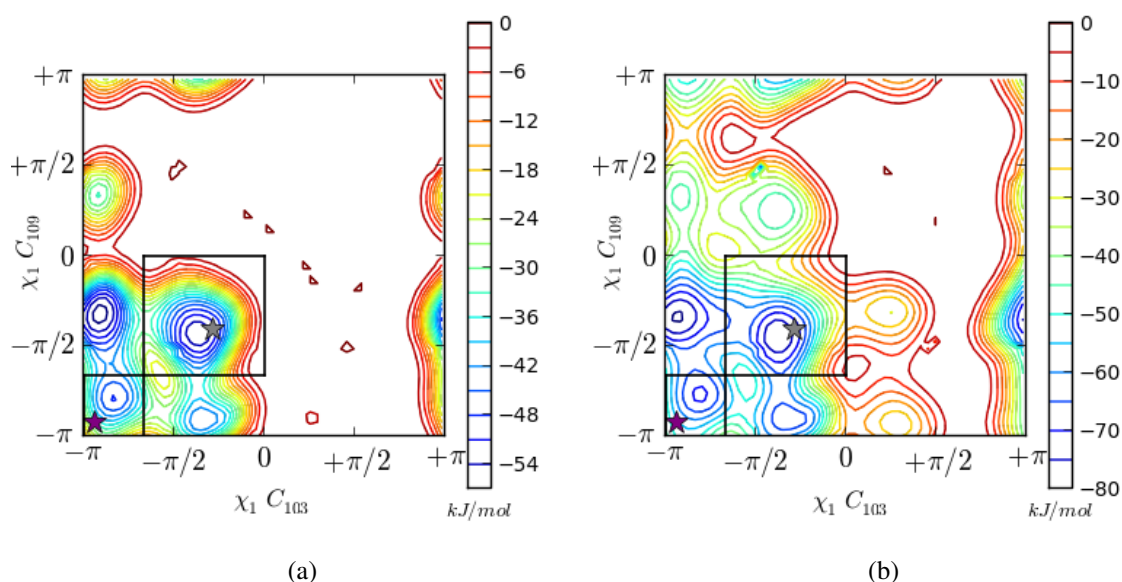


Fig. 4.16 A two-dimensional free-energy surface for the χ_1 dihedral angles of C103 and C109 was calculated using metadynamics. The free-energy surface after 100 ns and 300 ns of simulation time is shown in (a) and (b). The purple and grey stars indicate crystal structures of reduced (3PFU) and oxidised (1L6P) nDsbD on the two-dimensional free-energy surface

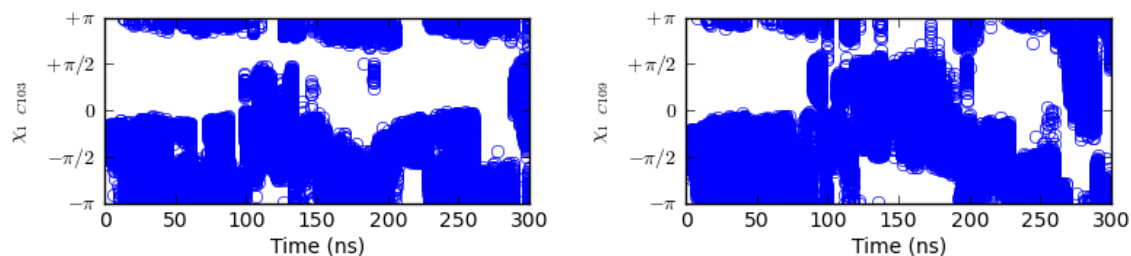


Fig. 4.17 The fluctuations of the C103 (a) and C109 χ_1 (b) dihedral angles in the metadynamics simulation.

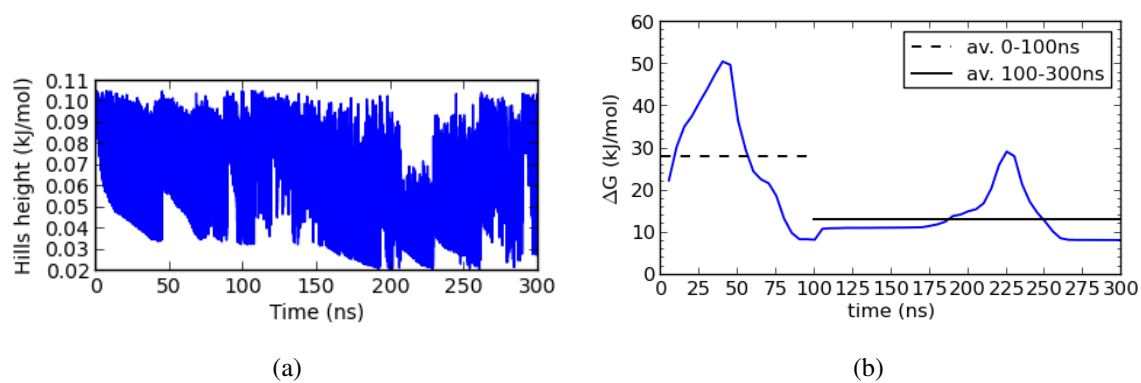


Fig. 4.18 More hills are added as the simulation progresses. The hill height is monitored in (a). The free-energy difference between the major conformer of oxidised nDsbD and a "reduced-like" state is monitored over the course of a total of 300 ns of metadynamics simulations in (b).

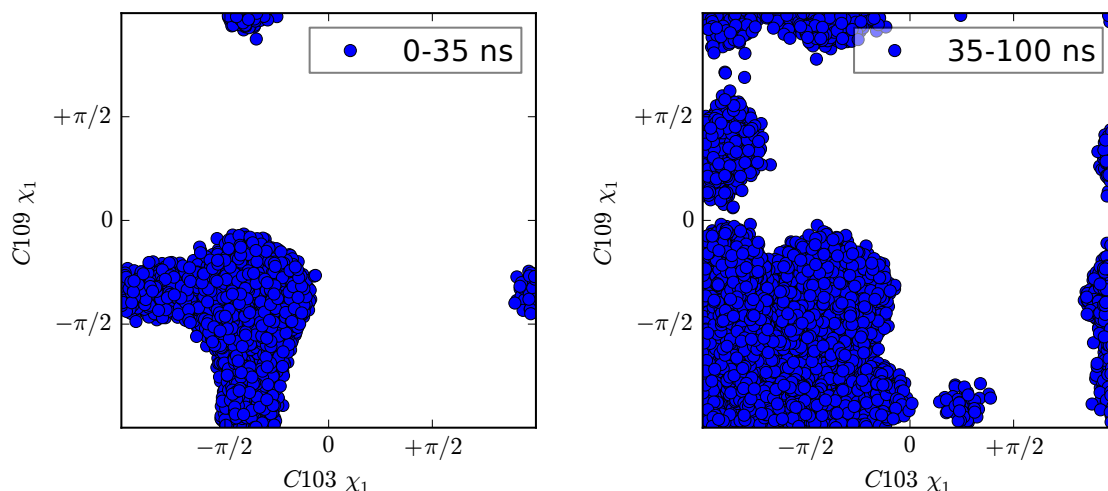


Fig. 4.19 The positions of the Gaussian potentials added in the metadynamics simulations are shown for the segment 0-35 ns and 35-100 ns for the metadynamics calculation.

well-tempered metadynamics algorithm used here, will converge to the correct free-energy profile for a collective variable if the sampling in the simulation is good enough.

One way to monitor the quality of the sampling is looking at the fluctuations of the collective variables. After adding a sufficient number of bias potentials in the collective variable space, the metadynamics potential fills up the free-energy surface and the system is diffusing freely in this space. Dense sampling, that may indicate relatively unrestricted motion was seen for both χ_1 of C103 and C109 between $-\pi$ and 0, which is the region of most interest for exploring the existence of a reduced-like minor state.

Another way to monitor the progress of the simulation is to look at the height of the Gaussian potentials added. The height of the potentials should decrease over time, which can be understood by looking at eqn. 2.18.

$$w = \omega e^{-V(s,t)/\Delta T} \tau_G \quad (2.18 \text{ revisited})$$

The bias w depends on the total bias at a point s in the collective variable space at a time t , *i.e.* $V(t,s)$ that has been built up. At places with a large bias potential w will be small. Hence w will decrease as the collective variable space is explored and decorated with bias potentials. Here, the hills height did not level off over the course of the 300 ns simulation (Fig. 4.18a). For example, a decrease of w around 200 ns, was followed by an increase as

a new region in the collective variable space was entered. The two dimensional coordinate space was not fully explored even after 300ns of simulation. This can be seen in Fig. 4.16b where the white regions show the absence of deposited bias potentials.

The most direct way to assess the quality of the simulation is by monitoring the free-energy differences over time. The time course of ΔG between the putative ground state, centred at C103 $\chi_1 = -\frac{\pi}{3}$ and excited state, close to C103 and C103 $\chi_1 = -\pi$ is shown in Fig. 4.18b. The graph suggests that ΔG tends to a value close to 10kJ/mol in the regions where the calculated ΔG is stable. There are two distinct spikes to much larger values. The first spike is straightforward to explain. Initially, the region around the starting structure was explored and bias potentials were added there (Fig. 4.19a). No bias potentials were added in the "reduced-like" state the during first 35 ns, as shown in Fig. 4.19a. Consequently, the reconstruction of the free-energy surface suggested that there is a large difference between the two states. As the simulation entered the putative excited state and Gaussian potentials were deposited there the free-energy differences became smaller. At 100ns, ΔG was around 10kJ/mol and both states were well explored as shown by Fig. 4.19b. This behaviour is expected and such fluctuations in ΔG are expected to decrease as the free-energy surface is better explored. Indeed, the second spike in ΔG was smaller.

Overall, the analysis of the metadynamics simulations suggested that while the whole free-energy surface was not explored in the 300ns of simulation time, the putative ground and excited states for oxidised nDsbD were well-explored and a free-energy difference of approximately 10 kJ mol⁻¹ between them was obtained. Therefore, computer simulations suggested that a "reduced-like" way of packing the cysteine side-chains is thermodynamically feasible in oxidised nDsbD.

4.4.8 Characterisation of the "reduced-like" conformation accessed by the metadynamics simulations

The "reduced-like conformation" of oxidised nDsbD can now be analysed in more detail. First, the side-chain conformation of F70 in the putative excited state structures from the metadynamics simulations was analysed. A different arrangement of this side-chain might explain the large chemical shift differences for E69 and F70 between the ground and excited states. The side-chain of F70 was found to switch between a gauche- and a trans conformation for oxidised nDsbD. For reduced nDsbD the side-chain of F70 is in a gauche- conformation (Fig. 3.18). Thus a reduced-like conformation would be expected to have F70 mostly in the gauche- conformation which was indeed observed (Fig. 4.20b) for frames that are in

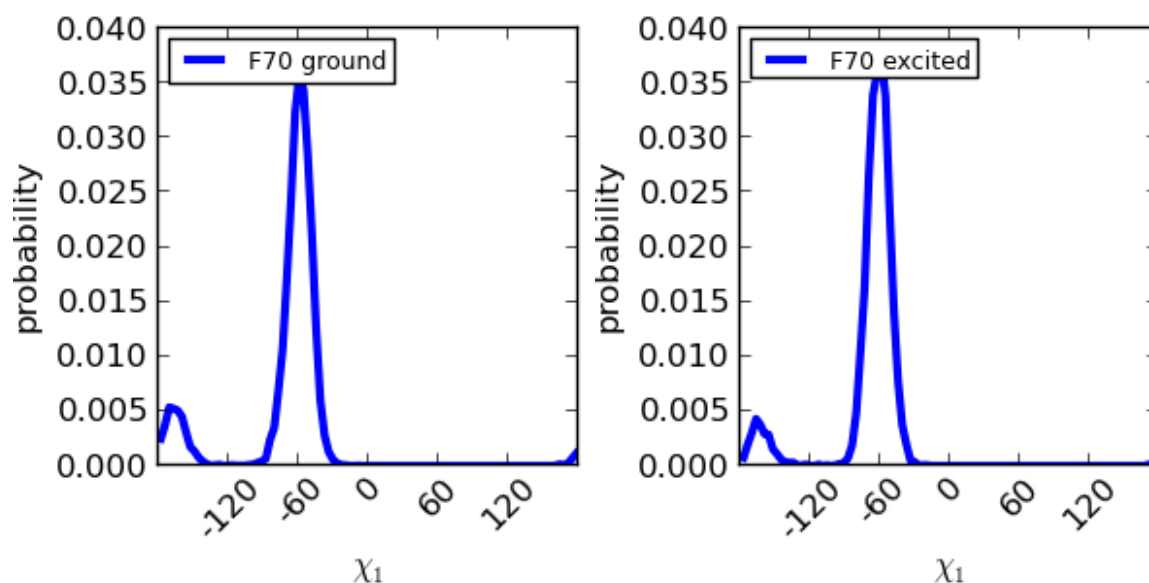


Fig. 4.20 The χ_1 dihedral angle of F70 is shown for the putative ground state and excited states from the 300ns metadynamics simulations.

the putative excited state. Frames were judged to be in the "reduced-like" conformation if their C103 and C109 dihedral angles were both between $-\pi$ and $-\frac{2}{3}\pi$. However, the ground state also primarily adopted a gauche- conformation (Fig. 4.20a), which is at variance with equilibrium MD simulations of oxidised nDsbD (Fig. 3.18b). The lack of conformational transitions for the χ_1 angle of F70 in the ground state may point at insufficient sampling in the metadynamics simulations.

Secondly, the state of the cap loop was assessed. The distances between the aromatic rings of F70 and Y71 were calculated for all frames of the metadynamics trajectories that accessed the "reduced-like" conformation. The cap loop stays closed in all structures associated with the reduced-like state as shown in Fig. 4.21b. For the major conformation of oxidised nDsbD, with χ_1 dihedral angles for C109 and C103 between $-\frac{2}{3}\pi$ and 0 however, transient opening of the cap loop is observed. Cap loop opening is evident in Fig. 4.21a, the line that surrounds the accessed region for the ground state of oxidised nDsbD encloses open conformations such as in the 1VRS crystal structure. While the difference between the ground and excited states of oxidised nDsbD in the metadynamics simulations may be caused by the limitations in sampling, it was encouraging that the ground and excited states behave as expected in the metadynamics simulations; the ground state of oxidised nDsbD is flexible and allows loop opening and the reduced-like state remains closed.

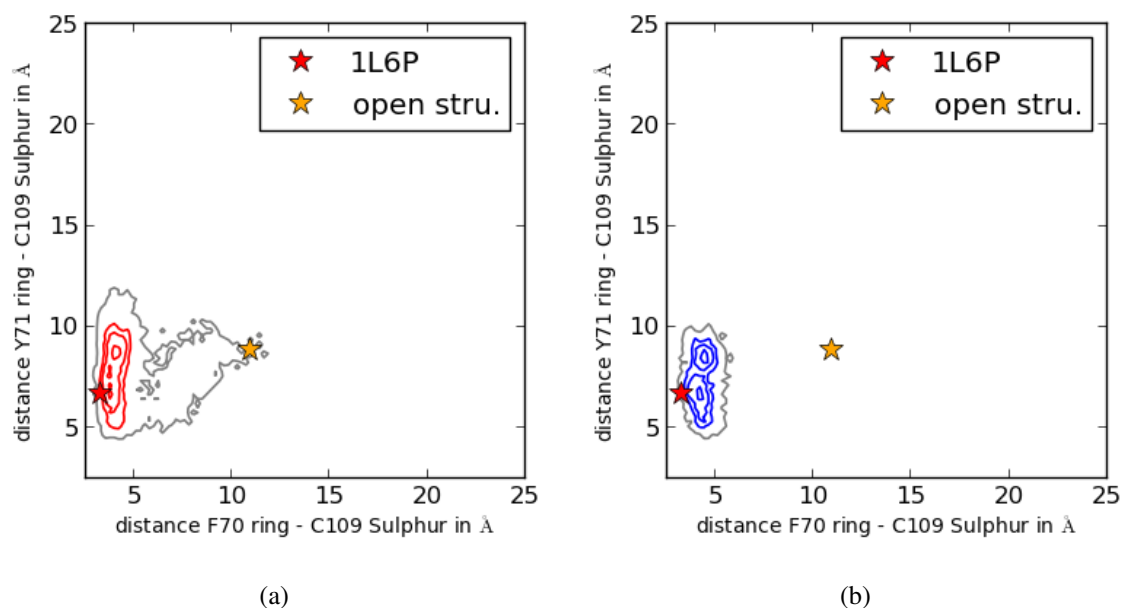


Fig. 4.21 The most populated and the reduced-like states of oxidised nDsbD in the 300 ns of metadynamics calculations are analysed using distance between the rings of F70 and Y71 and S_γ of C109 to detect cap loop opening. The cap-loop showed some opening in the putative major conformation (a), but stayed closed in the "reduced-like" excited state (b).

4.5 Discussion

The N-terminal domain of DsbD (nDsbD) was studied by NMR and computer simulations to understand the conformational equilibrium of the cap loop and the active site in both reduced and oxidised nDsbD. In previous chapters, experiments and simulations to characterise average structure and fast (ps-ns) fluctuations were discussed. In this chapter NMR relaxation dispersion experiments and metadynamics simulations were used to study relatively slow transitions that happen on average a couple of times every ms. NMR experiments and computer simulations were motivated by a set of questions that could not be answered based on static crystal structures alone:

1. The cap loop is thought to protect the active site of reduced nDsbD from non-cognate oxidation. The cap loop however needs to open for nDsbD to bind its cognate interaction partners in the *Dsb* reductant transfer pathway. It is therefore important to understand whether the loop is flexible and open conformations are ever accessed in the absence of the physiological interaction partners.
2. Differences in the conformational equilibrium between the reduced and oxidised iso-

forms of nDsbD might explain why oxidised nDsbD has a higher affinity for cDsbD, its cognate interaction partner, than reduced nDsbD.

4.5.1 Oxidised nDsbD has an alternative sparsely populated state

NMR relaxation dispersion experiments showed clear differences in the slow dynamics of reduced and oxidised nDsbD. For reduced nDsbD no excited state could be detected for the cap loop or the active site. For reduced nDsbD, spin-relaxation experiments probing fast-timescale dynamics and computer simulations already demonstrated that the cap loop is not disordered and that it protects the active-site cysteines at all times from the oxidising periplasm. By contrast, an alternative, sparsely populated state of the cap loop was found for oxidised nDsbD. This finding adds to the conclusion from the analysis of chemical shifts and the MD simulations that the cap loop is more dynamic in oxidised than in reduced nDsbD.

4.5.2 The structure of the excited state of oxidised nDsbD

What is the nature of the excited state of oxidised nDsbD? The relaxation dispersion profiles did not appear to depend strongly on the concentration of oxidised nDsbD which suggested that the excited state is accessed from the ground state by a uni-molecular reaction. This finding ruled out, for example, that the excited structure corresponds to an aggregated structure involving more than one nDsbD molecule.

An initial view of the structure of the alternative state of the cap loop in oxidised nDsbD was provided by its chemical shifts. The differences in chemical shift between the ground and excited states suggested that the excited state somewhat resembles reduced nDsbD. The cap loop is closed in reduced nDsbD, both in the 3PFU crystal structure and in MD simulations. The "reduced-like" excited state of oxidised nDsbD is therefore unlikely to feature an open cap loop.

If not an open conformation, the excited state may be due to the dynamics of the disulphide bond. The cysteine side-chains could adopt a conformation that allows closer packing of the cap loop onto the active site as in reduced nDsbD. In particular, the side-chains of C103 and C109 could adopt χ_1 dihedral angles similar to those in reduced nDsbD. This is of course consistent with the observed chemical shift differences.

The reduced-like conformation of oxidised nDsbD was studied in molecular detail in computer simulations. Sampling of slow dynamics was enhanced using the metadynamics algorithm. The metadynamics simulations indeed entered a reduced-like conformation with

cysteine side-chain dihedral angles similar to reduced nDsbD. In this "reduced-like" conformation, the cap loop was closed. Loop opening was observed in the dominant state of oxidised nDsbD in the metadynamics calculations as expected.

4.5.3 The free-energy landscape of nDsbD

The free-energy difference between the putative "reduced-like" and the major conformation of oxidised nDsbD in the metadynamics simulations can be compared to experiments. In experiments a free energy difference of about 10kJ mol^{-1} was found between the ground and excited states of oxidised nDsbD, which agreed with the simulations. The agreement between experiment and simulation should not be over-interpreted. The free-energy difference from the simulation neglects contributions from additional minor conformations that are likely be subsumed to one state or the other in the experiments. It is not clear how to partition them to either of the two states when comparing experiment and simulation. Nevertheless, it is encouraging that the free-energy differences are of similar magnitude in experiment and simulation- both might indeed report on the same process.

The experiments showed that the reduced-like excited state of oxidised nDsbD is enthalpically but not entropically more favourable than the ground-state. The simulations discussed were not extensive enough for splitting free energies into entropic and enthalpic contributions. In absence of the ΔH and ΔS , calculated from a rigorous theoretical model such as an extensively sampled MD simulation, the interpretation is qualitative at best. Based on the study of Grey *et al.* of the disulphide bond conformational equilibrium of BPTI, an explanation can nonetheless be attempted. For BPTI an enthalpically favourable minor conformation was described which was explained by a significant reduction in the torsional energy for the disulphide bond (Grey et al., 2003). An energetically more favourable structure of the disulphide bond could also underlie the excited state of oxidised nDsbD.

Alternatively, the enthalpy difference between the ground and excited state could be caused by differences in the hydrogen bonding in their water shells. The fluctuations of the cap loop in the major state of oxidised nDsbD increased the solvent exposure of F70 and this can disrupt the favourable hydrogen bonding in the solvent shell. The reduced-like conformation of oxidised nDsbD, with close packaging of the cap loop onto the active site minimises exposure of F70 and maximises hydrogen bonding in the surrounding solvent.

The NMR experiments detected no enthalpic barrier between the major and the minor states of oxidised nDsbD. The barrier between the ground state and the excited state

is entropic. The free-energy determined for the barrier could in principle be compared to the free-energy landscape from metadynamics, but it is not clear whether this comparison would be meaningful given that the value for $\Delta G^\ddagger$ depends on the choice of the transmission coefficient κ . The enthalpy difference between ground and transition states does not depend on κ , but a quantitative comparison of the experimental value and simulations is not possible as the sampling in the simulations is insufficient. The entropic character of the barrier is nonetheless interesting (Henzler-Wildman and Kern, 2007). No enthalpically favourable interaction, such as a hydrogen bond, appears to be disrupted in the transition state ensemble. The rate is instead directly limited by the speed of the diffusion between the conformational states (Dill and Bromberg, 2011).

4.5.4 Implications for biological activity

Overall, the differences in the flexibility of the cap loop help to rationalise the functional differences between reduced and oxidised nDsbD. The cap loop of oxidised nDsbD is flexible. There could be an element of conformational selection (Boehr et al., 2009) for structures with an open cap loop when oxidised nDsbD binds reduced cDsbD to receive reductant. For reduced nDsbD, the cap loop protects the active site at all times and the association with physiological partners such as CcmG of the cytochrome c maturation and DsbC of the disulphide bond isomerisation pathways, probably involves induced fit.

The physiological importance of the excited state of oxidised nDsbD is less clear. On the one hand, it resembles the next step in the physiological reaction, the reduced state of nDsbD that is formed after receiving reductant from cDsbD. Previous NMR relaxation dispersion experiments have highlighted excited states that resemble the next structure in a series of reactions, for example for Dihydrofolate reductase (Boehr et al., 2006). For oxidised nDsbD, the physiological reaction involves opening of the cap loop and the formation of a complex with reduced cDsbD. A reduced-like conformation with a closed cap loop would not be involved in the reductant transfer reaction that converts oxidised to reduced nDsbD and might be an "off-pathway" reaction intermediate.

However, the excited state identified shows, once more, that the disulphide bond disrupts the packing of the cap loop onto the active site. The major populations of oxidised nDsbD features a cap loop that undergoes more complicated dynamics than in reduced nDsbD and occasionally opens. Only a minor conformation shows the rigid packing of the cap loop onto the active site that protects reduced nDsbD from non-cognate side-reactions.

4.5.5 Future work

The conclusions from this chapter could be strengthened by repeating the sign determination experiments. This could confirm again that the minor conformation of oxidised nDsbD resembles the reduced protein. Strain energies could be calculated for representative structures of the major and minor states of oxidised nDsbD (Grey et al., 2003). This might help to rationalise why the excited state is enthalpically more favourable than the ground state.

Chapter 5

The association reaction of nDsbD with cDsbD

5.1 Introduction

In the previous chapters, the conformational equilibrium of the cap loop of nDsbD was studied to understand how it might open during its association with cDsbD and other cognate binding partners. Here the association reaction itself is probed by NMR and computational approaches. To study the mechanism of the association of cDsbD and nDsbD by NMR, relaxation dispersion experiments were used. The NMR relaxation dispersion experiments revealed the kinetic basis of the redox-state dependent interaction of the two periplasmic domains (Mavridou et al., 2011). *In silico*, the mechanism was investigated firstly, by using a simplified model for protein-protein association, transient complex theory (Alsallaq and Zhou, 2008), secondly by electrostatics calculations and thirdly by molecular dynamics (MD) simulations.

5.1.1 Studying protein-protein association reactions

NMR relaxation dispersion experiments enable the detailed study of protein-binding reactions and has given important mechanistic insight (Li et al., 2013; Sugase et al., 2007a). To study protein-protein association with relaxation dispersion experiments, one typically adds a small amount of unlabelled ligand to the isotopically labelled unbound protein (Sugase et al., 2007b). The exchange between the small bound fraction and the unbound state of the protein then gives rise to relaxation dispersion that enables the detailed characterisation of the reaction. In this way, protein-association reactions that are hard to study by NMR due

to extensive exchange broadening can be investigated. Perhaps the most interesting mechanistic information that relaxation dispersion experiments can provide is the identification of reaction intermediates. For example, such experiments showed how target binding is linked to folding for the transcription factor CREB (Sugase et al., 2007a) in a three-state reaction. The experiments can also reveal the association and dissociation rates, which can also inform on mechanism. For instance, a very fast association rate, that is close to the diffusion limited rate, would suggest that conformational changes, such as in the cap loop, do not significantly slow down the reaction. Furthermore, the association rate limits the speed of many chemical reactions in biological systems. The rate for reductant transfer between cDsbD and nDsbD was found to be on the order of $10^6 \text{ M}^{-1} \text{ s}^{-1}$ (Rozhkova et al., 2004). It will be interesting to see whether reductant transfer between the periplasmic domains of DsbD is limited by the association rate and would thus be "catalytically perfect" (Schreiber et al., 2009).

Computer models have also made crucial contributions to the study of protein-protein association reactions (Gabdouline and Wade, 1998). Brownian dynamics simulations that capture the diffusional motions of proteins have emphasised the importance of the shape of proteins in determining association rates. Only by taking into account shape information can one decide whether a reaction is diffusion controlled. The often quoted rate of $10^9 \text{ M}^{-1} \text{ s}^{-1}$ for a truly diffusion controlled reaction was derived for spheres and is not generally applicable to proteins (Schreiber et al., 2009) and typical rates are often on the order of $10^6 \text{ M}^{-1} \text{ s}^{-1}$ (Northrup and Erickson, 1992). Some association reactions are indeed very fast such as the binding of the green mamba toxin toxin fasciculin to acetylcholinesterase, with a k_{on} on the order of $10^8 \text{ M}^{-1} \text{ s}^{-1}$ (Radić et al., 1994). Electrostatic interactions in addition to passive diffusion drives the fast binding of fasciculin (Radić et al., 1997).

Computer simulations revealed that "macro-collisions" between proteins are a general feature of protein-protein association reactions (Northrup and Erickson, 1992; Ubbink, 2009). Two protein molecules encounter each other. Water molecules constantly push against the proteins. However, the interface between the two proteins is largely shielded from such collisions with water molecules. The random collisions with water therefore push the proteins towards each other which prolongs their encounters.

5.1.2 Encounter complexes

Such "encounter complexes" are critical intermediates in going from unbound to bound configurations (Ubbink, 2009) as emphasised in eqn. 5.1, where AB^* denotes the set of

conformations that form the encounter complex.



The macroscopic rate constants k_{on} and k_{off} are then given by eqn. 5.2 and eqn. 5.3.

$$k_{on} = \frac{k_1 k_2}{k_{-1} + k_2} \quad (5.2)$$

$$k_{off} = \frac{k_{-1} k_{-2}}{k_{-1} + k_2} \quad (5.3)$$

After forming an initial complex, conformational changes are often required to form the native complex. Slow conformational changes and small k_2 can thus limit the overall reaction. In such cases the population of the encounter complex would increase. Ubbink in his 2009 review also highlights that rapid turnover is facilitated by encounter complexes. Electron-transferring proteins, for example, associate rapidly via encounter complexes to form their productive complexes. The native complex is however in rapid exchange with the encounter complex, which then facilitates fast dissociation. The k_{off} rates for electron-transfer complexes are often in the range of 1000s^{-1} . Fast association and dissociation is hence achieved by having a large population of encounter complexes. In the reaction between nDsbD and cDsbD, conformational changes are important and turn-over may be fast (Mavridou et al., 2011). Encounter complexes may play a role in this.

Before proceeding it is important to consider the microscopic nature of encounter complexes. Initial studies have emphasised electrostatic interactions between proteins forming an encounter complex. The proteins would be largely solvated in the encounter complexes. Interfacial water would facilitate interactions between the charged groups of the binding partners (Ahmad et al., 2011; Schreiber et al., 2009). NMR relaxation dispersion experiments have however identified hydrophobic interactions that stabilise encounter complexes of the KIX domain of the transcription factor CREB and the phosphorylated peptide pKID (Sugase et al., 2007a). Computational studies based on NMR experiments have highlighted that encounter complex ensembles are structurally very diverse (Kim et al., 2008). Some conformations resemble the native state and overlap with the cognate binding site whereas other complexes are only loosely interacting. A recent study (Kozakov et al., 2014) suggested that non-specific interactions might enable proteins to search for their binding sites by diffusion on the surfaces of their binding partners. Here MD simulations will be used to characterise encounter complex structures for the association of cDsbD and nDsbD.

5.1.3 The association of cDsbD and nDsbD

Previous work in the Redfield and Ferguson groups has already yielded many clues about the mechanism of the association of cDsbD and nDsbD (Mavridou et al., 2011). Site-directed mutagenesis has highlighted the importance of V462 of cDsbD. The side-chain of V462 appears to stabilise the open cap loop of nDsbD in the 1VRS crystal structure. Exchanging this residue for an alanine greatly reduces the affinity of cDsbD for nDsbD. This has been taken to support an induced-fit mechanism for cap loop opening. It should be noted that this finding is also compatible with a conformational selection mechanism. In a conformational selection mechanism V462 would stabilise the cap loop once it has opened spontaneously.

Moreover, previous work by the groups also established that the interaction between cDsbD and nDsbD is redox-state dependent. Reduced cDsbD forms a complex with oxidised nDsbD. Reductant is transferred via a thiol-disulphide exchange reaction. After reductant transfer the affinity decreases, and the complex of reduced nDsbD and oxidised cDsbD would fall apart. The redox-state dependence of the interaction affinity does not only prevent the thermodynamically feasible transfer of reductant from nDsbD back to cDsbD, but also frees up the two domains for further reductant transfer steps. cDsbD can then receive reductant from the trans-membrane domain of DsbD, whereas nDsbD can provide reductant to the disulphide bond isomerisation and the cytochrome *c* maturation pathways. The redox-state dependence of the interaction affinities may enhance turnover in the *Dsb* pathway. Yet the rates for complex formation and dissociation are not known. It has been assumed that the complex would fall apart rapidly after the reductant transfer. The determination of the binding kinetics in this work will complete the picture.

The work on isolated nDsbD, discussed in the preceding chapters, showed how the redox-state of nDsbD could contribute to the redox-state dependent interaction of the periplasmic domains of DsbD. Particularly, differences in flexibility of the cap loop might explain why oxidised but not reduced nDsbD forms a stable complex with cDsbD. The cap loop of oxidised nDsbD is flexible and can open occasionally. The loop is rigid and protects the active site at all times in reduced nDsbD.

The redox-state of cDsbD also affects the affinity for the two domains. For the interaction between oxidised cDsbD and oxidised nDsbD a K_d of 800 μ M was determined. The interactions of oxidised cDsbD and reduced nDsbD cannot be easily studied directly, as the domains react rapidly in a thiol-disulphide exchange reaction. The resulting mixture of reactants, products and intermediates complicates the interpretation of experiments. Nonre-active C461A-cDsbD, where one of the active site cysteines is mutated to an alanine is used to study the interactions of reduced cDsbD. C461A-cDsbD forms a significantly stronger

interactions with oxidised nDsbD with a K_d of 86 μM . It is not clear how the redox-state of cysteines in the active site of cDsbD determine the affinity. The pK_a s of several residues differ somewhat between oxidised and reduced cDsbD (Mavridou et al., 2007). This might alter the surface electrostatics of cDsbD (Mavridou et al., 2011) which could affect the equilibrium for example by changing the association rate k_{on} . The redox-state of cDsbD affects its hydrogen bonding network (Mavridou et al., 2014) which could then influence the stability of the complex, *i.e.* k_{off} . The stability might also be influenced more directly by the redox-state of the domains. Two thiol groups fill more space than a disulphide bond. Packaging constraints introduced by converting a disulphide bond into two thiols may determine whether the hydrogen bonds between cDsbD and nDsbD can be maintained (Mavridou et al., 2011). Such a direct role of the active site cysteines resembles the situation found for nDsbD in Chapter 3.

5.1.4 Aims

Here the association of oxidised cDsbD, as well as C461A-cDsbD, with oxidised nDsbD was studied by NMR relaxation dispersion experiments. By determining the k_{on} and k_{off} of both pairs of domains, one might discover more about the molecular basis of the redox-state dependence of the cDsbD-nDsbD complex and the kinetic consequences for the pathway. The association rate will be compared to the rate constant for reductant transfer between the two domains. In addition, it will be interesting to see whether any intermediate states between bound and free conformations of cDsbD-nDsbD are detected by NMR. The association will also be investigated computationally using simplified models and atomistic MD simulations.

5.2 Materials and methods

5.2.1 Sample for NMR measurements

Dr Despoina Mavridou kindly provided samples for studying the association reaction between cDsbD and nDsbD. The first sample, contained ^{15}N labelled c461A-cDsbD with a final concentration of 0.55 mM. The concentration of oxidised nDsbD was 0.022 mM. The nDsbD concentration was determined based on 1D spectra shown in Fig. 5.2. The 1D spectra had suggested the concentration of the nDsbD stock sample had been overestimated and therefore the nDsbD concentration had to be estimated based on proton spectra. There was about 4 % D_2O in the sample. The pH was adjusted to 6.5.

In the second sample, the concentration of oxidised cDsbD was 0.53 mM. The final concentration of oxidised nDsbD was 0.04 mM. The nDsbD concentration was determined based on the known K_d . 1D spectra shown in Fig. 5.7 suggested a nDsbD concentration of 0.03 mM which roughly agreed with this estimate. The pH was adjusted to 6.5 and 15 μ L D₂O was added. There was about 4 % D₂O in the final sample volume.

5.2.2 Spectrometers

NMR experiments discussed in this chapter were recorded on home-built 500 MHz, 750 MHz and 950 MHz spectrometers. The spectrometers are equipped with Oxford Instruments Co magnets, home-built triple-resonance probes with triple axis gradients. The spectrometers were controlled by GE/Omega data acquisition computers and software.

5.2.3 NMR relaxation dispersion experiments

To study the association of ¹⁵N labelled C461A-cDsbD and unlabelled oxidised nDsbD, amide ¹⁵N relaxation dispersion experiments were collected at magnetic field strengths corresponding to proton resonance frequencies of 500 MHz and 950 MHz. To study the binding of ¹⁵N labelled oxidised cDsbD to unlabelled oxidised nDsbD, experiments were conducted at 500 MHz and 750 MHz. The pulse sequences introduced in the previous chapter were used. The length of the spin-echo train T_{CPMG} was 40 ms.

5.2.4 Analysis of the relaxation dispersion experiments

Experiments were processed by standard methods using NMRpipe (Delaglio et al., 1995). The intensities of the amide resonances were followed in CCPNMR Analysis (Vranken et al., 2005). The intensities were exported and converted to $R_{2,eff}$ using custom NumPy scripts. Relaxation dispersion curves were analysed using CATIA (Hansen, 2008). CATIA fits the data by numerically integrating over a representation of the pulse sequence and is the most general way of analysing relaxation dispersion experiments. For both samples, a single global exchange process was assumed.

5.2.5 Analysis of 1D NMR experiments

1D ¹H NMR spectra for the association of C461A-cDsbD as well as oxidised cDsbD with oxidised nDsbD were processed using NMRPipe (Delaglio et al., 1995). The resonances for the methyl groups of A458 and A17 of cDsbD and nDsbD were used to estimate the

fraction of nDsbD in the sample and the extent of complex formation. For analysis of association of C461A-cDsbD with nDsbD a somewhat wider region (between -0.22 ppm and -1.05 ppm) than for the sample with oxidised cDsbD and nDsbD was analysed and more extensive baseline correction was required. The baseline correction is required for the analysis of spectra by fitting Lorentzians to some of the resonances. The choices made for correcting the baseline will affect the results and hence these are at best semi-quantitative. The fitting to Lorentzians was enabled by importing the NMR data into python scripts using the NmrGlue library (Helmus and Jaroniec, 2013).

5.2.6 Electrostatics calculations

The Poisson Boltzmann equation was solved using APBS (Baker et al., 2001). The charges were assigned using pdb2pqr (Dolinsky et al., 2004). The D455 side-chain is protonated in the complex of cDsbD and nDsbD (Mavridou et al., 2009) and a proton was consequently added to this side-chain. Histidine side-chains were assumed to be positively charged. The calculation was run at zero ionic strength which is a good representation of the conditions for the NMR measurements. A control calculation where all histidine side-chains were neutral and D455 was de-protonated was also run (*not shown*). The conclusions about the native interaction interfaces did not change.

5.2.7 Molecular dynamics simulations of the association of reduced cDsbD and oxidised nDsbD

Simulations of the association reaction between cDsbD and nDsbD were setup based on the 1VRS B crystal structure (Rozhkova et al., 2004). This crystal structure was used to orientate nDsbD and cDsbD. The 1L6P and 3PFU crystal structures (Goulding et al., 2002; Mavridou et al., 2011) were aligned to the 1VRS B structure and subsequently translated, using MDAnalysis (Michaud-Agrawal et al., 2011). The D455 side-chain is protonated in the complex of cDsbD and nDsbD and a proton was consequently added to this side-chain in the setup of the MD simulations. Histidine side-chains were assumed to be positively charged.

The length of chemical bonds involving hydrogen atoms was constrained, as is typically done in simulations with CHARMM potential energy functions, using the PLINCS algorithm Hess (2008). The equations of motion were integrated with a 2 fs time step. The two domains were then solvated in a rhombic dodecahedral box. The minimum distance

Table 5.1 MD simulations of the association of reduced cDsbD and oxidised nDsbD. The cap loop was closed in the starting structures.

	salt concentration	length	centre of mass distance
run1	0.1 mM	200 ns	+22Å, 50Å in total
run2	no added salt	200 ns	+22Å, 50Å in total

between the protein and the box edges was 1.5 nm. Ions were added to equilibrate the simulation system. Na^+ and Cl^- ions up to a concentration of 0.1 mM were added to *run1*, as listed in Table 5.1. For *run2*, no ions beyond what was needed to neutralise the system were added. The system was energy minimised. To equilibrate the simulation system a 4 ns MD simulation with position restraints was used before the production runs.

MD simulations were run in GROMACS 4.5 (Bjellmar et al., 2010; van Der Spoel et al., 2005) using the CHARMM 22 potential energy function with the CMAP correction (Mackerell et al., 2004) following the approach outlined by Bjellmar *et al.*. Electrostatics was treated using the Particle Mesh Ewald approach with a short range cut-off of 12Å. The Lennard-Jones potential was switched to zero between 10Å and 12Å. The neighbour list was updated every five integration steps using a cut-off radius of 12Å. The length of covalent bonds to hydrogen atoms was constrained using the PLINCS algorithm (Hess, 2008). An integration time step of 2 fs was used.

5.3 Results: NMR experiments

5.3.1 Association of C461A-cDsbD with oxidised nDsbD

The association between C461A-cDsbD and oxidised nDsbD was followed by NMR relaxation dispersion experiments. A small amount of unlabelled nDsbD was added to ^{15}N labelled C461-cDsbD. 1D NMR experiments showed that there was about 25 fold more cDsbD than nDsbD in the sample (Fig. 5.2). Based on previous studies it was clear that adding too much nDsbD would complicate the measurements by severely increasing peak broadening that would render several key residues undetectable. Too low a concentration of nDsbD would make the bound conformation hard to detect by the relaxation dispersion experiments. Based on the K_d about 3 % of cDsbD would be in the complex under these conditions. This should be a good trade-off between detecting all key residues and having a large enough population of the minor, nDsbD bound state to observe dispersion. Good relaxation dispersion curves were indeed obtained under these conditions as shown Fig. 5.3.

The reaction is expected to be in slow exchange (Mavridou et al., 2011). Inspection of the relaxation dispersion curves demonstrated that the reaction is indeed in relatively slow exchange. Representative curves are shown for A461 and F531. The curves level off at refocusing fields around 300 Hz. As the curves appeared to have levelled-off completely, good fits should be obtained.

Fitting the relaxation dispersion curves for the association between C461A-cDsbD and oxidised nDsbD

The relaxation dispersion curves for the association of C461A-cDsbD with nDsbD were analysed using CATIA. A single global exchange process was found to describe the experimental curves (Table 5.2). All the residues that showed significant dispersion $|\Delta ppm| > 1$ ppm are in the interaction interface with nDsbD as shown in Fig. 5.5. The exchange is slow as expected, with an exchange rate k_{ex} of 220 s^{-1} . About 4 % of cDsbD is in complex with nDsbD. nDsbD bound cDsbD is visible in 1D ^1H spectra. Fitting a Lorentzian to the peaks for the bound and free protein suggested that about 2 % of cDsbD would be in the complex. The population of bound cDsbD from the fits also compared favourably to the estimate based on the dissociation constant ($K_d = 86 \mu\text{M}$) and the concentrations of cDsbD and nDsbD, which suggested that about 3 % of cDsbD would be bound to nDsbD.

To further validate the interpretation of the relaxation dispersion experiment, the chemical shift differences between the bound and free forms of cDsbD from the fitting and from previous experiments (Mavridou et al., 2011) were compared. Mavridou *et al.* determined the shift differences for complex formation based on the covalent complex of the two domains. Since the complexes in the two studies are not identical, a good, but not a perfect, correlation would be expected. A good correlation between the chemical shift differences from the two methods was indeed obtained as shown in Fig. 5.4. This further confirms the validity of the fit parameters and one can conclude that a single global, two-state, exchange process describes the data well.

The kinetics of association of C461A-cDsbD with oxidised nDsbD

Having validated the analysis of the fits, the macroscopic on and off rates for the complex formation between C461A-cDsbD and oxidised nDsbD can be determined (Table. 5.2) (Palmer et al., 2001). The off rate, k_{off} , is given simply by $k_{off} = p_A k_{ex}$ and is 210 s^{-1} . Using the dissociation constant previously determined by surface plasmon resonance, $K_d = 86 \mu\text{M}$, an association rate of $2.4 \times 10^6 \text{ M}^{-1} \text{ s}^{-1}$ (Table. 5.2) was calculated.

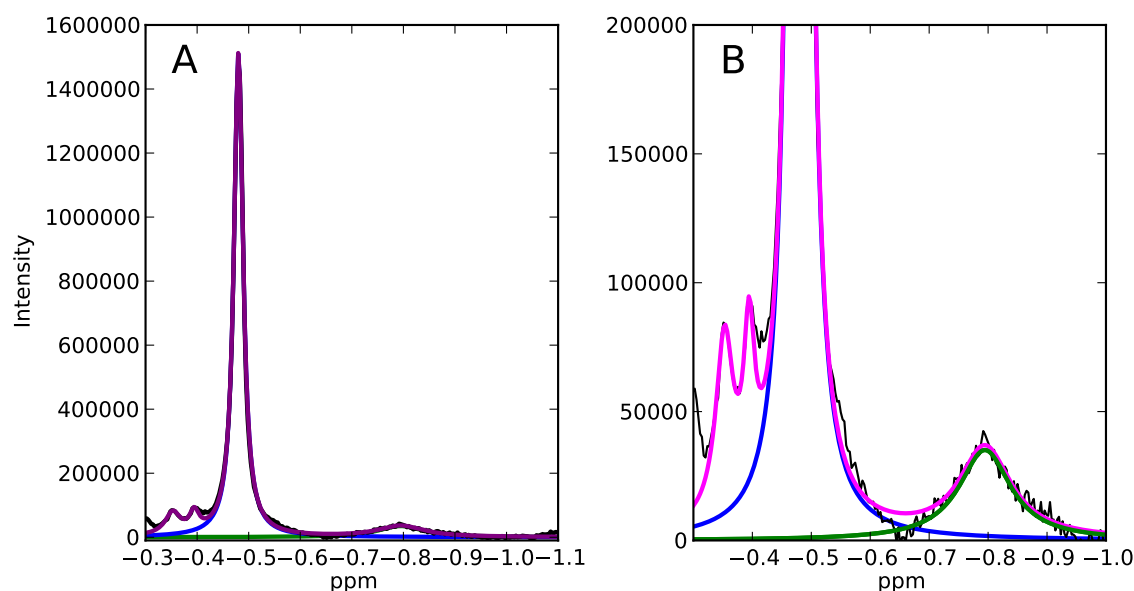


Fig. 5.1 A segment of the 1D ^1H NMR spectrum for C461A-cDsbD in the presence of nDsbD is shown (in black). The spectrum was base-lined corrected as described in section 5.2.5. Panel (a) demonstrates that cDsbD is in slow exchange between its bound and free forms. The side-chain methyl protons of A459 give rise to a free peak close to -0.5 ppm and a peak of the complex close to -0.8 ppm. To get an estimate of the population of the complex, the two peaks were approximated as Lorentzians (b) which are drawn in blue and green. About 2% of cDsbD is in the complex according to this analysis. The overlap of A458 with A17 was taken into account. The purple line shows the sum of all Lorentzians which compares well with the experimental data.

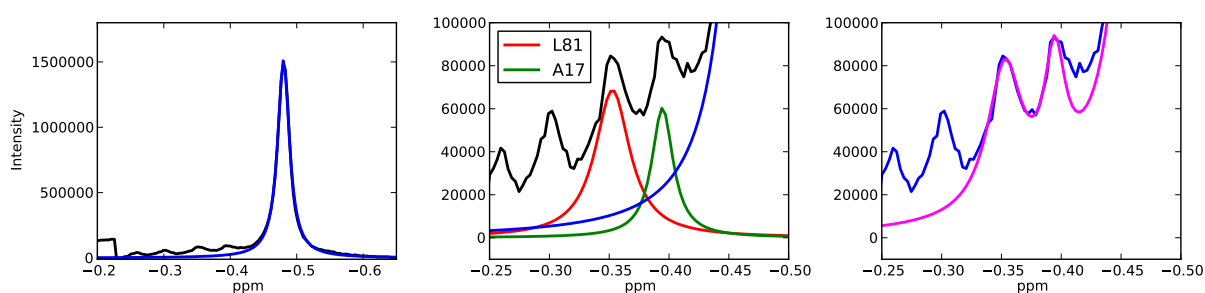
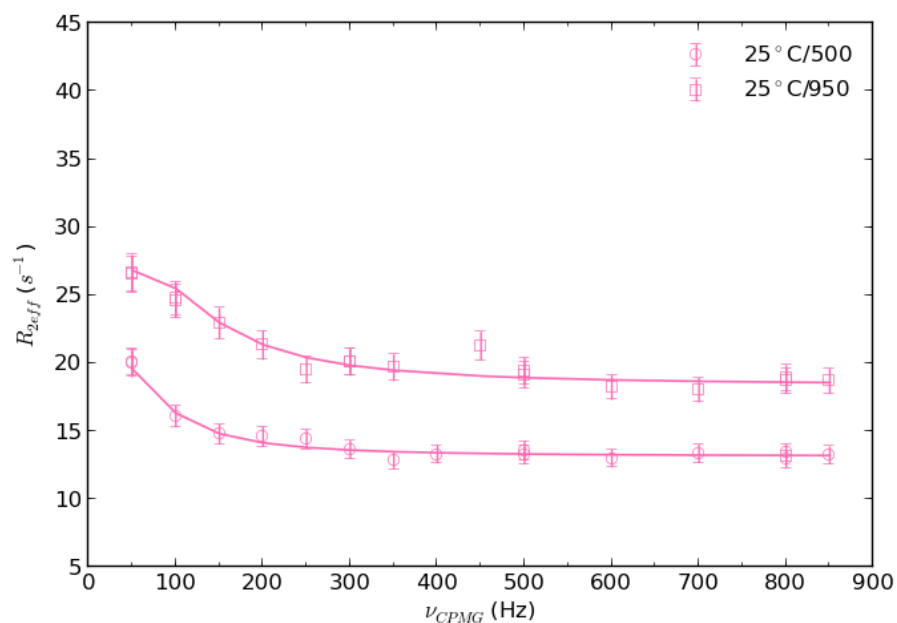
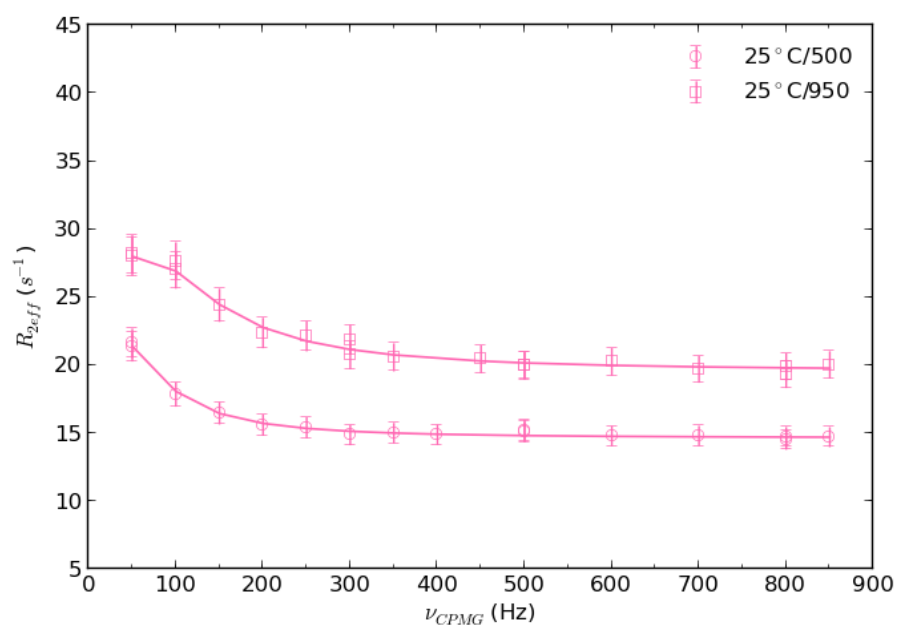


Fig. 5.2 The ratio of cDsbD to nDsbD was estimated from 1D spectra. The methyl proton peak for A17 was approximated by a Lorentzian. Overlap with A458 and L41 was taken into account. There was about 25 fold more cDsbD than nDsbD in the sample by comparing the peak areas for A17 and A458 (in blue). The sum of all fitted peaks is shown in purple.



(a) C461



(b) F531

Fig. 5.3 Relaxation dispersion experiments detect the slow exchange between C461A cDsbD and oxidised nDsbD

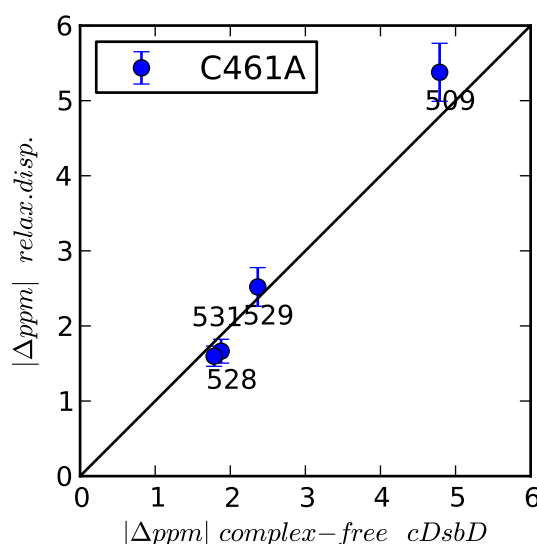


Fig. 5.4 The absolute chemical shifts from the relaxation dispersion experiments of the association of C461A-cDsbD and oxidised nDsbD reproduces the previously established absolute chemical shift differences between cDsbD on its own and in the complex with nDsbD (Mavridou et al., 2011) as shown by the blue markers.

Table 5.2 Relaxation dispersion fit parameters for the association of cDsbD with oxidised nDsbD and the macroscopic on and off-rates determined from them. The fit parameters for the association reactions with C461A-cDsbD, which mimics the reduced isoform, and with oxidised cDsbD are given in the table.

	C461A-cDsbD	oxidised cDsbD	oxidised cDsbD (p_B fixed)
k_{ex}	$220 \pm 26 \text{ s}^{-1}$	$2500 \pm 200 \text{ s}^{-1}$	$1720 \pm 130 \text{ s}^{-1}$
p_B	$4 \pm 0.4 \%$	$2.9 \pm 0.3 \%$	3.0%
k_{off}	$210 \pm 33 \text{ s}^{-1}$	$2430 \pm 200 \text{ s}^{-1}$	$1670 \pm 120 \text{ s}^{-1}$
k_{app}	$10 \pm 1 \text{ s}^{-1}$	$70 \pm 10 \text{ s}^{-1}$	$50 \pm 5 \text{ s}^{-1}$
$k_{on} = \frac{k_{off}}{K_d}$	$2.4 \times 10^6 \text{ M}^{-1} \text{ s}^{-1}$	$3.0 \times 10^6 \text{ M}^{-1} \text{ s}^{-1}$	$2.1 \times 10^6 \text{ M}^{-1} \text{ s}^{-1}$

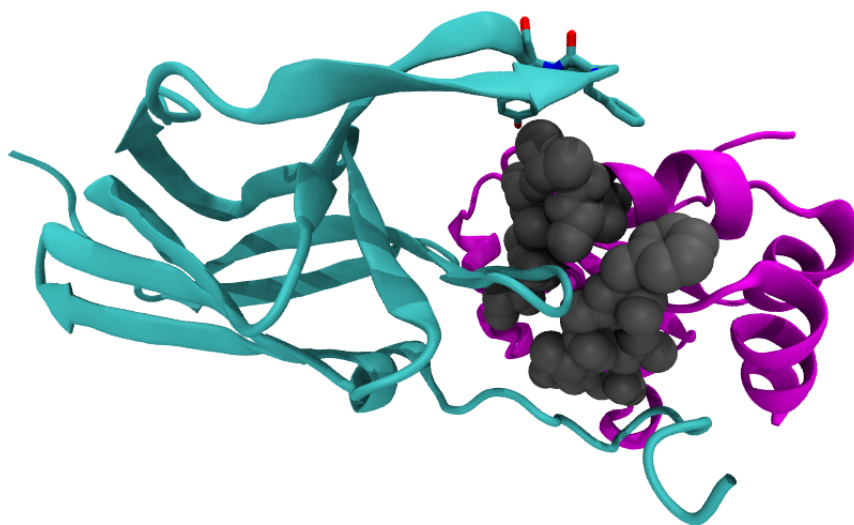


Fig. 5.5 The residues in C461A-cDsbD that show significant dispersion in the association reaction with nDsbD are highlighted. Here only residues for which $|\Delta ppm| > 1.0 ppm$ are shown in grey. These residues are G509, L510, R527, V528, T529, G530, F531, C461, V462, A463 and C464. cDsbD (1VRS E) is shown in magenta and nDsbD (1VRS E) in cyan.

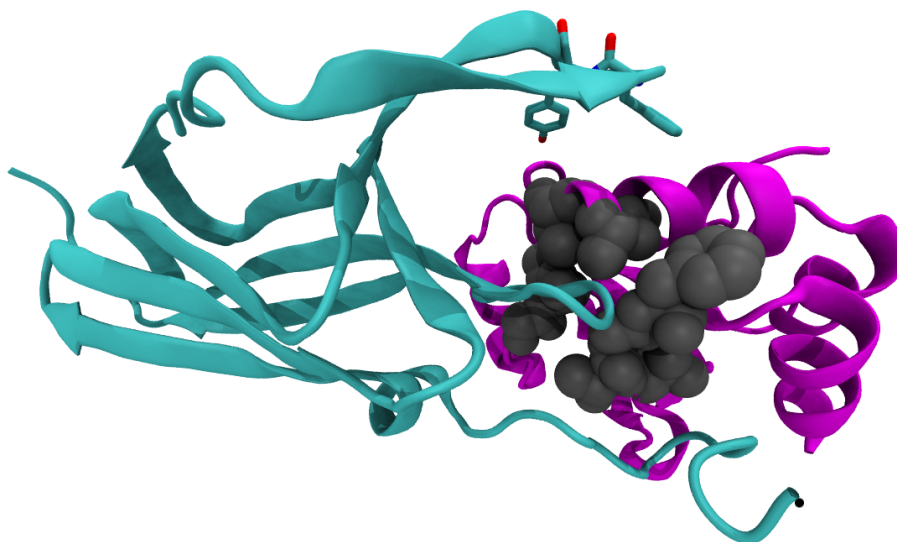


Fig. 5.6 The residues in oxidised cDsbD that show significant dispersion in the association reaction with nDsbD are highlighted. Here only residues for which $|\Delta ppm| > 1.0 ppm$ are shown in grey. These residues are C461, A463, C464, G509, L510, V528, T529, G530 and F531.

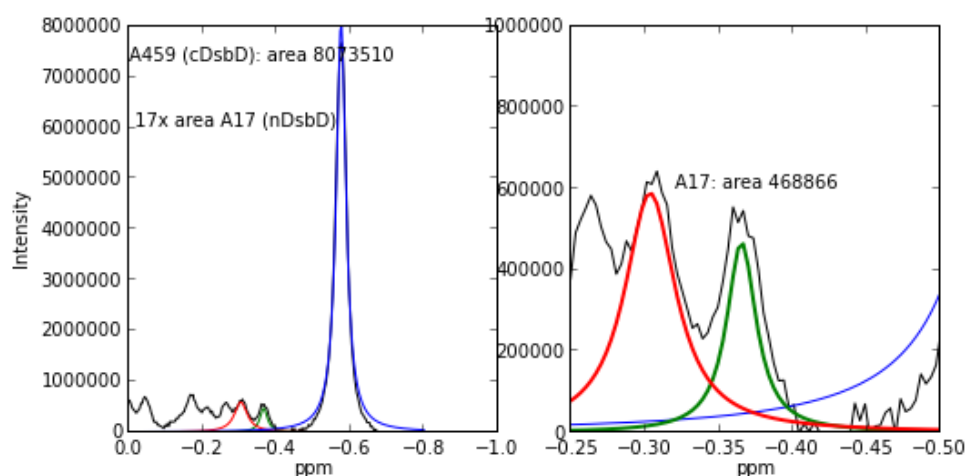


Fig. 5.7 1D ^1H NMR experiments on sample used to study the association of oxidised cDsbD with oxidised nDsbD. Fitting Lorentzians to the peaks for A17 (green) from nDsbD and A458 of cDsbD demonstrated that there was about 17 fold more oxidised cDsbD than oxidised nDsbD in the sample.

5.3.2 Association oxidised cDsbD with oxidised nDsbD

Relaxation dispersion experiments were recorded for the association of oxidised cDsbD with oxidised nDsbD. Previous experiments have highlighted that adding small amounts of oxidised nDsbD to oxidised cDsbD leads to significant peak broadening and some peaks become unobservable, almost as soon as ligand was added. Here a very small amount of unlabelled nDsbD is added to ^{15}N labelled cDsbD. 1D NMR measurements suggested that there was about 17 fold more cDsbD than nDsbD in the sample (Fig. 5.7). 2D HSQC spectra showed that key residues such as G509 and V462 that broaden significantly upon adding nDsbD (Mavridou et al., 2011) were still observable (*not shown*).

For several residues of cDsbD clear relaxation dispersion effects were observed. Here the curves for C464, which is in the active site as well as in nDsbD binding interface, and F531 that is also in the binding site are shown. For V462, which a key residue stabilising complexes of cDsbD and nDsbD, noisy dispersion data was obtained at 500MHz. At 750MHz peaks could not be assigned with confidence for this residue, likely due to extensive exchange broadening. The curves suggested a much faster exchange process than for the association of C461A cDsbD with nDsbD. The curves only appeared to level-off around 800Hz. CPMG relaxation dispersion curves for such fast exchange processes can be challenging to interpret. Furthermore the exchange process was likely not fully refocused at the ν_{CPMG} frequencies up to 850Hz, that could routinely accessed on the spectrometers. This

will further complicate the data analysis and will increase errors in the fit parameters.

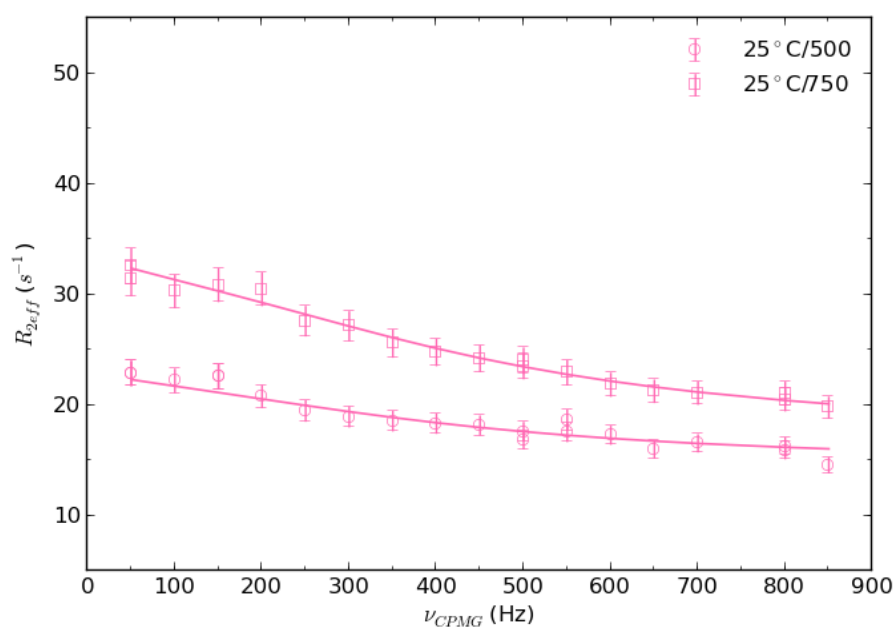
Fitting the relaxation dispersion curves for the association of oxidised cDsbD with oxidised nDsbD

All residues for which clear dispersion effects were identified were included in the fitting using CATIA. A global exchange process was assumed. The fitting, suggested a process with an exchange rate of 2500 s^{-1} and a population of the bound state of 2.9% (Table 5.2). Based on the concentrations of the two domains and the previously determined dissociation constant about 3% of cDsbD are expected to be in the complex. As for C461A-cDsbD residues for which the fit revealed large chemical shift changes between bound and free conformations are in the interface with nDsbD as depicted in Fig. 5.6. The chemical shift differences between the bound and the free states of cDsbD agree well with the previously determined chemical shift differences for the formation of the complex (Fig. 5.9). Considering that both the population of the excited state and the chemical shift differences agree with previous knowledge it appears that the relaxation dispersion experiments capture the exchange process well. In general p_B and Δppm cannot be determined from relaxation dispersion curves for such a fast process (Luz and Meiboom, 1963; Vallurupalli et al., 2011). Nevertheless enough residues have large chemical shift differences between the bound and the unbound state and are outside the fast-exchange limit to enable the determination of p_B and Δppm in the global fitting. Yet, the fitted value for k_{ex} is rather large, given that curves appeared to level-off around 800 Hz.

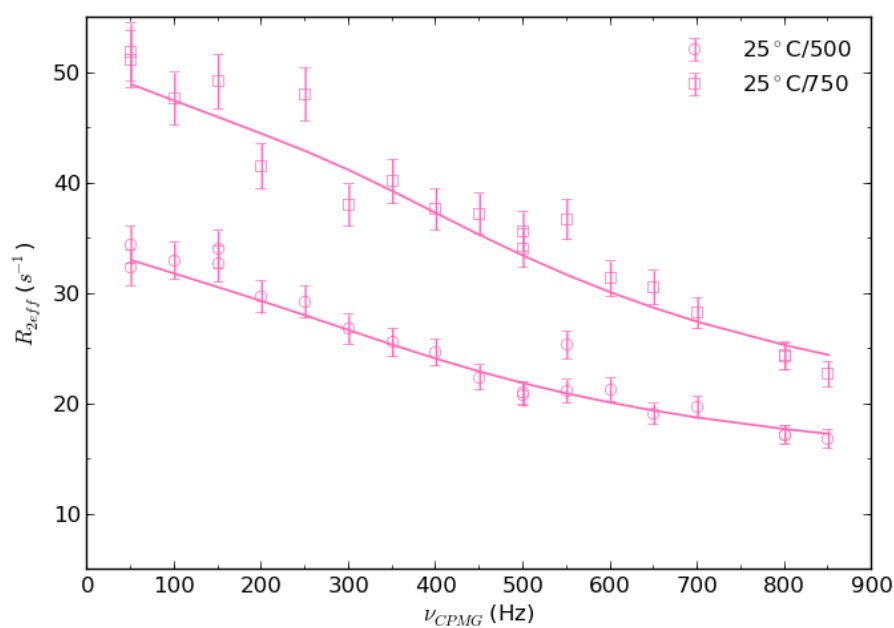
To explore whether the exchange is really that fast, the fits were repeated with a more constrained parameter space. For this p_B was fixed to the previously determined value of 3%. The chemical shift differences were still reproduced as shown in Fig. 5.9. The exchange rate is now 1720 s^{-1} . The large difference in k_{ex} between the two sets of fits puts the initial value in doubt. Further work is likely required to accurately determine the exchange rate. For instance, the exchange rate in absence of exchange, $R_{2,0}$ is known for oxidised cDsbD, as relaxation dispersion experiments were recorded before adding nDsbD (*not shown*). Known values for $R_{2,0}$ have been used before to more accurately determine k_{ex} (Gardino and Kern, 2007) for fast reactions.

The kinetics of association of oxidised cDsbD with oxidised nDsbD

The macroscopic on and off rates are also obtained for the association of oxidised cDsbD with oxidised nDsbD. Dissociation rates of 2430 s^{-1} and 1670 s^{-1} were obtained with p_B



(a) C464



(b) G509

Fig. 5.8 Relaxation dispersion curves at 500 MHz and 750 MHz for C464 and G509. In the fit presented in this figure p_B was allowed to vary. Very similar curves were obtained with $p_B = 3\%$ (not shown).

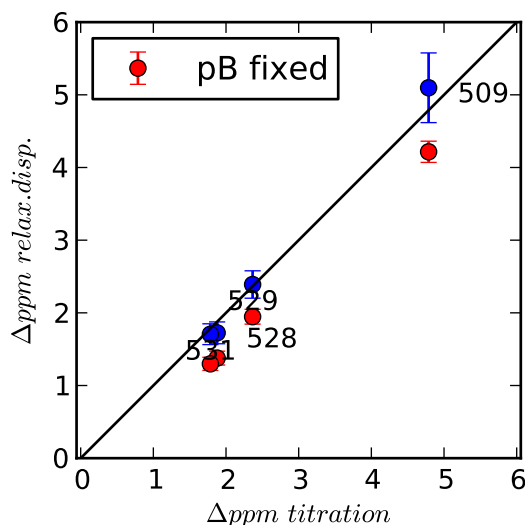


Fig. 5.9 The chemical shifts differences from the relaxation dispersion experiments of the association between oxidised cDsbD and oxidised nDsbD reproduces the previously established absolute chemical shift differences between cDsbD on its own and in the complex with nDsbD as shown by the blue markers. Fit with p_B fixed to 3% also reproduced the chemical shift differences.

either free to vary in the fitting or fixed to 3%. While there is a considerable difference between the values, it is clear that the dissociation rate is fast, even compared to electron-transferring proteins ($k_{off} = 1000 \text{ s}^{-1}$) that are known for their fast kinetics (Ubbink, 2009). For both values of k_{ex} an association rate on the order of $10^6 \text{ M}^{-1} \text{ s}^{-1}$ was determined as shown in Table 5.2. The association rates for C461A and oxidised cDsbD are of similar magnitude, with a larger value for C461A-cDsbD. A more careful consideration of uncertainties may allow discerning whether the rate for C461A-cDsbD is really larger than for oxidised cDsbD.

The temperature dependence of the association of oxidised cDsbD and oxidised nDsbD

To further explore the association reaction between oxidised cDsbD and oxidised nDsbD experiments were repeated at 15 °C. The analysis of the relaxation dispersion curves collected at 25 °C did not reveal any intermediate states in the association reaction, but such states may exist. Comparing relaxation dispersion at different temperatures can be used to understand complicated multi-state exchange processes (Grey et al., 2003). For example, slow exchange processes can be "frozen out" at lower temperatures. Fast exchange processes can sometimes be speeded up further by increasing the temperature, so that they

no longer contribute to the relaxation dispersion curves. This strategy might be applicable for the association of oxidised cDsbD with nDsbD. For example, the curves for C464 in Fig. 5.10b at 15 °C and 25 °C appear to show a slow decay, that levels off around 400 Hz. In addition there is a faster decay that only starts to plateau round 800 Hz.

So far experiments at 15 °C were only recorded at a single magnetic field strength, corresponding to a ^1H resonance frequency of 500 MHz. Before acquiring data at a second higher field strength, no fitting can be attempted. Some tentative qualitative conclusions can be drawn. First, the reaction still occurs at 15 °C as can be seen by comparing the curves at the two temperatures (Fig. 5.10). Secondly, no obvious slowing down was apparent. The barrier might be entropic.

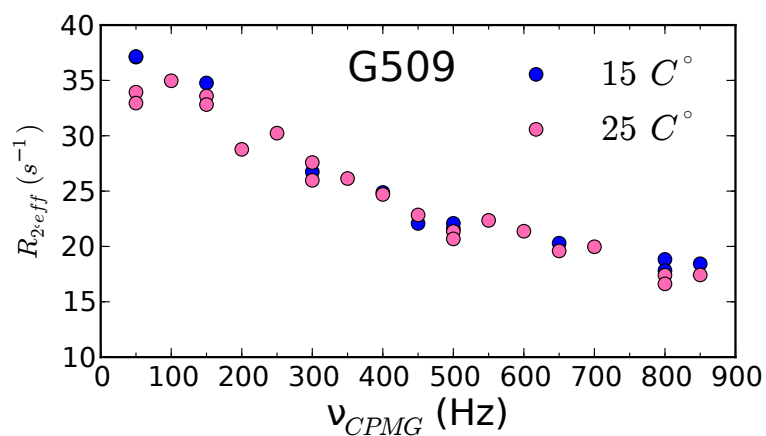
It is not clear whether reducing the temperature would be a good strategy to test for an intermediate state. Detecting intermediate encounter complexes may be hard given the very low concentration of nDsbD. Collecting the relaxation dispersion experiments at different nDsbD concentrations might be a better way to test for reaction intermediates (Sugase et al., 2007a).

5.4 Results: Computational study of the association reaction

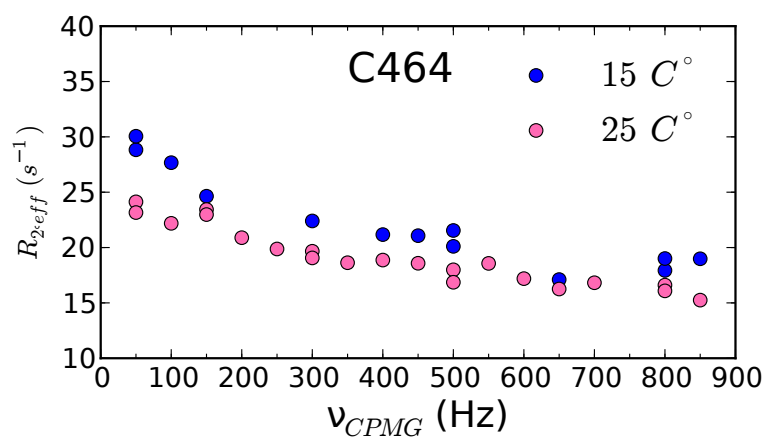
To gain further insight into the association reaction between cDsbD and nDsbD computational methods were employed. First, a relatively simple model for association between two proteins was used. Secondly, the association mechanism was investigated using electrostatics calculations. Thirdly, atomistic MD calculations enabled a highly detailed look at the recognition process between the two domains.

5.4.1 A simplified model of the association reaction

To explore the association reaction computationally a simplified computational model that ignores molecular flexibility was used (Schreiber et al., 2009). In the model by Zhou *et al.*, which they call "transient complex theory" (Alsallaq and Zhou, 2008; Schreiber et al., 2009), the initial association of two proteins is simulated using Brownian dynamics simulations. The stability of structures formed after the initial association, the encounter complex, are then assessed by Poisson-Boltzmann electrostatics calculations. Transient complex calculations were run on the TransComp server (Qin et al., 2011).



(a)



(b)

Fig. 5.10 Relaxation dispersion curves reporting on the association of oxidised ^{15}N labelled cDsbD with unlabelled nDsbD were also collected at 15 $^{\circ}\text{C}$. The curves are compared to experiments conducted at 25 $^{\circ}\text{C}$. Here curves for G509 (a) and C464 (b) are shown.

Using the model, the association rate between an open conformation of oxidised nDsbD and reduced cDsbD was estimated. A basal association rate of $5 \times 10^5 \text{ M}^{-1} \text{ s}^{-1}$ was obtained. This basal rate was calculated using only Brownian dynamics and electrostatics was not included. This may suggest that the shape of cDsbD and nDsbD makes this reaction relatively slow compared to some rates on the order of $10^8 \text{ M}^{-1} \text{ s}^{-1}$ (Radić et al., 1994) or higher (Schreiber et al., 2009) that have been reported in the literature for other protein-protein association reactions.

To test the conclusion, the calculation was repeated for a stable encounter complex found in atomistic MD simulations (final frame of *run1* described below, Section 5.4.3). The basal rate for the formation of the encounter complex should not depend strongly on the molecular detail. In the model by Zhou *et al.* the boundary between the encounter complex and the native complex is defined as the structures that have lost the restrictions on translation and rotation posed in the native state. The overall shape of the two domains rather than molecular detail would be expected to determine the definition of the encounter complex and therefore the rate of encounter complex formation. The loop is closed in the encounter complex structure used and arrangements of the N-terminus of nDsbD, that is flexible according to NMR measurements and MD simulations differs between the two input structures. Note that flexibility is not taken into account in this calculation. A basal association rate of around $1 \times 10^6 \text{ M}^{-1} \text{ s}^{-1}$ was found. This however shows that structural detail, such as whether the loop is open or the structure of a disordered linker, can significantly affect the results from the TransComp server. Nonetheless the results suggest that shape alone might make the association reaction between cDsbD and nDsbD relatively slow. A conformational change is not required to explain why the reaction is relatively slow.

Furthermore electrostatic "steering" did not appear to speed up the reaction (Fig. 5.11). To confirm this however the calculation has to be repeated with different protonation states of D455 and the histidine side-chains. In the calculation presented D455 was de-protonated and the histidine side-chains were charge neutral. D455 is more likely to be protonated in the complex rather than in free cDsbD (Mavridou et al., 2009). It might already be protonated during the association reactions and one might have to take this into account to model the electrostatically driven encounter complex formation. At the experimental pH of 6.5, some of the histidine side-chains will be protonated. One might have to repeat the calculation with different combination of histidine side-chain protonation states to rule out electrostatic rate enhancement.

$$k_a = k_{a0} \exp(-\Delta G_{el}^*/k_B T) / [1 + 10^{-4} \exp(-\Delta G_{el}^*/k_B T)], \text{ at } T = 298 \text{ K}$$

Ionic Strength	k_{a0}	ΔG_{el}^*	k_a
(M)	($M^{-1} s^{-1}$)	(kcal/mol)	($M^{-1} s^{-1}$)
0.100	5.07e+05	0.194	3.65e+05

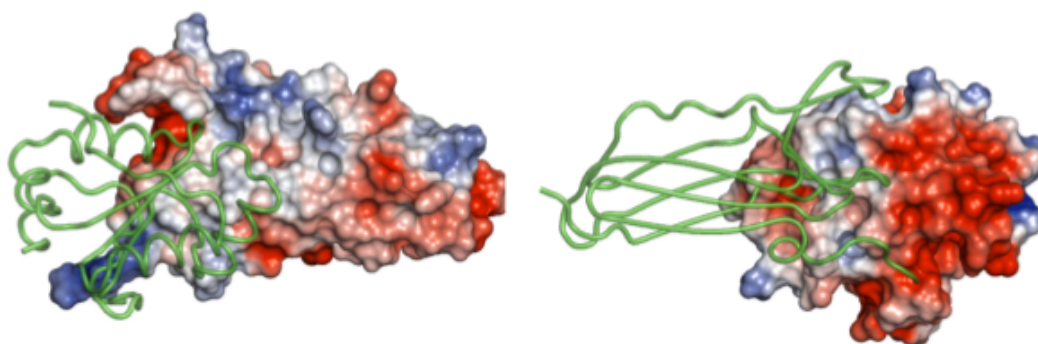


Fig. 5.11 The output from the TransComp server (Qin et al., 2011) of the association of oxidised nDsbD with an open cap loop and reduced cDsbD is shown.

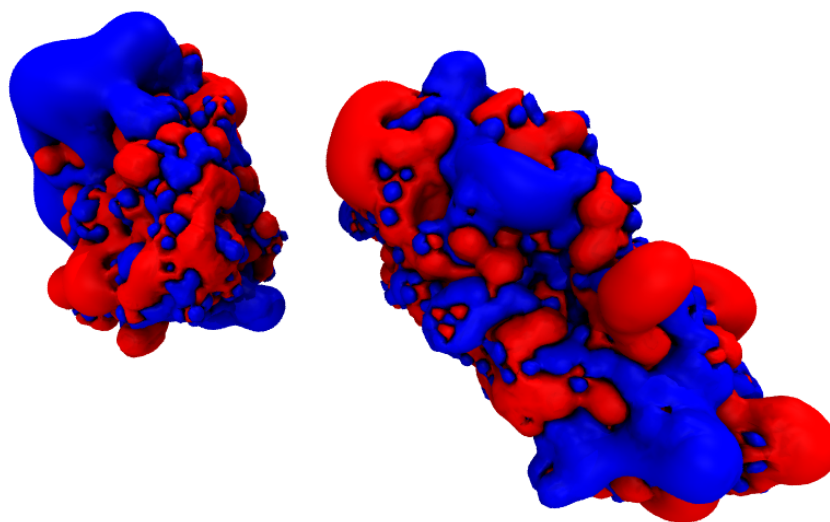


Fig. 5.12 Electrostatic potential surfaces are shown for cDsbD (to the left) and nDsbD (to right). The iso-contour surfaces are shown at +ve and -ve $2kTe^-$ in blue and red, where kT is the thermal energy at 25 °C and e is the charge of an electron.

5.4.2 Electrostatics calculations

Electrostatics calculations were used to probe the association mechanism between cDsbD and nDsbD. Here the focus was on the later stages of the association process when the domains are already fairly close. These latter stages can be investigated by MD, as described below. The electrostatic potential surface was calculated for reduced cDsbD and oxidised nDsbD with the centre of masses separated by about 50 Å. The calculated surfaces (Fig. 5.12) already showed that there is a large negatively charged patch around the cap loop region. At the same time, there is not a clear complementary positively charged surface on cDsbD. The nDsbD binding site of cDsbD contains both negatively and positively charged regions. This calculation also suggested that the reaction is not strongly steered by electrostatics.

5.4.3 Molecular dynamics simulations of the association of cDsbD and nDsbD

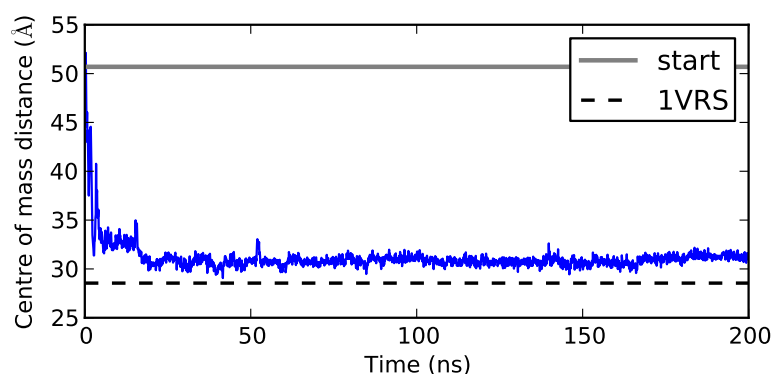
To explore the mechanism of the association between cDsbD and nDsbD, and especially the opening of the cap loop in this reaction, molecular dynamics were used. The starting struc-

ture for MD calculations was generated based on the 1VRS crystal structure. Importantly, the open structure of nDsbD in the crystal structure was replaced by a closed structure of nDsbD. The centre of masses of nDsbD and cDsbD were moved apart in the z direction by 22 Å. At the beginning of the simulations, the centre of masses were hence separated about 50 Å (Fig. 5.13a). The fairly large separation between the two domains was chosen deliberately. This allows the proteins to associate not just by head-on collision, closing the separation introduced by translating the domains, but also by more complicated association pathways. More complicated association pathways and intermediate structures that lie on them might be critical in the opening of the cap loop.

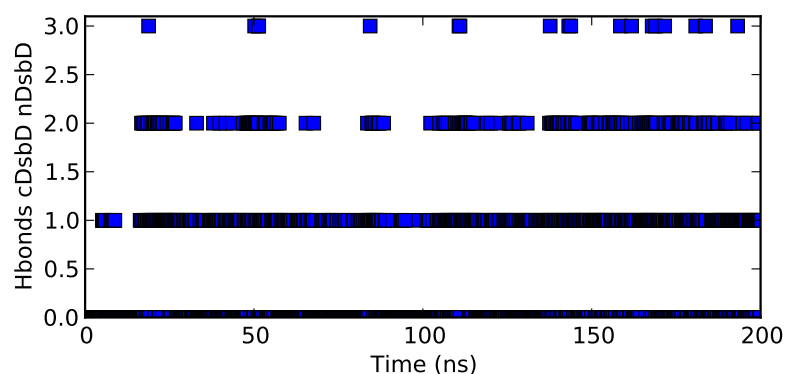
In the simulations, cDsbD and nDsbD started to interact. Differences were observed between two simulations, *run1* and *run2* (Table 5.1). In *run1*, nDsbD and cDsbD quickly associate, within about 30 ns, as shown in Fig. 5.13a. The two domains maintain the interactions for the remainder of this 200 ns trajectory. The complex formed resembles the 1VRS crystal structure. Residues that show large chemical shift changes in the relaxation dispersion experiments of the association reaction are in the interface of the two domains (Fig. 5.14). The active site of cDsbD faces the catalytic site of nDsbD and the cap loop. The cap loop however remained closed during the entire trajectory. This complex is stabilised by up to three hydrogen bonds (Fig. 5.13b) between the two domains. A stable hydrogen bond was for example formed between the amine of C109 of nDsbD and the backbone carbonyl oxygen of L509 (Fig. 5.13c). A similar interaction is seen in the 1VRS crystal structure between the amine of C109 and the carbonyl of L510. Yet the three native hydrogen bonds were not formed in the 200 ns trajectory (Rozhkova et al., 2004). This complex might constitute a native-like encounter complex.

The second simulation showed a different picture. After an initial association, which reduced the centre of mass distance to 40 Å, nDsbD and cDsbD appeared to repel each other (Fig. 5.15). In *run2*, no salt was added to the simulation system and electrostatic interaction are thus not shielded by the ions in the solvent. The repulsion may be due to electrostatic interactions. Electrostatic calculations for the two domains have already shown that they are both largely negatively charged. Nonetheless an encounter complex is eventually formed after about 50 ns of simulation.

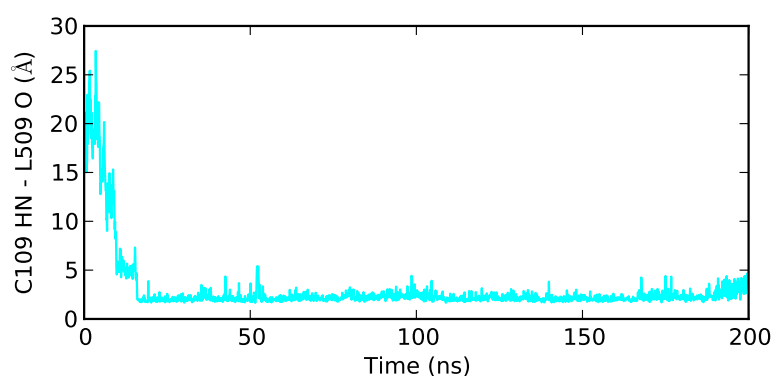
Interestingly, this complex behaves very differently from the one in *run1*. The distance between the two domains fluctuates even after the initial formation of the complex as shown by Fig. 5.15. The complex is very dynamic. Closely packed configurations, such as highlighted in Fig. 5.17a exchange with much looser arrangements, such as the one presented in Fig. 5.17b where the inter-facial region between the two domains is solvated. Moreover,



(a) The centre of mass distance between nDsbD and cDsbD in *run1*.



(b) The total number of inter domain hydrogen bonds fluctuates in *run1*. The main hydrogen bonds are between HN C109 and O L509 , T529 and E69, F531 and E69, as well as between R8 and V507.



(c) A stable hydrogen bond is formed between C109 and L509.

Fig. 5.13 An encounter complex is formed in *run1*. The centre of mass distance between the two domains is shown in (a). The hydrogen bond between C109 and L509 is analysed in (c). The distance between HN C109 and O L509 is on average 2.3 Å over the last 150 ns.

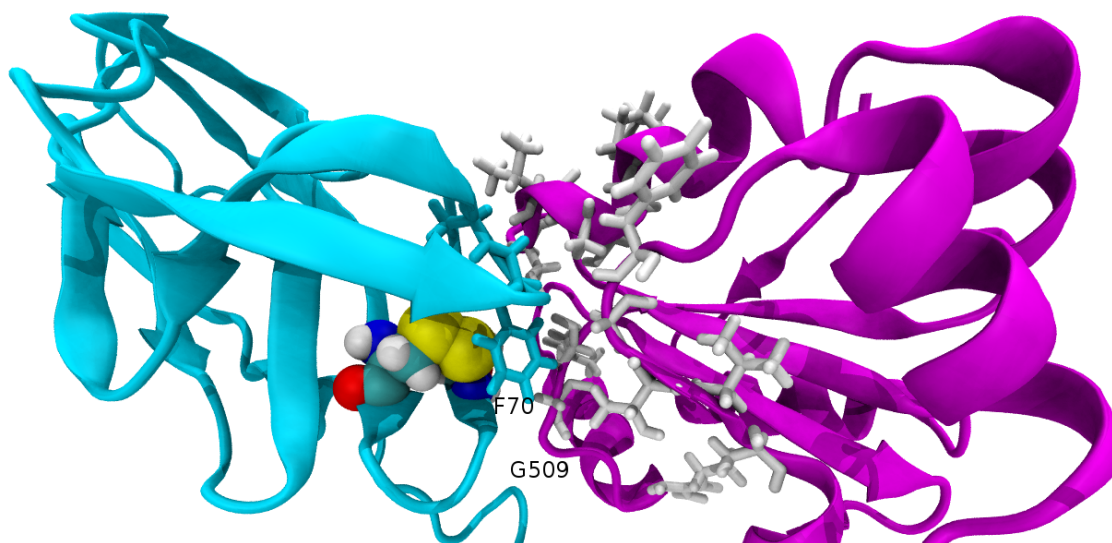


Fig. 5.14 The final structure from *run1* is shown, cDsbD is rendered in magenta and nDsbD in cyan. Residues of cDsbD for which chemical shift differences larger than 1 ppm were found in the relaxation dispersion experiments are highlighted in white.

visual inspection of the trajectory showed that different parts of the two proteins interact with each other over time. For example at 78 ns the C-terminus of cDsbD is in contact with the segment preceding the cap loop of nDsbD (Fig. 5.16a). About 10 ns later, the C-terminus of cDsbD has moved along the surface of nDsbD to interact with its C-terminal region (Fig. 5.16b).

Non-native interactions were formed in these encounter complexes. For example at 88 ns, the side-chain of Q521 and the carbonyl oxygen of G63 (Fig. 5.16a) form a close interaction that is stable for several ns. The interaction was broken as cDsbD moves away from the cap loop region. This interaction can be considered a non-native hydrogen bond. Previous studies have highlighted (Schreiber et al., 2009; Sugase et al., 2007a; Ubbink, 2009) electrostatic and hydrophobic interactions that stabilise encounter complexes. The simulations presented here show that non-native hydrogen bonds could stabilise encounter complexes.

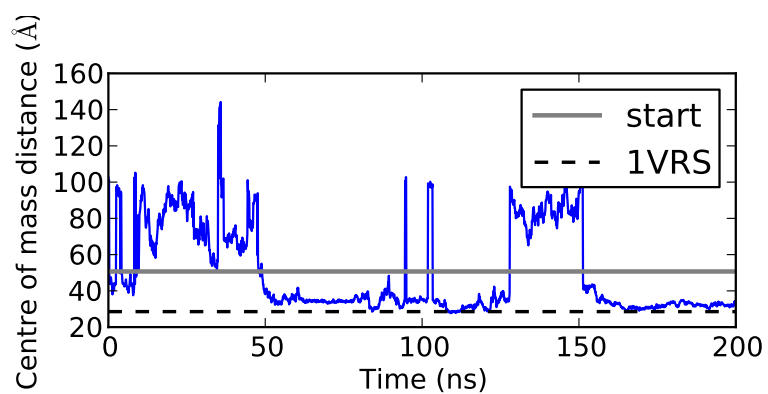


Fig. 5.15 The centre of mass distance between nDsbD and cDsbD in *run2*.

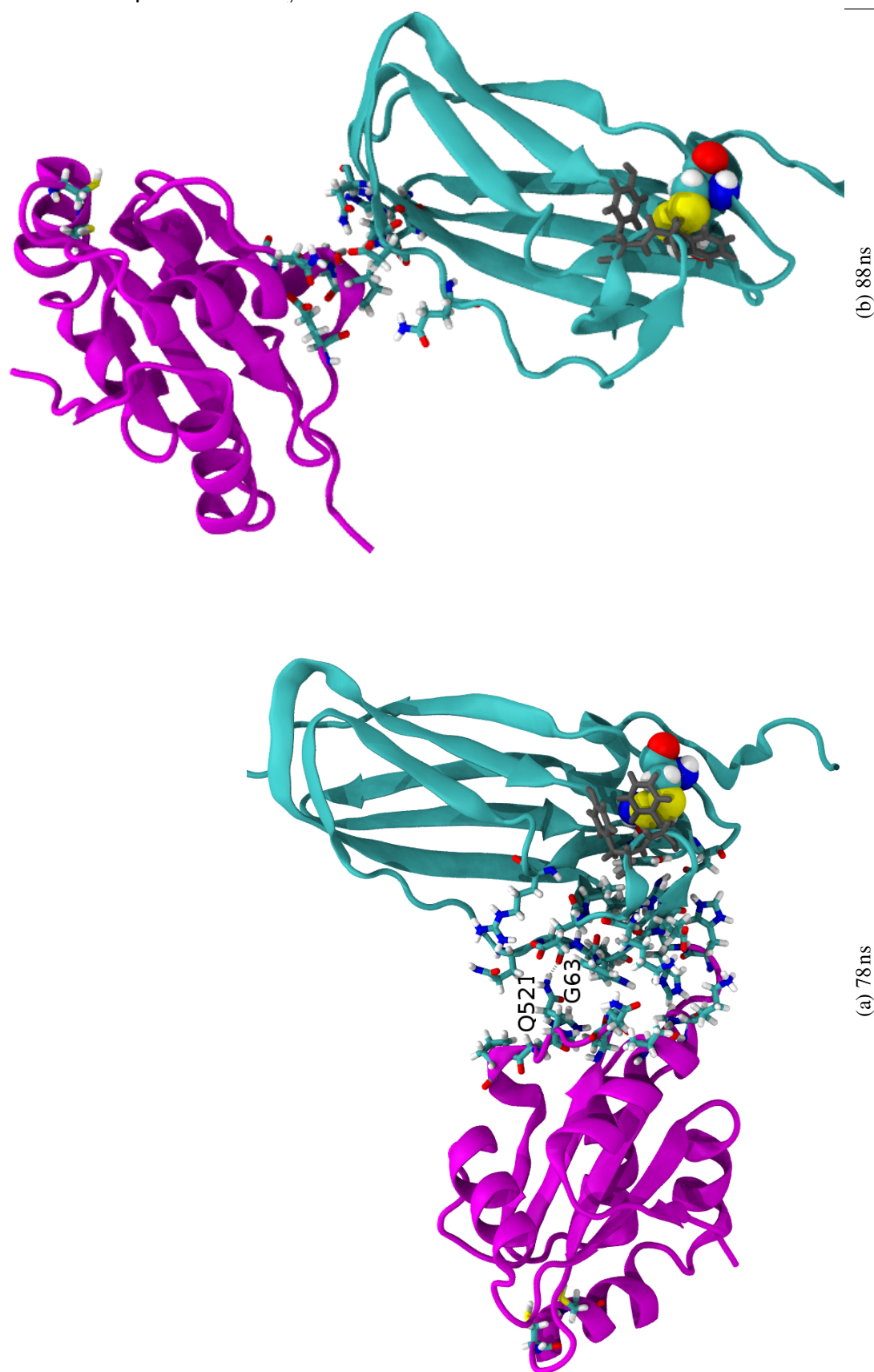


Fig. 5.16 This image highlights close, non-native interactions between the C-terminus of cDsbD and the region preceding the cap loop of nDsbD at 78 ns of simulation time. At 88 ns, the C-termini of both domains interact. The two snapshots illustrate the dynamic nature of the encounter complex in *run2*. The cap loop (in grey) remained packed onto the disulphide bond of nDsbD (vdW representation).

5.5 Discussion

Here, the association of cDsbD and nDsbD was probed by NMR and computer simulations. The transmembrane domain of DsbD moves reductant from the cytoplasm into the oxidising periplasm of Gram-negative bacteria. The reductant transported across the membrane is first received by cDsbD which then associates with oxidised nDsbD. The complex falls apart after reductant transfer and nDsbD distributes electrons to the disulphide bond isomerisation as well as the cytochrome *c* maturation pathways.

5.5.1 The association of cDsbD with nDsbD

NMR relaxation dispersion experiments gave important insight into the physiological reaction between reduced cDsbD and oxidised nDsbD. Here C461A-cDsbD was used to mimic reduced cDsbD. This prevents thiol-disulphide reactions which would complicate the analysis of the association kinetics. No intermediate states had to be evoked to explain the data. The reaction might be a simple two-state reaction, but one should note that the low concentrations of nDsbD used might mean that any intermediates would be sparsely populated and therefore remained undetected. Intermediates could thus not be ruled out conclusively.

Macroscopic on and off rates were determined from the relaxation dispersion experiments. The association rates for both C461A and oxidised cDsbD with oxidised nDsbD were relatively fast, around $2 \times 10^6 \text{ M}^{-1} \text{ s}^{-1}$. The association by C461A appeared to be more rapid, but a more careful analysis of experimental and fitting uncertainties is required before one can be sure of this. The k_{on} rates are approximately an order of magnitude faster than the rate constant for the reductant transfer reaction between reduced cDsbD and oxidised nDsbD (Rozhkova et al., 2004). This might suggest that the reductant transfer between the two periplasmic domains is not limited by the kinetics of the complex formation but by the covalent chemistry of the thiol-disulphide exchange reaction. The reductant transfer between cDsbD and nDsbD might not be "catalytically perfect" (Schreiber et al., 2009).

The off rates from the relaxation dispersion experiments gave further mechanistic insight. C461A dissociates from the complex with a k_{off} rate of 210 s^{-1} . Oxidised cDsbD dissociates much more rapidly from the complex with oxidised nDsbD with a k_{off} rate of approximately 1670 s^{-1} . The rate is fast even in comparison with the typical dissociation rates of electron transfer proteins ($k_{off} = 1000 \text{ s}^{-1}$) which are known for their very rapid dissociation and turnover (Ubbink, 2009). The physiologically relevant rate is likely even larger than the one determined for the complex of oxidised cDsbD and oxidised nDsbD. After the

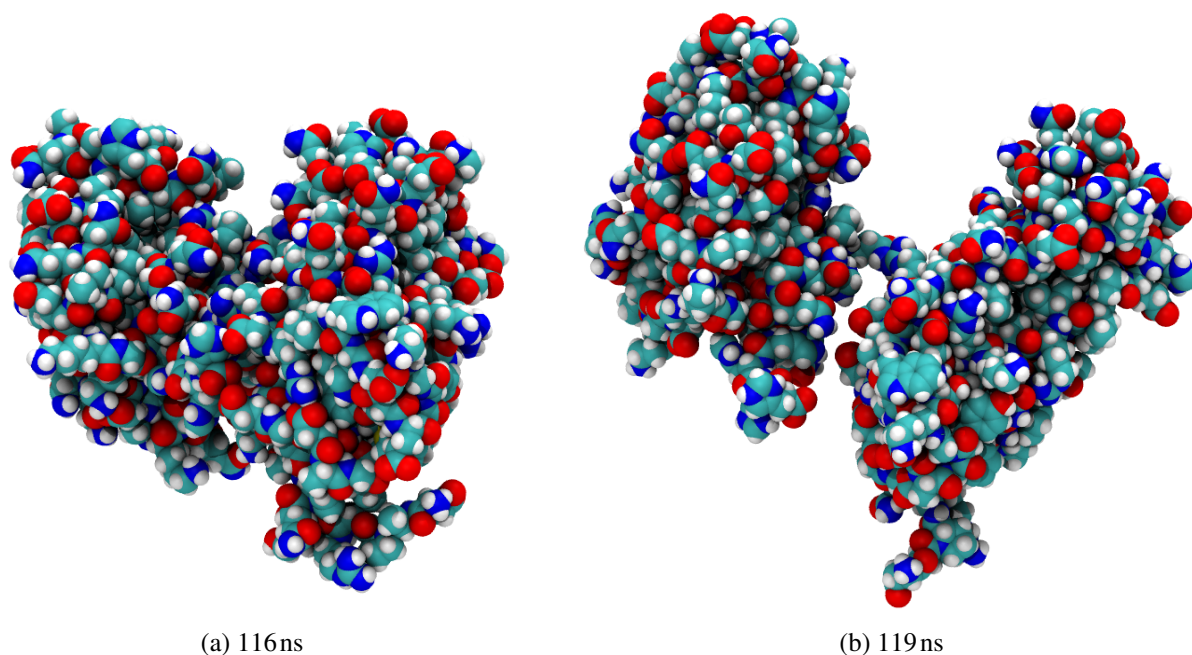


Fig. 5.17 The encounter complex in *run2* switches between closely packed (a) and loose configurations (b) where the interface between the two domains is largely solvated. The domain to the left is cDsbD and the one to the right is nDsbD, in both (a) and (b).

reductant transfer, oxidised cDsbD and reduced nDsbD interact with an even weaker affinity than the complex of the two oxidised domains, with a K_d in the range of 2.4 mM–3.5 mM instead of 800 μ M (Mavridou et al., 2011). This reduction in affinity may imply an even larger k_{off} given that $K_d = \frac{k_{off}}{k_{on}}$. By determining the rates it becomes clear that observed redox-state dependence in the interactions between the periplasmic domains of cDsbD indeed lead to rapid dissociation of the complex after reductant transfer. This will limit the extent to which the thermodynamically feasible back reaction can occur. The redox-state dependent increase in the dissociation rate supports the hypothesis that the redox-state dependent interactions between cDsbD and nDsbD enhance turnover in the pathway.

Computer models gave direct insight into the mechanism for the association of oxidised and reduced nDsbD. A simplified model, "transient complex theory" (Alsallaq and Zhou, 2008), gave rates of similar magnitude to the experiments for the formation of encounter complexes. This does not rule out that cap loop opening slows down the reaction. Based on this, the finding that the cap loop can open spontaneously in oxidised nDsbD and that no long-lived intermediates were detected so far, one can however speculate that cap loop opening does not slow down the reaction significantly. The cap loop might open frequently enough before the association reaction or within the encounter complex and the native com-

plex would then be formed. The loop opening in the encounter complex might be spontaneous as in oxidised nDsbD on its own. The loop did not open in the MD simulations of the association reaction which might be due to insufficient sampling. The native interaction between F70 and V462 was not formed. This interaction is crucial for the stability of the complex, but its effect on the mechanism, whether V462 opens the cap loop as proposed (Mavridou et al., 2011) or stabilises an open conformation, remains unclear.

The electrostatic interactions might not steer the two domains towards forming the complex and may not significantly increase the reaction rate. The electrostatic potential surfaces for reduced cDsbD and oxidised nDsbD demonstrated that the binding interfaces on the two proteins both contain significant negatively charged regions. *In vivo*, 10-15 amino acid long linkers keep the domains close to each other (Mavridou et al., 2011; Rozhkova and Glockshuber, 2007) and the trans-membrane domain of DsbD, and this rather than the electrostatics of the two periplasmic domains might speed up the reaction.

5.5.2 The encounter complex

The MD simulations presented here give microscopic insight into the encounter complex between cDsbD and nDsbD. Stable encounter complexes that somewhat resemble the native complex (Rozhkova et al., 2004) were found in one trajectory. However the cap loop remained closed. Very dynamic encounter complexes with rapidly changing arrangements were also observed in a second simulation. Both types of encounter complexes were found (Kim et al., 2008) in a previous simulation study based on NMR paramagnetic relaxation enhancement measurements. The dynamic encounter complexes presented here sample both arrangements where the domains are largely solvated and configurations where they are packed together closely. Non-native interactions such as hydrogen bonds may stabilise these complexes in addition to electrostatics (Ubbink, 2009) and hydrophobic interactions (Sugase et al., 2007a).

A recent docking study (Kozakov et al., 2014) highlighted that proteins can move along the surfaces of binding partners, searching for their binding sites. Non-specific interactions facilitate this and would thus enhance the protein-protein association. Here, such a diffusion on the surface of the target protein was observed directly in MD simulations which unlike docking calculation capture the dynamics and kinetics of molecular processes.

5.5.3 Further experimental work

An important goal for further experimental work will be the characterisation of any intermediate states on the path to complex formation. Intermediates were not detected experimentally so far. One approach would be to revisit the association reaction between oxidised cDsbD and oxidised nDsbD. The complex is weaker than between the C461A-cDsbD and oxidised nDsbD. Thus a larger fraction of the ensemble might be in an encounter complex. One could record experiments at multiple temperatures or different nDsbD concentrations to dissect a putative multi-state association reaction mechanism (Grey et al., 2003; Sugase et al., 2007b).

For the association reaction of C461A-cDsbD, 2D line-shape analysis (Cavanagh et al., 2007; Waudby and Christodoulou, 2014) could be used to rule out the long-lived intermediate structures conclusively. A suitable dataset was already collected by Prof. Redfield. In these experiments the whole titration was followed. As the titration series progresses the population of the bound-state increases. It might be possible to conclusively test for the existence of long-lived intermediate structures under these conditions.

Encounter complexes could be directly detected by NMR in paramagnetic relaxation enhancement (PRE) experiments (Tang et al., 2006). A nitroxide spin-label will be introduced on C464 of cDsbD. At this position, the spin label will not prevent binding to nDsbD. The high sensitivity of PRE experiments enables the detection of even very weak interactions. For example, a complex between cDsbD and reduced nDsbD has not been detected previously. Transient non-specific encounter complexes might yet be formed and these will be detected by the PRE measurements. Adding spin-labelled cDsbD to oxidised nDsbD will lead not only to the formation of non-specific complexes as for reduced nDsbD but also to the stereo-specific complexes. Stereo-specific complexes and non-specific binding can be distinguished. PRE effects due to specific binding level off as the binding site becomes saturated (Fawzi et al., 2010). Stereo-specific encounter complexes would be critical if V462 forces the closed cap loop to open in an induced-fit mechanism as previously proposed. It might however be difficult to distinguish stereoscopic encounter complexes from the native complex.

5.5.4 Further computational work

There is also scope for further computational studies of the association reaction. In the MD simulations the association of nDsbD with an open cap loop and reduced cDsbD could be studied to directly test whether the complex forms through an induced fit or a

conformational-selection mechanism. However, very long simulations would probably be required to draw conclusions from this with confidence. For instance, observing spontaneous loop closure would not rule out a conformational-selection mechanism in the absence of exhaustive sampling. The redox-state dependence of the interaction affinities could be probed more easily in MD simulations. Previous modelling has suggested that steric conflicts and packaging constraints may contribute to the redox-state dependent interaction between the periplasmic domains of DsbD (Mavridou et al., 2011). Two thiols groups fill more space than a disulphide bond. As a consequence the three native hydrogen bonds between cDsbD and nDsbD were fully maintained for the complex of reduced cDsbD and oxidised nDsbD, but not for other redox-state combinations. To explore this further simulations will be setup based on the 1VRS crystal structure of the covalent intermediate in the reductant transfer between cDsbD and nDsbD. The redox state of the active site will be varied and following simulations will be started with the complex in the following combination of redox states that have been studied experimentally:

1. reduced cDsbD and oxidised nDsbD, the complex preceding the reductant transfer;
2. oxidised cDsbD and oxidised nDsbD;
3. oxidised cDsbD and reduced nDsbD, the complex after the reductant transfer;
4. reduced cDsbD and reduced nDsbD.

The simulation may give structural explanations for why the stability of the complex follows the observed trend (1.>2.>3.>4.) and might confirm the importance of the packing constraints.

Chapter 6

Studying conformational changes in the transport cycle of LacY

6.1 Introduction

Here, large-scale conformational changes are studied in the Major Facilitator Superfamily (MFS) transporter Lactose permease (LacY). Using a novel method, Dynamics Importance Sampling (DIMS), I was able to test proposed mechanisms for the conformational change between cytoplasmic and periplasmic open states that underlies the transport mechanism. The simulations were validated by comparison to experimental Double Electron-Electron Resonance experiments and MFS transporter crystal structures.

6.1.1 LacY is a paradigm for membrane transport

Lactose permease (LacY) is the best studied member of the major facilitator superfamily (MFS) of secondary transporters (Forrest et al., 2011). The MFS is the largest family of transporters (Hunter et al., 2012). LacY of *Escherichia coli* is a symporter and uses the proton (H^+) electrochemical gradient to transport galactosides, such as lactose, from the periplasm into the cytoplasm (Guan and Kaback, 2006). The structural elements that allow or prevent exchange of sugar and H^+ between the central binding site and either the cytoplasm or the periplasm are often referred to as “gates” (Gadsby, 2009). LacY transports substrate by switching from a conformation with an open cytoplasmic gate and a closed periplasmic gate, as depicted in Fig. 6.1, to a periplasmic open conformation in which the states of the gates are reversed. This alternating access mechanism (Jardetzky, 1966; Mitchell, 1967) has been detected (Kaback et al., 2011) in site-directed alkylation (Kaback

et al., 2007) and cross-linking (Zhou et al., 2008), tryptophan quenching (Smirnova et al., 2009), single molecule FRET (Majumdar et al., 2007) and double electron-electron spin resonance (DEER) experiments (Smirnova et al., 2007).

6.1.2 The transport cycle of LacY

The transition between the cytoplasmic and periplasmic open states of LacY is not understood in full atomic detail. Partly this reflects a lack of structural data. There are several crystal structures for cytoplasmic open LacY (Abramson et al., 2003; Chaptal et al., 2011; Mirza et al., 2006), but no crystal structure for periplasmic open LacY was available at the time this project was carried out. Kumar et al., 2014 recently published the first periplasmic open crystal structure which will lead to interesting new investigations. Before that the periplasmic open state was studied based on the periplasmic open crystal structure of the distantly related MFS transporter FucP from *E. coli* (Dang et al., 2010). Radestock *et al.* have built two models of periplasmic open LacY: the first is based on its homology to the FucP structure and the second uses the internal inverted repeat symmetry of the transporter (Radestock and Forrest, 2011). These models were used to represent the periplasmic open state in this study.

Biochemical studies have identified key events in the transport (Guan and Kaback, 2006). First, cytoplasmic open LacY releases the transported sugar to the cytosol. The subsequent deprotonation of E325 initiates the transition to the periplasmic open conformation (Andersson et al., 2012) (The positioning of the residues crucial for the transport mechanism is highlighted in Fig. 6.1B). Once E325 is de-protonated it can interact with R302 and this salt-bridge may be important in the transition. Periplasmic open LacY is a minor conformer in the absence of substrate (Guan and Kaback, 2006). Protonation of H322 and substrate binding then stabilises the periplasmic open state and initiates the return to the cytoplasmic open conformation, thereby carrying sugar and H^+ across the membrane.

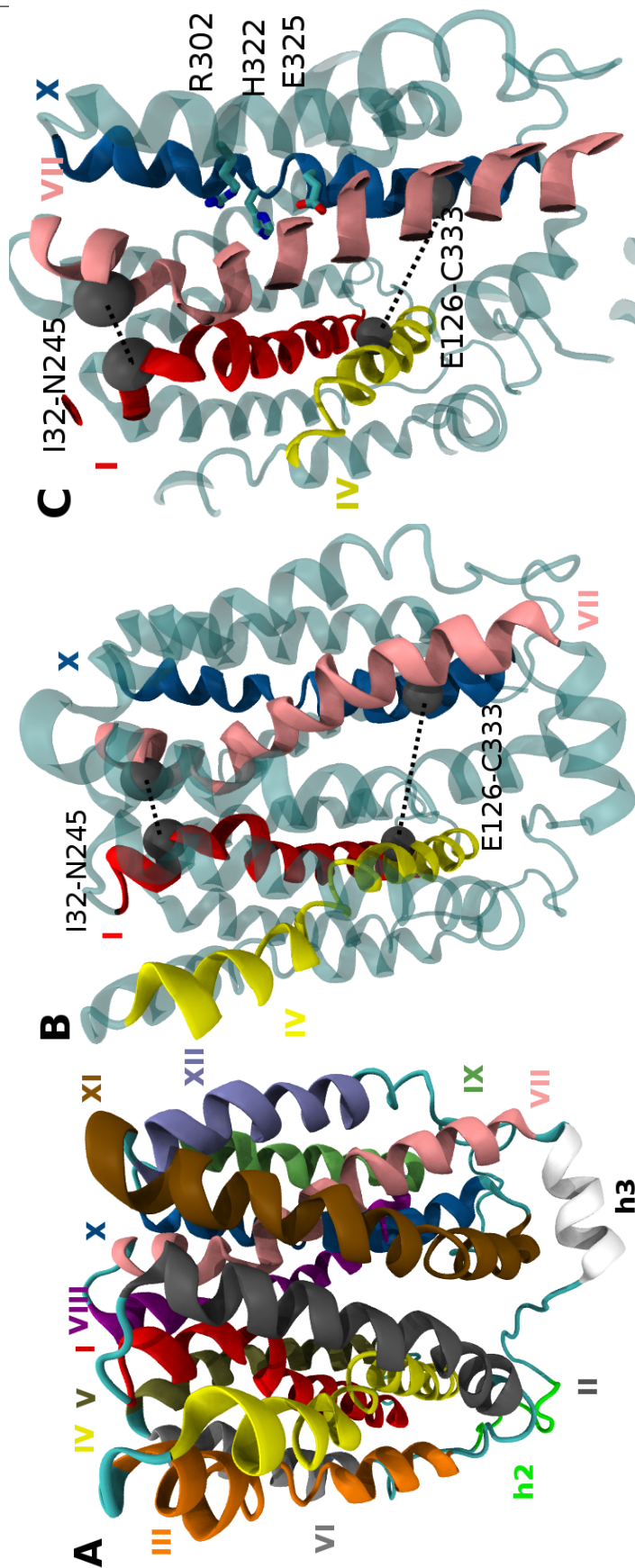


Fig. 6.1 The wild-type cytoplasmic open crystal structure of LacY (PDB: 2V8N) was the starting point for this study. LacY has two domains, with six transmembrane helices each, that surround the central cavity (A). The N-terminal domain is made up of transmembrane helices I (red), II (grey), III (orange), IV (yellow), V (tan) and VI (silver). Transmembrane helices VII (pink), VIII (purple), IX (lime green), X (blue), XI (ochre) and XII (iceblue) constitute the C-terminal domain. A linker connects the two domains and contains two segments, h2 and h3, that can adopt helical conformations. The structure is closed to the periplasmic side (B) and open to the cytoplasm. The periplasmic gate is formed by the close approach of helices I and VII. The state of the periplasmic gate in MD simulations can be monitored by tracking the $C_{\alpha} - C_{\alpha}$ distance between I32 of helix I and N245 of helix VII. Helices IV and X form the cytoplasmic gate that was tracked *via* the $C_{\alpha} - C_{\alpha}$ distance between E126 (helix IV) and C333 (helix X). The C_{α} atoms used to follow the state of the gates are shown as grey spheres. The side chains of R302, H322 and E325 are critical for the transport mechanism of LacY and are highlighted in C, which provides a cut-through view of the crystal structure.

6.1.3 The mechanism of the conformational transition in LacY

In addition to biochemical evidence, functional considerations provide some clues about the mechanism of the conformational transition. A transporter with both gates open would dissipate vital trans-membrane gradients, such as the proton gradient in bacteria or the sodium gradient in higher eukaryotes to the detriment of the cell that has expended 20% of its energy to establish the gradient (Rolfe and Brown, 1997). The two gates of a transporter should therefore never be open at the same time (Gadsby, 2009) and it is thought that, to prevent this, the movement of the gates must occur in a precise order, so that one gate closes before the other starts to open as in the operation of an airlock (Fig. 6.2A) (Forrest, 2013). Such an “airlock model” for transport predicts an occluded intermediate conformation that is closed to ion and substrate permeation from both cytoplasm and periplasm. The occluded conformation is thus fundamental to the transport mechanism. Preventing both gates from being open at the same time necessitates coupling between the movements of the two gates, so that only the closure of one gate allows opening of the other gate. The case where one gate must close fully before the other can begin to open can be described as a *sequential motion*. In this case the transporter passes through intermediates that are fully occluded.

It has been hypothesised that a rigid-body movement of the N- and C-terminal domains of LacY relative to one another (Fig. 6.2B) would allow the cytoplasmic and periplasmic open states to interconvert (Dang et al., 2010). This model is called the “rocker switch mechanism” (Abramson et al., 2003; Forrest et al., 2011). The rigidity of the domains would provide a direct mechanical linkage between the two gates so that the closure of one gate would occur simultaneously with the opening of the other. This can be thought of as a *simultaneous motion*. A truly rigid motion of the two domains would probably not be able to form a fully occluded conformation. Allowing the domains some flexibility would allow the formation of an occluded conformation, ruling out the dissipation of transmembrane potentials through channel-like conformations. Hydrogen/deuterium exchange experiments have established that LacY is more dynamic than many other membrane proteins and is therefore unlikely to contain large, rigid domains (le Coutre et al., 1998; Patzlaff et al., 1998; Sayeed and Baenziger, 2009). The conformational flexibility of LacY has also been highlighted in NMR experiments and in molecular dynamics (MD) simulations (Bennett et al., 2006; Holyoake and Sansom, 2007).

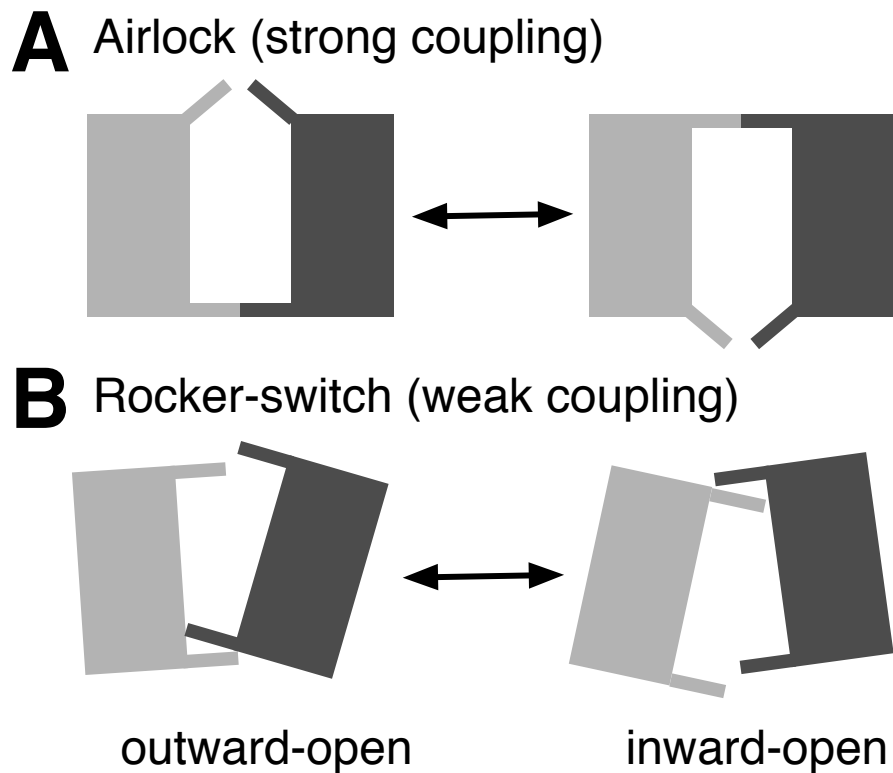


Fig. 6.2 Transport in a MFS transporter such as LacY might occur in an "air-lock" mechanism (A) or in a rocker-switch mechanism (B). A transporter consisting of two domains (light and dark grey) may undergo conformational transitions that expose a central substrate binding site alternatively to the extra- and the intracellular side. (A) In the airlock model, two gates open and close sequentially to form an outward open and inward open conformation. In order to prevent formation of a leak state, the gates must be strongly coupled so that they never open at the same time. (B) In the rocker-switch model the two domains move simultaneously as rigid bodies relative to each other and so form the outward open and inward open states. No additional coupling is required to prevent formation of a leaky transporter.

6.1.4 Evidence for an occluded intermediate state

In the absence of direct evidence one cannot be certain if LacY passes through an occluded intermediate and, if it does, what such an intermediate would look like. Double electron-electron spin resonance (DEER, also known as PELDOR) measurements for a range of residue pairs on LacY have been interpreted to demonstrate the existence of an intermediate occluded state (McHaourab et al., 2011; Smirnova et al., 2009) and have recently been used to propose a model of an occluded conformation of LacY (Madej et al., 2012). To date, occluded conformations of five MFS transporters at atomic resolution have been published. These are EmrD (Yin et al., 2006), NarU (Yan et al., 2013), Xyle (Quistgaard et al., 2013; Yan et al., 2013; all from *E. coli*), PepT_{So} (Newstead et al., 2010, from *Shewanella oneidensis*) and PiPT (Pedersen et al., 2013, from the fungus *Piriformospora indica*). In addition, an occluded conformation of the MFS transporter OxIT from *Oxalobacter formigenes* was resolved by electron microscopy, albeit at a lower resolution of 6.5 Å (Hirai and Subramaniam, 2004).

6.1.5 Studying conformational changes in the transport cycle of LacY

Computer simulations can suggest intermediate conformations and plausible transition mechanisms which may then be compared to experimental data (Stansfeld and Sansom, 2011). The crystal structures of LacY and GlpT transporters are in cytoplasmic open conformations. The cytoplasmic gate of both structures has been observed to undergo partial closure in equilibrium MD simulations, thereby forming partially occluded conformations (Anderson et al., 2012; Enkavi and Tajkhorshid, 2010; Holyoake and Sansom, 2007). The turnover number for LacY is 16 s^{-1} to 21 s^{-1} (Viitanen et al., 1984), thus on average one transition every $50 \times 10^3\text{ }\mu\text{s}$ would be expected. At present standard atomic-resolution equilibrium MD simulations typically yield trajectories of the length of up to $1\text{ }\mu\text{s}$ and hence one cannot readily characterise the kinetics of transport using such simulations. Reducing the level of detail modelled is one way to overcome this challenge and this approach enabled the transition between cytoplasmic and periplasmic open FucP to be studied using an elastic network model (Chang et al., 2011), which hinted at the existence of an occluded intermediate.

Alternatively, fully atomistic approaches that bias the sampling, such as stratified umbrella sampling (Fowler and Sansom, 2013), dynamic importance sampling (DIMS) MD (Beckstein et al., 2009; Perilla et al., 2011; Woolf, 1998), or self-guided Langevin dynamics (SGLD) (Wu and Brooks, 2003) can be used to study such conformational changes. Pendse *et al.* used SGLD to simulate the transition from cytoplasmic open to periplasmic open

LacY (Pendse et al., 2010). Although transition intermediates obtained from the SGLD simulations contained both cytoplasmic and periplasmic constrictions, no detailed study of the likely functional relevance of these structures was presented. To study the functional implications of such intermediate structures and the transition mechanism itself, I have used DIMS MD to generate transitions and transition intermediates for the conformational change between cytoplasmic and periplasmic open LacY. DIMS MD has been successfully applied to study how other membrane proteins, such as a potassium ion channel (Denning and Woolf, 2010) and a transporter (Shimamura et al., 2010), change conformation.

6.1.6 Aims

In this chapter a critical step in the transport cycle of LacY, the switch between the cytoplasmic open and the periplasmic open conformation is investigated. Equilibrium MD simulations were used to understand the intrinsic flexibility of the cytoplasmic and periplasmic open conformations. To understand the conformational change itself Dynamics Importance Sampling (DIMS) MD calculations were employed to study this slow process. A model for the occluded state was obtained from DIMS MD and subsequent equilibrium MD simulations. The simulations and the occluded state model in particular were validated by comparison to experimental data. The trajectories were compared to experimental DEER measurements. A distance-based order parameter was derived to compare the simulations and the occluded state model to the known crystal structures of MFS transporters.

6.2 Methods

6.2.1 Membrane protein insertion

The LacY wild-type crystal structure (pdb code 2V8N) was downloaded from the Protein Data Bank (Guan et al., 2007). The homology model from the Forrest group for periplasmic open LacY (PM0077183), which is based on the periplasmic open structure of FucP, was obtained from the protein model database (Castrignanò et al., 2006; Radestock and Forrest, 2011). In all structures His322 was uncharged as suggested by the proposed mechanism (Guan and Kaback, 2006); the remaining hydrogen, bonded to the imidazole nitrogens, was attached to N_ε, except for simulations *CopenB1* and its continuation *CopenB2* in which N_δ was selected by the pdb2gmx tool (Hess et al., 2008) as a more favourable position. These structures, as well as a model for a fully occluded structure of LacY obtained from DIMS simulations, were inserted into 1-palmitoyl,2-oleoyl-sn-glycero-3-phosphocholine (POPC)

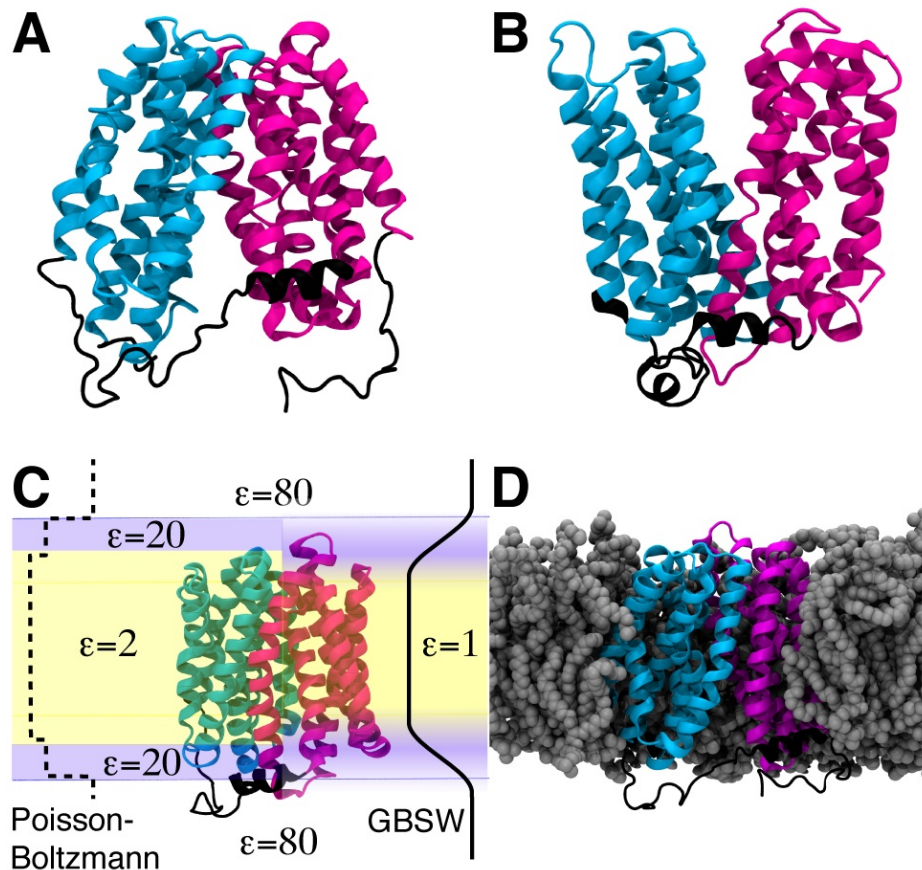


Fig. 6.3 The atomistic simulations setup employed in this study of the conformational changes of LacY. The wild-type crystal structure of cytoplasmic open LacY (PDB 28VN) (A) (Guan et al., 2007) and a FucP based homology model for periplasmic open LacY (PMDB PM0077183) (B) (Radestock and Forrest, 2011) used in atomistic simulations are shown with the periplasmic face at the top and the cytoplasmic face at the bottom. The N-terminal and C-terminal domains, residues 7-185 and 220-399 are coloured cyan and magenta respectively. The linker connecting the domains and flexible N- and C-terminal stretches are shown in black. (C) The membrane and solvent environment was represented implicitly as regions with differing dielectric constants ϵ in electrostatics calculations (using the Poisson-Boltzmann approach (Beckstein et al., 2004)) and simulations of rare events (with the GBSW solvent model (Feig et al., 2004; Im et al., 2003a,b)). (D) In equilibrium MD simulations both the protein and the membrane and solvent environment were represented in full atomistic detail. Cytoplasmic open LacY inserted into a POPC membrane, highlighted in grey, is shown here. The solvent molecules and ions are omitted for clarity. The structures were rendered in VMD (Humphrey et al., 1996).

bilayers using coarse-grained self-assembly simulations (Bond and Sansom, 2006; Scott et al., 2008). The protein structures were converted into a coarse-grained representation for the self-assembly simulations. Typically four atoms map to one coarse-grained particle. The overall structure of the protein was maintained through an elastic network model. A distance cut-off of either 7 Å or 10 Å was used for the elastic network model that preserves protein structure in the coarse-grained insertion protocol. Self-assembly simulations were run for 200 ns which led to the insertion of the protein into the membrane. The coarse-grained systems were then converted back into atomistic systems, by superimposing the original atomistic structures onto the coarse-grained protein structure (Stansfeld and Sansom, 2011). Membrane insertion by the self-assembly approach ensures that the position of the membrane depends on the interactions between the protein and the membrane. No manual intervention is required and as such the approach is superior to methods where protein is embedded into membrane based on an initial guess of its position.

The overall system charge was neutralised and additional Na^+ and Cl^- ions were added, resulting in a physiological free NaCl concentration of 100 mM.

6.2.2 Equilibrium MD simulations

Atomistic simulations were run in Gromacs 4.5 (Hess et al., 2008), either with the GROMOS96 43a1 (van Gunsteren et al., 1996) force field with Berger's lipid parameters ("GROMOS"), or the CHARMM22 protein force field69, including the CMAP correction (Mackerell et al., 1998; Mackerell et al., 2004) and CHARMM36 lipid parameters (Klauda et al., 2010) ("CHARMM"). All atomistic simulations were performed at 315 K and 1 bar. The Parrinello-Rahman barostat was used for semi-isotropic pressure coupling (Parrinello and Rahman, 1981). Temperature coupling was achieved by velocity rescaling with a time constant of 0.1 ps (Bussi et al., 2007). The P-LINCS algorithm was used to constrain bonds involving hydrogen atoms (Hess, 2008). The SPC (Hermans et al., 1984) and the CHARMM TIP3P (Jorgensen et al., 1983) water models were used for atomistic simulations with GROMOS and CHARMM respectively. The radius for neighbour list searches was set to 10 Å and 12 Å respectively for simulations with GROMOS and CHARMM force fields. Long-range electrostatics were simulated with the Particle Mesh Ewald method (Essmann et al., 1995), either with a real space cut-off of 10 Å (GROMOS) or a switching function between 8 Å and 10 Å (CHARMM). Van der Waals forces were cut off at 10 Å (GROMOS) or switched off between 8 Å and 10 Å (CHARMM). Production runs were preceded by energy minimisation and 5 ns equilibration using position restraints with a harmonic force constant

of $1000 \text{ kJ mol}^{-1} \text{ nm}^{-2}$ on the protein heavy atoms.

Table 6.1 Overview of the equilibrium MD simulations of LacY

Name	Starting structure	Number of lipids	Length (ns)	Force field
<i>CopenA1</i>	2V8N crystal structure	126	100	GROMOS96 43a1
<i>CopenA2</i>	final structure <i>COpenA1</i>	126	100	GROMOS96 43a1
<i>CopenB1</i>	2V8N crystal structure	256	100	GROMOS9643a1
<i>CopenB2</i>	final structure <i>COpenB1</i>	256	200	GROMOS96 43a1
<i>CopenC</i>	2V8N crystal structure	152	100	CHARMM22 + CHARMM36
<i>PopenA</i>	FucP-based homology model	254	100	GROMOS96 43a1
<i>PopenB</i>	FucP-based homology model	254	100	CHARMM22 + CHARMM36
<i>Occ</i>	End structure DIMS closure	256	100	GROMOS96 43a1

A list of the equilibrium MD simulations analysed in this chapter is given in Table. 6.1. The discussion of the results will be focused on the simulations with the GROMOS force field and in particular *CopenA1*, *PopenA* and *Occ*. The other simulations are analysed in Appendix B.1 and will be referred in the text to make specific points.

6.2.3 Analysis of simulation stability

C_α RMSDs were calculated using MDAnalysis (Michaud-Agrawal et al., 2011). The stability of secondary structure elements was investigated using the *do dssp* program which is part of GROMACS (Hess et al., 2008). The program is based on a standardised definition of secondary structure (Kabsch and Sander, 1983). The data was then plotted using a script provided by Prof. Oliver Beckstein.

6.2.4 Cluster analysis

The simulation of the occluded conformer of LacY was analysed using the algorithm by Daura *et al.* (Daura et al., 1999) in order to pick relevant frames for further analysis. All protein atoms were considered in the geometric clustering and a distance cut-off of 0.25 nm was employed. The structure used for electrostatics calculations – the final structure of the trajectory *Occ* – was part of the dominant cluster with a cut-off of 0.25 nm.

6.2.5 DIMS simulations

DIMS MD (Perilla et al., 2011; Woolf, 1998) was used to sample transitions between cytoplasmic and periplasmic open LacY. Simulations were run with the CHARMM 36a1 program using the CHARMM22 all atom force field (Brooks et al., 1983; MacKerell et al., 1998). A soft ratcheting bias $\Delta\phi = 10 \times 10^{-4}$ was used in *DIMS-c.open-p.open* and $\Delta\phi = 10 \times 10^{-6}$ in *DIMS-closure*. Langevin dynamics were run at 300 K with a timestep of 2 fs and a friction coefficient of 25.0ps^{-1} .

For efficiency, the protein environment was modelled with the GBSW implicit solvation model (Im et al., 2003a,b), which has previously been used to obtain DIMS transitions for the KcsA ion channel in a membrane environment (Denning and Woolf, 2010). Dr Elizabeth Denning (University of Maryland) kindly provided example scripts for the DIMS GBSW calculations which facilitated the simulation setup. The CMAP dihedral correction and atomic radii optimised for GBSW were used (Chen et al., 2006). The GBSW model approximates the bilayer as an exclusion zone from the aqueous phase ($\epsilon = 80$) (Im et al.,

2003a). Here a GBSW membrane thickness of $h_{\text{memb}} = 40 \text{ \AA}$ and a smoothing length $w_m = 6.5 \text{ \AA}$ (Fig. 6.3C) were employed to approximate the POPC membrane geometry seen in the explicit solvent simulations (Fig. 6.3D). The dielectric changed smoothly from $\epsilon = 1$ to $\epsilon = 80$ in the headgroup region. The protein was positioned in the implicit membrane by structural alignment to a fully atomistic system generated by the membrane insertion protocol described above using PyMOL (Schrödinger, LLC) and energy minimised.

Standard parameters were employed for the evaluation of the energies in the GBSW calculations as detailed here. The half-smoothing length sw for the dielectric boundary was set to the default of $0.3 \text{ \AA}$ and the non-polar solvation contribution $s\gamma$ was $0.04 \text{ kcal}/(\text{mol}^2)$. The default non-bonded options for GBSW simulations were employed, with an inner cut-off of $16 \text{ \AA}$ as well as a cut-off for the energy calculation of $16 \text{ \AA}$. A cut-off for non-bonded list generation of $20 \text{ \AA}$ was used.

The RMSD of the non-hydrogen backbone atoms for N- (residues 7-185) and C-terminal domains (residues 220-399) to the target conformation was used as a progress variable in DIMS. The target structure was reoriented relative to the current structure every 10 MD steps. Non-hydrogen backbone atoms of the N- and C-terminal domain were used for rotational and translational fit. To close just the cytoplasmic or periplasmic gates only atoms with $z < 0 \text{ \AA}$ or $z > 0 \text{ \AA}$ (where $z = 0 \text{ \AA}$ marks the midplane of the membrane) were used for the RMSD calculation assessing the change in progress variable that drives the soft ratcheting bias. Eighty simulations were performed with differing seeds for the random number generator that drives the stochastic term of the Langevin equation so that each simulation sampled a different microscopic pathway.

6.2.6 Pore profile analysis

Centred and fitted trajectories were used for pore radii analysis of equilibrium MD trajectories. Conformations were written out every 0.1 ns and their radial pore profiles calculated with the program HOLE (Smart et al., 1993). NumPy (Oliphant, 2007) was used to histogram the pore profiles, which allowed the calculation of average pore profiles as well as standard deviations and figures were plotted with Matplotlib (Hunter, 2007). All frames from the 80 DIMS trajectories between cytoplasmic and periplasmic open LacY were analysed by HOLE. For each structure, the backbone RMSD distance of the N- (residues 7-185) and C-terminal (220-399) domains to the periplasmic open target was calculated with MD-Analysis (Liu et al., 2010; Michaud-Agrawal et al., 2011; Theobald, 2005). The HOLE profiles and the RMSD distances were gridded and interpolated with Matplotlib to be shown

as contour plots. Python code for the pore radius analysis is available as part of the HOLE analysis module of the MDAnalysis library (Michaud-Agrawal et al., 2011).

6.2.7 Water density analysis

Water densities were calculated for centred and RMS-fitted trajectories using the program *gridcount* (Beckstein and Sansom, 2003, available from <https://github.com/orbeckst/gridcount>). The densities were represented using the volume viewer in UCSF Chimera (Goddard et al., 2007; Pettersen et al., 2004) and measured relative to the density of SPC bulk water at standard conditions, 0.970 g cm^{-3} (Hermans et al., 1984).

6.2.8 DEER distance analysis

Comparison of the simulations to DEER experiments was enabled by an approach developed by Dr Philip W. Fowler (available from <http://mdanalysis.googlecode.com>). The simulations were compared to measurements that used (1-oxyl-2,2,5,5-tetramethylpyrroline-3-methyl)-methanethiosulfonate (MTSL) spin labels (Smirnova et al., 2007). MTSL has a linker with five rotatable bonds and is therefore very flexible. The distance distributions between the two spin labels measured by the experiment for LacY are typically broad and often multi-modal. The distributions are therefore a convolution of the flexibility of the MTSL spin label and the conformational spread of the proteins in the sample. The method developed by Dr Fowler takes the flexibility of the spin-label into account. Firstly, a rotamer library of MTSL spin labels is mapped onto each position (Polyhach et al., 2011). Secondly rotamers that would lead to steric clashes with the protein are excluded from further analysis. In general a distance cut-off of $2\text{ \AA}$ is used to find steric clashes. For spin-spin distance pairs 73-340, 136-340 and 137-340 a clash distance cut-off of $1.5\text{ \AA}$ was used. Thirdly, the distance distributions are calculated for the remaining rotamers. Distance distributions were calculated for entire trajectories. Finally, the estimated distance-distance distributions could be compared to the previously reported DEER measurements of LacY.

6.2.9 Poisson-Boltzmann electrostatics calculations

The Poisson-Boltzmann electrostatics calculations build upon previous work by Prof. Oliver Beckstein that led to the BornProfiler software package (Beckstein et al., 2004) and scripts based on it. The aim was to calculate the position resolved electrostatic free-energy for a

proton interacting with LacY, $\Delta W_{elec}(r)$.

$$\Delta W_{elec}(r) = G_{protein+ion} - G_{protein} - G_{ion} \quad (6.1)$$

$G_{protein+ion}$ is the electrostatic free energy for the system when the protein and the ion interact. To obtain $W_{elec}(r)$ the electrostatic free-energy for the protein $G_{protein}$ and the ion G_{ion} when there is no interaction between the two species has to be subtracted from $G_{protein+ion}$. The proton is represented as a cavity in the bulk solvent with an effective radius a and a charge of +1. This Born description is a very simple description for the complex behaviour of a proton in water. Proton transport in aqueous solution involves a complicated “fluxional complex” in which the Eigen (H_3O_4^+) and Zundel cation (H_5O_2^+) hydration structures of the proton rapidly interconvert (Woutersen and Bakker, 2006). The use of such a simple model is motivated by the insight that the energetics are likely dominated by the electrostatics of the charge in the low-dielectric membrane compared to smaller differences due to the detailed proton-water interactions. The effective Born radius of the solvated proton was estimated as $a=1.4848 \text{ \AA}$ by solving the Born equation $a = -166 \text{ \AA kcal mol}^{-1} \frac{q^2}{G_{solv}^*} (1 - \frac{1}{80})$ with the experimental hydration free energy of the oxonium ion (H_3O^+) at the 1 M standard state, $G_{solv}^* = -110.4 \pm 0.7 \text{ kcal/mol}$ (Pliego and Riveros, 2000). In the case of membrane proteins, the reference energy $G_{protein}$ is calculated for the protein embedded in a membrane-like environment as described below.

The linearised Poisson-Boltzmann (PB) equation was then solved with APBS (Baker et al., 2001). The ionic strength was 100 mM NaCl and the temperature was 298.15 K. Ion positions r were generated with a customised version of HOLLOW (Ho and Gruswitz, 2008) at a point spacing of 1 $\text{\AA}$. A dielectric constant of $\epsilon = 80$ was assigned to bulk water and $\epsilon = 10$ to the protein. An implicit membrane was modelled after the membrane observed in the MD simulations as a low dielectric slab with a hydrophobic core of thickness 33.5 $\text{\AA}$ with $\epsilon = 2$ and a head group region of thickness 6.5 $\text{\AA}$ with $\epsilon = 20$ (Fig. 6.3C). The water accessible interior of the transporter was modelled as a high dielectric ($\epsilon = 80$). Manipulation of the dielectric maps was carried out with software based on the *draw membrane* tool in APBSmem (Callenberg et al., 2010) and an updated version of the BornProfiler analysis package (Beckstein et al., 2004).

PB calculations employed a manual focusing strategy in three steps on a computational grid of 129x129x129, starting from a cubic bounding box of length 250 $\text{\AA}$, via a 100- $\text{\AA}$ box, to a final dimension of 50 $\text{\AA}$, centred on the position of the ion. Increasing the computational grid to 161x161x161 did not change the energies and hence the results can be

considered converged with respect to the grid. Starting structures for simulations *CopenA1* and *PopenB* - which had been equilibrated in atomistic membranes and are fully cytoplasmic and periplasmic open respectively - as well as final structures from MD simulations of *CopenA1*, which showed partial closure, and fully closed *Occ* were used for PB calculations. Partial charges and radii from the CHARMM force field were assigned with *pdb2pqr* (Dolinsky et al., 2004). All ionisable residues were simulated in their default charge state as either suggested by PROPKA (Li et al., 2005) or because the default protonation state better described a transporter in that part of the transport cycle that is believed to occur without involvement of substrate or proton transport 2.

6.2.10 Order parameter analysis

Two order parameters that report on whether a MFS transporter structure is cytoplasmic or periplasmic open were defined for EmrD (PDB code 2GFP; Yin et al., 2006), FucP (3O7Q, 3O7P; Dang et al., 2010), GlpT (1PW4; Huang et al., 2003), cytoplasmic open LacY (1PV6, 1PV7, 2CFQ, 2CFP, 2V8N, 2Y5Y; Abramson et al., 2003; Chaptal et al., 2011; Guan et al., 2007; Mirza et al., 2006); or periplasmic open LacY (4OAA, homology model: PM0077183, repeat swapped model: PM0076824; Kumar et al., 2014; Radestock and Forrest, 2011), NarU (4IU8; Yan et al., 2013), NarK (4JR9; Zheng et al., 2013), PepT_{So} (2XUT; Newstead et al., 2010), PepT_{St} (4APS)95, PiPT (4J05; Pedersen et al., 2013) POT (4IKV; Doki et al., 2013), XylE (4GBY, 4JA3, 4JA4; Quistgaard et al., 2013; Sun et al., 2012) and YajR (3WDO; Jiang et al., 2013).

Distance matrices were calculated for C_{α} atoms of the pairs of transmembrane helices 1 and 7 for the closed periplasmic gates of LacY, GlpT, EmrD, NarK, NarU, PepT_{So}, PiPT, POT and XylE using MDAnalysis (Michaud-Agrawal et al., 2011). The transmembrane helices were defined by STRIDE, DSSP and visual inspection in VMD (Frishman and Argos, 1995; Humphrey et al., 1996; Kabsch and Sander, 1983). Helix definitions are listed in Table 6.2. The closest contact defined the $C_{\alpha} - C_{\alpha}$ distance that serves as the order parameter. Similarly the cytoplasmic order parameter was calculated for the closed cytoplasmic gates of FucP, EmrD, NarU, PepT_{So}, PiPT, XylE and YajR as well as the homology and the repeat swapped model for periplasmic open LacY.

No structure featuring the closed cytoplasmic gate exists for GlpT, NarK, PepT_{St} and POT. Similarly the structures of the closed periplasmic gates of FucP and YajR are unknown. The order parameters were defined via structural alignments. GlpT, NarK, PepT_{St} and POT were aligned with cytosolic open LacY by the DALI pairwise alignment server

Table 6.2 Gate helix definitions of MFS transporters with known structures.

	Structure	TM4 (cyt. gate)	TM10 (cyt. gate)	TM1 (peri. gate)	TM7 (peri. gate)
ErmD	2GFP:A	100-130	296-316	9-36	202-230
LacY	2V8N:A	104-136	312-340	7-35	220-250
	PM0077183				
NarU	4IU8:A	128-153	345-386	21-59	255-282
PepT _{So}	2UXT:A	109-137	405-434	15-41	292-328
PiPT	4J05:A	128-161	412-441	31-63	307-342
XylE	4GBY	122-153	370-398	8-46	277-307
YajR	3WDO	101-127	302-330	-	-
FucP	3O7Q	117-144	347-373	-	-
GlpT	1PW4	-	-	20-54	253-279
PepT _{St}	4APs	-	-	13-46	284-314
POT	4KIV	-	-	23-51	280-323
NarK	4JR9:A	-	-	23-65	257-284

(Hasegawa and Holm, 2009), which was also used to align YajR to the periplasmic open homology model of LacY. For FucP the refined structural alignment with cytoplasmic open LacY obtained by Radestock *et al.* (Radestock and Forrest, 2011), that underlies the homology model for periplasmic open LacY, was used.

6.3 Results: Molecular dynamics of cytoplasmic and periplasmic open LacY

6.3.1 Stability of the simulations

Analysis of C_α Root-mean square deviation in the simulations

Before analysing the simulations of cytoplasmic and periplasmic open LacY in more detail the structural stability of the simulations was assessed. Drastic changes or loss of structure in simulations can point at deficiencies in the simulation parameters, the force field, or in the starting structures. This assessment is particularly important in this study as the simulations of the periplasmic open state were based on a homology model. A homology model might not capture all the stabilising interactions that would be represented in an experimental structure. Stability of the simulations was assessed firstly by calculating C_α RMSD deviations to follow structural drift over the course of the simulations and secondly by tracking the stability of the secondary structure elements of this transporter.

To assess structural drift in the equilibrium MD simulations the C_{α} RMSDs for the N- and C-terminal domains (residues 7-185 and 220-399), to the starting structures of simulations *CopenA1* and *PopenA* were calculated. Flexible regions such as the inter-domain linker were excluded from the analysis. The C_{α} RMSDs calculated for the simulation *CopenA1* stabilised around 3 Å (Fig. 6.4A), as the protein relaxed away from the crystal structure in an atomistic lipid membrane. The homology model for periplasmic open LacY was less stable with the RMSD rising to values close to 5 Å (Fig. 6.4 B). Not all interactions stabilising periplasmic open LacY might have been captured in the homology model, which would not be surprising given the low sequence identity between LacY and FucP (10 %) (Radestock and Forrest, 2011). For example, the side-chains of E325 R302 are thought to interact in the periplasmic open transporter but cannot interact in the current model.

Secondary structure in the MD simulations

The stability of the simulations was further assessed by analysing changes in secondary structure over the course of the simulations. The secondary structure as assigned by the DSSP algorithm (Kabsch and Sander, 1983) shows some local breaking of α -helices (Fig. 6.5). Specifically, in *CopenB2*, helix II broke at its end as can be seen in Fig. 6.5A, but is more stable in *CopenC* with CHARMM (Fig. 6.5D). At the timescales accessed a degree of helix breaking and subsequent reforming is to be expected, especially for a protein with many distorted helices such as a LacY. Overall the cytoplasmic open transporter retained its trans-membrane helices in simulations with either force field. In the simulations with the periplasmic open homology model there is also some loss of helical structure. In particular the beginning of helix VIII appears to unwind in *PopenA*. Interestingly, this helix is stable in the simulations with CHARMM Fig. 6.5E. Nonetheless, the periplasmic protein maintained most of its helical structures in simulations with GROMOS and to a slightly higher degree with CHARMM. This together with the RMSD calculation shows that the homology model is reasonably stable, which is therefore an excellent starting point for this study. Having assessed the quality of the simulations, a more detailed analysis of the conformations accessed in the simulations will be provided.

6.3.2 Functional analysis of the simulations of cytoplasmic and periplasmic open LacY

The MD calculation were initiated with the the apo conformation of both end states, *i.e.* there was no sugar bound and E325 was not protonated. The apo conformations should

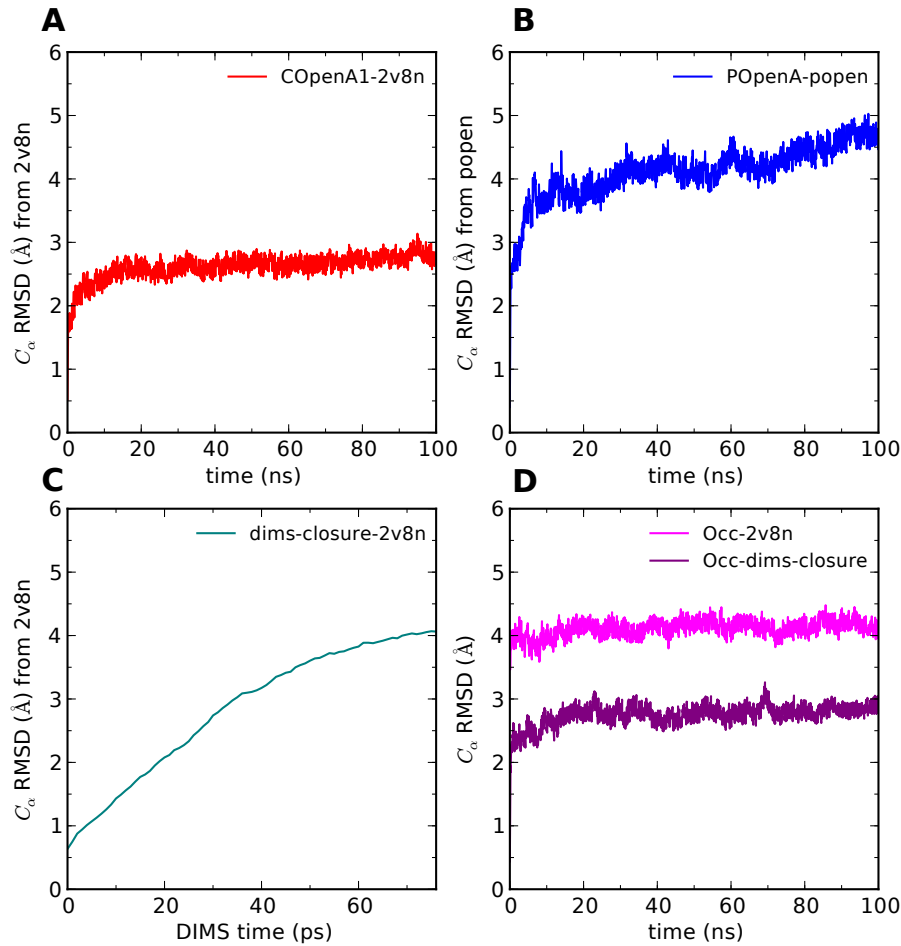


Fig. 6.4 C_{α} RMSDs for representative simulations *CopenA1*, *Occ*, *PopenA*, and *DIMS* – closure. The C_{α} RMSD was calculated for the structured N- (7-185) and C-terminal domains (220-399) of LacY omitting the flexible linker, using the MDAnalysis library (Michaud-Agrawal et al., 2011). The simulations were compared to their respective starting structures, the 2V8N crystal structure for *CopenA1* (A) and *DIMS* – closure (C), the periplasmic open homology model based on FucP for *PopenA* (B). The simulation of the occluded structure, *Occ*, was compared to its starting conformation (D), the cytoplasmic and periplasmic closed final frame of *DIMS* – closure (in purple). To emphasise that *Occ* is in a sense a continuation of *DIMS* – closure, the trajectory of its C_{α} RMSD relative to the 2V8N crystal structure is shown in D (in magenta).

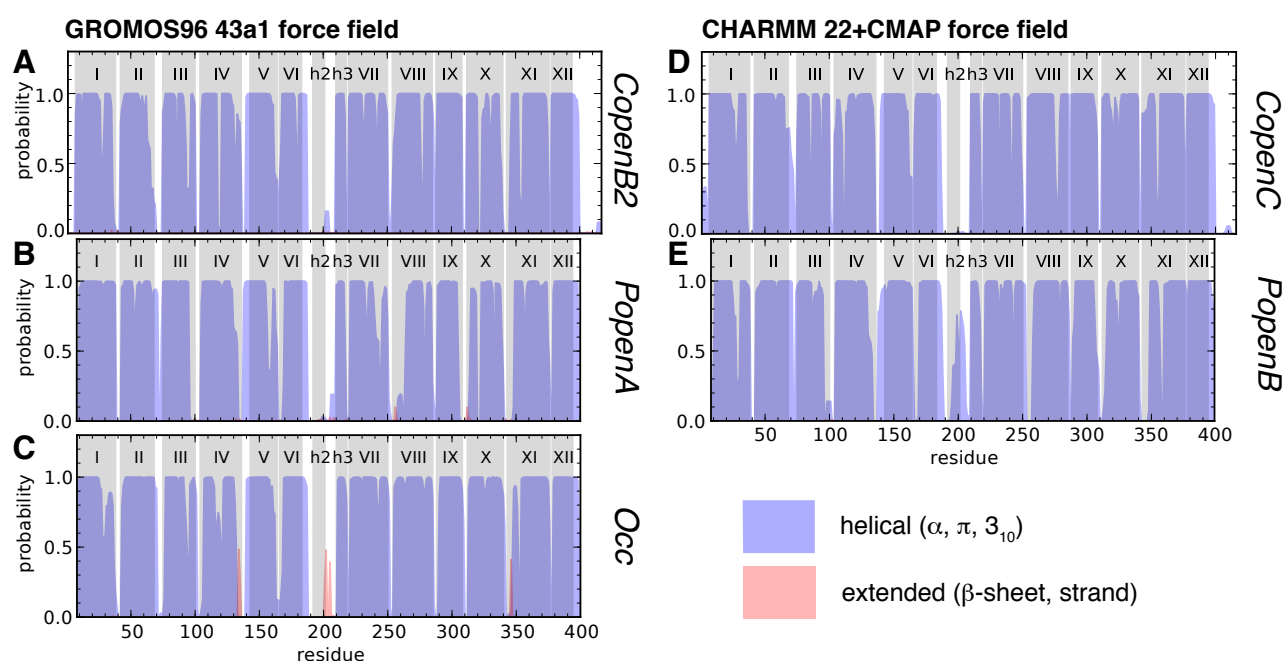


Fig. 6.5 Secondary structure in equilibrium MD simulations. Secondary structure elements assigned by Abramson et al., 2003 are labelled and shown as grey bars (transmembrane helices I-XII and solvent-exposed helices h2–h3). The probability to observe a particular secondary structure for a given residue (as defined by DSSP (Kabsch and Sander, 1983)) over an equilibrium MD trajectory is shown as blue (any helix) or red (strand/sheet) filled curves. The figure was generated from data obtained by the Gromacs tool *do dssp* (Hess et al., 2008) using a script written by Prof. Oliver Beckstein as discussed in the methods (section 6.2.3).

be on the cusp of changing conformation (Guan and Kaback, 2006) and thus simulation might reveal the initial stages of the transition. Since simple visual inspection of a static structure is not sufficiently accurate to determine if the central binding site is accessible to both substrate and protons a range of computational methods will be applied to assess the functional state of the generated protein conformations (Fig. 6.6), *i.e.* whether the gates are open or closed. First, the functional state of the transporter in the simulations will be characterised by calculating the radius of the translocation pathway through the transporter. Then the density of water in the gates and the electrostatic free energy of a permeating H^+ will be examined. These are more sensitive and together these methods will allow assigning the likely functional state of any structure.

6.3.3 Pore profile analysis

Pore profiles calculated from the equilibrium MD simulations of the model of periplasmic open LacY, showed that the periplasmic gate remained open and the cytoplasmic gate stayed closed. This was seen for both *PopenA* (Fig. 6.6C) and for the repeat *PopenC* (Fig. B.1 and appendix B.1). These simulations demonstrated that both the conformation and its functional state were maintained over the duration of the simulations.

In contrast, during simulations of cytoplasmic open LacY the periplasmic gate remained closed but the cytoplasmic gate started to close. In simulation *CopenA1*, the minimum radius on the cytoplasmic side of the transporter decreased from 4 Å in the crystal structure to 1 Å (Fig. 6.6A). The same effect was seen in multiple simulations with different box dimensions and force fields (as evidenced by the pore radius profiles in Fig. B.1). The observed closure did not therefore depend on the dimensions of the box or the force field used. The variation of pore radius along the membrane normal suggests that sugars would be sterically blocked from entering the central cavity from both cytoplasm and periplasm. The cytoplasmic open apo-structure of LacY, as resolved by X-ray crystallography, quickly relaxes to a more occluded structure that would prevent sugar permeation. Such a structure might be an early intermediate in the transition between cytoplasmic and periplasmic open LacY (Holyoake and Sansom, 2007).

6.3.4 Water density analysis

It was assessed if this intermediate is functionally occluded, *i.e.* whether solvated protons could still permeate, despite the closure of the cytoplasmic gate. Water densities were calculated from the MD simulations to define aqueous pathways that could act as proton conduits.

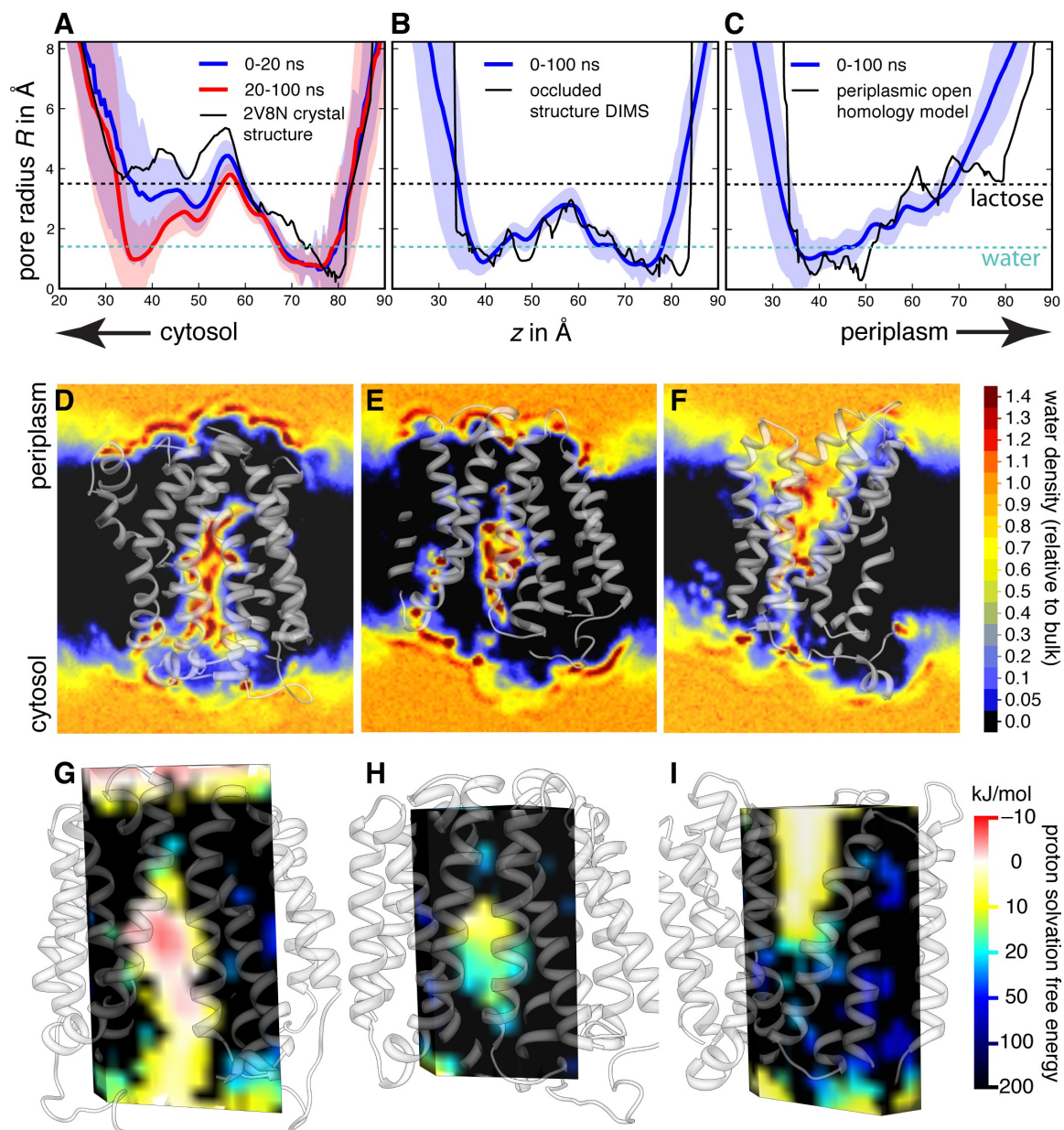


Fig. 6.6 Functional characterisation of LacY in MD simulations of multiple conformational states. Pore radii as determined by HOLE (Smart et al., 1993) are shown along the membrane normal z for the central cavity in simulations of cytoplasmic open *CopenA1* (A), occluded *Occ* (B) and periplasmic open *POpenA* (C). Shaded regions indicate one standard deviation of the radius from the mean value. Water densities relative to bulk SPC water (0.970 g cm⁻³) were calculated with gridcount (Beckstein and Sansom, 2003; Hermans et al., 1984) and visualised as 3D densities in UCSF Chimera (Goddard et al., 2007; Pettersen et al., 2004). Cuts through the densities for simulations *CopenA1* (D), *Occ* (E) and *POpenA* (F) are shown. For *CopenA1* and *Occ* the last frame from 100 ns of equilibrium MD is depicted. For *POpenB* the equilibrated starting structure for the 100 ns simulation is shown. The position-resolved electrostatic free energy $W_{elec}(r)$ for solvated H⁺ interacting with cytoplasmic open (G; 2V8N structure, relaxed after membrane insertion by position-restrained MD), fully occluded (H; final structure from *Occ*) and periplasmic open LacY (I; homology model from Radestock and Forrest, 2011, relaxed after membrane insertion by position-restrained MD), calculated with BornBrofiler (Beckstein et al., 2004) and visualised in UCSF Chimera (Goddard et al., 2007; Pettersen et al., 2004) is shown.

Since the desolvation of a charged species is energetically unfavourable, in the absence of compensating interactions, dehydration of a gate will prevent ion permeation (Beckstein and Sansom, 2003; Beckstein et al., 2004). The water density for the simulation of cytoplasmic open LacY (Fig. 6.6D, Fig. B.2A,B) indicated that the constriction between the cytoplasm and the central cavity remained partially solvated even after the initial closure of the cytoplasmic gate, indicating protons could still be expected to permeate and thus the conformation cannot be considered a functionally occluded state.

Analysis of the water densities also provided further insight. No solvent filled pathway connected the cavity and the periplasm in the simulation of cytoplasmic open LacY and thus the periplasmic gate remained completely shut. The cytoplasmic gate in the simulations of periplasmic open LacY remained dehydrated (Fig. 6.6F) whereas a wide aqueous pathway connected the central cavity to the periplasmic solution, suggesting the conformation to represent a functional periplasmic open state.

6.3.5 Electrostatic free-energy for proton permeation

Poisson-Boltzmann calculations of the electrostatic free energy of a solvated proton offer another way of assessing whether a gate is open since the heights of the energetic barriers between the cavity and bulk can be quantified (Beckstein et al., 2004). The energies calculated for the cytoplasmic open (Fig. 6.6G) and periplasmic open (Fig. 6.6I) conformations suggest that the cytoplasmic and the periplasmic gates, respectively, offer no barrier to the movement of protons and so would permit relatively free exchange of H^+ between the bulk solvent and the central cavity. Electrostatic free energies in the cytoplasmic and periplasmic vestibules and the central cavity were only slightly less favourable than bulk (0-10 kJ/mol). The periplasmic gate of cytoplasmic open LacY is closed because H^+ exchange between the central cavity and the periplasm would involve crossing a large volume with unfavourable electrostatic free energies (>100 kJ/mol, Fig. 6.6G). By the same reasoning, the cytoplasmic gate of the periplasmic open transporter is also closed to H^+ permeation (Fig. 6.6I). The initial closure of the cytoplasmic gate seen in the pore profile calculations (Fig. 6.6A) led to only a limited increase in the electrostatic free energy barrier in the space connecting the cavity to the cytoplasm, from 0-10 kJ/mol to 10-20 kJ/mol (Fig. 6.7). This modest increase in the electrostatic free energy for proton permeation together with the water density calculations suggests that the observed closure of the cytoplasmic gate resulted in a partially functionally occluded intermediate. The gate appears therefore impermeable to sugar but not to H^+ exchange between the cytoplasm and the internal cavity.

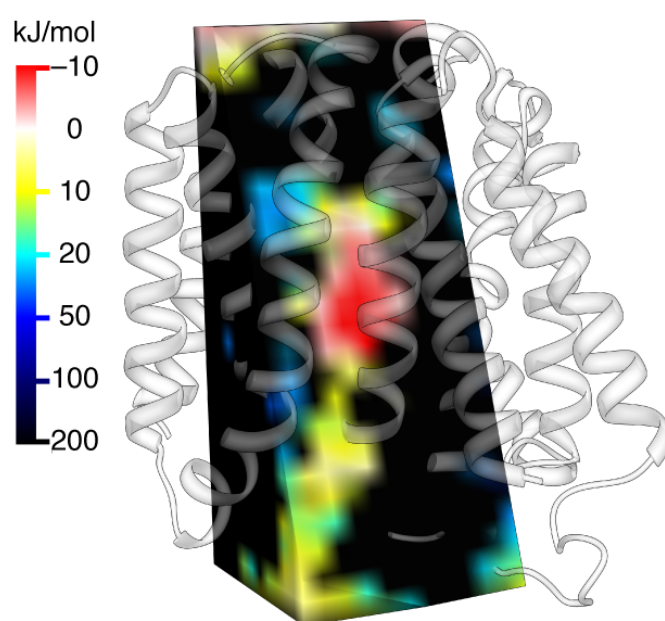


Fig. 6.7 Electrostatic free-energy for H^+ permeation through the cytoplasmic open LacY after initial closure in MD. The position-resolved free energy $W_{elec}(r)$ for a solvated proton interacting with the final snapshot from simulation *CopenA1* was calculated with BornProfiler and visualised in UCSF Chimera is shown (Goddard et al., 2007; Pettersen et al., 2004). Energies are given in kJ/mol.

6.4 Results: DIMS simulations of the conformational changes in LacY

Eight equilibrium MD simulations of cytoplasmic and periplasmic open LacY with an aggregated total time of $0.8\mu\text{s}$ yielded neither a transition between the end states nor a functionally fully occluded intermediate. This is not surprising given that, on average, one transition would be expected to be observed every $50 \times 10^3\mu\text{s}$ for a typical turnover number of 16s^{-1} to 21s^{-1} (Viitanen et al., 1984). Therefore DIMS MD was used to investigate whether LacY passes through an occluded state as it changes from a cytoplasmic open conformation to a periplasmic open conformation (Perilla et al., 2011). In DIMS MD an importance sampling procedure is used to generate transitions between an initial and a target structure.

DIMS MD was used to generate trajectories starting from the cytoplasmic open structure of LacY during which the cytoplasmic gate would close. It was assumed that the cytoplasmic and periplasmic gates are located on the cytoplasmic and periplasmic sides of the transporter, respectively. Therefore, only the cytoplasmic half of periplasmic open LacY, *i.e.* the one featuring the closed cytoplasmic gate, was used as the target structure. The DIMS importance sampling procedure then used this target structure to select pathways that led to cytoplasmic closure. The variation in the pore radius profile along the membrane normal of the final structure from the DIMS MD simulation showed constrictions on both the cytoplasmic and periplasmic sides (Fig. 6.6B), consistent with it being an occluded structure. This trajectory will be referred to as *DIMS-closure* from hereon. A similarly constricted structure was generated from periplasmic open LacY using the periplasmic half of cytoplasmic open LacY, *i.e.* the closed periplasmic gate, as the target structure (*not shown*).

RMSD calculations confirmed that the model for the occluded conformation is a distinct conformational state. The C_α RMSD curve for DIMS-closure rises to $4\text{ \AA}$ (Fig. 6.4C). The more limited changes associated with the closure of the cytoplasmic gate gave a C_α RMSD profile that stabilises close to $3\text{ \AA}$ as depicted in Fig. 6.4C. The structural changes that accompany the closure of the cytoplasmic gate go beyond this more limited change.

6.4.1 Stability of the model for the occluded state

To help assess whether the DIMS method had produced a functionally fully occluded conformation standard equilibrium MD was used. The final structure from DIMS MD was simulated for 100 ns (simulation *Occ*). Based on the trajectory the stability of the model

was assessed. As the occluded structure is relaxed into an atomistic bilayer in an equilibrium simulation (*Occ*), the C_α RMSD to the crystal structure stays close to 4 Å (Fig. 6.4D), which means that no further drift away from the cytoplasmic open structure occurred. The C_α RMSD to the final frame of the DIMS simulation of the closure of the cytoplasmic gate, *DIMS-closure*, levels off around 3 Å (Fig. 6.4D). The occluded structure derived from molecular simulations is hence comparable in stability to the crystal structure of cytoplasmic open LacY as judged by C_α RMSD. Furthermore the model maintained regular secondary structure as shown in Fig. 6.5C. Some helices, such as transmembrane helix VIII were much more stable in *Occ* than in simulation *PopenA* of the periplasmic open homology model. Visual inspection showed that the central water-filled internal cavity kept its shape, which is also apparent from the pore radius calculations presented in Fig. 6.6B. The functional states of the cytoplasmic and periplasmic gates in *Occ* were then analysed in detail.

6.4.2 Function characterisation of the occluded conformation

The permeability of the cytoplasmic and periplasmic gates in the occluded structure were investigated in (1) pore profile, (2) water density and (3) electrostatics calculations. Pore radius profiles calculated with HOLE (Fig. 6.6B) showed that the minimum pore radii at the cytoplasmic and periplasmic constrictions remained smaller than the radius of a water molecule (around 1.4 Å) during the entire simulation. The radius of gyration of lactose is 3.5 Å and therefore sugar substrates would not be able to access the central cavity.

The water density computed from the simulation of the occluded structure, *Occ*, was also consistent with both gates being closed. Slices through the 3D water density showed that regions with very low density separate the central cavity from both the cytoplasm and periplasm. The internal cavity remained solvated (Fig. 6.6E) at between 0.7 to 1.4 times the bulk water density. Some water, however, remained in the interfacial regions. For example, on the cytoplasmic side water density at 70 % of bulk water is apparent in the slice shown in Fig. 6.6E but only regions with low density connect it to the cytoplasm and the internal cavity. While ions are likely to be stable in the water filled central cavity, there are no continuous permeation pathways joining either the periplasm or the cytoplasm to the central cavity.

To probe the electrostatic free-energy of proton permeation, a representative structure from was chosen from *Occ*. The structures generated in the trajectory were clustered (see Methods). The last frame of the simulation was selected as a representative structure (Fig. 6.8A) since it belonged to the dominant cluster. This structure is of comparable quality

to the wild-type crystal structure of cytoplasmic open LacY as discussed in Appendix B.1.2. The electrostatic free energy of placing a proton in this structure was calculated as before. Energies in the internal cavity ranged from 0 kJ/mol to 20 kJ/mol (Fig. 6.6H) but importantly, very high electrostatic free energies (>100 kJ/mol) were obtained in the spaces on the periplasmic and cytoplasmic sides of the internal cavity. Some regions with more favourable electrostatic free energies (20–50 kJ/mol) exist between the internal cavity and the cytoplasm and the periplasm, as can be seen on the periplasmic side of the slice through the 3D density. Given that such regions were surrounded by space very unfavourable to protons it is reasonable to conclude that electrostatic barriers to ion permeation from both the cytoplasm and the periplasm exist. The electrostatics and water density calculations are thus consistent with this structure of the transporter being not only sterically closed to sugar substrates but also occluded to protons and thus preventing a proton leak across the membrane. One can then conclude that a true functionally occluded state was accessed in the simulations.

6.4.3 The R302-E325 salt-bridge

The model for the occluded state is also consistent with the proposed mechanism for the transport cycle (Guan and Kaback, 2006). The residues crucial for function E269, R302, H322 and E325 (highlighted in Fig. 6.8) as well as E126 and R144 (*not shown*) surround the central cavity as expected. In particular, it was found that E325 and R302 form a stable salt bridge in simulation *Occ* and in the occluded state model (Fig. 6.8C) but neither in the cytoplasmic (Fig. 6.8B) nor the periplasmic open (Fig. 6.8D) structures that were the inputs for this study. Plotting the minimum distance between nitrogen atoms of the guanidinium group of R302 and the oxygen atoms of the E325 side-chain, for simulations *CopenA1* and *PopenA*, demonstrated that no stable salt-bridge was formed (Fig. 6.9) during the course of the simulations. The average separation remained of 5.1 ± 0.6 Å in *CopenA1*. E325-R302 remained at a large separation of 14.0 ± 0.9 Å in *PopenA*. Closing the cytoplasmic gate in a DIMS simulation (*DIMS-closure*) brought the two side-chains closer together (Fig. 6.9B), from 5.6 Å to 4.7 Å, but the E325-R302 salt-bridge was only fully formed upon relaxing the occluded structure in the *Occ* equilibrium simulation (Fig. 6.9C). The minimum distance between the carboxylate oxygen atoms of E325 and the nitrogen atoms of R302 was found to have an average value of 3.3 ± 0.5 Å in *Occ*. The formation of the E325-R302 salt bridge is a vital step in the transition from the cytoplasmic to periplasmic open conformations and the computer simulations capture this important aspect of the molecular interactions in the

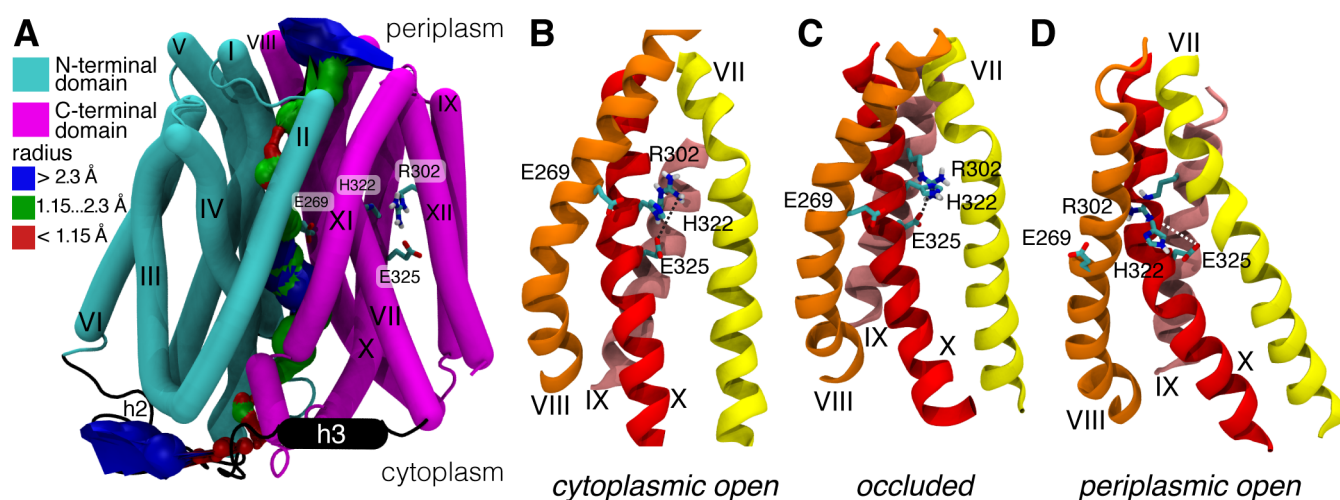


Fig. 6.8 Structural characterisation of the occluded state model. (A) Helix representation of the occluded state model from the end of the *Occ* equilibrium simulation (rendered with Bendix and VMD Dahl et al., 2012; Humphrey et al., 1996). The N- and C-terminal domain are colour-coded. The internal surface of the cavity as calculated by HOLE is shown. Residues essential for conformational changes and transport are shown as sticks. (B) Close-up of residues E269, R302, H322, and E325 in the cytoplasmic open conformation. The average distance of R302–E325 (dashed line) is 5.1 ± 0.6 Å (mean $\pm$ s.d.) over the course of the equilibrium MD simulation. (C) The average R302–E325 distance is 3.3 ± 0.5 Å in the occluded state. (D) The FucP-based homology model of the periplasmic state places R302 and E325 far apart and they remain separated over the MD simulation (14.0 ± 0.9 Å). The time course of the interaction between E325 and R302 is shown in Fig. 6.9 for simulations *CopenA1*, *PopenA* and *Occ*.

transporter.

6.4.4 Comparison to experimental Double Electron Electron resonance data

Having established that a functionally fully occluded state was generated by DIMS MD and equilibrium MD, this model and the simulations were validated by comparing them with DEER experiments. These have been previously interpreted as reporting on the presence of cytoplasmic and periplasmic open states as well as occluded conformations (Smirnova et al., 2007). To better interpret these data I have used an approach developed by Dr Fowler (see section 6.2.8 in the methods segment of this chapter) that takes into account the flexibility of the spin label and so predicts the electron spin-spin distance distribution that would be observed for a single structure or an entire MD trajectory. This is important as it allows

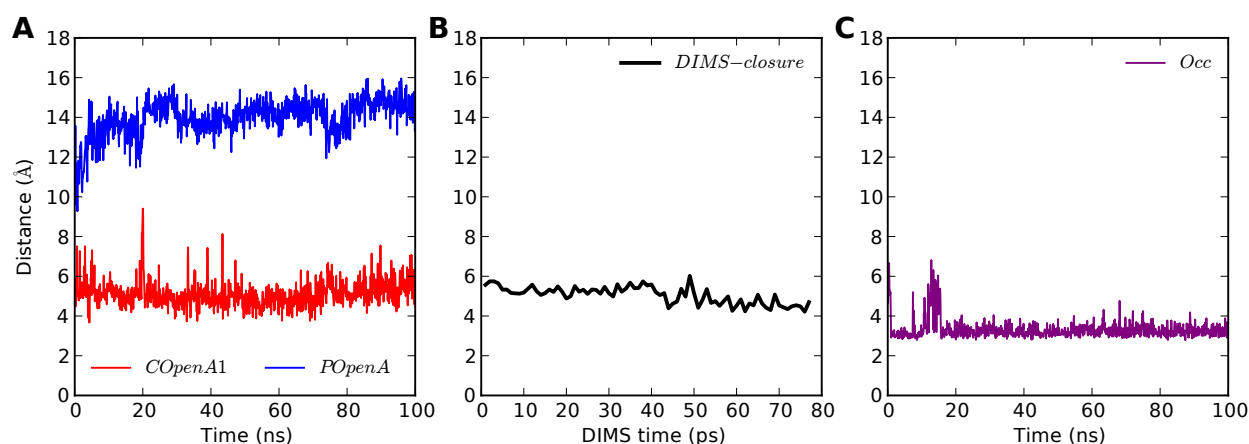


Fig. 6.9 The interaction between E325 and R302 was followed in the simulations. The minimum distance between the two carboxylate oxygen atoms of E325 and the nitrogen of the guanidinium group of R302 was tracked for simulations *CopenA1*, *PopenA*, *Occ* and *DIMS-closure*. The distances are shown in red and blue for *CopenA1* and *PopenA* in A. In B the salt-bridge is tracked in *DIMS-closure*. In C the E325-R302 interaction during simulation *Occ* is depicted in purple.

comparing like with like. Nine experimental distances were compared to the simulations; three on the periplasmic side of the protein (residue pairs 105-310, 164-310 and 164-375; Fig. 6.10A) and six on the cytoplasmic side (73-401, 73-340, 136-340, 137-340 136-401 and 137-401; Fig. 6.10B). As the periplasmic open homology model (Radestock and Forrest, 2011) lacks some C-terminal residues, spin-spin distances 73-401, 136-401 and 137-401 can only be compared for the simulations of cytoplasmic open LacY.

The distance distributions calculated from the simulations of cytoplasmic open LacY agreed well with the DEER experiments on LacY in the presence of NPGlc (4-nitrophenyl- α -D-glucopyranoside). This non-binding sugar ensures the protein is found predominately in cytoplasmic open conformations (Smirnova et al., 2007). In general, there was substantial overlap between the distributions calculated from simulations and the positions of the major peaks from the DEER experiments as is apparent when comparing the red lines for the MD of cytoplasmic open LacY with the grey histograms for the experimental data in Fig. 6.10. Distributions from MD also compared favourably to experiments for the spin-spin distances of residues 73-401, 136-401 and 137-340 (Fig. B.3). The absolute heights and widths of the experimental peaks are not expected to quantitatively agree with the simulations because the three independent simulations of individual conformational states do not capture the relative distribution of different states observed in experiments.

The distance distributions estimated from the simulation of periplasmic open LacY (blue

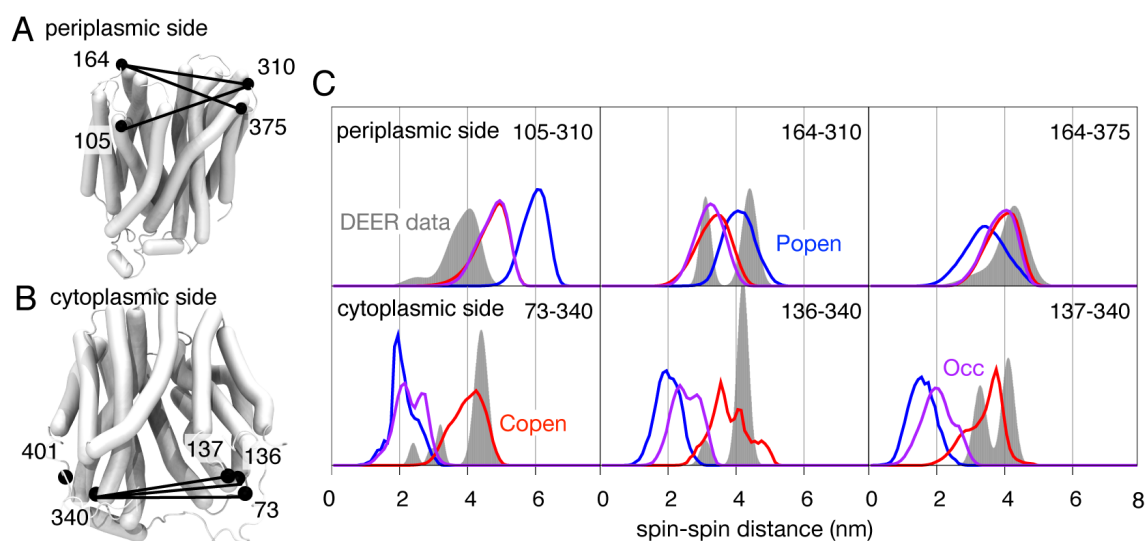


Fig. 6.10 Experimental and simulated spin-spin distance distributions. (A) Location of spin labels on residues (black spheres: C_{α} atoms) and experimentally measured distances (black lines) on the periplasmic face of LacY (periplasmic open conformation). (B) Spin-spin distances on the cytoplasmic face (cytoplasmic open conformation). (C) Comparison of experimental DEER spin-spin distributions from Smirnova et al., 2007 (grey histograms, “DEER data”) to computed distributions from equilibrium MD simulations of LacY in three different conformational states: periplasmic open (blue lines, “Popen”), cytoplasmic open (red lines, “Copen”), occluded (purple lines, “Occ”). DEER experiments were performed in the presence of the non-binding sugar NPGlc, which shifts the conformational equilibrium towards the cytoplasmic open state (Smirnova et al., 2007).

lines in Fig. 6.10C) show the pattern expected for a periplasmic open state, *i.e.* the spin label distances on the periplasmic side are shifted to larger values and the distances on the cytosolic side are shifted to smaller values compared to the equivalent distances estimated from the simulations of the cytoplasmic open transporter. One exception is distance 164-375 where the end of helix 5 with residue 164 moved closer to 375 on helix 11 at the beginning of the simulation and so inverted the expected pattern. The distributions from the simulations generally overlap with the positions of the major experimental peaks for predominately periplasmic open LacY (Fig. 6.10C, blue lines, and Fig. B.3B).

Finally, the model of the occluded state was validated against the DEER data. The spin-spin distance distributions calculated from the simulation *Occ* of the occluded model (Fig. 6.10C, purple lines) were consistent with the experimental data. On the cytoplasmic side of the protein, the predicted spin-spin distances were shorter than for cytoplasmic open LacY. The distributions for 73-340 and 136-340 coincided with minor peaks in the DEER experiments with NPGlc. These peaks are indicative of cytoplasmic closure as they overlap with the spin-spin distance distribution from the simulation of periplasmic open LacY where the cytoplasmic gate is closed. Peaks for 73-340, 137-340 and 136-340 for *Occ* also overlap with peaks from experiments recorded in presence of NPGal that coincide with peaks for the simulation of the periplasmic open transporter. NGal, 4-nitrophenyl- α -D-galactopyranoside (Fig. B.3B) and additional discussion are presented in appendix B.1.

On the periplasmic side the distances closely matched the distances for cytoplasmic open LacY (Fig. 6.10C, red lines), which is not surprising given that the periplasmic gate is closed. The peak for 164-310 for *Occ* closely follows the distribution from the simulations of cytoplasmic open LacY where the periplasmic gate is closed. For residues 164-310, the experimental DEER peaks at 31 Å and 37 Å in the presence of binding and non-binding sugar have been attributed to conformations with both gates closed (Madej et al., 2012). Encouragingly, the two peaks intersect the spin-spin distribution calculated from *Occ*.

6.4.5 Comparison to Major Facilitator Superfamily crystal structures

In order to compare the model of fully occluded LacY to known occluded MFS transporter structures such as PepT_{So} (Newstead et al., 2010), XylE (Quistgaard et al., 2013; Sun et al., 2012), EmrD (Yin et al., 2006), NarU (Yan et al., 2013), and PiPT (Pedersen et al., 2013), I derived a pair of geometric order parameters. Putative gates were identified by comparing the experimental structures of cytoplasmic open LacY and the models of periplasmic open LacY (Radestock and Forrest, 2011). The key feature of the closed periplasmic gate is

the tight packing of helices 1 and 7, which seals off the central cavity to the periplasm. A similarly close packing of helices 4 and 10 is the crucial element of the cytoplasmic gate. Other helices have also been implicated in the gating of MFS transporters (Newstead et al., 2010). To capture the essence of the mechanism, I focused on helices 1 and 7 and 4 and 10. The order parameter for the periplasmic gate was defined as the distance between the pair of C_α atoms in helices 1 and 7 with the shortest inter-helical distance in the wild-type crystal structure of LacY: this is between I32 and N245. The order parameter for the cytoplasmic gate is defined in a similar way. The pair of residues on helices 4 and 10 that shows the smallest $C_\alpha - C_\alpha$ distance in the homology model of periplasmic open LacY was determined. The cytoplasmic order parameter is therefore defined as the distance between the $C_\alpha - C_\alpha$ atoms of residues E126 and C333. The choice of order parameters is consistent with previous studies of gating, as discussed in more detail in section B.1.4.

When the periplasmic and cytoplasmic gates are closed, the transmembrane helices 1 and 7 as well as 4 and 10 are kinked, enabling them to pack tightly against each other, sealing off any pathways across the membrane. Sequence motifs that facilitate helix bending by disrupting the trans-membrane helices, such as prolines, are conserved across the MFS family (Radestock and Forrest, 2011; Vardy et al., 2005). This observation suggests that key features of the gating mechanism could be shared between MFS transporters. By comparing the conformations of the gates in both the available crystal structures and the simulations the hypothesis was tested that the overall conformational changes in the transport cycle are conserved across the MFS superfamily (Radestock and Forrest, 2011). For this order parameters for 12 different MFS transporters were derived using the definitions for transmembrane helices 1 and 7 (for the periplasmic gate) and 4 and 10 (of the cytoplasmic gate) reported in Table 6.2. The determined order parameters for the gates are listed in Table 6.3.

The order parameters track the state of the gates

The order parameters tracked the state of the gates in the simulations (Fig. 6.11A). Cytoplasmic open LacY was predominately in a partially cytoplasmic occluded conformation as shown in Fig. B.4A which highlights the most populated structures. Periplasmic open LacY remained largely in this functional state as judged by the order parameters for the two gates. The order parameter analysis agreed with both the pore profile and water density analysis (A more detailed comparison is provided in Appendix B.1). The equilibrium simulation of the occluded structure, *Occ*, sampled a region of order parameter space (Fig. 6.11A) intermediate between cytoplasmic and periplasmic open LacY. The cytoplasmic order parameter

Table 6.3 Order parameters for MFS transporters. The cytoplasmic gate was tracked by the $C_\alpha - C_\alpha$ distances between TM4 and TM10. The order parameter for the periplasmic gate was defined by a $C_\alpha - C_\alpha$ between TM1 and TM7. The residues defining the cytoplasmic and periplasmic order parameters are listed below.

Transporter	Cyt. gate TM4-TM10	Peri. gate TM1-TM7
ErmD	Thr 119 and Phe 311	Ala 32 and Cys 229
FucP	Thr 143 and Ser 370	Gln 51 and Ile 282
GlpT	Trp 138 and Ile 368	Pro 53 and Thr 278
LacY	Glu 126 and Cys 333	Ile 32 and Asn 245
NarU	Met 149 and Phe 367	Ser 54 and Ala 273
NarK	Met 151 and Phe 370	Ser 56 and Ala 275
PepT _{So}	Ser 131 and Leu 427	Thr36 and Ser 320
PepT _{St}	Ser 130 and Gly 407	Ala 34 and Ser 303
PiPT	Ala 154 and Ala 439	Asn 53 and Leu 332
POT	Asn 138 and Gly420	Val 47 and Thr 314
Xyle	Met 149 and Ser 396	Ser 32 and Tyr 298
YajR	Ser 125 and Ser 328	Thr 37 and Pro236

remained predominately <13 Å while the periplasmic order parameter was < 8 Å. Both gates were shut which is consistent with the calculations of the pore radius (Fig. 6.6B).

The order parameters classify the MFS transporter crystal structures

The order parameters intuitively classified all the available MFS transporter crystal structures (Fig. 6.11A) even though the pairwise sequence identity between members of the MFS is typically low, *e.g.* no more than 12% between LacY and any of the other transporters studied here (Prlić et al., 2010) up to a maximum of 46% for the closely related peptide transporters PepT_{St} and POT (Doki et al., 2013). The cytoplasmic open crystal structures for LacY, GlpT, PepT_{St}, POT and the inward open structure of Xyle mapped onto the region of order parameter space sampled by the simulations of cytoplasmic open LacY. The two virtually identical FucP crystal structures were close to the region of order parameter space explored by the simulations of periplasmic open LacY, while the recently solved periplasmic open structure of YajR was located slightly farther outside. The crystal structures of NarK and PepT_{So} border on the region of order parameter space sampled by the simulations of cytoplasmic open LacY during which it partially closed, while NarU B is only slightly removed, all of which is consistent with the partially cytoplasmic closed structure of LacY constituting a transition intermediate (Holyoake and Sansom, 2007). Importantly, the occluded crystal structures of PepT_{So}, EmrD, NarU (A and B), NarK, and cytoplasmic

partially occluded structure of Xyle mapped to the intermediate region accessed by *Occ*, with the occluded structure of PiPT close by. Classifying the crystal structure of periplasmic partially occluded Xyle (Sun et al., 2012) using the order parameters put it outside the region sampled by *Occ*. While the periplasmic order parameter for Xyle was similar to EmrD (8.0 Å), the cytoplasmic order parameter for Xyle was very low, 5.1 Å. This is much lower than the cytoplasmic order parameters seen in the simulations or any other MFS transporter structure and probably reflects the presence of the additional intracellular helices that pack close to the cytoplasmic gate and that are not present in the other MFS transporter structures.

Comparison to intermediate models and structures for LacY

Recently, more structural information has become available for MFS transporters such as LacY. A model for the occluded state of LacY was generated by homology to the crystal structure of partially occluded PepT_{So} (Madej et al., 2012). Even though the model by Madej *et al.* and my model were derived by completely different methods, they are generally quite similar with a C_α RMSD of 3.5 Å. The similarity is also emphasised by the order parameters for the gates of the occluded homology model. The cytoplasmic as well as the periplasmic gate are closed in the homology model and the model maps onto the region of the 2D order parameter plot (Fig. 6.11A) populated by occluded structures from the equilibrium simulation of the occluded state, *Occ*.

Very recently a new crystal for LacY has become available (Kumar et al., 2014). Two mutations enabled the trapping of LacY in a partially periplasmic occluded state. As shown in Fig. 6.11 the conformations from this new X-ray structure, 4OAA A and 4OAA B, map onto the region in the order parameter space sampled by the simulations of periplasmic open LacY. The periplasmic gate is partially closed in this region which agreed with the classification of this structure. This region is also close to the intermediate area occupied by *Occ*.

The cytoplasmic and periplasmic gates in simulations with CHARMM

In order to assess the robustness of the conclusions a fairly drastic change was introduced into equilibrium simulations by changing the underlying force field from GROMOS96 to CHARMM22. As discussed in Appendix B.1, both sets of simulations behave very similarly and in particular cover similar regions of order parameter space (Fig. B.4B). Thus, the order parameter analysis and the simulations are robust against changes in the simula-

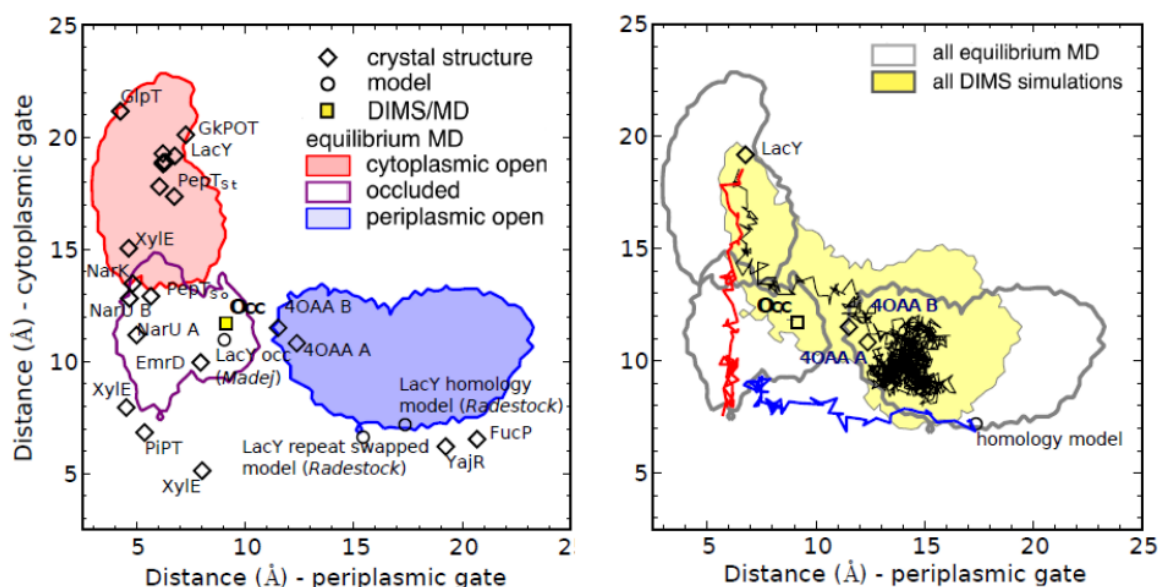


Fig. 6.11 MFS transporter crystal structures can be characterised by two order parameters based on their cytoplasmic and periplasmic gates. (A) Individual structures are shown together with equilibrium MD simulations: Experimental crystal structures (diamonds with labels), models of LacY (circles with labels, Madej et al., 2012; Radestock and Forrest, 2011), and the occluded model generated by DIMS and equilibrium MD (square). The red area indicates the range of order parameter values sampled in the simulations started from the cytoplasmic open state (LacY crystal structure 2V8N, Guan et al., 2007). Similarly, the purple outline corresponds to simulation *Occ* started from the occluded DIMS structure and the blue area shows the range of all simulations started from the FucP-based homology model from Radestock and Forrest, 2011 in the periplasmic open state. (B) The range of order parameters covered by an ensemble of DIMS MD simulations from the cytoplasmic open state (LacY) towards the periplasmic open state (homology model) is shown as a filled yellow area (simulations *DIMS-c.open-p.open*). One example DIMS transition is drawn as a thin black line. Occluded conformations were generated by special DIMS simulation that either started from the cytoplasmic open state and closed the periplasmic gate (red line, simulation *DIMS-closure*) or began from the periplasmic open state and closed the cytoplasmic gate (blue line). The endpoint of *DIMS-closure* was simulated for 100 ns with equilibrium MD (simulation *Occ*, purple area in A) and resulted in the structural model for the occluded state labelled “Occ” (square). The ranges of order parameters sampled by the equilibrium simulations in A are shown as grey outlines.

tion parameters and are therefore likely to reflect upon a real property of the LacY protein, namely the flexibility of its cytosolic and periplasmic gates.

Taken together the analysis suggests that key features of the gating mechanism are conserved across the MFS transporters and a meaningful structural comparison is possible even in cases such as Xyle where additional structural elements interact with the gates. In particular, the model for the occluded conformation of LacY is consistent with the known structural evidence for occluded conformations in the MFS superfamily, as revealed by the order parameter analysis.

6.4.6 DIMS simulations of the transition between cytoplasmic and periplasmic open LacY

In order to test the hypothesis that the conformational change proceeds via an intermediate occluded state we generated an ensemble of 80 transitions between the end states. These DIMS MD simulations (*DIMS-c.open-p.open*) started with the cytoplasmic open wild-type crystal structure of LacY and used as the target structure the entire N- and C-terminal domains (residues 7-185 and 220-399) of the homology model of periplasmic open LacY. During the DIMS MD simulations, the structure of the transporter changed from a cytoplasmic open/periplasmic closed state to a cytoplasmic closed/periplasmic open conformation. The transitions can be followed in Fig. 6.12. Pore profiles were calculated for the 80000 frames that constitute this ensemble. The pore profiles were averaged using the backbone RMSD to the target structures. The contour map in Fig. 6.12 then shows the average pore radius at a given z position and RMSD to the target state. At large RMSD to the target structure, in the initial phase of the transitions, the cytoplasmic gate was open. The periplasmic gate remained closed. Around 4 Å distance to the target state, both cytoplasmic and periplasmic gate were constricted on average, as shown by the yellow contours. This region is indicated by a dashed line on Fig. 6.12 and a representative intermediate structure is shown. The structure demonstrates constrictions on the cytoplasmic and periplasmic sides. From there on the cytoplasmic side appeared to close further, while the periplasmic side opened up increasingly.

These transitions were also followed using the order parameters for the cytoplasmic and periplasmic gates. The trajectories connect the broad wells sampled by the equilibrium MD simulations of cytoplasmic and periplasmic open LacY in the order parameter space (Fig. 6.11B). Intermediate structures in the DIMS transitions between cytoplasmic and periplasmic open LacY were occluded according to this analysis which agreed with the

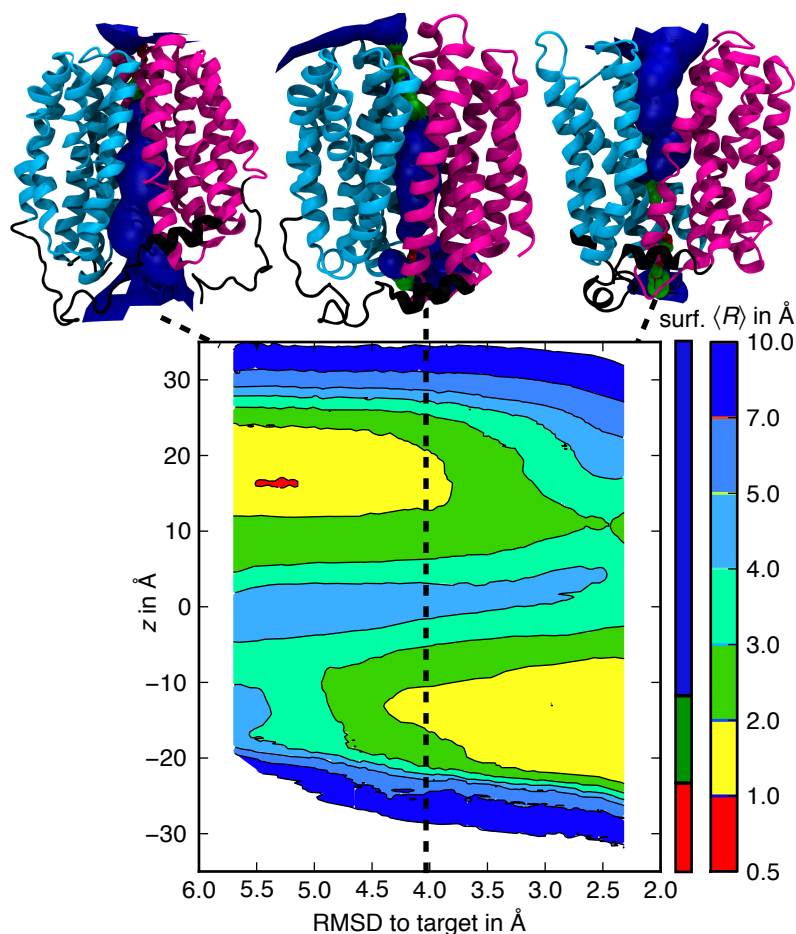


Fig. 6.12 Average pore profiles for DIMS transitions between cytoplasmic and periplasmic open LacY. The cytoplasmic open starting, occluded DIMS intermediate, and the periplasmic open target structure are shown. The N-terminal (residues 7-185) and C-terminal (residues 220-399) domains are shown in cyan and magenta with the flexible linker connecting the domains in black. Pore surfaces calculated with HOLE (Smart et al., 1993) for cytoplasmic open, an occluded intermediate, and periplasmic open LacY are shown. Red, green and blue pore surfaces indicate pore positions with radii where no water (<1.15 Å), single file of water (between 1.15 Å and 2.3 Å) and multiple water molecules (>2.3 Å) could be accommodated, as specified by the left colour bar. Averaged pore profiles are plotted against the backbone RMSD distance of the N and C-terminal domains to the periplasmic open target structure. The average pore radius over the whole DIMS trajectory ensemble at a given z position and progress variable (backbone RMSD to target) is indicated by the coloured contours. The radius corresponding to colours is specified by the right colour bar.

pore radius calculations. These structures occupy a region in order parameter space that is sampled by *Occ* in equilibrium MD simulation (Fig. 6.11A,B). Importantly, a model for a partially cytoplasmic occluded conformation of LacY mapped to the region where *Occ* and *DIMS-c.open-p.open* overlapped, as would be expected for such a structure. Moreover the crystal structures of partially periplasmic occluded LacY (4OAA A and B) are located right next to this area and overlays with the transitions in *DIMS-c.open-p.open*. Thus, two different DIMS MD simulations produced structures with both their gates impermeable to H^+ and sugar (1) a full transition between cytoplasmic and periplasmic open LacY (*DIMS-c.open-p.open*) and (2) a transition using partial structural information to close the cytoplasmic half of cytoplasmic open LacY (*DIMS-closure*) together with subsequent equilibrium MD. The intermediate structures agreed with a model derived by a very different method and are consistent with a new structure from X-ray crystallography.

The results of the simulations lie somewhere between a strict rocker-switch mechanism, where the gates open simultaneously (Dang et al., 2010; Forrest et al., 2011), and a strict air-lock mechanism (Forrest, 2013), where the gates open sequentially (Fig. 6.2B). A single, diagonal straight line would be expected to connect the cytoplasmic and periplasmic open states in the order parameter plot (Fig. 6.11) in the case of a strict-rocker-switch mechanism, and an “L”-shape for the latter. The simulated transitions between the cytoplasmic open and periplasmic open conformations using DIMS (*DIMS-c.open-p.open*) did not fall on either set of hypothetical lines.

6.4.7 Structural comparison of cytoplasmic, occluded and periplasmic open LacY

In a strict rocker-switch mechanism (Dang et al., 2010) a rigid-body motion of the N- and C-terminal domains of MFS transporters would inter-convert cytoplasmic open to periplasmic open structures. Relatively small changes within the domains would hence be expected when comparing the structures of the cytoplasmic open and periplasmic open states. Aligning the N-terminal domains of the 2V8N crystal structure with the N-terminal domain of the homology model for periplasmic open LacY gave C_α RMSD of 3.6Å. A similar results was obtained for the C-terminal domains. In addition, the N- and C-terminal domains of cytoplasmic open LacY and the occluded model were compared. The considerable differences within the domains of structures representing the cytoplasmic open, occluded and periplasmic open states argues against a strict rocker-switch mechanism.

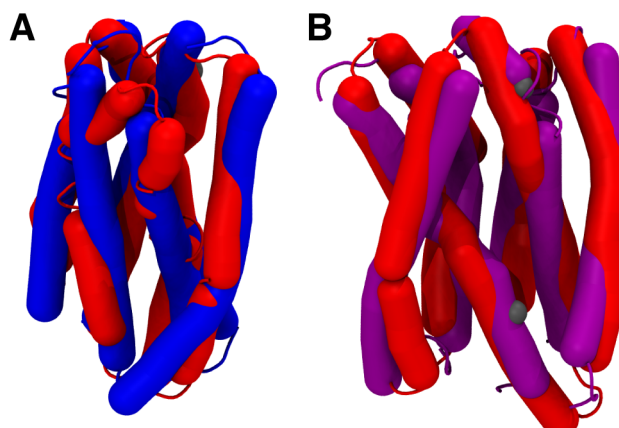


Fig. 6.13 Structural comparison of the N-terminal domains of cytoplasmic and periplasmic open LacY (A) The C_{α} atoms of the N-terminal domains (residues 7 to 185) of cytoplasmic (red) and periplasmic open (blue) LacY were aligned using VMD6 with a C_{α} RMSD of 3.6 Å. (B) The C_{α} atoms of the N-terminal domains (residues 7 to 185) of cytoplasmic and occluded LacY were aligned with a C_{α} RMSD of 3.2 Å. The transmembrane helices of the cytoplasmic open transporter are shown in red and the helices of the periplasmic open one in purple. The C_{α} atoms for I32 (at the top) and E126 (at the bottom) are shown in van der Waals representation (in grey) to indicate the location of the periplasmic and cytoplasmic gates. The transmembrane helices are depicted using Bendix (Dahl et al., 2012).

6.5 Discussion

In this chapter the conformational change between cytoplasmic and periplasmic open LacY was investigated by molecular simulations. A novel method, Dynamics Importance Sampling (DIMS) MD enabled the investigation of key aspects of the transition which happens on a much slower time scale than the one accessible with current equilibrium MD simulations. Structures from the simulations were validated by comparison to DEER experiments and crystal structures across the MFS superfamily.

6.5.1 Conformational changes in LacY

In MD simulations the X-ray crystal structure of cytoplasmic open LacY quickly relaxed to a more occluded structure that would prevent sugar permeation. The results are consistent with earlier studies that observed a similar closure of the cytoplasmic gate in multiple equilibrium MD simulations (Andersson et al., 2012; Holyoake and Sansom, 2007) or a more limited closure after 10 ns of simulation (Yin et al., 2006). Considering possible proton permeation pathways through the computation of water densities as well as electrostatics

calculations suggested that the initial closure would not lead to a fully occluded structure. Protons would still be able to permeate the cytoplasmic gate. The order parameter analysis highlighted the similarity of such structures to MFS transporter crystal structures with the cytoplasmic gate partially closed, such as PepT_{So}, NarK and NarU B. Taken together, these results indicate that the initial closure observed in this and other MD studies leads to a functionally relevant intermediate.

6.5.2 The occluded state of LacY

To investigate the transition between cytoplasmic and periplasmic open LacY in atomistic simulations a considerable enhancement in sampling is required. I have used DIMS MD to probe the rare transition events. In DIMS MD simulations the cytoplasmic gate of the cytoplasmic open X-ray structure of LacY could be closed using the cytoplasmic half of the periplasmic open homology model as the target structure. The generated structure is stable in an un-restrained equilibrium MD simulation and both the cytoplasmic and the periplasmic gates are closed as judged by pore radius calculations. The pore profile is similar to the one observed in MD simulations of the occluded structure of the related MFS transporter EmrD (Baker et al., 2012). Water density and electrostatics calculations further confirmed the conclusion that the structure would prevent both sugar and proton permeation between the central cavity and the cytoplasm and periplasm.

The model of the occluded state is consistent with the available experimental data. The simulations were compared to DEER experiments. Direct comparison was enabled by taking into account the flexibility of the spin-label through an approach developed by Dr Philip W. Fowler. Spin-spin distributions for residue pairs 73-340 and 136-340 coincided with minor peaks in the experimental distributions that could originate from transporters with the cytoplasmic gate closed. The calculated distribution for 164-310 is overlapped with experimental peaks that have been attributed to conformations with both gates closed. The occluded structure is relatively similar to the homology model of LacY that was built based on the cytoplasmic partially occluded structure of PepT_{So} (Madej et al., 2012). Encouragingly, the order parameter analysis showed that my occluded model sampled a region of the 2D order parameter space that is close to the occluded crystal structures of PepT_{So}, EmrD, NarU (A and B), NarK, the cytoplasmic partially occluded structure of XyleE and the partially periplasmic closed crystal structure of LacY. This crystal structure only became available after the simulation study had been published.

6.5.3 The mechanism of the conformational transition

Two different mechanisms have been proposed for the conformational changes in LacY. In the rocker-switch mechanism, a rigid-body motion of the N- and C-terminal domains would interconvert the cytoplasmic and periplasmic open states. The gates would move simultaneously. Alternatively, the gates might open and close sequentially, as in the operation of an airlock. The intradomain flexibility observed for LacY argues against the rigid-body motion expected in a literal rocker-switch mechanism. A simple structural comparison between the N- and C-terminal domains of cytoplasmic open, periplasmic open and occluded LacY suggested that there are considerable differences within the domains in the three conformational states. However based on the recent crystal structure much smaller difference were found between the domains of cytoplasmic open and partially periplasmic occluded LacY (Kumar et al., 2014). The crystal structure for partially periplasmic occluded LacY was solved by molecular replacement with the cytoplasmic open LacY structure. One can then speculate whether the bias introduced by using the phase information from the N- and C-terminal domains of the cytoplasmic open structure (2V8N) might make the two structures appear more similar than they really are. Both the 2V8N and the periplasmic open 4OAA crystal structures are relatively low resolution structures with resolutions of 3.6Å and 3.5Å respectively, which would compound any model bias. MD simulation of the new crystal structure would probe the similarity of the N- and C-terminal domains in the two LacY structures. Nonetheless there is much evidence for flexibility in LacY, which has been highlighted in H/D exchange experiments (le Coutre et al., 1998; Patzlaff et al., 1998; Sayeed and Baenziger, 2009) and intra-domain cross-linking experiments. Furthermore, equilibrium MD simulations of LacY have suggested its individual helices are flexible (Andersson et al., 2012; Holyoake and Sansom, 2007) during the initial stages of the conformational switch. SGLD simulations (Pendse et al., 2010) of LacY have shown intra-domain re-arrangements of the helices as well as global motions during the transition itself. Based on the trajectories generated in this work, some conclusions about the transition mechanism can be drawn.

First, in DIMS MD simulations the two gates, e.g in *DIMS-closure*, could be closed independently of opening the other gate. This argues against a strict rocker-switch mechanism. However more quantitative work such as calculating the free-energy changes for this transformation would be required to be sure of the thermodynamic feasibility of sequential gate closure. As the final structure from *DIMS-closure* relaxed in an atomistic membrane in *Occ*, the packing of both the cytoplasmic and the periplasmic gates became looser, which would not be expected for a true "airlock" mechanism. The DIMS trajectories for the transition from cytoplasmic open to periplasmic open LacY also showed this trend. They neither

conformed to a strict rocker-switch mechanism, as the closure of one gate and the opening of the other gate did not occur simultaneously as judged by the order parameters. Yet the transition did not resemble a true "airlock" mechanism either. The gates would be expected to be more closely packed in an intermediate structure for a strict "airlock" mechanism than observed in *DIMS-closure*.

6.5.4 Conclusions

To conclude, molecular simulations gave insight into the transition mechanism of LacY. The cytoplasmic and the periplasmic gate are flexible. The flexible gates enable the formation of occluded transition intermediates that are consistent with the available crystal structures and DEER spectroscopy data. A true "rocker-switch" mechanism based on rigid body motions of the N- and C-terminal domains can be ruled out.

Appendix A

Relaxation dispersion experiments on nDsbD

A.1 Relaxation dispersion in the secondary binding site of nDsbD

For both oxidised and reduced nDsbD, relaxation dispersion effects were found for F22. The aromatic side-chain is solvent exposed and might be flexible. This may explain why relaxation dispersion is detected for the amide of F70. For both oxidised and reduced nDsbD the curves appear to level off close to 500Hz. The residues in the cap loop region appear to require a stronger refocusing field; the curves stabilise at higher ν_{CPMG} values (around 700Hz-800Hz according to Fig. 4.6).

F22 is far removed from the cap loop region. It was thus excluded from the analysis of the dynamics in the cap loop using CATIA. F22 is the secondary binding site for DsbC binding (Haebel et al., 2002). The conformational flexibility of F22 may be important for DsbC binding.

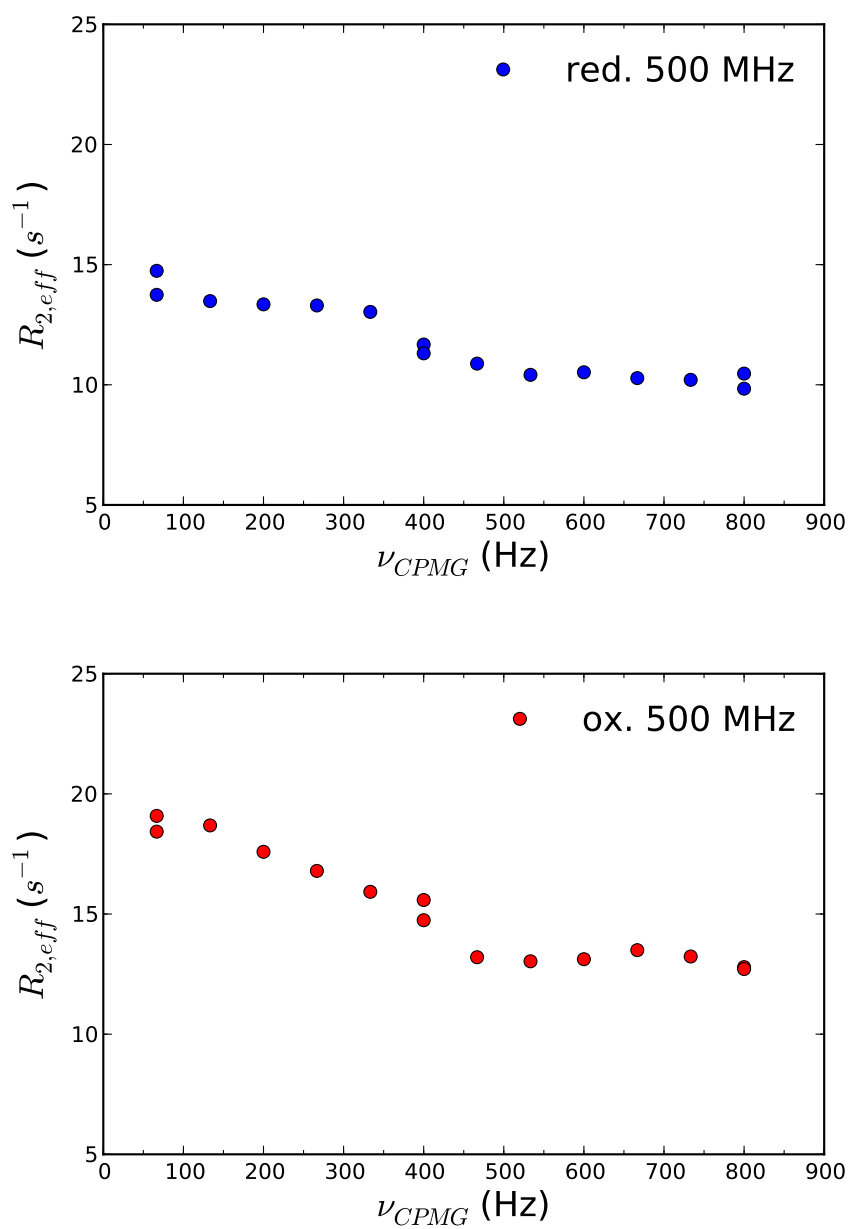


Fig. A.1 For reduced and oxidised nDsbD relaxation dispersion was found for F22.

Appendix B

Additional analysis of the simulations of LacY

B.1 Additional analysis of the simulations of LacY

Here additional analysis of the molecular simulations with LacY are provided.

B.1.1 Additional pore profile and water density analysis

The partial closure of the cytoplasmic gate from a radius of 4 Å in the crystal structure to about 1 Å, which was observed in simulation *CopenA1* (Fig. 6.6A in the main text) is maintained over another 100 ns (*CopenA2*, an extension of simulation *CopenA1*) as seen from the pore radius profile in Fig. B.1A. A repeat of the simulation (starting again from the crystal structure) but using a larger membrane patch (Fig. B.1B) and also when using the CHARMM force field (Fig. B.1C) confirmed the partial closure, although only to a minimum radius of 2 Å. The simulation of the periplasmic open state with the CHARMM force field showed a similar partial closure of the periplasmic gate from 4 Å to slightly above 2 Å. Taken together, these simulations all indicate that the gates of LacY are rather flexible and at least partially mobile on the 100-ns time scale.

Despite the reduction of the radius of the access pathway from 4 Å to about 2 Å as observed in the simulations *CopenB2*, *CopenC*, and *PopenB* (Fig. B.1C,B,D) there was still continuous water density visible, as shown by the water density analysis (Fig. B.2C,B,D). Thus, even in cases where partial gate closure was observed, a continuous water pathway existed from the bulk region to the central cavity. Therefore, neither the repeat simulations nor the ones using a different force field produce a true occluded state.

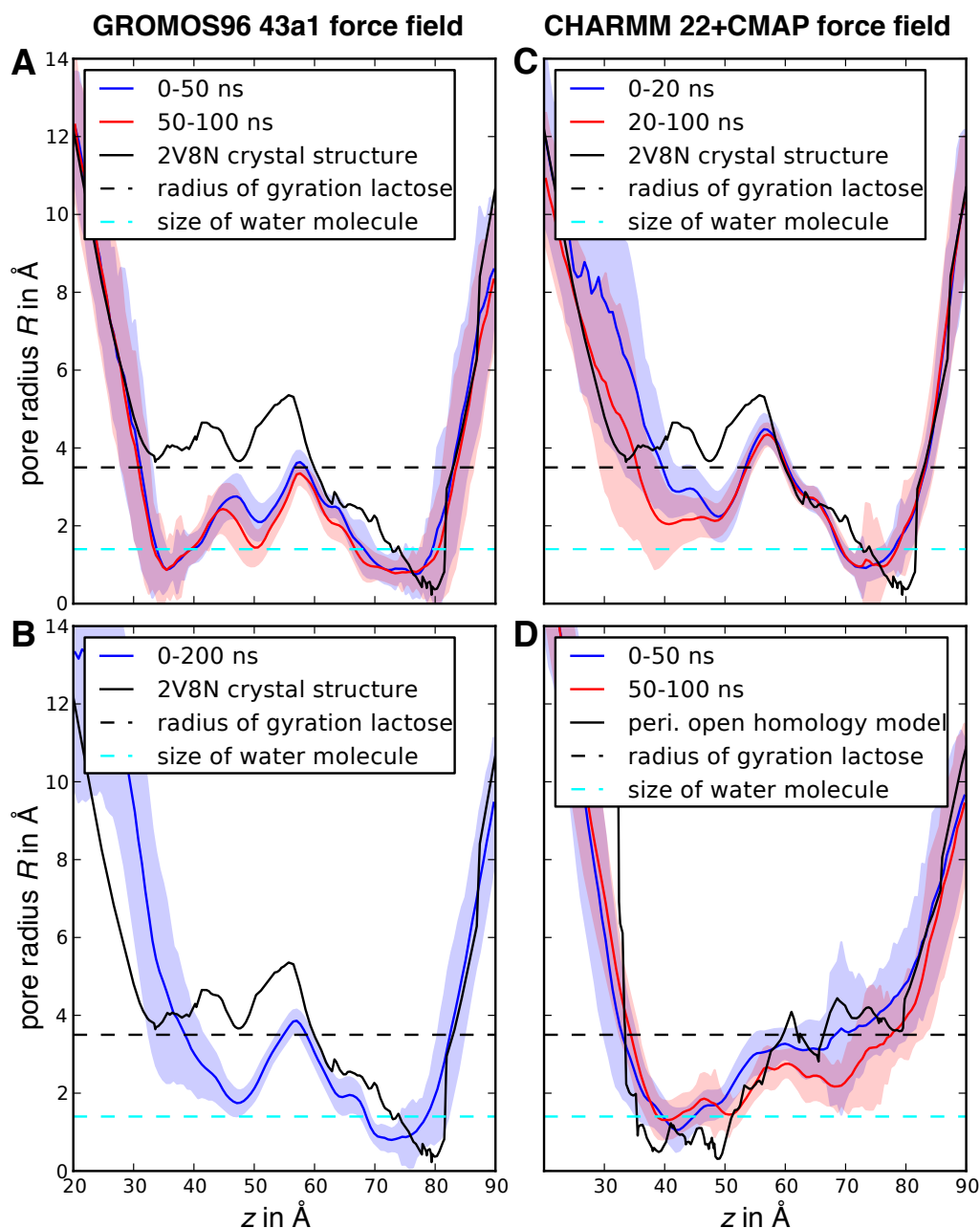


Fig. B.1 Pore profiles for *CopenA2*, *CopenB2*, *CopenC* and *PopenB*. Pore radii determined by HOLE (Smart et al., 1993) are plotted against z -coordinates of the cavity for each simulation and the crystal structure (2V8N) or the FucP-based homology model of the periplasmic open state. The average of the pore radius profile over the indicated time of the trajectory is plotted as a solid line together with the standard deviation as a shaded band. For scale, the size of a water molecule (cyan dashed line) and a lactose molecule (black dashed line) are indicated. (A) Simulation *CopenA2* (100-ns extension of *CopenA1*, shown in Fig. 6.6A in the main text) (B) Simulation *CopenB2*. (C) *CopenC* (CHARMM force field, cytoplasmic open). (D) *PopenB* (CHARMM force field, periplasmic open).

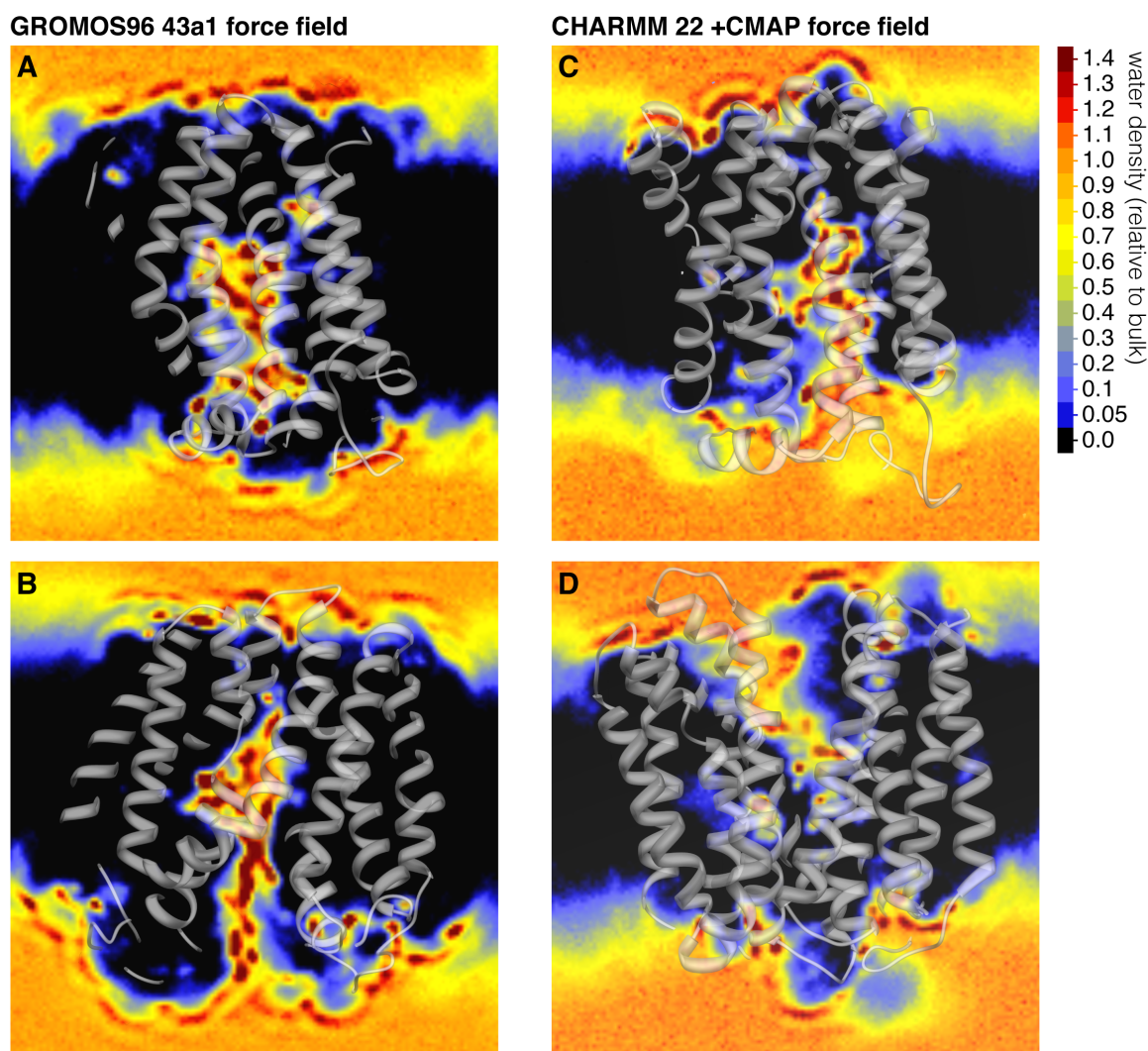


Fig. B.2 Water densities for *CopenA2*, *CopenB2*, *CopenC* and *PopenB*. Slices through the water density (relative to bulk SPC water Hermans et al., 1984) are shown together with the protein in grey cartoon representation. Selected simulations performed with the GROMOS96 43a1 force field (see also Fig. 6.6D–F in the main text): (A) *CopenA2*. (B) *CopenB2*. Simulations performed with the CHARMM 22 + CMAP force field: (C). *CopenC*. (D) *PopenB*. Densities were visualised in USCF Chimera (Goddard et al., 2007; Pettersen et al., 2004).

Table B.1 Molprobit quality evaluation of the 2V8N crystal structure and the model for occluded LacY.

	2V8N:A, 1-417	Model from <i>Occ</i> , 7-399	Goal
Poor rotamers	53 (15.32 %)	25 (8.56 %)	<1 %
Ramachandran outliers	74 (17.83 %)	16 (4.09 %)	<0.05 %
Ramachandran favoured	228 (54.94 %)	314 80.31 %	>98 %
C_β deviation >0.25 Å	0 (0 %)	62 (17.27 %)	0
Bad backbone bonds	0/1667 (0 %)	5/1570 (0.32 %)	0 %
Bad backbone angles	8/2082 (0.38 %)	41/1961 (2.09 %)	<0.1 %

B.1.2 Evaluation of the occluded model from DIMS and MD simulation

The structural quality of the model for the occluded conformer of LacY was evaluated using the Molprobit server (Chen et al., 2010). The quality of the model is comparable to the 3.6 Å crystal structure (2V8N, Guan and Kaback, 2006) of cytoplasmic open wild-type LacY (Table B.1). It performs better than the 2V8N crystal structure in respect to rotameric states and Ramachandran plot statistics, while C_β and backbone geometry is worse than in the wild-type crystal structure although it needs to be remembered that this model is a snapshot from an equilibrium MD simulation at finite temperature and hence bonds and angles are expected to slightly diverge from ideal geometry due to thermal fluctuations.

B.1.3 Additional analysis of DEER distances

Fig. B.3A shows an analysis of DEER distances that were only present in the simulations based on the crystal structure (labelled residues 73–401, 136–401, and 137–401); the FucP-based homology model of periplasmic open LacY (Radestock and Forrest, 2011) misses part of the C-terminus including residue 401 so that we could not compute DEER distance distributions for simulations based on the model (*PopenA*, *Occ*). The computed distributions from the cytoplasmic open simulation agree well with the locations of the primary peaks from the experiments with the non-binding sugar NPGlc, which shifts the conformational equilibrium of LacY towards the cytoplasmic open state (Smirnova et al., 2007). When the same distances are compared to experimental data with the binding sugar NPGal, which shifts the equilibrium away from the cytoplasmic open state, it is observed that the computed peaks now agree with secondary or tertiary experimental peaks (Fig. B.3B).

Comparisons of DEER distances from the occluded and periplasmic open simulation with the experimental distributions from the experiments with NPGal (Fig. B.3B) generally

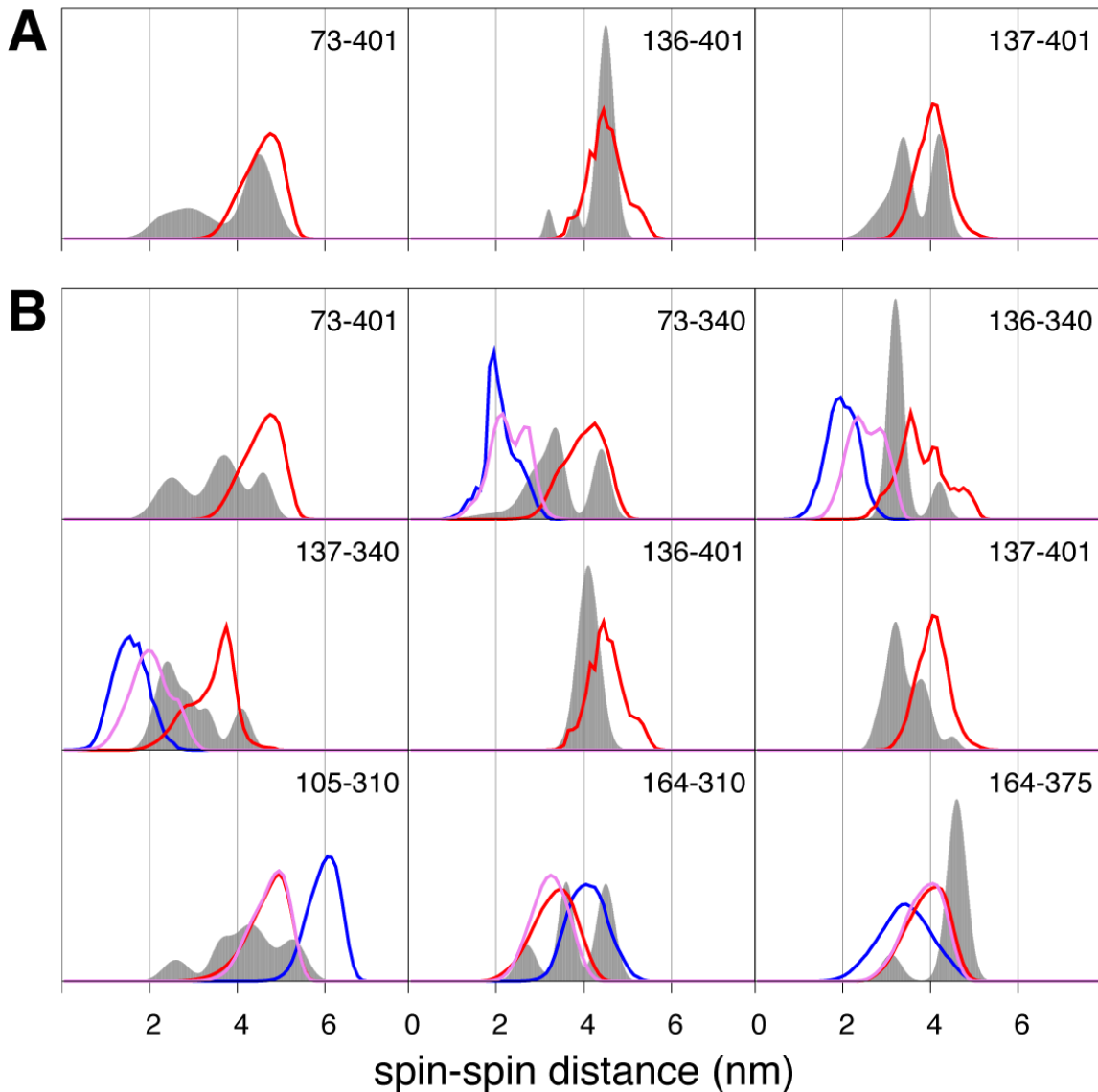


Fig. B.3 Comparison of DEER distance distributions from simulations and experiments. Experimental distributions are shown in grey. Distance distributions from simulations of cytoplasmic and periplasmic open as well as occluded LacY are indicated by red, blue and magenta lines. (A) Comparing additional residues pairs (73-401, 136-401 and 137-401) to the experimental distributions of predominately cytoplasmic open LacY. Experiments were in presence of non-binding sugar NPGlc. For some distances 73-401, 136-401 and 137-401, only the simulations of cytoplasmic open LacY can be compared to the experiments. The homology model of the periplasmic open state and the model for the occluded state derived from DIMS MD are missing some residues at the extreme C-terminus such as residue 401. (B) Comparing simulations to the DEER distributions of predominately periplasmic open LacY. Experiments were performed in presence of binding sugar NPGal that shifts the conformational equilibrium (Smirnova et al., 2007).

exhibit the expected pattern where calculated peaks of distances on the cytoplasmic side (involving residues 73, 136, 137) are shifted towards smaller distances, indicating cytoplasmic gate closure. The calculated peak centres co-locate in the vicinity of the highest experimental peaks because NPGal shifts the conformational equilibrium towards the periplasmic open (and the occluded) state where the cytoplasmic gate is closed. Similarly, calculated distributions from the occluded and cytoplasmic open simulations coincide on the periplasmic side (distances involving residues 105 and 164) although those peaks agree less well with the experimental data. As discussed in the main text, the distance 164–375 (on the periplasmic side) suffers from an unexpected local conformational change in the simulation *PopenA* that brings helices 5 and 11 closer together and thus leads to a distance that is smaller in the periplasmic open state than in the cytoplasmic open one.

B.1.4 Order parameter analysis of the MD simulations

The choice for the cytoplasmic and periplasmic order parameters agrees with previous studies of the gating mechanism of MFS transporters. The order parameter for the periplasmic gate of LacY was defined as the C_α distance between I32 and N245 on helices 1 and 7. Cross linking experiments based on computer simulations (Jensen et al., 2007) implicated N245 and nearby I40 (in the helix 1–2 linker) in periplasmic gating (Zhou et al., 2009). The state of the cytoplasmic gate of LacY was tracked using the C_α distance between E126 and C333 on helices 4 and 10. These are close to residues F140, F334 and Y350, on helices 5, 10 and 11, respectively, which are conserved and have been implicated in cytoplasmic closure (Kasho et al., 2006). In addition the order parameter for the periplasmic gate (S32–Y298) of Xyle agrees with previous work. Y298 is thought to prevent D-xylose permeation (Sun et al., 2012).

The order parameter analysis for the simulations with cytoplasmic and periplasmic open LacY agrees with the functional characterisation based on water densities and pore profiles. In the equilibrium MD simulations of cytoplasmic open LacY the cytoplasmic order parameter, 19 Å for the crystal structure, decreased in value, indicating closure of the cytoplasmic gate (as shown by the pore radius calculations), but not far enough to suggest a fully closed gate (as evidenced by the water densities). The periplasmic gating distances were mostly < 8 Å and so the periplasmic gate remained closed during the course of the simulations as indicated by the contour lines in Fig. B.4A. The order parameters for the simulations of periplasmic open LacY showed that it remained in this state. Limited periplasmic closure was detected for simulation *PopenB* (Fig. B.4B), conducted with the CHARMM force field,

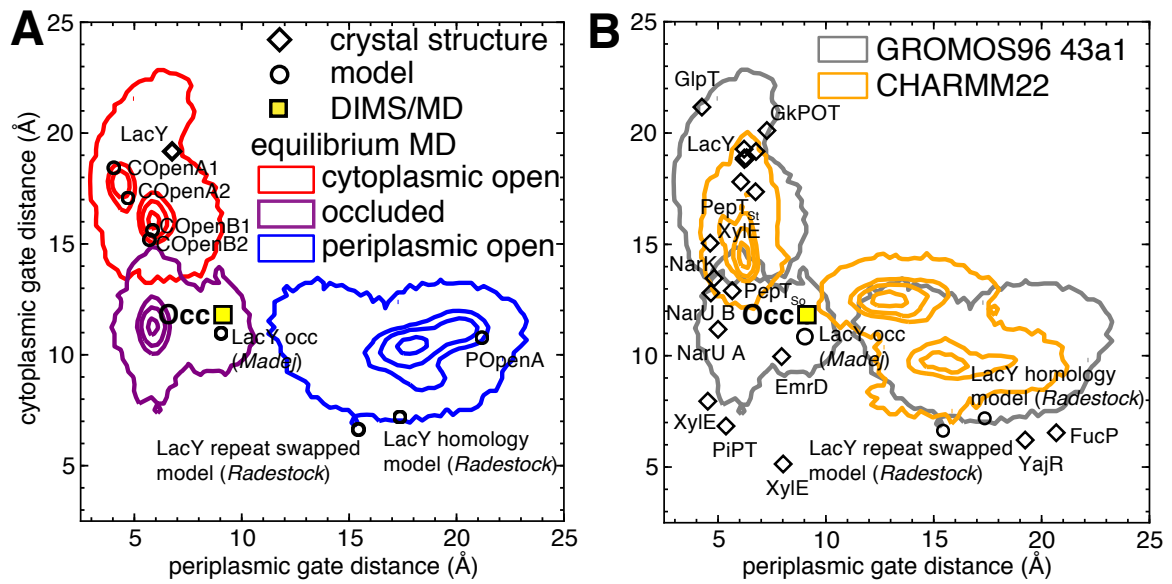


Fig. B.4 . In A the GROMOS trajectories are analysed in more detail. The most populated regions in order parameter space are indicated by contour lines shown at 20%, 40% and 80% of the maximum bin occupancy. The order parameters for the final frames of the simulations with the GROMOS force field (*CopenA1*, *CopenA2*, *CopenB2*, *Occ* and *PopenA*) are depicted. The order parameter space visited in simulations started with cytoplasmic and periplasmic open as well as occluded structures are highlighted in red, blue and purple. Simulations with the CHARMM force field (B) are outlined in orange while the GROMOS trajectories from Fig. 6.11A and Fig. B.4A are shown in grey. Experimental crystal structures are indicated as diamonds.

Table B.2 Order parameters for initial and final frames.

	2V8N	<i>CopenA1</i>	<i>CopenA2</i>	<i>CopenB1</i>	<i>CopenC2</i>	<i>CopenC</i>
average c. gate (Å)	-	17.3 ± 1.2	17.8 ± 0.8	16.4 ± 0.9	16.1 ± 0.6	15.1 ± 1.1
final frame c. gate (Å)	19.2	18.4	17.1	15.2	15.6	13.4

which agreed with the pore radius analysis.

Equilibrium simulations using an alternative force field (CHARMM; orange lines in Fig. B.4) showed very similar order parameter values to the GROMOS96 simulations discussed in the main text. The overlap in the order parameter regions in Fig. B.4B indicates that the observed conformational flexibility in LacY is a robust output from the simulations as it is not strongly dependent on the force field.

The order parameters also enable a more detailed analysis of the initial closure of the cytoplasmic open gate in simulations of cytoplasmic open LacY. Structures with a significantly closed cytoplasmic gate (16 Å as opposed to 19 Å) make up the most populated region in the 2D order parameter space in simulations with the GROMOS force field (Fig. B.4A), as shown by the contour lines. Closed conformations are even more populated in the simulations with the CHARMM force field (Fig. B.4B). The order parameters also reveal differences between the simulations in respect to the initial closure of the cytoplasmic gate. Simulations *CopenA1* and *CopenA2* show a less drastic closure than *CopenB1*, *CopenB2* and *CopenC*, as judged by both their average (Table B.2) and their final structures which are highlighted for the GROMOS simulations in Fig. B.4A. The relatively high order parameter for the final frame for *CopenA1* illuminates the result from Poisson-Boltzmann electrostatics calculation for this frame. While the minimum pore radius dropped to 1 Å in *CopenA1*, a value consistent with an functionally occluded structure, the cytoplasmic gate itself remained open and proton permeation appeared possible as shown by the absence of a significant electrostatic free energy barrier (Fig. 6.7). The water densities calculated from this trajectory also show that no functionally occluded structure was formed (see Fig. 6.6D in the main text).

References

- Abramson, J., Smirnova, I., Kasho, V., Verner, G., Kaback, H. R., and Iwata, S. (2003). Structure and mechanism of the lactose permease of *Escherichia coli*. *Science*, 301:610–615.
- Ahmad, M., Gu, W., Geyer, T., and Helms, V. (2011). Adhesive water networks facilitate binding of protein interfaces. *Nat. Commun.*, 2:261.
- Akke, M., Brueschweiler, R., and Palmer Arthur G, I. I. I. (1993). NMR order parameters and free energy: an analytical approach and its application to cooperative calcium (2+) binding by calbindin D9k. *J. Am. Chem. Soc.*, 115:9832–9833.
- Alsallaq, R. and Zhou, H. X. (2008). Electrostatic rate enhancement and transient complex of protein-protein association. *Proteins Struct. Funct. Genet.*, 71:320–335.
- Andersson, M., Bondar, A. N., Freites, J. A., Tobias, D. J., Kaback, H. R., and White, S. H. (2012). Proton-coupled dynamics in lactose permease. *Structure*, 20:1893–1904.
- Baker, J., Wright, S. H., and Tama, F. (2012). Simulations of substrate transport in the multidrug transporter EmrD. *Proteins Struct. Funct. Bioinforma.*, 80:1620–1632.
- Baker, N. A., Sept, D., Joseph, S., Holst, M. J., and McCammon, J. A. (2001). Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. U. S. A.*, 98(18):10037–41.
- Balbach, J., Forge, V., van Nuland, N. A., Winder, S. L., Hore, P. J., and Dobson, C. M. (1995). Following protein folding in real time using NMR spectroscopy. *Nat. Struct. Biol.*, 2:865–870.
- Baldwin, A. J. and Kay, L. E. (2009). NMR spectroscopy brings invisible protein states into focus. *Nat. Chem. Biol.*, 5(11):808–14.
- Barducci, A., Bonomi, M., and Parrinello, M. (2011). Metadynamics. *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, 1(5).
- Barducci, A., Bussi, G., and Parrinello, M. (2008). Well-tempered metadynamics: A smoothly converging and tunable free-energy method. *Phys Rev Lett*, 100.
- Beckstein, O., Denning, E. J., Perilla, J. R., and Woolf, T. B. (2009). Zipping and unzipping of adenylate kinase: atomistic insights into the ensemble of open $\leftrightarrow$ closed transitions. *J. Mol. Biol.*, 394(1):160.

- Beckstein, O. and Sansom, M. S. P. (2003). Liquid-vapor oscillations of water in hydrophobic nanopores. *Proc. Natl. Acad. Sci. U. S. A.*, 100(12):7063–8.
- Beckstein, O., Tai, K., and Sansom, M. S. P. (2004). Not ions alone: Barriers to ion permeation in nanopores and channels. *J. Am. Chem. Soc.*, 126:14694–14695.
- Bennett, M., D’Rozario, R., Sansom, M. S. P., and Yeagle, P. L. (2006). Asymmetric stability among the transmembrane helices of lactose permease. *Biochemistry*, 45:8088–8095.
- Best, R. B., Hummer, G., and Eaton, W. a. (2013). Native contacts determine protein folding mechanisms in atomistic simulations. *Proc. Natl. Acad. Sci. U. S. A.*, 110:17874–9.
- Bjelkmar, P., Larsson, P., Cuendet, M. A., Hess, B., and Lindahl, E. (2010). Implementation of the CHARMM Force Field in GROMACS: Analysis of Protein Stability Effects from Correction Maps, Virtual Interaction Sites, and Water Models. *J. Chem. Theory Comput.*, 6(2):459–466.
- Boehr, D. D., McElheny, D., Dyson, H. J., and Wright, P. E. (2006). The dynamic energy landscape of dihydrofolate reductase catalysis. *Science (80-.)*, 313:1638–1642.
- Boehr, D. D., Nussinov, R., and Wright, P. E. (2009). The role of dynamic conformational ensembles in biomolecular recognition. *Nat. Chem. Biol.*, 5:789–796.
- Bond, P. J. and Sansom, M. S. P. (2006). Insertion and assembly of membrane proteins via simulation. *J. Am. Chem. Soc.*, 128:2697–2704.
- Bonomi, M., Branduardi, D., Bussi, G., Camilloni, C., Provasi, D., Raiteri, P., Donadio, D., Marinelli, F., Pietrucci, F., Broglia, R. A., and Parrinello, M. (2009). PLUMED: A portable plugin for free-energy calculations with molecular dynamics. *Comput. Phys. Commun.*, 180(10):1961–1972.
- Bouvignies, G., Bernadó, P., Meier, S., Cho, K., Grzesiek, S., Brüschweiler, R., and Blackledge, M. (2005). Identification of slow correlated motions in proteins using residual dipolar and hydrogen-bond scalar couplings. *Proc. Natl. Acad. Sci. U. S. A.*, 102:13885–13890.
- Bouvignies, G., Vallurupalli, P., Hansen, D. F., Correia, B. E., Lange, O., Bah, A., Vernon, R. M., Dahlquist, F. W., Baker, D., and Kay, L. E. (2011). Solution structure of a minor and transiently formed state of a T4 lysozyme mutant. *Nature*.
- Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., and Karplus, M. (1983). CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.*, 4:187–217.
- Brot, N., Collet, J. F., Johnson, L. C., Jönsson, T. J., Weissbach, H., and Lowther, W. T. (2006). The thioredoxin domain of *Neisseria gonorrhoeae* PilB can use electrons from DsbD to reduce downstream methionine sulfoxide reductases. *J. Biol. Chem.*, 281:32668–32675.
- Buck, M., Bouguet-Bonnet, S., Pastor, R. W., and MacKerell, A. D. (2006). Importance of the CMAP correction to the CHARMM22 protein force field: dynamics of hen lysozyme. *Biophys. J.*, 90(4):L36–L38.

- Bussi, G., Donadio, D., and Parrinello, M. (2007). Canonical sampling through velocity rescaling. *J. Chem. Phys.*, 126(1):014101.
- Bustamante, C., Cheng, W., and Meija, Y. X. (2011). Revisiting the central dogma one molecule at a time. *Cell*, 144:480–497.
- Callenberg, K. M., Choudhary, O. P., de Forest, G. L., Gohara, D. W., Baker, N. A., and Grabe, M. (2010). APBSmem: A Graphical Interface for Electrostatic Calculations at the Membrane. *PLoS One*, 5(9):e12722.
- Castrignanò, T., De Meo, P. D., Cozzetto, D., Talamo, I. G., and Tramontano, A. (2006). The PMDB Protein Model Database. *Nucleic Acids res.*, 34:D306–9.
- Cavanagh, J., Fairbrother, W., Palmer III, A., Rance, M., and Skelton, N. (2007). *Protein NMR spectroscopy. Principles and Practice.*, volume 2nd.
- Chandler, D. (2005). Interfaces and the driving force of hydrophobic assembly. *Nature*, 437(7059):640–7.
- Chandrasekhar, I., Clore, G. M., Szabo, A., Gronenborn, A. M., and Brooks, B. R. (1992). A 500 ps molecular dynamics simulation study of interleukin-1 beta in water. Correlation with nuclear magnetic resonance spectroscopy and crystallography. *J. Mol. Biol.*, 226(1):239–250.
- Chang, S., Li, K.-s., Hu, J.-p., Jiao, X., and Tian, X.-h. (2011). Allosteric and transport behavior analyses of a fucose transporter with network models. *Soft Matter*, 7:4661.
- Chaptal, V., Kwon, S., Sawaya, M. R., Guan, L., Kaback, H. R., and Abramson, J. (2011). Crystal structure of lactose permease in complex with an affinity inactivator yields unique insight into sugar recognition. *Proc. Natl. Acad. Sci. U. S. A.*, 108:9361–9366.
- Chen, J., Im, W., and Brooks III, C. L. (2006). Balancing Solvation and Intramolecular Interactions: Toward a Consistent Generalized Born Force Field. *J. Am. Chem. Soc.*, 128(11):3728–3736.
- Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., and Richardson, D. C. (2010). MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallogr. Sect. D Biol. Crystallogr.*, 66:12–21.
- Cho, S.-H. and Beckwith, J. (2009). Two snapshots of electron transport across the membrane: insights into the structure and function of DsbD. *J. Biol. Chem.*, 284:11416–11424.
- Cho, S.-H. and Collet, J.-F. (2012). The many roles of the bacterial envelope reducing pathways. *Antioxid. Redox Signal.*, 18:1690–1698.
- Cho, S.-H., Porat, A., Ye, J., and Beckwith, J. (2007). Redox-active cysteines of a membrane electron transporter DsbD show dual compartment accessibility. *EMBO J.*, 26:3509–3520.
- Chodera, J. D. and Mobley, D. L. (2013). Entropy-enthalpy compensation: role and ramifications in biomolecular ligand recognition and design. *Annu. Rev. Biophys.*, 42:121–42.

- Choy, W.-Y., Zhou, Z., Bai, Y., and Kay, L. E. (2005). An ^{15}N NMR spin relaxation dispersion study of the folding of a pair of engineered mutants of apocytochrome b562. *J. Am. Chem. Soc.*, 127(14):5066–5072.
- Dahl, A. C. E., Chavent, M., and Sansom, M. S. P. (2012). Bendix: Intuitive helix geometry analysis and abstraction. *Bioinformatics*, 28:2193–2194.
- Dang, S., Sun, L., Huang, Y., Lu, F., Liu, Y., Gong, H., Wang, J., and Yan, N. (2010). Structure of a fucose transporter in an outward-open conformation. *Nature*, 467(7316):734–738.
- Darden, T., York, D., and Pedersen, L. (1993). Particle mesh Ewald: An $N \log(N)$ method for Ewald sums in large systems. *J. Chem. Phys.*, 98(12):10089–10092.
- Daura, X., Gademann, K., Jaun, B., Seebach, D., Van Gunsteren, W. F., and Mark, A. E. (1999). Peptide Folding: When Simulation Meets Experiment. *Angew. Chemie Int. Ed.*, 38:236–240.
- Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995). NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR*, 6:277–293.
- Denning, E. J. and Woolf, T. B. (2010). Cooperative nature of gating transitions in $\text{K}(+)$ channels as seen from dynamic importance sampling calculations. *Proteins*, 78:1105–1119.
- Denoncin, K., Vertommen, D., Paek, E., and Collet, J. F. (2010). The protein-disulfide isomerase DsbC cooperates with SurA and DsbA in the assembly of the essential β -barrel protein LptD. *J. Biol. Chem.*, 285:29425–29433.
- Depuydt, M., Leonard, S. E., Vertommen, D., Denoncin, K., Morsomme, P., Wahni, K., Messens, J., Carroll, K. S., and Collet, J.-F. (2009). A periplasmic reducing system protects single cysteine residues from oxidation. *Science*, 326(5956):1109–11.
- Depuydt, M., Messens, J., and Collet, J.-F. (2011). How proteins form disulfide bonds. *Antioxid. Redox Signal.*, 15:49–66.
- Diehl, C., Engström, O., Delaine, T., Håkansson, M., Genheden, S., Modig, K., Leffler, H., Ryde, U., Nilsson, U. J., and Akke, M. (2010). Protein flexibility and conformational entropy in ligand design targeting the carbohydrate recognition domain of galectin-3. *J. Am. Chem. Soc.*, 132:14577–14589.
- Dill, K. A. (1999). Polymer principles and protein folding. *Protein Sci.*, 8:1166–1180.
- Dill, K. A. and Bromberg, S. (2011). *Molecular Driving Forces*.
- Doki, S., Kato, H. E., Solcan, N., Iwaki, M., Koyama, M., Hattori, M., Iwase, N., Tsukazaki, T., Sugita, Y., Kandori, H., Newstead, S., Ishitani, R., and Nureki, O. (2013). Structural basis for dynamic mechanism of proton-coupled symport by the peptide transporter POT. *Proc. Natl. Acad. Sci. U. S. A.*, 110:11343–8.

- Dolinsky, T. J., Nielsen, J. E., McCammon, J. A., and Baker, N. A. (2004). PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids res.*, 32:W665–7.
- Dutton, R. J., Boyd, D., Berkmen, M., and Beckwith, J. (2008). Bacterial species exhibit diversity in their mechanisms and capacity for protein disulfide bond formation. *Proc. Natl. Acad. Sci. U. S. A.*, 105:11933–11938.
- Edeling, M. A., Guddat, L. W., Fabianek, R. A., Thöny-Meyer, L., and Martin, J. L. (2002). Structure of CcmG/DsbE at 1.14 Å resolution: High-fidelity reducing activity in an indiscriminately oxidizing environment. *Structure*, 10:973–979.
- Enkavi, G. and Tajkhorshid, E. (2010). Simulation of spontaneous substrate binding revealing the binding pathway and mechanism and initial conformational response of GlpT. *Biochemistry*, 49:1105–1114.
- Essmann, U., Perera, L., Berkowitz, M. L., Darden, T., Lee, H., and Pedersen, L. G. (1995). A smooth particle mesh Ewald method. *J. Chem. Phys.*, 103(19):8577–8593.
- Fawzi, N. L., Doucleff, M., Suh, J.-Y., and Clore, G. M. (2010). Mechanistic details of a protein-protein association pathway revealed by paramagnetic relaxation enhancement titration measurements. *Proc. Natl. Acad. Sci. U. S. A.*, 107(4):1379–1384.
- Feig, M., Im, W., and Brooks III, C. L. (2004). Implicit solvation based on generalized Born theory in different dielectric environments. *J. Chem. Phys.*, 120(2):903–911.
- Fersht, A. (1999). *Structure and mechanism in protein science: a guide to enzyme catalysis and protein folding*. Macmillan.
- Forrest, L. R. (2013). (Pseudo-)Symmetrical Transport. *Science*, 339(6118):399–401.
- Forrest, L. R., Krämer, R., and Ziegler, C. (2011). The structural basis of secondary active transport mechanisms. *Biochim. Biophys. Acta*, 1807(2):167–88.
- Fowler, P. W. and Sansom, M. S. P. (2013). The pore of voltage-gated potassium ion channels is strained when closed. *Nat. Commun.*, 4:1872.
- Frenkel, D. and Smit, B. (2002). *Understanding Molecular Simulation*.
- Frishman, D. and Argos, P. (1995). Knowledge-based protein secondary structure assignment. *Proteins Struct. Funct. Genet.*, 23:566–579.
- Gabdoulline, R. R. and Wade, R. C. (1998). Brownian dynamics simulation of protein-protein diffusional encounter. *Methods*, 14:329–341.
- Gadsby, D. C. (2009). Ion channels versus ion pumps: the principal difference, in principle. *Nat. Rev. Mol. Cell Biol.*, 10(5):344–52.
- Gardino, A. K. and Kern, D. (2007). Functional dynamics of response regulators using NMR relaxation techniques. *Methods Enzym.*, 423:149–165.

- Gardino, A. K., Villali, J., Kivenson, A., Lei, M., Liu, C. F., Steindel, P., Eisenmesser, E. Z., Labeikovsky, W., Wolf-Watz, M., Clarkson, M. W., and Kern, D. (2009). Transient non-native hydrogen bonds promote activation of a signaling protein. *Cell*, 139(6):1109–18.
- Goddard, T. D., Huang, C. C., and Ferrin, T. E. (2007). Visualizing density maps with UCSF Chimera. *J. Struct. Biol.*, 157:281–287.
- Goulding, C. W., Sawaya, M. R., Parseghian, A., Lim, V., Eisenberg, D., and Missiakas, D. (2002). Thiol-disulfide exchange in an immunoglobulin-like fold: Structure of the N-terminal domain of DsbD. *Biochemistry*, 41(22):6920–6927.
- Grey, M. J., Wang, C., and Palmer, A. G. (2003). Disulfide bond isomerization in basic pancreatic trypsin inhibitor: multisite chemical exchange quantified by CPMG relaxation dispersion and chemical shift modeling. *J. Am. Chem. Soc.*, 125(47):14324–35.
- Guan, L. and Kaback, H. R. (2006). Lessons from lactose permease. *Annu. Rev. Biophys. Biomol. Struct.*, 35(1):67–91.
- Guan, L., Mirza, O., Verner, G., Iwata, S., and Kaback, H. R. (2007). Structural determination of wild-type lactose permease. *Proc. Natl. Acad. Sci. U. S. A.*, 104:15294–15298.
- Haebel, P. W., Goldstone, D., Katzen, F., Beckwith, J., and Metcalf, P. (2002). The disulfide bond isomerase DsbC is activated by an immunoglobulin-fold thiol oxidoreductase: crystal structure of the DsbC-DsbD α complex. *EMBO J.*, 21(18):4774–84.
- Hansen, D. F. (2008). CATIA (Cpmg, Anti-trosy, and Trosy Intelligent Analysis).
- Hansen, D. F., Vallurupalli, P., and Kay, L. E. (2009). Measurement of methyl group motional parameters of invisible, excited protein states by NMR spectroscopy. *J. Am. Chem. Soc.*, 131:12745–12754.
- Hasegawa, H. and Holm, L. (2009). Advances and pitfalls of protein structural alignment.
- Hekkelman, M. L., te Beek, T. A. H., Pettifer, S. R., Thorne, D., Attwood, T. K., and Vriend, G. (2010). WIWS: A protein structure bioinformatics web service collection. *Nucleic Acids Res.*, 38(SUPPL. 2).
- Helmus, J. J. and Jaroniec, C. P. (2013). Nmrglue: An open source Python package for the analysis of multidimensional NMR data. *J. Biomol. NMR*, 55:355–367.
- Hemmis, C. W., Berkmen, M., Eser, M., and Schildbach, J. F. (2011). TrbB from Conjugative Plasmid F Is a Structurally Distinct Disulfide Isomerase That Requires DsbD for Redox State Maintenance. *J. Bacteriol.*, 193(18):4588–97.
- Hemmis, C. W. and Schildbach, J. F. (2013). Thioredoxin-like proteins in F and other plasmid systems. *Plasmid*, 70:168–189.
- Henzler-Wildman, K. and Kern, D. (2007). Dynamic personalities of proteins. *Nature*, 450:964–72.
- Hermans, J., Berendsen, H. J. C., van Gunsteren, W. F., and Postma, J. P. M. (1984). A consistent empirical potential for water–protein interactions. *Biopolymers*, 23:1513–1518.

- Hess, B. (2008). P-LINCS: A Parallel Linear Constraint Solver for Molecular Simulation. *J. Chem. Theory Comput.*, 4(1):116–122.
- Hess, B., Kutzner, C., van der Spoel, D., and Lindahl, E. (2008). GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.*, 4(3):435–447.
- Hirai, T. and Subramaniam, S. (2004). Structure and transport mechanism of the bacterial oxalate transporter OxlT. *Biophys. J.*, 87:3600–3607.
- Ho, B. K. and Gruswitz, F. (2008). HOLLOW: generating accurate representations of channel and interior surfaces in molecular structures. *BMC Struct. Biol.*, 8:49.
- Hodges, C., Bintu, L., Lubkowska, L., Kashlev, M., and Bustamante, C. (2009). Nucleosomal fluctuations govern the transcription dynamics of RNA polymerase II. *Science*, 325:626–628.
- Holyoake, J. and Sansom, M. S. P. (2007). Conformational change in an MFS protein: MD simulations of LacY. *Structure*, 15(7):873–884.
- Huang, Y., Lemieux, M. J., Song, J., Auer, M., and Wang, D.-N. (2003). Structure and Mechanism of the Glycerol-3-Phosphate Transporter from Escherichia coli. *Science*, 301(5633):616–620.
- Humphrey, W., Dalke, A., and Schulten, K. (1996). VMD: Visual molecular dynamics. *J. Mol. Graph.*, 14:33–38.
- Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Comput. Sci. Eng.*, 9:99–104.
- Hunter, S., Jones, P., Mitchell, A., Apweiler, R., Attwood, T. K., Bateman, A., Bernard, T., Binns, D., Bork, P., Burge, S., De Castro, E., Coggill, P., Corbett, M., Das, U., Daugherty, L., Duquenne, L., Finn, R. D., Fraser, M., Gough, J., Haft, D., Hulo, N., Kahn, D., Kelly, E., Letunic, I., Lonsdale, D., Lopez, R., Madera, M., Maslen, J., McAnulla, C., McDowall, J., McMenamin, C., Mi, H., Mutowo-Muellenet, P., Mulder, N., Natale, D., Orengo, C., Pesseat, S., Punta, M., Quinn, A. F., Rivoire, C., Sangrador-Vegas, A., Selengut, J. D., Sigrist, C. J. A., Scheremetjew, M., Tate, J., Thimmajananathan, M., Thomas, P. D., Wu, C. H., Yeats, C., and Yong, S. Y. (2012). InterPro in 2011: New developments in the family and domain prediction database. *Nucleic Acids Res.*, 40.
- Igumenova, T. I., Brath, U., Akke, M., and Palmer III, A. G. (2007). Characterization of Chemical Exchange Using Residual Dipolar Coupling. *J. Am. Chem. Soc.*, 129:13396–13397.
- Im, W., Feig, M., and Brooks, C. L. (2003a). An implicit membrane generalized born theory for the study of structure, stability, and interactions of membrane proteins. *Biophys. J.*, 85:2900–2918.
- Im, W., Lee, M. S., and Brooks, C. L. (2003b). Generalized Born Model with a Simple Smoothing Function. *J. Comput. Chem.*, 24:1691–1702.

- Ishima, R. and Torchia, D. A. (2006). Accuracy of optimized chemical-exchange parameters derived by fitting CPMG R2 dispersion profiles when $R_2^{0a} \neq R_2^{0b}$. *J. Biomol. NMR*, 34:209–219.
- Ito, K. and Inaba, K. (2008). The disulfide bond formation (Dsb) system. *Curr. Opin. Struct. Biol.*, 18(4):450–8.
- Jardetzky, O. (1966). Simple allosteric model for membrane pumps. *Nature*, 211:969–970.
- Jarymowycz, V. A. and Stone, M. J. (2006). Fast time scale dynamics of protein backbones: NMR relaxation methods, applications, and functional consequences. *Chem. Rev.*, 106(5):1624–71.
- Jensen, M. O., Yin, Y., Tajkhorshid, E., and Schulten, K. (2007). Sugar transport across lactose permease probed by steered molecular dynamics. *Biophys. J.*, 93:92–102.
- Jiang, D., Zhao, Y., Wang, X., Fan, J., Heng, J., Liu, X., Feng, W., Kang, X., Huang, B., Liu, J., and Zhang, X. C. (2013). Structure of the YajR transporter suggests a transport mechanism based on the conserved motif A. *Proc. Natl. Acad. Sci. U. S. A.*, 110:14664–9.
- Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983). Comparison of simple potential functions for simulating liquid water.
- Juraszek, J. and Bolhuis, P. G. (2010). (Un)Folding mechanisms of the FBP28 WW domain in explicit solvent revealed by multiple rare event simulation methods. *Biophys. J.*, 98(4):646–56.
- Kaback, H. R., Dunten, R., Frillingos, S., Venkatesan, P., Kwaw, I., Zhang, W., and Ermolova, N. (2007). Site-directed alkylation and the alternating access model for LacY. *Proc. Natl. Acad. Sci. U. S. A.*, 104:491–494.
- Kaback, H. R., Smirnova, I., Kasho, V., Nie, Y., and Zhou, Y. (2011). The Alternating Access Transport Mechanism in LacY. *J. Membr. Biol.*, 239:85–93.
- Kabsch, W. and Sander, C. (1983). Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*, 22:2577–2637.
- Kasho, V. N., Smirnova, I. N., and Kaback, H. R. (2006). Sequence Alignment and Homology Threading Reveals Prokaryotic and Eukaryotic Proteins Similar to Lactose Permease. *J. Mol. Biol.*, 358(4):1060–1070.
- Katzen, F., Deshmukh, M., Daldal, F., and Beckwith, J. (2002). Evolutionary domain fusion expanded the substrate specificity of the transmembrane electron transporter DsbD. *EMBO J.*, 21:3960–3969.
- Kim, Y. C., Tang, C., Clore, G. M., and Hummer, G. (2008). Replica exchange simulations of transient encounter complexes in protein-protein association. *Proc. Natl. Acad. Sci. U. S. A.*, 105(35):12855–12860.
- Klauda, J. B., Venable, R. M., Freites, J. A., O'Connor, J. W., Tobias, D. J., Mondragon-Ramirez, C., Vorobyov, I., MacKerell, A. D., and Pastor, R. W. (2010). Update of the CHARMM all-atom additive force field for lipids: validation on six lipid types. *J. Phys. Chem. B*, 114:7830–7843.

- Kovriggin, E. L., Kempf, J. G., Grey, M. J., and Loria, J. P. (2006). Faithful estimation of dynamics parameters from CPMG relaxation dispersion measurements. *J. Magn. Reson.*, 180(1):93–104.
- Kozakov, D., Li, K., Hall, D. R., Beglov, D., Zheng, J., Vakili, P., Schueler-Furman, O., Paschalidis, I. C., Clore, G. M., and Vajda, S. (2014). Encounter complexes and dimensionality reduction in protein-protein association. *Elife*, 2014(3).
- Krishnamurthy, H., Piscitelli, C. L., and Gouaux, E. (2009). Unlocking the molecular secrets of sodium-coupled transporters. *Nature*, 459(7245):347–355.
- Kumar, H., Kasho, V., Smirnova, I., Finer-Moore, J. S., Kaback, H. R., and Stroud, R. M. (2014). Structure of sugar-bound LacY. *Proc. Natl. Acad. Sci. U. S. A.*, 111:1784–8.
- Kumar, P., Sannigrahi, S., Scoullar, J., Kahler, C. M., and Tzeng, Y.-L. (2011). Characterization of DsbD in *Neisseria meningitidis*. *Mol. Microbiol.*, 79(6):1557–73.
- Laio, A. and Parrinello, M. (2002). Escaping free-energy minima. *Proc. Natl. Acad. Sci. U. S. A.*, 99:12562–12566.
- Lakomek, N. A., Carlomagno, T., Becker, S., Griesinger, C., and Meiler, J. (2006). A thorough dynamic interpretation of residual dipolar couplings in ubiquitin. *J. Biomol. NMR*, 34:101–115.
- le Coutre, J., Kaback, H. R., Patel, C. K., Heginbotham, L., and Miller, C. (1998). Fourier transform infrared spectroscopy reveals a rigid alpha-helical assembly for the tetrameric *Streptomyces lividans* K⁺ channel. *Proc. Natl. Acad. Sci. U. S. A.*, 95:6114–6117.
- Leach, A. R. (2001). *Molecular Modelling: Principles and Applications*.
- Lee, M.-K., Gal, M., Frydman, L., and Varani, G. (2010). Real-time multidimensional NMR follows RNA folding with second resolution. *Proc. Natl. Acad. Sci. U. S. A.*, 107:9192–9197.
- Levitt, M. (2001). The birth of computational structural biology. *Nat Struct Mol Biol*, 8(5):392–393.
- Li, H., Robertson, A. D., and Jensen, J. H. (2005). Very fast empirical prediction and rationalization of protein pK_a values. *Proteins Struct. Funct. Genet.*, 61:704–721.
- Li, Y., Altorelli, N. L., Bahna, F., Honig, B., Shapiro, L., and Palmer, A. G. (2013). Mechanism of E-cadherin dimerization probed by NMR relaxation dispersion. *Proc. Natl. Acad. Sci. U. S. A.*, 110:16462–16467.
- Lindorff-Larsen, K., Piana, S., Dror, R. O., and Shaw, D. E. (2011). How fast-folding proteins fold. *Science*, 334:517–20.
- Lipari, G. and Szabo, A. (1982a). Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity. *J. Am. Chem. Soc.*, 104(17):4546–4559.

- Lipari, G. and Szabo, A. (1982b). Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 2. Analysis of experimental results. *J. Am. Chem. Soc.*, 104:4559–4570.
- Lippens, G., Caron, L., and Smet, C. (2004). A microscopic view of chemical exchange: Monte Carlo simulation of molecular association. *Concepts Magn. Reson. Part A Bridg. Educ. Res.*, 21(1):1–9.
- Liu, P., Agrafiotis, D. K., and Theobald, D. L. (2010). Fast determination of the optimal rotational matrix for macromolecular superpositions. *J. Comput. Chem.*, 31:1561–1563.
- Long, D., Bouvignies, G., and Kay, L. E. (2014). Measuring hydrogen exchange rates in invisible protein excited states. *Proc. Natl. Acad. Sci. U. S. A.*, 111:8820–5.
- Loria, J. P., Rance, M., and Palmer, A. G. (1999). A Relaxation-Compensated Carr-Purcell-Meiboom-Gill Sequence for Characterizing Chemical Exchange by NMR Spectroscopy. *J. Am. Chem. Soc.*, 121(10):2331–2332.
- Luz, Z. and Meiboom, S. (1963). Nuclear Magnetic Resonance Study of the Protolysis of Trimethylammonium Ion in Aqueous Solution—Order of the Reaction with Respect to Solvent. *J. Chem. Phys.*, 39(2):366.
- MacKerell, A. D., Bashford, D., Bellott, M., Dunbrack, R. L., Evanseck, J. D., Field, M. J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F. T. K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D. T., Prodhom, B., Reiher, W. E., Roux, B., Schlenkrich, M., Smith, J. C., Stote, R., Straub, J., Watanabe, M., Wiorkiewicz-Kuczera, J., Yin, D., and Karplus, M. (1998). All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B*, 102:3586–3616.
- Mackerell, A. D., Feig, M., and Brooks, C. L. (2004). Extending the treatment of backbone energetics in protein force fields: limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *J. Comput. Chem.*, 25(11):1400–15.
- Madej, M. G., Soro, S. N., and Kaback, H. R. (2012). PNAS Plus: Apo-intermediate in the transport cycle of lactose permease (LacY).
- Maity, H., Maity, M., Krishna, M. M. G., Mayne, L., and Englander, S. W. (2005). Protein folding: the stepwise assembly of foldon units. *Proc. Natl. Acad. Sci. U. S. A.*, 102:4741–4746.
- Majumdar, D. S., Smirnova, I., Kasho, V., Nir, E., Kong, X., Weiss, S., and Kaback, H. R. (2007). Single-molecule FRET reveals sugar-induced conformational dynamics in LacY. *Proc. Natl. Acad. Sci. U. S. A.*, 104:12640–12645.
- Mandel, A. M., Akke, M., and Palmer, A. G. (1995). Backbone dynamics of Escherichia coli ribonuclease HI: correlations with structure and function in an active enzyme. *J. Mol. Biol.*, 246:144–163.
- Massi, F., Grey, M. J., and Palmer, A. G. (2005). Microsecond timescale backbone conformational dynamics in ubiquitin studied with NMR R1rho relaxation experiments. *Protein Sci.*, 14:735–742.

- Mavridou, D. A., Stelzl, L. S., Ferguson, S. J., and Redfield, C. (2012a). ^1H , ^{13}C and ^{15}N resonance assignments for the oxidized and reduced states of the N-terminal domain of DsbD from *Escherichia coli*. *Biomol. NMR Assign.*, 6(2):163–167.
- Mavridou, D. A. I., Ferguson, S. J., and Stevens, J. M. (2012b). The interplay between the disulfide bond formation pathway and cytochrome c maturation in *Escherichia coli*. *FEBS Lett.*, 586:1702–1707.
- Mavridou, D. A. I., Saridakis, E., Kritsiligkou, P., Goddard, A. D., Stevens, J. M., Ferguson, S. J., and Redfield, C. (2011). Oxidation State-dependent Protein-Protein Interactions in Disulfide Cascades. *J. Biol. Chem.*, 286(28):24943–56.
- Mavridou, D. A. I., Saridakis, E., Kritsiligkou, P., Mozley, E. C., Ferguson, S. J., and Redfield, C. (2014). An extended active-site motif controls the reactivity of the thioredoxin fold. *J. Biol. Chem.*, 289:8681–8696.
- Mavridou, D. A. I., Stevens, J. M., Ferguson, S. J., and Redfield, C. (2007). Active-site Properties of the Oxidized and Reduced C-terminal Domain of DsbD Obtained by NMR Spectroscopy. *J. Mol. Biol.*, 370(4):643–658.
- Mavridou, D. A. I., Stevens, J. M., Goddard, A. D., Willis, A. C., Ferguson, S. J., and Redfield, C. (2009). Control of periplasmic interdomain thiol:disulfide exchange in the transmembrane oxidoreductase DsbD. *J. Biol. Chem.*, 284:3219–3226.
- McHaourab, H. S., Steed, P. R., and Kazmier, K. (2011). Toward the fourth dimension of membrane protein structure: Insight into dynamics from spin-labeling EPR spectroscopy.
- Michaud-Agrawal, N., Denning, E. J., Woolf, T. B., and Beckstein, O. (2011). MDAnalysis: A Toolkit for the Analysis of Molecular Dynamics Simulations. *J. Comput. Chem.*
- Millet, O., Loria, J. P., Kroenke, C. D., Pons, M., and Palmer, A. G. (2000). The Static Magnetic Field Dependence of Chemical Exchange Linebroadening Defines the NMR Chemical Shift Time Scale. *J. Am. Chem. Soc.*, 122(12):2867–2877.
- Mirza, O., Guan, L., Verner, G., Iwata, S., and Kaback, H. R. (2006). Structural evidence for induced fit and a mechanism for sugar/ H^+ symport in LacY. *EMBO J.*, 25:1177–1183.
- Mitchell, P. (1967). Translocations through natural membranes. In Nord, F. F., editor, *Adv. Enzym. Relat. Areas Mol. Biol.*, volume 29, pages 33–87. John Wiley & Sons, Inc.
- Mulder, F. A., Mittermaier, A., Hon, B., Dahlquist, F. W., and Kay, L. E. (2001a). Studying excited states of proteins by NMR spectroscopy. *Nat. Struct. Biol.*, 8(11):932–5.
- Mulder, F. A. A., Skrynnikov, N. R., Hon, B., Dahlquist, F. W., and Kay, L. E. (2001b). Measurement of Slow (us-ms) Time Scale Dynamics in Protein Side Chains by ^{15}N Relaxation Dispersion NMR Spectroscopy: Application to Asn and Gln Residues in a Cavity Mutant of T4 Lysozyme. *J. Am. Chem. Soc.*, 123(5):967–975.
- Neudecker, P., Robustelli, P., Cavalli, A., Walsh, P., Lundstrom, P., Zarrine-Afsar, A., Sharpe, S., Vendruscolo, M., and Kay, L. E. (2012). Structure of an Intermediate State in Protein Folding and Aggregation.

- Newstead, S., Drew, D., Cameron, A. D., Postis, V. L. G., Xia, X., Fowler, P. W., Ingram, J. C., Carpenter, E. P., Sansom, M. S. P., McPherson, M. J., Baldwin, S. A., and Iwata, S. (2010). Crystal structure of a prokaryotic homologue of the mammalian oligopeptide-proton symporters, PepT1 and PepT2. *EMBO J.*, 30(September):417–426.
- Noé, F., Schütte, C., Vanden-Eijnden, E., Reich, L., and Weikl, T. R. (2009). Constructing the equilibrium ensemble of folding pathways from short off-equilibrium simulations. *Proc. Natl. Acad. Sci. U. S. A.*, 106:19011–19016.
- Northrup, S. H. and Erickson, H. P. (1992). Kinetics of protein-protein association explained by Brownian dynamics computer simulation. *Proc. Natl. Acad. Sci. U. S. A.*, 89(8):3338–3342.
- Oliphant, T. E. (2007). Python for scientific computing. *Comput. Sci. Eng.*, 9:10–20.
- Palazzesi, F., Barducci, A., Tollinger, M., and Parrinello, M. (2013). The allosteric communication pathways in KIX domain of CBP. *Proc. Natl. Acad. Sci.*, 110:14237–14242.
- Palmer, A. G. (2004). NMR characterization of the dynamics of biomacromolecules. *Chem. Rev.*, 104(8):3623–3640.
- Palmer, A. G., Kroenke, C. D., and Loria, J. P. (2001). Nuclear magnetic resonance methods for quantifying microsecond-to-millisecond motions in biological macromolecules. *Methods Enzym.*, 339(1997):204–238.
- Parrinello, M. and Rahman, A. (1981). Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.*, 52(12):7182–7190.
- Patzlaff, J. S., Moeller, J. A., Barry, B. A., and Brooker, R. J. (1998). Fourier Transform Infrared Analysis of Purified Lactose Permease: A Monodisperse Lactose Permease Preparation Is Stably Folded, alpha - Helical, and Highly Accessible to Deuterium Exchange. *Biochemistry*, 37(44):15363–15375.
- Pedersen, B. r. P., Kumar, H., Waight, A. B., Risenmay, A. J., Roe-Zurz, Z., Chau, B. H., Schlessinger, A., Bonomi, M., Harries, W., Sali, A., Johri, A. K., and Stroud, R. M. (2013). Crystal structure of a eukaryotic phosphate transporter. *Nature*, 496:533–6.
- Pendse, P. Y., Brooks, B. R., and Klauda, J. B. (2010). Probing the periplasmic-open state of lactose permease in response to sugar binding and proton translocation. *J. Mol. Biol.*, 404:506–521.
- Perilla, J. R., Beckstein, O., Denning, E. J., and Woolf, T. B. (2011). Computing ensembles of transitions from stable states: Dynamic importance sampling. *J. Comput. Chem.*, 32(2):196–209 ST – Computing ensembles of transitions f.
- Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004). UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.*, 25:1605–1612.
- Piana, S., Klepeis, J. L., and Shaw, D. E. (2014). Assessing the accuracy of physical models used in protein-folding simulations: Quantitative evidence from long molecular dynamics simulations. *Curr. Opin. Struct. Biol.*, 24:98–105.

- Pliego, J. R. and Riveros, J. M. (2000). New values for the absolute solvation free energy of univalent ions in aqueous solution.
- Polyhach, Y., Bordignon, E., and Jeschke, G. (2011). Rotamer libraries of spin labelled cysteines for protein studies. *Phys. Chem. Chem. Phys.*, 13:2356–2366.
- Premkumar, L., Kurth, F., Neyer, S., Schembri, M. A., and Martin, J. L. (2014). The multidrug resistance IncA/C transferable plasmid encodes a novel domain-swapped dimeric protein-disulfide isomerase. *J. Biol. Chem.*, 289:2563–2576.
- Prlić, A., Bliven, S., Rose, P. W., Bluhm, W. F., Bizon, C., Godzik, A., and Bourne, P. E. (2010). Pre-calculated protein structure alignments at the RCSB PDB website. *Bioinformatics*, 26:2983–2985.
- Qin, S., Pang, X., and Zhou, H. X. (2011). Automated prediction of protein association rate constants. *Structure*, 19:1744–1751.
- Quinternet, M., Tsan, P., Selme-Roussel, L., Jacob, C., Boschi-Muller, S., Branlant, G., and Cung, M. T. (2009). Formation of the Complex between DsbD and PilB N-Terminal Domains from *Neisseria meningitidis* Necessitates an Adaptability of nDsbD. *Structure*, 17:1024–1033.
- Quistgaard, E. M., Löw, C., Moberg, P., Trésaugues, L., and Nordlund, P. (2013). Structural basis for substrate transport in the GLUT-homology family of monosaccharide transporters. *Nat. Struct. Mol. Biol.*, 20:766–8.
- Radestock, S. and Forrest, L. R. (2011). The alternating-access mechanism of MFS transporters arises from inverted-topology repeats. *J. Mol. Biol.*
- Radić, Z., Duran, R., Vellom, D. C., Li, Y., Cervenansky, C., and Taylor, P. (1994). Site of fasciculin interaction with acetylcholinesterase. *J. Biol. Chem.*, 269:11233–11239.
- Radić, Z., Kirchhoff, P. D., Quinn, D. M., McCammon, J. A., and Taylor, P. (1997). Electrostatic influence on the kinetics of ligand binding to acetylcholinesterase: Distinctions between active center ligands and fasciculin. *J. Biol. Chem.*, 272:23265–23277.
- Rinnenthal, J., Klinkert, B., Narberhaus, F., and Schwalbe, H. (2010). Direct observation of the temperature-induced melting process of the *Salmonella* fourU RNA thermometer at base-pair resolution. *Nucleic Acids Res.*, 38(11):3834–3847.
- Rinnenthal, J., Klinkert, B., Narberhaus, F., and Schwalbe, H. (2011). Modulation of the stability of the *Salmonella* fourU-type RNA thermometer. *Nucleic Acids Res.*, 39(18):8258–8270.
- Robustelli, P., Stafford, K. A., and Palmer, A. G. (2012). Interpreting Protein Structural Dynamics From NMR Chemical Shifts. *J. Am. Chem. Soc.*, 134(14):6365–6374.
- Rolfe, D. F. and Brown, G. C. (1997). Cellular energy utilization and molecular origin of standard metabolic rate in mammals. *Physiol Rev*, 77(3):731–758.
- Rozhkova, A. and Glockshuber, R. (2007). Kinetics of the Intramolecular Disulfide Exchange Between the Periplasmic Domains of DsbD. *J. Mol. Biol.*, 367:1162–1170.

- Rozhkova, A. and Glockshuber, R. (2008). Thermodynamic Aspects of DsbD-Mediated Electron Transport. *J. Mol. Biol.*, 380:783–788.
- Rozhkova, A., Stirnimann, C. U., Frei, P., Grauschopf, U., Brunisholz, R., Grütter, M. G., Capitani, G., and Glockshuber, R. (2004). Structural basis and kinetics of inter- and intramolecular disulfide exchange in the redox catalyst DsbD. *EMBO J.*, 23(8):1709–19.
- Ruiz, N., Chng, S.-S., Hiniker, A., Kahne, D., and Silhavy, T. J. (2010). Nonconsecutive disulfide bond formation in an essential integral outer membrane protein. *Proc. Natl. Acad. Sci. U. S. A.*, 107:12245–12250.
- Salonen, L. M., Ellermann, M., and Diederich, F. (2011). Aromatic rings in chemical and biological recognition: Energetics and structures. *Angew. Chem., Int. Ed.*, 50(21):4808–4842.
- Sayed, W. M. H. and Baenziger, J. E. (2009). Structural characterization of the osmosensor ProP. *Biochim. Biophys. Acta, Biomembr.*, 1788(5):1108–1115.
- Schreiber, G., Haran, G., and Zhou, H.-X. (2009). Fundamental aspects of protein-protein association kinetics. *Chem. Rev.*, 109(3):839–60.
- Scott, K. A., Bond, P. J., Ivetac, A., Chetwynd, A. P., Khalid, S., and Sansom, M. S. P. (2008). Coarse-Grained MD Simulations of Membrane Protein-Bilayer Self-Assembly. *Structure*, 16:621–630.
- Sevier, C. S. and Kaiser, C. A. (2002). Formation and transfer of disulphide bonds in living cells. *Nat. Rev. Mol. Cell Biol.*, 3(11):836–847.
- Shen, Y. and Bax, A. (2013). Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. *J. Biomol. NMR*, 56(3):227–241.
- Shimamura, T., Weyand, S., Beckstein, O., Rutherford, N. G., Hadden, J. M., Sharples, D., Sansom, M. S., Iwata, S., Henderson, P. J., and Cameron, A. D. (2010). Molecular basis of alternating access membrane transport by the sodium-hydantoin transporter Mhp1. *Science*, 328(5977):470–473.
- Skrynnikov, N. R., Dahlquist, F. W., and Kay, L. E. (2002). Reconstructing NMR spectra of "invisible" excited protein states using HSQC and HMQC experiments. *J. Am. Chem. Soc.*, 124(41):12352–12360.
- Smart, O. S., Goodfellow, M. J., and Wallace, B. A. (1993). The pore dimensions of gramicidin A. *Biophys. J.*, 65:2455–2460.
- Smirnova, I., Kasho, V., Choe, J.-Y., Altenbach, C., Hubbell, W. L., and Kaback, H. R. (2007). Sugar binding induces an outward facing conformation of LacY. *Proc. Natl. Acad. Sci. U. S. A.*, 104:16504–16509.
- Smirnova, I., Kasho, V., Sugihara, J., and Kaback, H. R. (2009). Probing of the rates of alternating access in LacY with Trp fluorescence. *Proc. Natl. Acad. Sci. U. S. A.*, 106:21561–21566.

- Smith, L. J., Mark, A. E., Dobson, C. M., and van Gunsteren, W. F. (1995). Comparison of MD simulations and NMR experiments for hen lysozyme. Analysis of local fluctuations, cooperative motions, and global changes. *Biochemistry*, 34(34):10918–10931.
- Stansfeld, P. J. and Sansom, M. S. P. (2011). Molecular simulation approaches to membrane proteins. *Structure*, 19(11):1562–1572.
- Stelzl, L. S. (2010). *Studying the structure and dynamics of the N-terminal domain of the oxidoreductase DsbD by NMR*. PhD thesis, University of Oxford.
- Stirnemann, C. U., Grütter, M. G., Glockshuber, R., and Capitani, G. (2006). nDsbD: a redox interaction hub in the Escherichia coli periplasm. *Cell. Mol. Life Sci.*, 63(14):1642–8.
- Stirnemann, C. U., Rozhkova, A., Grauschopf, U., Grütter, M. G., Glockshuber, R., and Capitani, G. (2005). Structural basis and kinetics of DsbD-dependent cytochrome c maturation. *Structure*, 13:985–993.
- Sugase, K., Dyson, H. J., and Wright, P. E. (2007a). Mechanism of coupled folding and binding of an intrinsically disordered protein. *Nature*, 447:1021–1025.
- Sugase, K., Lansing, J. C., Dyson, H. J., and Wright, P. E. (2007b). Tailoring relaxation dispersion experiments for fast-associating protein complexes. *J. Am. Chem. Soc.*, 129:13406–13407.
- Sun, L., Zeng, X., Yan, C., Sun, X., Gong, X., Rao, Y., and Yan, N. (2012). Crystal structure of a bacterial homologue of glucose transporters GLUT1–4.
- Tang, C., Iwahara, J., and Clore, G. M. (2006). Visualization of transient encounter complexes in protein-protein association. *Nature*, 444(7117):383–386.
- Theobald, D. L. (2005). Rapid calculation of RMSDs using a quaternion-based characteristic polynomial. *Acta Crystallogr. Sect. A Found. Crystallogr.*, 61:478–480.
- Ubbink, M. (2009). The courtship of proteins: understanding the encounter complex. *FEBS Lett.*, 583(7):1060–6.
- Vallurupalli, P., Bouvignies, G., and Kay, L. E. (2011). Increasing the exchange time-scale that can be probed by CPMG relaxation dispersion NMR. *J. Phys. Chem. B*, 115:14891–14900.
- Vallurupalli, P., Bouvignies, G., and Kay, L. E. (2012). Studying "invisible" excited protein states in slow exchange with a major state conformation. *J. Am. Chem. Soc.*, 134:8148–8161.
- Vallurupalli, P., Hansen, D. F., and Kay, L. E. (2008). Structures of invisible, excited protein states by relaxation dispersion NMR spectroscopy. *Proc. Natl. Acad. Sci. U. S. A.*, 105:11766–11771.
- Vallurupalli, P., Hansen, D. F., Stollar, E., Meirovitch, E., and Kay, L. E. (2007). Measurement of bond vector orientations in invisible excited states of proteins. *Proc. Natl. Acad. Sci. U. S. A.*, 104:18473–18477.

- van Der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A. E., and Berendsen, H. J. C. (2005). GROMACS: fast, flexible, and free. *J. Comput. Chem.*, 26(16):1701–18.
- van der Spoel, D., Lindahl, E., Hess, B., Kutzner, C., van Buuren, A. R., Apol, E., Meulenhoff, P. J., Tieleman, D. P., Sijbers, A. L. T. M., Feenstra, K. A., van Drunen, R., and Berendsen, H. J. C. (2005). Gromacs User Manual version 4.0.
- van Gunsteren, W. F., Billeter, S. R., Eising, A. A., Hünenberger, P. H., Krüger, P., Mark, A. E., Scott, W. R. P., and Tironi, I. G. (1996). *Biomolecular Simulation: The GRO-MOS96 Manual and User Guide*. Vdf Hochschulverlag AG an der ETH Zürich, Zürich, Switzerland.
- Vardy, E., Steiner-Mordoch, S., and Schuldiner, S. (2005). Characterization of bacterial drug antiporters homologous to mammalian neurotransmitter transporters. *J. Bacteriol.*, 187:7518–7525.
- Viitanen, P., Garcia, M. L., and Kaback, H. R. (1984). Purified reconstituted lac carrier protein from Escherichia coli is fully functional. *Proc. Natl. Acad. Sci. U. S. A.*, 81:1629–1633.
- Vranken, W. F., Boucher, W., Stevens, T. J., Fogh, R. H., Pajon, A., Llinas, M., Ulrich, E. L., Markley, J. L., Ionides, J., and Laue, E. D. (2005). The CCPN data model for NMR spectroscopy: Development of a software pipeline. *Proteins Struct., Funct., Genet.*, 59(4):687–696.
- Vriend, G. (1990). WHAT IF: a molecular modeling and drug design program. *J. Mol. Graph.*, 8(1):52–56, 29.
- Waudby, C. A. and Christodoulou, J. (2014). Direct 2D Lineshape Analysis for Improved Accuracy in the Analysis of NMR Titrations , with Particular Application to Macromolecular Interactions The quantitative analysis of lineshapes across NMR titrations can be an extremely powerful probe of complex ..
- Whittier, S. K., Hengge, A. C., and Loria, J. P. (2013). Conformational motions regulate phosphoryl transfer in related protein tyrosine phosphatases. *Science*, 341:899–903.
- Woolf, T. B. . (1998). Path corrected functionals of stochastic trajectories: towards relative free energy and reaction coordinate calculations. *Chem. Phys. Lett.*, page 433.
- Woutersen, S. and Bakker, H. J. (2006). Ultrafast vibrational and structural dynamics of the proton in liquid water. *Phys Rev Lett*, 96(13):138305.
- Wu, X. and Brooks, B. R. (2003). Self-guided Langevin dynamics simulation method. *Chem. Phys. Lett.*, 381:512–518.
- Xue, Y., Ward, J. M., Yuwen, T., Podkorytov, I. S., and Skrynnikov, N. R. (2012). Microsecond time-scale conformational exchange in proteins: Using long molecular dynamics trajectory to simulate NMR relaxation dispersion data. *J. Am. Chem. Soc.*, 134:2555–2562.
- Yan, H., Huang, W., Yan, C., Gong, X., Jiang, S., Zhao, Y., Wang, J., and Shi, Y. (2013). Structure and Mechanism of a Nitrate Transporter. *Cell Rep.*, 3:716–723.

- Yin, Y., Jensen, M. O., Tajkhorshid, E., and Schulten, K. (2006). Sugar binding and protein conformational changes in lactose permease. *Biophys. J.*, 91:3972–3985.
- Zeiske, T., Stafford, K. a., Friesner, R. a., and Palmer, A. G. (2013). Starting-structure dependence of nanosecond timescale intersubstate transitions and reproducibility of MD-derived order parameters. *Proteins*, 81(3):499–509.
- Zheng, H., Wisedchaisri, G., and Gonen, T. (2013). Crystal structure of a nitrate/nitrite exchanger. *Nature*, 497:647–51.
- Zhou, Y., Guan, L., Freites, J. A., and Kaback, H. R. (2008). Opening and closing of the periplasmic gate in lactose permease. *Proc. Natl. Acad. Sci. U. S. A.*, 105:3774–3778.
- Zhou, Y., Nie, Y., and Kaback, H. R. (2009). Residues gating the periplasmic pathway of LacY. *J. Mol. Biol.*, 394(2):219–225.