

https://dspace.org/

The role of a conserved interdomain interaction in Escherichia coli glutaredoxin-2

Parbhoo, Nishal

2010-08-26
http://hdl.handle.net/10539/8557

Downloaded from a DSpace repository. https://dspace.org/

The role of a conserved interdomain interaction in *Escherichia coli* Glutaredoxin-2

Nishal Parbhoo

A dissertation submitted to the Faculty of Science, University of the Witwatersrand, Johannesburg, in fulfillment of the requirements for the degree of Master of Science.

Johannesburg, 2010

Declaration

I declare that this dissertation is my own, unaided work. It is being submitted for the degree of Master of Science in the University of the Witwatersrand, Johannesburg. It has not been submitted for any other degree or examination at any other University.

_____

Nishal Parbhoo

_____ day of_____, 2010

ABSTRACT

Domain interfaces play an important role in protein stability and folding. A major structural feature at the interdomain interface of the GST class of proteins is the conserved hydrophobic 'lock-and-key' motif. In a monomeric homologue of the GSTs, Grx2, the hydrophobic interdomain 'lock-and-key' motif is formed by insertion of the side-chain of methionine 17 (Met17) from domain 1 into a hydrophobic pocket in domain 2. This study evaluates the contribution of the Met17 residue to the stability of Glutaredoxin-2 (Grx2). Protein engineering techniques were employed to generate a Met17 to Alanine (M17A) mutant protein and comparative studies with wild-type and M17A Grx2 were performed. The spectral properties of M17A Grx2 monitored using far and near-ultraviolet circular dichroism and tryptophan fluorescence indicated no significant changes in secondary or tertiary structure in the native state. Conformational stability studies were performed to determine the contribution of the 'lock-and-key' motif to protein stability. Equilibrium unfolding studies, displayed significant impact on the conformational stability of the protein with a $\Delta\Delta G(H_2O)$ of 4 kcal.mol^{-1} as a result of the replacement of the Met17 residue with alanine. The co-operativity of unfolding is slightly decreased, with the *m*-value being reduced by 0.3 kcal.mol^{-1}.M^{-1} suggesting an intermediate formation. This intermediate becomes more prominent during equilibrium unfolding in the presence of ANS which showed an increase in intensity in the unfolding transition for M17A Grx2 but was absent for wild-type Grx2. The kinetics of unfolding of both Grx2 proteins are complex, both displaying two observable phases (fast and slow) which occur in parallel as confirmed by performing initial conditions test. The slow phase involves structural rearrangements that expose small amounts of surface area while the fast phase represents gross structural unfolding exposing large amounts of surface area. The rate of the fast unfolding phase is increased for M17A Grx2, as the time constant decreased from 2.4s (wild-type) to 830ms, however there is negligible change in the rate of the slow phase. The increase in the unfolding rate of the fast phase is in agreement with the equilibrium studies which highlights the destabilisation as a result of the mutation.

"*There are known knowns. There are known unknowns. There are also unknown unknowns*"

Donald H. Rumsfield

ACKNOWLEDGMENTS

Professor H.W. Dirr for his support and guidance throughout my studies and for the opportunity for working in his laboratory.

My co-supervisor and mentor, Dr. Samantha Gildenhuys, without whom none of this work would have been possible. She has taught me how to think critically, take criticism and to be a good scientist.

All of my current and past colleagues of the Protein-Structure Function Research Unit.

Professor H.W. Dirr and the NRF for financial assistance.

Finally, all of my family and friends who have put up with me during this entire endeavour.

TABLE OF CONTENTS

Declaration ii
ABSTRACT iii
ACKNOWLEDGMENTS v
TABLE OF CONTENTS vi
LIST OF FIGURES ix
LIST OF TABLES x
ABBREVIATIONS xi
CHAPTER 1. INTRODUCTION 1
1.1.1 Forces responsible for protein stability 1
1.1.1.1 van der Waals forces 2
1.1.1.2 Hydrogen bonds 3
1.1.1.3 Hydrophobic interactions 4
1.1.1.4 Electrostatic interactions 4
1.1.2 Protein folding models and pathways 5
1.1.3 Protein folding intermediates and the molten globule state 8
1.2 Domains 10
1.3 Domain-domain interfaces: role in protein stability and folding and function 12
1.3.1 Features of protein-protein interfaces 12
1.4 Stability and unfolding of the GST family of proteins 13
1.4.1. Role of the domain interface in GSTs 17
1.4.2 Glutaredoxin-2: A GST monomeric homologue 22
1.4.2.1 Grx2 Structure 22
1.4.2.2 Domain interface of Grx2 24
1.5 Aims and objectives 26
CHAPTER 2. EXPERIMENTAL PROCEDURE 27
2.1 Materials 27
2.2 Experimental 27
2.2.1 Construction of Grx2 mutants 27
2.2.2 Transformation, over-expression and purification of mutant and wild-type Grx2 29
2.2.3 SDS-PAGE 32
2.2.4 SE - HPLC 32

2.2.5 Absorbance spectroscopy 33
2.2.6 Circular dichroism spectroscopy 34
2.2.7 Fluorescence spectroscopy 35
2.2.7.1 Intrinsic fluorescence-tryptophan fluorescence 35
2.2.7.2 Extrinsic fluorescence - ANS binding 36
2.2.8 DTNB assay 36
2.2.9 Thermal denaturation studies 37
2.2.10 Urea - induced equilibrium unfolding studies of wild-type and M17A Grx2 37
2.2.10.1 Reversibility of folding 37
2.2.10.2 Urea - induced equilibrium unfolding/refolding 38
2.2.11 Data fitting 38
2.2.12 Unfolding kinetics 41
2.2.12.1 Kinetic studies 41
2.2.12.2 Single-jump unfolding studies 42
CHAPTER 3. RESULTS 43
3.1 Sequence identity 43
3.2 Over-expression and purification 43
3.3 Size and purity determination 46
3.4 Quantification of free thiol groups 46
3.5 Structural Characterisation 46
3.5.1 Secondary structure characterisation 46
3.5.2 Tertiary structure characterisation 50
3.5.2.1 Intrinsic fluorescence 50
3.5.2.2 Near-UV circular dichroism 52
3.6. Conformational stability 54
3.6.1 Thermal – induced unfolding 54
3.6.2 Reversibility of urea-induced unfolding 56
3.6.3 Urea-induced equilibrium unfolding 59
3.6.4 Urea-induced equilibrium unfolding in the presence of ANS 62
3.7 Unfolding kinetics 65
3.7.1 Single-jump unfolding kinetics 65
3.7.2 Initial conditions test 69
CHAPTER 4. DISCUSSION 71
4.1 Role of Met17 in the stability of Grx-2 71

4.2 Unfolding kinetics of Grx2 72
4.3 Domain architecture (Grx2 vs GSTs) 75
4.4 Molten globule intermediate or hydrophobic region exposed? 77
4.5 Domain co-operativity in Grx2 78
CHAPTER 5. REFERENCES 79

LIST OF FIGURES

Figure 1-1. Protein folding energy landscapes. 7
Figure 1-2. Schematic model comparison of the molten globule state. 9
Figure 1-3. 'Lock-and-key' motifs found at the dimer and domain interface of GSTs. 16
Figure 1-4. Structure based sequence alignment of domain 1 of GSTs 18
Figure 1-5. The 'lock-and-key' motif found at the domain interface of hGSTA1-1 and Clic1. 20
Figure 1-6. Bar-graph depicting interdomain interface area among the GSTs. 21
Figure 1-7. Ribbon representation of Glutaredoxin-2 23
Figure 1-8. The 'lock-and-key' motif at the domain interface of Grx2 25
Figure 2-1. Vector map of pET24a+ 31
Figure 2-2. Urea denaturation curve 40
Figure 3-1. Wild-type (A) and M17A Grx2 (B) plasmid sequencing results 44
Figure 3-2. SDS-PAGE analysis of M17A Grx2 overexpression. 44
Figure 3-3 Elution profile of Grx2 proteins from purification 45
Figure 3-4. SDS-PAGE analysis of wild-type and M17A Grx2 48
Figure 3-5. SE-HPLC of the Grx2 proteins 48
Figure 3-6. Far-UV circular dichroism spectra of Grx2 49
Figure 3-7. Fluorescence spectra of Grx2 51
Figure 3-8. Near UV CD spectra of Grx2 53
Figure 3-9. Thermal unfolding of Grx2. 55
Figure 3-10. Reversibility of unfolding of wild-type Grx2 monitored by CD and fluorescence 57
Figure 3-11. Reversibility of unfolding of M17A Grx2 monitored by CD and fluorescence 58
Figure 3-12. Urea-induced equilibrium unfolding of the Grx2 proteins 60
Figure 3-13. Grx2 urea-induced equilibrium unfolding in the presence of ANS 63
Figure 3-14. Grx2 urea-induced equilibrium unfolding expressed as maximum emission wavelength 64
Figure 3-15. Unfolding kinetic traces for the Grx2 proteins 66
Figure 3-16. Effect of urea on the unfolding phases of the Grx2 proteins 67
Figure 3-17. Unfolding initial conditions test 70
Figure 4-1. Interdomain 'lock–and-key' motif in Grx2 76

LIST OF TABLES

Table 1-1. List of GST proteins used in alignment 18
Table 3-1. Conformational stability parameters 61
Table 3-3. Unfolding kinetic parameters 68

ABBREVIATIONS

A_{280}	absorbance at 280 nm
Å	Ångström
ANS	8-anilino-1-naphthalene sulphonate
CD	circular dichroism
C_m	the denaturant concentration at the midpoint of the unfolding curve
Clic	chloride intracellular channel
Da	Dalton
DEAE	diethylaminoethyl
ΔASA	change in (solvent) accessible surface area
Δ*G*	the change in Gibbs free energy
Δ*G*(H_2O)	change in Gibbs free energy of unfolding in the absence of denaturant
Δ*H*	the change in enthalpy
Δ*S*	the change in entropy
dH_2O	distilled water
DLS	dynamic light scattering
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
ds	double-stranded
DTNB	5,5'-dithiobis(2-nitrobenzoic acid)
DTT	dithiothreitol
EDTA	ethylenediaminetetra-acetic acid
ε	molar extinction coefficient
Far-UV CD	far-ultraviolet circular dichroism
G	Gibbs free energy
Grx	Glutaredoxin
GSH	reduced glutathione
GST	glutathione *S*-transferase
hGST A1-1	human Alpha class glutathione *S*-transferase
HT	voltage the voltage applied to the circular dichroism photomultiplier tube

IPTG	isopropyl-β-D-thiogalactopyranoside
K_{eq}	equilibrium constant
LB	Luria-Bertani
λ_{max}	fluorescence emission wavelength maximum
mdeg	millidegrees
mRNA	messenger ribonucleic acid
m-value	the dependence of the free energy of unfolding on denaturant concentration
m_u	change in accessible surface area during unfolding
M17A	Met17 replaced with Alanine
MWCO	molecular weight cut-off
N	native
NMR	nuclear magnetic resonance spectroscopy
OD_{600}	optical density at 600 nm
ORF	open reading frame
PDB	Protein Data Bank
pI	isoelectric point
PQS	Protein Quaternary Structure server
RNA	ribonucleic acid
rpm	revolutions per minute
SASA	solvent accessible surface area
SCOPPI	structural classification of protein-protein interfaces database
SDS - PAGE	sodium dodecyl sulphate polyacrylamide gel electrophoresis
SE-HPLC	size exclusion high performance liquid chromatography
*Sj*GST	*Schistosoma japonicum* glutathione *S*-transferase
T	temperature
TEMED	*N*, *N*, *N*', *N*'-tetramethylethylenediamine
τ	time constant, which is the inverse of the apparent rate constant
[θ]	mean residual ellipticity
$[\theta]_{222}$	mean residual ellipticity at 222 nm

U	unfolded
UV	ultraviolet
YT	yeast tryptone

The IUPAC-IUBMB one and three letter codes for the amino acids were used throughout.

CHAPTER 1. INTRODUCTION

1.1 Protein folding and stability

Inside the cell, proteins are synthesised as linear chains of amino acids, which fold into unique 3-D structures, known as the native (folded) state. The wide range of biochemical functions are specified by the proteins' detailed structures and has led to one of the great unsolved problems of science, which is the prediction of these 3-D structures from the amino acid sequence (the folding problem) (reviewed in Dill *et al*., 2008). This has been a concern since the initial experiments of protein folding performed by Anfinsen, who postulated that the three-dimensional structure of a protein is determined by its primary sequence (Anfinsen, 1973). Despite the large degrees of freedom, surprisingly, proteins fold into their native states in a very short time, which is known as Levinthal's Paradox (Levinthal, 1968). Structural changes and chemical interactions occur throughout the entire folding process, and strongly cooperative mechanisms are necessary to bring the protein in its native conformation within a very short time period (Nölting *et al*., 1995). Protein folding is therefore governed by two inter-related parameters, kinetic control, where the polypeptide chain has a time limit within which it must obtain its native conformation, and thermodynamic control where the native conformation is arguably the most stable conformation that the amino acid chain can obtain.

The thermodynamic stability of a protein is measured by the free-energy difference between the folded state and the unfolded state ($\Delta G = G_{\text{unfold}} - G_{\text{fold}}$). It determines the fraction of folded proteins, thereby having a profound effect on protein function. The energetic contributions from the favourable folding forces such as hydrophobic packing, hydrogen-bonding and electrostatic interactions are nearly offset by the entropic penalisation of folding (Dill and Chan 1997).

1.1.1 Forces responsible for protein stability

Protein stability refers to the difference between interactions that favour the native state and interactions that favour the denatured state. This difference is represented by the change in Gibbs free energy ($\Delta G°$) upon total unfolding. Under constant pressure, $\Delta G°$ is made up of two contributions:

$$\Delta G° = \Delta H° - T\Delta S° \qquad (1)$$

where $\Delta H°$ is the enthalpic (bond formation) and $\Delta S°$ is the entropic (freedom of a system to explore conformational space) contribution to $\Delta G°$. Contributions to the enthalpic term are primarily from the loss of intramolecular and protein-solvent hydrogen bonds and van der Waals contacts. The formation of van der Waals contacts, salt-bridges and hydrogen bonds, as well as the change in solvent-solvent interactions near protein surfaces contribute toward the binding enthalpy of a reaction (Dill, 1990). The breaking of bonds will be endothermic (positive $\Delta H°$) whereas the negative $\Delta H°$ refers to the formation of hydrogen bonds and van der Waals interactions (Sigurskjold *et al*., 1991). The entropy component is composed of various contributions (Murphy, 1999) which include: the entropy contribution from restructuring of solvent due to burial of hydrophobic groups (ΔS^{o}_{solv}); the changes in conformational degrees of freedom of the protein backbone and side chain groups (ΔS^{o}_{conf}), and the result from changes in translational, rotational and vibrational degrees of freedom (ΔS^{o}_{mix}).

Since the native state is more compact than the denatured state, conformational energy is lost opposing folding and destabilising the native state. On the other hand, inter- and intra-molecular interactions formed upon native state acquisition favour folding, resulting in native state stabilisation (Dill, 1990). In addition, certain enthalpic interactions can stabilise both the native and unfolded conformations so that their contribution to stability seems negligible (Dill, 1990). Globular proteins are only marginally stable, estimates of their conformational stabilities of small proteins range from 5 kcal. mol^{-1} to 15 $kcal.mol^{-1}$ (Pace, 1990).

1.1.1.1 van der Waals forces

van der Waals interactions, also known as London dispersion forces, result from transient dipoles that nonbonded atoms induce in each other. The strength of the van der Waals interactions is directly proportional to the ability of the atom to polarise and inversely proportional to the sixth power of the distance between them (Pace, 2001). Hence, van der Waals interactions are short-ranged and are only functional at very short distances. It is this strong distance dependence that will determine whether they will stabilise or destabilise the native state, due to the packing of atoms in the protein core, relative to their interaction with solvent (Dill, 1990). Steric complementarity in the interior, domain and/or subunit interface of proteins is needed to maximise the stabilising effects of these forces. Ratnaparkhi and Varadarajan, (2000), proposed that the loss of packing interactions rather than the

hydrophobic effect dominates protein stability. van der Waals interactions are ubiquitous, and although they are weak in proteins, they accumulate to a significant amount. The strength of the individual interaction depends on the types of interacting atoms, and varies with the chemical environment of the atoms involved, e.g. for carboxyl carbon atoms the interactions are usually stronger, and van der Waals distances are usually shorter than for tetrahedral carbon atoms (Pace, 2001).

1.1.1.2 Hydrogen bonds

A particularly important bond in biological systems is the hydrogen bond. Hydrogen bonds are noncovalent interactions that arise from the partial sharing of a hydrogen atom between a hydrogen bond donor group, such as a hydroxyl (-OH) group or an amino (-NH) group, and a hydrogen bond acceptor atom, such as oxygen or nitrogen. Many potential hydrogen bond donor and acceptor groups are present in proteins, namely the peptide backbone groups and the polar amino acid side chains (Stickle *et al*., 1992). Hydrogen bond strengths range from 2 - 14 kcal.mol^{-1}, depending on the geometry of the interactions (Hagler *et al*., 1979; Dauber and Hagler, 1980; Dill, 1990; Privalov and Makhatadze, 1993). A large majority of the hydrogen bonding interactions in globular proteins (68 %) are between the amide hydrogen and the carbonyl oxygen in the peptide group (Stickle *et al.*, 1992), and only 11 % of carbonyl oxygens and 12 % of amide nitrogens are not hydrogen bonded (Baker and Hubbard, 1984). Due to the polar nature of the peptide backbone, much of it is buried hence the requirement of hydrogen bonds.

In terms of domain interfaces, most hydrogen bond contributions are due to polar residues, since these contacts are mostly non-local, while backbone hydrogen bonds are mainly local. The fact that the peptide backbone and polar residues can form hydrogen bonds once again highlights the importance of correct packing in the interior and at domain and subunit interfaces of proteins. In the folded conformation, specific intra-molecular hydrogen bonds must form to replace the fluctuating intra- and inter-molecular bonds that form in the unfolded protein so that the native state is enthalpically favoured. Although hydrogen bonds can provide a significant contribution toward protein stability, they are not the dominant folding force (Dill, 1990). If they were, solvents that form strong hydrogen bonds with the protein should unfold it, while solvents that do not form or form weak hydrogen bonds with the protein should not affect or should not stabilise the native state (Dill, 1990). No such correlation has been observed.

1.1.1.3 Hydrophobic interactions

The absence of hydrogen bonding between water and non-polar groups rather than the presence of favourable interactions between the non-polar groups themselves constitutes an important source of the protein stability in aqueous solution, the hydrophobic interaction (Dill, 1990). The transfer of the sidechains of hydrophobic amino acid residues from an apolar environment into water is energetically costly, and thus, the burial of hydrophobic sidechains in the folding reaction is energetically favourable (Dill, 1990). The process (at room temperature) is entropically driven since the addition of non-polar molecules to water disrupts the hydrogen bonded structure of water (Dill, 1990). Thus water molecules order themselves around the non-polar molecules to maximise their contacts with each other and minimise their contacts with the non-polar substance (Geiger *et al*., 1979; Stillinger, 1980). The close packing of these hydrophobic residues, aided by direct interatomic van der Waals forces, minimises the surface area exposed to the solvent (Pace *et al*., 1996). The hydrophobic free energy contribution to protein stability is estimated to be about 60 $kcal.mol^{-1}$ (Dill *et al*., 1989).

1.1.1.4 Electrostatic interactions

Electrostatic contacts, in proteins, occur between charged residues with the strength of interactions related by Coulomb's law:

$$F = k \times q_1 \times q_2 / D \times r^2 \qquad (2)$$

where F is the force between the two electrical charges, q1 and q2 separated by distance r. k is the proportionality constant and D is the dielectric constant of the medium. The higher the dielectric constant of the medium, the weaker the force is between the two charges, hence, a more non-polar environment such as that in the protein interior, at domain and subunit interfaces, will strengthen electrostatic interactions. Charge contacts, in proteins, can be classified into two groups. 'Classical' electrostatic effects occur due to non-specific repulsions between surface residues with the same charge destabilising proteins (Dill, 1990). The extent of destabilisation is affected by ionic strength and pH. An increase in ionic strength in the protein's environment results in better shielding of opposite charges, which decreases repulsions and stabilises the protein. A pH increase or decrease will result in an increase of the protein's net charge, leading to more charge repulsions along its surface and destabilisation. Therefore, the majority of proteins are most stable at a pH that is close to their pI, where their net charge will be zero.

The second group of electrostatic interactions is the 'specific' charge contacts known as ion pairing or salt-bridges. Salt-bridges are formed between the acidic/negatively charged aspartic acid (Asp) and glutamic acid (Glu) residues and the basic/positively charged amino acids arginine (Arg), lysine (Lys) and histidine (His). The energy contribution of salt-bridges varies between 5-15 kcal.mol^{-1} per ion pair according to their geometry, location in the protein, whether they are isolated or networked, if they are hydrogen bonded or not (Kumar and Nussinov, 1999).

1.1.2 Protein folding models and pathways

A number of models have been proposed over the years to enlighten the understanding of highly cooperative process of the folding of proteins to their native states. The earlier models proposed that protein folding took place over a defined hierarchical (corresponding to protein structure) pathway of discrete steps, with distinct intermediates and transition states, analogous to classic organic chemistry where a small number of initial species interconvert, governed by energy barriers, to a final product via a series of small chemical steps that can be isolated much in the way chemical reactions occur (Ptitsyn, 1973; Honig *et al*., 1976; Brooks *et al*., 1988). A later model, the hydrophobic collapse model (Dill, 1985), sees the polypeptide initially collapse cooperatively due to long range hydrophobic interactions, followed by the formation of secondary and tertiary structures. Current models of protein folding introduced the concept of energy landscapes and folding funnels (Wolynes *et al*., 1995; Dill and Chan, 1997; Dobson and Karplus, 1999). The width of the funnel represents the conformational entropy of the polypeptide chain, while the depth represents the free energy of stabilisation (Figure 1-1). The unfolded chain compromises a large number of different low-energy conformations in rapid equilibrium with each other, represented by the wide mouth of the funnel. Folding proceeds (going down the funnel), through a progressive organisation of partially folded structures in which it may encounter intermediate states and/or kinetic traps, driven by the accumulation of favourable enthalpic interactions. The native structure exists at the bottom of the funnel, at a global energy minimum, having the most stable conformation. The ruggedness of the funnel's landscape represents the presence or absence of local energy minima. The folding landscape also implies that there are many alternate routes to the native state. A heterogeneous population of starting species often results in parallel folding pathways as opposed to multiple kinetic phases with uniform starting species where the additional rate limiting steps such as the

presence of intermediates result in sequential folding pathways (reviewed in Wallace and Matthews, 2002)

The folding pathway of a protein can only be fully understood when its native or unfolded state as well as intermediates formed during (un)folding can be positioned in the order in which they form and that they are fully characterised (Creighton, 1990; Jaenicke, 1999).

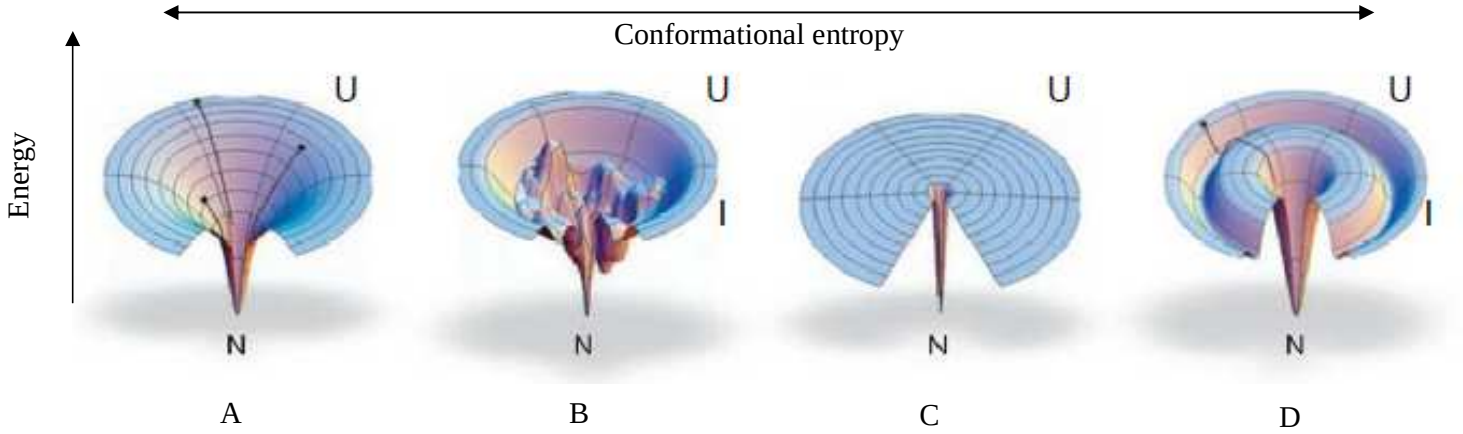


Figure 1-1. Protein folding energy landscapes.
Folding funnel cartoons illustrating (**A**) a smooth energy landscape for a fast folding polypeptide, (**B**) a rugged landscape that contains kinetic traps, (**C**) a smooth "golf course" energy landscape where the conformational search is diffusional, and (**D**) a moat style landscape, where the protein is forced to undergo an intermediate state (I) in order to get to the native conformation (N). U represents the unfolded state ensemble. Image adapted from (Dill *et al.*, 2008).

1.1.3 Protein folding intermediates and the molten globule state

The mechanism by which a protein folds into its biologically active state is an intricate process and it does so through well defined pathways which involve a limited number of intermediate species. One of the major difficulties of the folding pathway is to avoid falling down into the 'traps' (local energy minima) in the funnel. Point mutations in proteins sometimes block the folding pathway at the level of stable intermediate states thus disabling the protein to adopt its native conformation, thereby causing misfolded proteins or the formation of aggregates. The consequences are altered or lost biological activity and ultimately genetic diseases (Bychkova and Ptitsyn, 1995; Ptitsyn *et al.*, 1995). Thermodynamic and kinetic studies have shown the presence of stable intermediate states in a number of proteins (Privalov, 1996; Horwick 2002; Calamai *et al*., 2005). One such intermediate is called the 'molten globule' (Ohgushi and Wada, 1983), characterised by compact secondary structure, but fluctuating tertiary conformation. It also has dimensions slightly larger than those of the native state, but much smaller than those of the unfolded state (Ohgushi and Wada, 1983). During folding, the nascent synthesised protein collapses into a flexible compact species whose tertiary architecture lacks the tight packing typical of the native state (Ptitsyn, 1992) (Figure 1-2). The result of the loose packing makes the molten-globule state susceptible to binding a hydrophobic dye, e.g. ANS (Semisotnov *et al*., 1991; Chaffotte *et al*., 1992), as it contains an accessible hydrophobic surface. The native state is too rigid to allow this accessibility, and the loss of clusters of hydrophobicity in the unfolded state precludes binding of the dye.

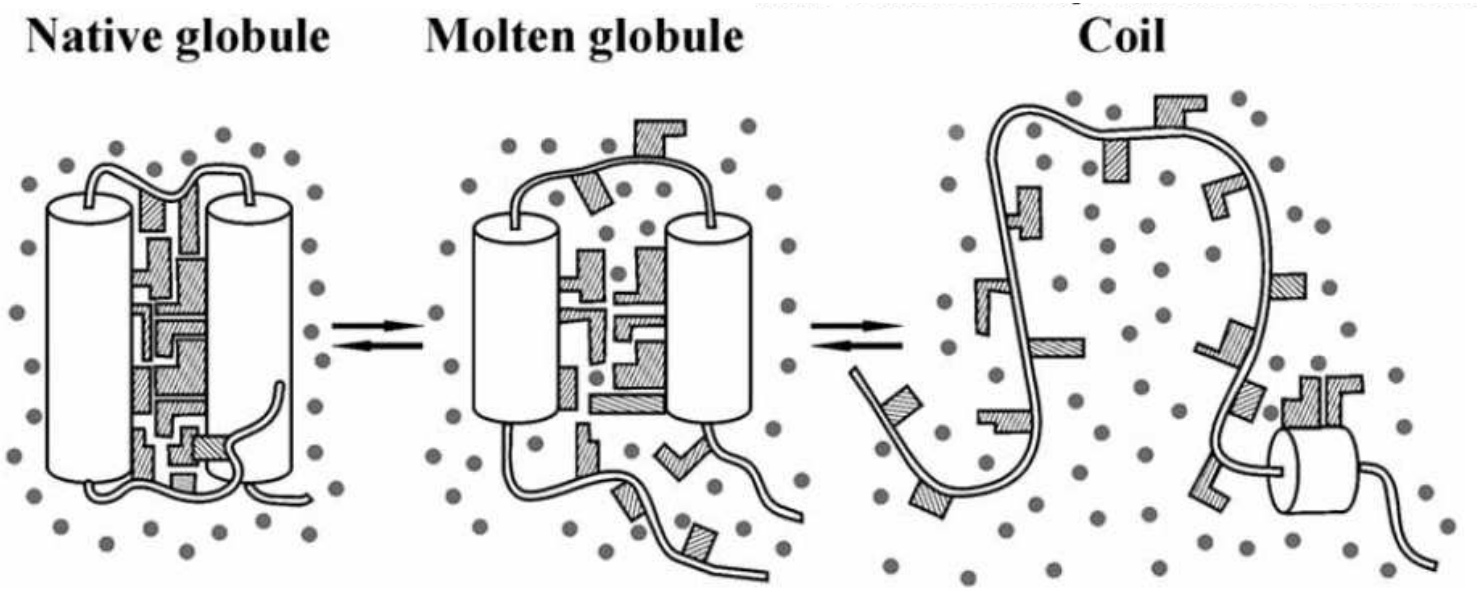


Figure 1-2. Schematic model comparison of the molten globule state.
Native state of the protein is well ordered and rigid with water (grey dots) being present only on the surface of the protein. The molten-globule state contains secondary structure (helical structures are represented by the tubes) but with loose packing of its tertiary contacts (side chains of the amino-acids are represented by the different shaped blocks) thus exposing its hydrophobic core. The unfolded protein is often considered to be a random coil but residual secondary structure may be still evident. Image adapted from (Finkelstein *et al.*, 2007).

Well studied proteins that form molten globular intermediate states include α- lactalbumin (Kuwajima, 1977; Dolgikh *et al*., 1981; Griko *et al*., 1994), equine lysozyme (van Dael *et al*., 1993; Griko *et al*., 1995), staphylococcal nuclease (Dill and Shortle, 1991; Shortle, 1993, 1995, 1996) and apomyglobin (Cocco *et al*., 1992; Barrick and Baldwin, 1993). These molten-globule states of these proteins were characterised and it was found that one domain or subdomain of the partially unfolded protein remains folded, while the other is unfolded (Freire *et al*., 1992; Freire, 1995; Privalov, 1996; Vreuls *et al*., 2004). In general, one of the domains is intrinsically less stable than the other. Molten-globule states are also involved in several biological or pathological processes such as membrane insertion, transmembrane trafficking, and chaperone-assisted refolding that require the protein to become partially unfolded (Hartl *et al*., 1994).

In contrast, other intermediate species that have been detected do not fit the criteria of a molten globule state (reviewed in Ptitsyn, 1995). These have been designated 'pre-molten globule' states or 'highly ordered molten globule' states (Uversky *et al*., 1992), depending on their level of native-like structural content. As a kinetic intermediate it was discovered with the use of ultra-fast (sub-millisecond) measuring techniques (Jones *et al*., 1993; Shastry and Roder, 1998). In addition, proteins with disulfide bonds allow for the trapping of various intermediates indicative of the order of disulphide bond formation (Creighton, 1991).

Intermediates are however generally unstable and poorly populated at equilibrium (Yon, 2001), but kinetic experiments on protein folding can detect transiently formed intermediates during the folding process because the conditions can be manipulated in order to populate the marginally stable species and enable their detection (Utiyama and Baldwin, 1986). A large majority of the work elucidating protein folding pathways has been conducted for small single domain proteins and it was thought that the information obtained from work on multi-domain proteins could be inferred from single-domain proteins by assuming that the domains folded as single units (Jaenicke, 1999).

1.2 Domains

The basic unit of proteins are domains and it is through these domains that proteins interact with each other. Approximately 65% of all proteins contain multiple domains of which

95% contain 2-5 domains (Han *et al.*, 2007). Domains have been shown to act as units of protein function by forming protein substructures with distinct functional properties or by completing active sites through domain interactions. Domains also act as units of protein evolution through domain 'stealing', swapping, and addition, as well as units of protein structure by acting as the building blocks of oligomeric proteins (Jaenicke, 1999). Based on these different properties, there are seven definitions used to classify a domain (Jaenicke, 1999). The definitions of domains are based on the folding and structural features of domains as described below.

According to Wetlaufer (1973), domains are stable units of protein structure that fold autonomously, thereby playing a central role as intermediates during folding. Note, however, that not all domains are autonomous folding units (AFUs). A more accurate definition is given by Richardson who describes domains as compact, local, semi independent structural units (Richardson, 1981). In terms of folding, domains have been defined as cooperative thermodynamic units, detectable by distinct folding/unfolding transitions and are separable by hydrodynamic and spectroscopic measurements (Privalov, 1979). Nearly all large proteins are built from domains (Wodak and Janin, 1981), and large relative movements of domains provide spectacular examples of protein flexibility. Studies have shown that although these domains may be able to fold independently (Teichmann *et al.*, 1999) they cannot reproduce a functionally active protein separately (Yon, 2001).
Combining Wetlaufer's, Richardson's and Privalov's definitions of a domain, it can be described as compact, semi-independent substructures of proteins that fold cooperatively and in some cases autonomously. Domains have hydrophobic cores that make more contacts with themselves than with the rest of the protein. They allow folding to occur at multiple sites along the polypeptide chain ('folding-by-parts') and thus enhance the folding rate. This mechanism may be considered to be an evolutionary advantage in three ways (Jaenicke, 1999):

(i) Domain folding is an efficient way of excluding wrong intramolecular interactions in the case of large protein molecules,
(ii) It protects the nascent polypeptide chain from proteolysis,
(iii) It may be considered a simple mechanism to proceed from monomeric to multimeric proteins by 'domain swapping' (Bennett *et al.*, 1995).

1.3 Domain-domain interfaces: role in protein stability and folding and function

The definition of domains allows for two possibilities in bi-/multi-domain proteins: (i) domains may be independent units or, (ii) they may strongly interact with each other. In the second case all interactions (section 1.1.1) would occur through domain interface residues. The domain interface describes the surface area buried upon domain association as well as the contacts formed between an interacting pair of domains. Therefore, in order to fully comprehend the folding mechanism of multi-domain proteins, in addition to studying the individual domains, one must analyse the features and characteristics of domain interfaces involved in stability and folding.

1.3.1 Features of protein-protein interfaces

The residues at protein interfaces are considered conserved (Valdar and Thornton, 2001; Elcock and McCammon, 2001) because of the evolutionary constraint to maintain interactions. However, it is also argued that the interfaces are conserved only marginally more than the other sequences (Grishin and Phillips, 1994; Caffrey *et al*., 2004). The conservation of interfaces is also used for the prediction of binding sites (Armon *et al*., 2001; Pupko *et al*., 2002).

The sequential N- to C-terminal order of domain combinations tends to be strongly conserved because different multidomain proteins evolved by duplication following a single ancestral recombination event (Bashton and Chothia, 2002). An analysis of protein-protein interfaces showed that there is a good correlation between structurally conserved residues and experimentally-identified amino acids that are important in stability and folding (Keskin *et al*., 2005). Structural alignment of domain interfaces can be consequently used as the basis for protein engineering experiments that look to target critical residues involved in protein structure maintenance. Previous bioinformatics-based studies of domain-domain and protein-protein interface anatomy (Jones and Thornton, 1995; Jones and Thornton, 1996; Tsai *et al*., 1996; Stites, 1997 Jones *et al*., 2000), have shown that the size of the interface is directly related to the extent of interface contacts and thus, to protein stability. As the interfaces are highly diverse in terms of size, affinity, and shape, no simple criterion is sufficient to discriminate specific and nonspecific interfaces such as crystal-packing artifacts (Lo Conte *et al*., 1999; Ponstingl *et al*., 2000; Nooren and Thornton, 2003;

Bahadur *et al*., 2004). However in general, the change in accessible surface area (ΔASA) is known as the most significant predictor, where the interfaces are categorised into large (ΔASA > 2 000 Å^2), medium (ΔASA < 2 000 Å^2), and small (ΔASA < 1 400 Å^2) (Vajda and Camacho, 2004). These interface area classifications as well as five distinct residue–residue contacts within 5Å are the criteria for the classification of interfaces database, SCOPPI (http://www.scoppi.org) (Winter *et al*., 2006), which parses domain sequences from the PQS server (http://pqs.ebi.ac.uk) (Henrick and Thornton, 1998).

The relationship between the conservation of interdomain geometry and protein sequence was investigated and it was found that more conserved domains interact with a more similar geometry (Aloy *et al*., 2003). It was also found that the second domain occupied the same position in 60% of the pairs of homologous domains from different proteins (Han *et al*., 2006) and this is achieved through the conservation of the domain interface. Analysis of the structural and biophysical properties of interfaces of several well-characterised domain pairs, both in terms of thermodynamics and kinetics, showed that large complementary interface not only allows domains to adopt specific conformations relative to each other but will also function to stabilise a protein (Han *et al*., 2007). For the same set of proteins, 55% - 75% of the residues involved at the domain interface were found to be hydrophobic. The interactions at the domain interface are also believed to decrease the probability of unfolding (by increasing unfolding half-lives) and to promote rapid reformation of structure by increasing the refolding rate of proteins that might undergo many rounds of unfolding and refolding (Batey *et al*., 2006).

1.4 Stability and unfolding of the GST family of proteins

The GSTs (EC 2.5.1.18) are a family of multi–functional dimeric enzymes involved in the cellular detoxification and excretion of many physiological and xenobiotic substances (Wilce and Parker, 1994). The GSTs share little sequence similarity but their tertiary structure has been conserved (Wilce and Parker, 1994). The exception is Kappa GST whose C-terminal domain is inserted into the N-terminal domain (Ladner *et al*., 2004). The dimeric structure also has been shown to be involved in stabilisation of tertiary structures of individual subunits (Erhardt and Dirr, 1995) as well as to provide a non-substrate ligand-binding site at the subunit interface (Sayed *et al*., 2000; Lyon and Atkins, 2002; Yassin *et al*., 2004). The GSTs displays two distinct types of subunit interactions (Armstrong, 1997).

The first is a ‘lock-and-key’ hydrophobic interaction, involving an aromatic ‘key’ residue from domain 1 of one subunit that inserts into several hydrophobic ‘lock’ residues of domain 2 in the other subunit (Figure 1-3) found in mammalian classes of Alpha (Sayed *et al.*, 2000), Mu (Hornby *et al.*, 2000) and Pi (Stenberg *et al.*, 2000). The inter-subunit ‘lock-and-key’ is also found in the Zeta class but the ‘key’ is not an aromatic side chain but the side chain of methionine (Wilce and Parker, 1994). The second type of intersubunit interaction is more hydrophilic and lacks the hydrophobic ‘lock-and-key’ motif and electrostatic forces predominate as in the classes of Sigma (Stevens *et al.*, 1998; Stevens *et al.*, 2000) and Theta (Rossjohn *et al.*, 1998). An interesting variant of the ‘lock-and-key’ motif exists in the Delta class of GSTs, known as the ‘clasp lock-and-key’ (Wongsantichon and Ketterman, 2006). The aromatic ‘key’ residue not only inserts into a hydrophobic ‘lock’ of the neighbouring subunit, but also acts as part of the ‘lock’ for the other subunit ‘key’. The ‘key’ residues from both subunits show aromatic ring stacking with each other in a pi– pi interaction, generating a ‘clasp’ in the middle of the subunit interface (Wongsantichon and Ketterman, 2006).

Studies of equilibrium folding revealed that dimer formation of GSTs such as Alpha (Wallace *et al.*, 1998), Pi (Dirr and Reinemer, 1991) and Sj (Andujar-Sanchez *et al.*, 2004) has significant impact on stabilisation of subunit tertiary structure, as these proteins unfold via a ‘two-state’ pathway with the absence of any stable monomeric intermediates. However, the dimerisation of GSTs such as Sigma (Stevens *et al.*, 1998) and Mu (Hornby *et al.*, 2000) has less influence on subunit stability, due to the presence of stable monomeric intermediates in an unfolding/refolding pathway. The lack of equilibrium intermediates in Alpha/Pi/Sj classes of GSTs was attributed to the hydrophobic nature of the subunit interface that would be unstable upon solvent exposure. In spite of the differences between the unfolding pathways of Alpha/Pi/Sj and Mu class GSTs, mutation of the intersubunit ‘lock-and-key’ motif showed that this interaction is important in the stabilisation of the two subunits (Hornby *et al.*, 2000; Sayed *et al.*, 2000; Codreanu *et al.*, 2005; Alves *et al.*, 2006).

Each subunit of the GSTs has two domains (Figure 1-3), an N-terminal domain, which is topologically similar to the thioredoxin fold, and a C-terminal domain is all α-helical with the number of helices varying between different classes. It is believed that the differences in the structure of this domain are responsible for the differences in substrate specificity

between the GST classes (Wilce and Parker, 1994). The two domains are connected via a short linker. The domain-domain interfaces of GST proteins have not been analysed in detail. However, in the case of human Alpha-class GST (hGSTA1-1) an inter-domain 'lock-and-key' motif, similar to the one found in most GST subunit interfaces, was shown to play an important role in stabilising the domain-domain interface (Wallace *et al.*, 2000).

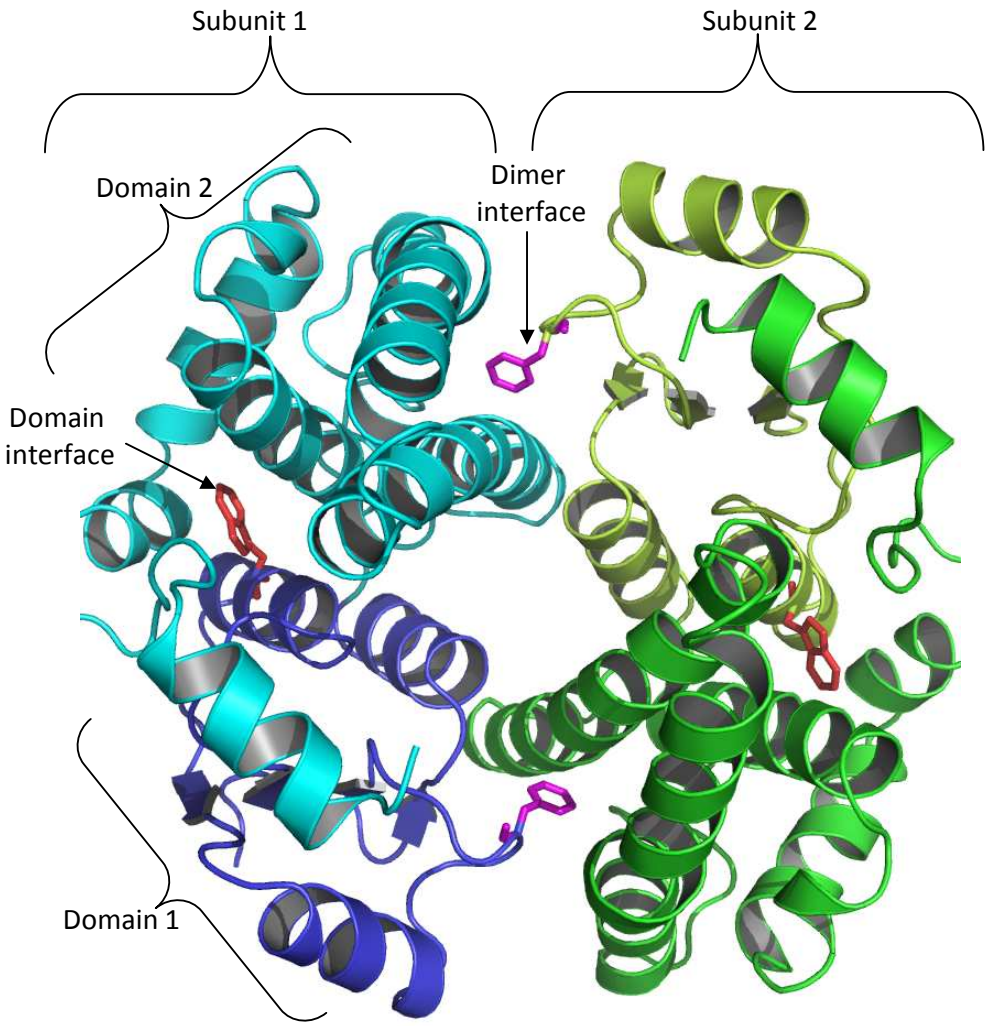


Figure 1-3. 'Lock-and-key' motifs found at the dimer and domain interface of GSTs.
The 'key' residue of the subunit interface (pink coloured residue shown in stick format) and domain interface (red coloured residue shown in stick format). Domain 1 (coloured in blue of subunit 1 and lime green of subunit 2) contains the thioredoxin fold. Domain 2 (coloured in cyan of subunit 1 and green in subunit 2) contains the all α-helical fold. Illustration of hGSTA1-1 is used here as an example (PDB code: 1GUH). Image rendered using PyMOL v0.99 (DeLano Scientific, 2006)

1.4.1. Role of the domain interface in GSTs

The N- and C-terminal domains of GST proteins have been shown to unfold cooperatively in the case of class Alpha (Wallace *et al.*, 1998), Sigma (Stevens *et al.*, 1998), Pi (Erhardt and Dirr, 1995) and Sj (Kaplan *et al.*, 1997). In addition, there is no evidence to suggest that the individual domains of GST proteins unfold independently as indicated by the equilibrium unfolding profiles which display monophasic, coincident transitions. Therefore, the burial of a significant amount of hydrophobic surface area upon domain association indicates that the domain interfaces of GSTs play an important role in the stability and folding of these proteins. A structural based sequence alignment of the GSTs (Figure 1-4) shows that there is a conserved hydrophobic interaction at the domain interface.

The details of the domain-domain interfaces of GST proteins are not as well studied as much as their subunit-subunit interfaces. Luo *et al.* (2002) investigated the conformational stability and equilibrium unfolding of two domain-exchanged chimeric isoenzymes. The domains of the class Mu isoenzymes M1-1 and M2-2 were exchanged resulting in chimeras (M12 and M21) with one domain from M1 and one domain from M2 (Luo *et al.*, 2002). It was shown that the M12 and M21 monomers are less stable then the wild-type monomers (Luo *et al.*, 2002). This indicates that domain interface complementarity is critical for correct domain-domain packing which in turn plays a role in protein stability. The effect of domain packing on stability and function was also investigated in the case of human class alpha GST (hGSTA1-1) (Wallace *et al.*, 2000). The residue chosen for mutagenesis was tryptophan 20 (Trp 20), a conserved amino acid in the class Alpha GST proteins (Wallace *et al.*, 2000). Similarly to the 'lock-and-key' motif found in the subunit interface of Alpha/Pi/Mu/Sj class proteins, the indole ring of Trp 20 protrudes from the N-terminal domain into a hydrophobic pocket of the C-terminal domain where it is completely buried (Wallace *et al.*, 2000) (Figure 1-5A). The W20A mutation, a cavity forming mutation, was both disruptive and destabilising with the equilibrium unfolding results pointing to the accumulation of one or more intermediate species (Wallace *et al.*, 2000). However, the unfolding kinetic data from a less disruptive mutation on hGSTA1-1, W20F, showed that

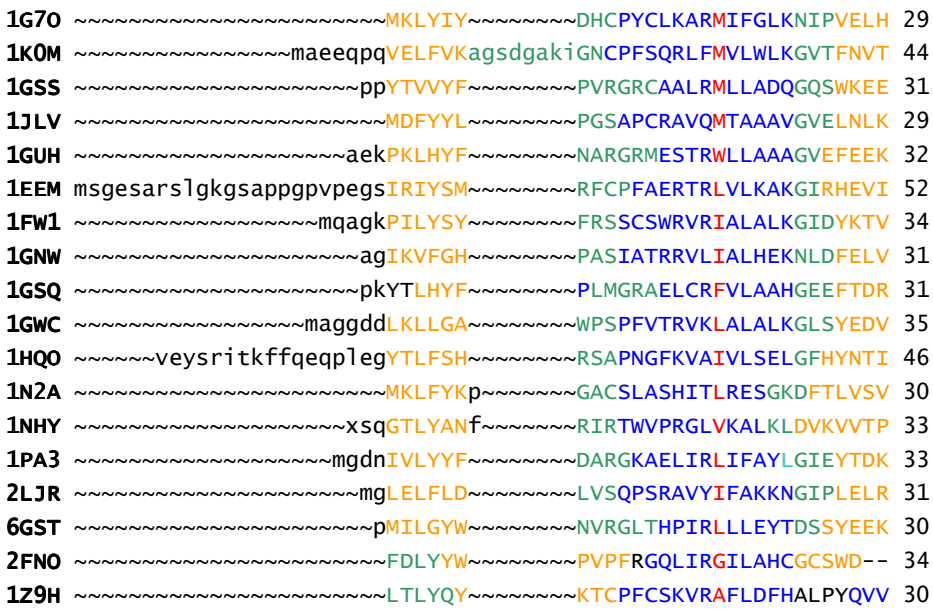


Figure 1-4. Structure based sequence alignment of domain 1 of GSTs
The alignment was performed using sequences of one member of domain 1 of the 18 classes GSTs. The structure-based alignment tool, VAST was used (Gibrat *et al.*, 1996). The resulting alignment displays the secondary structural elements of each protein; helices are in blue, β-strands in green and other structures in orange. Amino acid sequences were obtained from the NCBI (http://www.ncbi.nlm.nih.gov) from the Molecular Modeling Database (MMDB) (Chen *et al.*, 2003) using the PDB accession codes (Table 1-1) obtained from SCOP (http://scop.mrc-lmb.cam.ac.uk/scop/data/scop.b.d.fh.b.html; Murzin *et al.*, 1995). The conserved hydrophobic 'key' residues are highlighted in red.

Table 1-1. List of GST proteins used in alignment
Accession codes of the GSTs were obtained from SCOP database (http://scop.mrc-lmb.cam.ac.uk/scop/data/scop.b.d.fh.b.html, Murzin *et al.*, 1995).

PDB Code	GST Class	Reference
1g7o	Grx2	Xia *et al.*, 2001
1k0m	Clic1	Harrop *et al.*, 2001
1guh	Alpha	Sinning *et al.*, 1993
1n2a	Beta	Rife *et al.*, 2003
1jlv	Delta	Oakley *et al.*, 2001
6gst	Mu	Ji *et al.*, 1994
1eem	Omega	Board *et al.*, 2000
1gnw	Phi	Reinemer *et al.*, 1996
1gss	Pi	Reinemer *et al.*, 1992
1gsq	Sigma	Ji *et al.*, 1995
1gwc	Tau	Thom *et al.*, 2002
2ljr	Theta	Rossjohn *et al.*, 1998
1fw1	Zeta	Polekhina *et al.*, 2001
1nhy	GST-like domain of EF-1	Jeppesen *et al.*, 2003
1pa3	*Plasmodium falciparum* GST	Perbandt *et al.*, 2004
1hqo	yeast prion protein ure2p	Umland *et al.*, 2001
2fno	Hypothetical protein Atu5508	Kosloff *et al.*, 2006
1z9h	Microsomal prostaglandin E synthase-2	Yamada *et al.*, 2005

the global unfolding of the protein during which the unfolding event was not affected by amino acid replacement. A topologically equivalent interaction was found in a monomeric homologue of the GSTs, Clic1 (Stoychev, 2008, PhD Thesis), where the side-chain of Met32 protrudes into a hydrophobic pocket of the C-terminal domain (Figure 1-5B) analogous to the indole ring of Trp20 of hGSTA1-1. Equilibrium unfolding studies show that the mutation of Met32 to Ala was also disruptive and de-stabilising with the accumulation of one or more stable intermediates (Stoychev, 2008, PhD Thesis). These two proteins, hGSTA1-1 and Clic1, also display relatively large interface areas to the other GST classes of protein (Figure 1-6). This was determined by calculating the ΔASA of the domains on each of the GST class of proteins using NACCESS v2.1.1 (http://wolf.bms.umist.ac.uk/naccess) which is an implementation of the Lee and Richards method (1971).

These results show that domain-domain contacts and their correct packing contribute significantly toward protein stability and function. In addition, domain and subunit interfaces seem to have similar roles in the folding and maintenance of the native conformation.

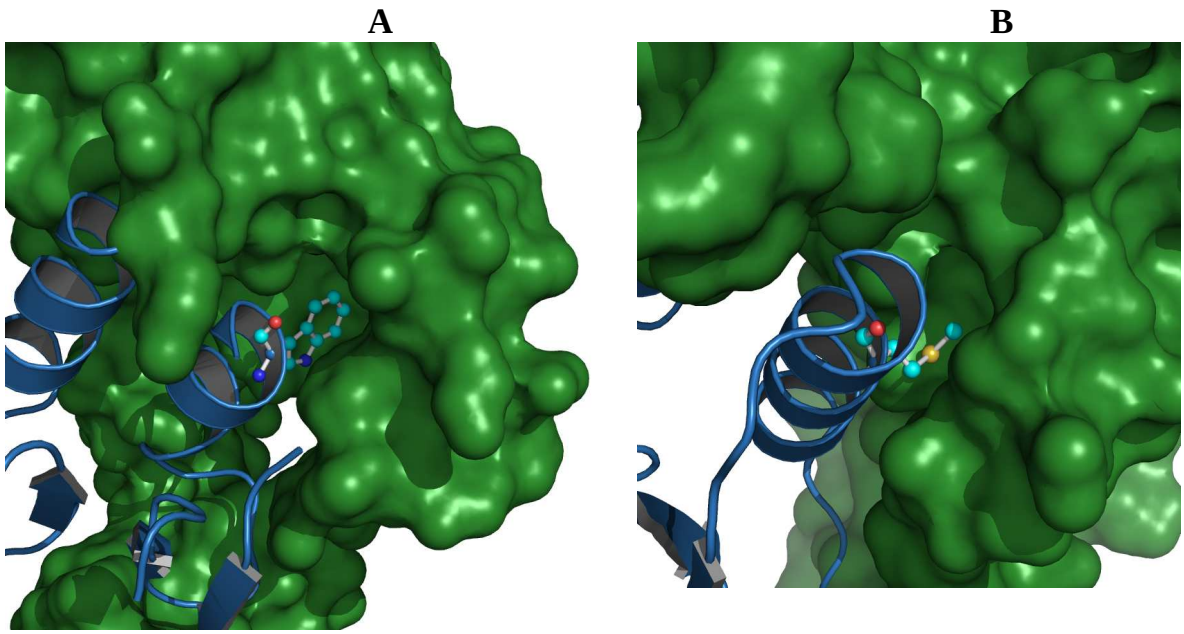


Figure 1-5. The 'lock-and-key' motif found at the domain interface of hGSTA1-1 and Clic1.
(A) The 'key',Trp20 (shown in 'ball and stick' format) of α1 of domain 1 (blue) into its 'lock' of α6 of domain 2 (green). (B) The 'key' residue of Clic1 is Met32 of α1 of domain 1 (blue) into its hydrophobic 'lock' in α8 of domain 2 (green). Image rendered using PyMOL™ v0.99 (DeLano Scientific, 2006) (PDB codes: 1guh and 1k0m).

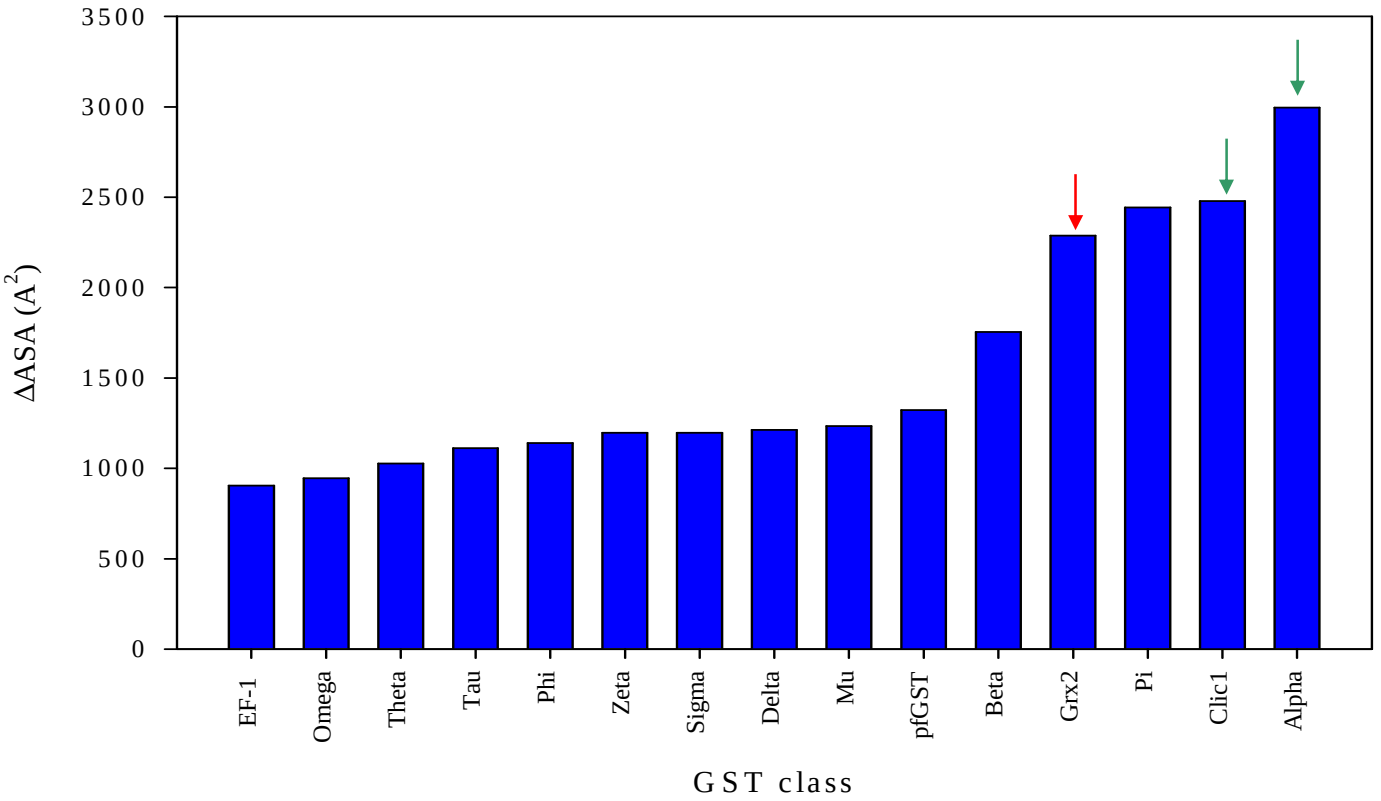


Figure 1-6. Bar-graph depicting interdomain interface area among the GSTs.
Values were computed using NACCESS v2.1.1 (http://wolf.bms.umist.ac.uk/naccess) by calculating the difference between the surface area of the individual domains and surface areas of the domains as a unit. The ΔASA of hGSTA1-1 and Clic1 are shown by the green arrows and Grx2 is shown by the red arrow.

1.4.2 Glutaredoxin-2: A GST monomeric homologue

The elucidation of how domain interactions contribute to the stability of the subunits of GSTs, is complicated by the presence of quaternary interactions, which include residues involved in interdomain contacts (Luo *et al.*, 2002). This study will therefore look at a monomeric homologue of the GSTs, Glutaredoxin-2 (Grx2) from *Escherichia coli.*

The glutaredoxin proteins comprise of three proteins, Grx1-3, that form part of the hydrogen donor system in *E. coli* (Laurent *et al.*, 1964; Holmgren, 1976). Glutaredoxins are reduced by GSH, which is in turn reduced by NADPH and glutathione reductase (Holmgren and Åslund, 1995). Grx2 is an atypical glutaredoxin in size, sequence and function. Unlike its Grx counterparts , it is not a hydrogen donor for ribonucleotide reductase (Åslund *et al.*, 1994). Its size (24.3 kDa) is abnormally large to the other Grx proteins (~10 kDa) (Holmgren and Åslund, 1995). The only similarity it shares with the other Grx proteins is the conserved active site sequence of Cys-Pro-Tyr-Cys (Holmgren and Åslund, 1995).

1.4.2.1 Grx2 Structure

Grx2 has a three-dimensional structure highly similar to GSTs (Xia *et al.*, 2001) (Figure 1-7). Grx2 has the common structural characteristics of GSTs. These are, the N-terminal domain (domain 1) (residues 1-72) having the thioredoxin fold (Holmgren *et al.*, 1975) containing the 4-residue active-site (C-P-Y-C) connected to the all α-helical C-terminal domain (domain 2) (residues 84-215), via an 11-residue linker (residues 73-83) (Xia *et al.*, 2001). The thioredoxin fold in domain 1 of Grx2 is composed a four strand mixed β-sheet (β1, β2, β3 and β4) and this mixed β-sheet is then flanked by three α-helices (α1, α3 and α2). The all α-helical domain 2 consists of six α-helices which are all connected by loops. There are two tryptophan residues in Grx2, Trp89 and Trp190, and both are located in the domain 2. The two cysteine residues (Cys9 and Cys12) of the active site are located at the domain interface. In contrast to the enzymatic activity of the GSTs, Grx2 is involved in the reduction of disulphides with high catalytic activity in resolving the mixed disulfide between β-hydroxyethyl disulfide (HED) and GSH (Åslund *et al.*, 1994; Vlamis-Gardikas *et al.*, 1997; Lillig *et al.*, 1999).

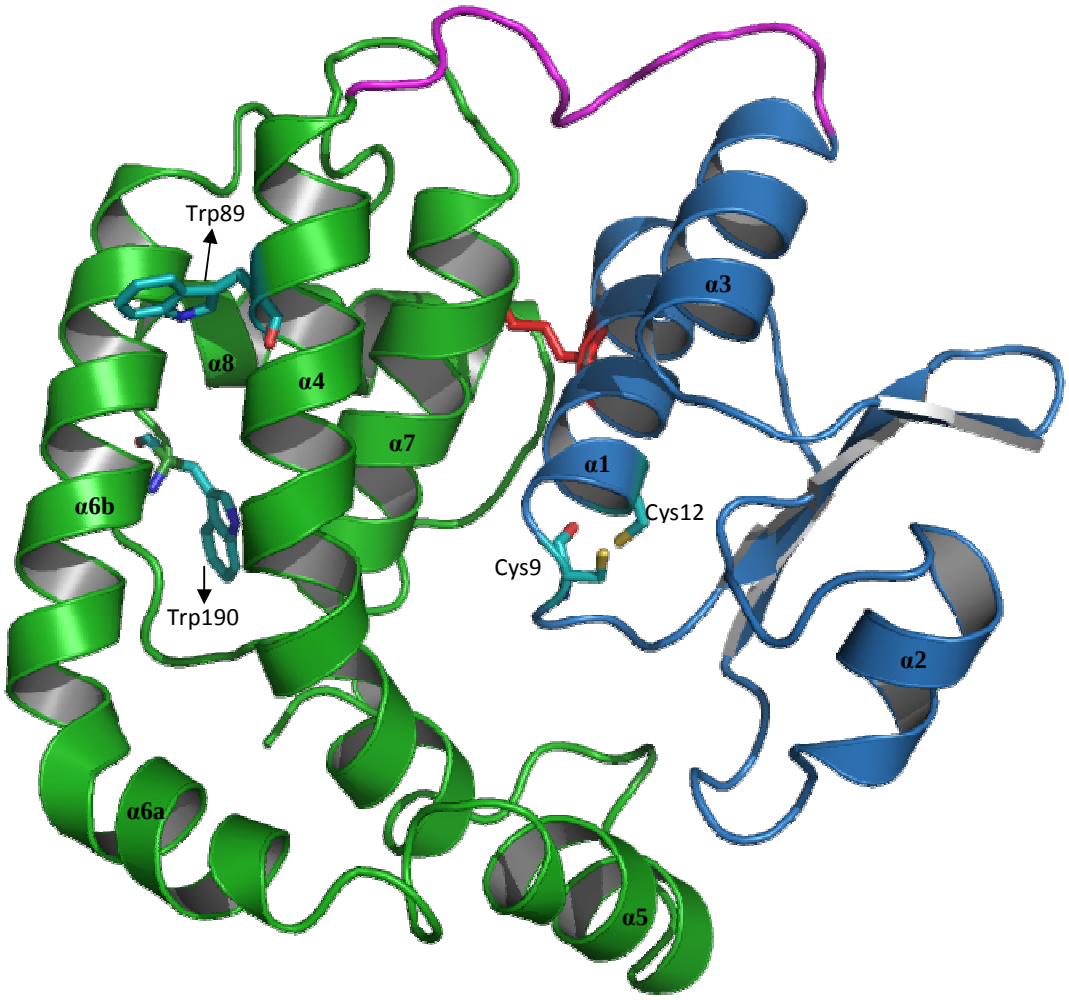


Figure 1-7. Ribbon representation of Glutaredoxin-2
NMR solution structure of reduced Grx2 showing the two domains: N-terminal domain (blue), and C-terminal domain (green). The two domains are connected by the domain linker (magenta). Tryptophan residues (Trp 89 and Trp190) are located on the C-terminal domain. The 'key' residue, Met17 (red) as well as the cystine residues (Cys9 and Cys12) of the active site is located at the domain interface. Image rendered using PyMOL™ v0.99 (DeLano Scientific, 2006) (PDB code: 1g7o).

1.4.2.2 Domain interface of Grx2

Grx2 has been shown to unfold via an 'all-or-none' pathway with no detectable equilibrium intermediates (Gildenhuys *et al.*, 2008). The unfolding kinetics of Grx2 are complex as a result of native-state heterogeneity and are characterised by two observable unfolding reactions that occur in parallel (Gildenhuys *et al.*, 2008). There is also no evidence indicating that unfolding proceeds via a high-energy intermediate that might suggest independent unfolding of the two non identical domains in Grx2 (Gildenhuys *et al.*, 2008). Thus, it seems that the extensive domain interface of Grx2 has a major role in stabilising the individual domains, questioning the intrinsic stability of the N- and C-terminal domains. An attempt was made to isolate the two domains of Grx2 by creating two truncated mutants (Vlamis-Gardikas *et al.*, 1997), but they expressed as inclusion bodies and were less than 50% pure after purification. This can be attributed to the large hydrophobic surfaces that are exposed to the solvent upon domain dissociation. The domain interface of Grx2 has a total accessible surface area buried upon domain association of approximately 2200Å^2 (Figure 1-6). The interactions at the domain interface of Grx2 are predominantly hydrophobic and there are 3 hydrogen bonds between, His8, Asn33 and Glu112 (Xia *et al.*, 2001). A 'lock-and-key' motif is also present at the domain interface analogous to the Trp20 of human class alpha GST and Met32 of Clic1. In Grx2, Met17 is the 'key' residue of domain 1 which locks into a hydrophobic pocket of domain 2 (Figure 1-8).

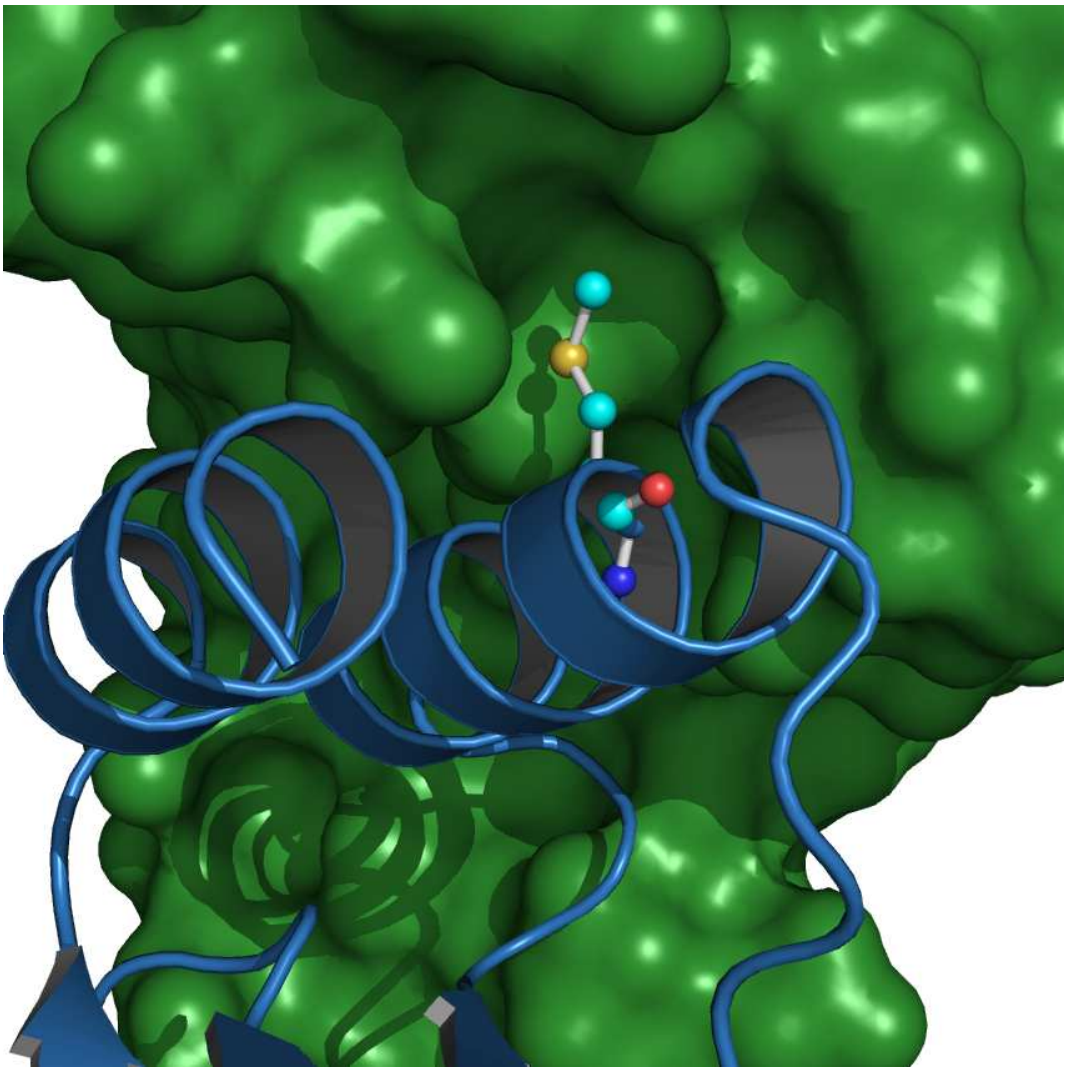


Figure 1-8. The 'lock-and-key' motif at the domain interface of Grx2
The 'key' residue, Met17 (depicted in a 'ball and stick' format), of α1 in domain 1 (blue) into the 'lock' of domain 2 (surface representation in green). Image rendered using PyMOL™ v0.99 (DeLano Scientific, 2006) (PDB code: 1g7o).

1.5 Aims and objectives

The elucidation of protein folding mechanisms requires the study both intra-molecular interactions as the polypeptide chain associates with itself, and inter-molecular interactions as it associates with another polypeptide subunit. This study is aimed at investigating the role and contribution of the domain interface interactions in protein stability and folding. Due to the complexity of the subunit interactions of the dimeric GSTs, this study will focus on a monomeric homologue of GST, Grx2.

This endeavour will use Grx2 as a model protein to investigate the role of the domain interface in generating and maintaining the native conformation. Met 17 of Grx2, a structurally conserved hydrophobic residue in $\alpha 1$ of domain 1 in the GST protein family, will be mutated to an alanine (M17A) in order to investigate the importance of packing at the domain interface. The removal of this inter-domain 'lock-and-key' motif will indicate whether its function is analogous to the function of the 'lock-and-key' motif found in the domain interface of hGSTA1-1.

The objectives to elucidate the role of the 'lock-and-key' motif are threefold. Firstly, to characterise the proteins with respect to their secondary and tertiary structures. Secondly, to determine the conformational stability of the protein using denaturant-induced equilibrium unfolding. Thirdly, single-jump unfolding kinetics will then be used to compare the amplitudes and rates of unfolding as this method has proven to be a powerful means for revealing transient yet important conformational states and interactions that are not accessible by equilibrium measurements. This technique will also be used to perform an initial conditions test to determine native state heterogeneity.

CHAPTER 2. EXPERIMENTAL PROCEDURE

2.1 Materials

The cDNA encoding wild-type Grx2, cloned at the *Bam*H I and *Nde* I restriction sites of the pET24a+ plasmid, was a generous gift from Dr. J. Dyson (The Scripps Institute, CA, USA) (Xia *et al.*, 1999). The QuikChange® Site-Directed Mutagenesis II and StrataPrep plasmid miniprep kits were purchased from Stratagene (La Jolla, CA, USA). Kanamycin and chloramphenicol were purchased from Roche Diagnostics (Mannheim, Germany). 8-Anilino-1-naphthalenesulfonic acid (ANS) was purchased from Sigma-Aldrich (St. Louis, MO USA). SDS-PAGE molecular weight marker (SM0431), DTT and IPTG were purchased from Fermentas Life Sciences (St. Leon-Rot, Germany). DEAE-Sepharose was purchased from GE Healthcare Life Sciences (Uppsala, Sweden). Ultrapure (99.5%) urea was purchased from Merck chemicals (Darmstadt, Germany). All other reagents were of analytical grade. Inqaba Biotech (Pretoria, South Africa) conducted synthesis of oligonucleotide primers and performed all sequencing to confirm the identity of the plasmids.

2.2 Experimental

2.2.1 Construction of Grx2 mutants

Site-directed mutagenesis was employed to create the plasmid DNA encoding the M17A Grx2 mutant. Briefly, the method requires dsDNA (parental plasmid DNA) and two oligonucleotide primers containing the desired mutation complementary to opposite strands of the plasmid. During thermal cycling, these primers are entended, generating the mutant plasmid. The parental plasmid is digested with Dpn I endonuclease which is specific for methalyated DNA (Nelson and McClelland, 1992).

Oligonucleotide primers were designed in accordance with the published wild-type nucleotide sequence of Grx2 (Vlamis-Gardikas *et al.*, 1997), under the guidelines prescribed by the Stratagene QuikChange® II Site-Directed Mutagenesis kit (La Jolla, CA USA) manual, with the aid of the Primer-X software (http://bioinformatics.org/primerx) and Gene Runner software v3.01 (Hastings Software Inc., NY, USA). The sequences of the primers synthesised used to construct the M17A Grx2 mutant are listed below with the respective codons for Ala

highlighted in bold, two silent mutations (italicised) were incorporated in the sequence to avoid hairpins and loops from forming:

Forward: 5' CT TAC TGT CTC AAA *GCT* CGC **GCA** ATT TTC GGC CTG *AAG* AAT ATC 3'

Reverse: 5' GAT ATT *CTT* CAG GCC GAA AAT **TGC** GCG *AGC* TTT GAG ACA GTA AG 3'

The DNA encoding for the M17A Grx2 protein was generated by following the protocol described in the QuikChange® II Site-Directed Mutagenesis kit from Stratagene (La Jolla, CA USA) (Braman *et al.*, 1996). The sample reaction had a final volume of 51 µl which comprised of 5 µl (10x) reaction buffer, 1 µl (50 ng) double stranded DNA template, 1 µl (125 ng) forward primer, 1 µl (125 ng) reverse primer, 1 µl dNTP mix, 41 µl milli-Q water and 1 µl (2.5 U/µl) Pfu DNA polymerase. The product was generated by 16 amplification cycles of 30 seconds at 95°C to denature the wild-type dsDNA, 60 seconds at 55°C to anneal the mutant primers and 60 seconds at 68°C for DNA extension. Parental DNA template was digested with 1 µl (10 U/µl) Dpn I for one hour at 37°C and one hour at 20°C. The reaction products were then used to transform *E. coli* XL1-Blue Supercompetent cells supplied with the mutagenesis kit. The cells were plated onto LB agar plates (1% (w/v) tryptone, 0.5% (w/v) yeast extract, 1.0% (w/v) NaCl, 1.5% (w/v) agar) supplemented with kanamycin ($30 \mu g.ml^{-1}$) and were incubated for 12-16 hours at 37°C. Colonies were chosen and overnight cultures were made in LB medium (1% (w/v) tryptone, 0.5% (w/v) yeast extract, 1.0% (w/v) NaCl). Plasmid DNA was then extracted from the overnight culture of cells, using the Strataprep plasmid miniprep kit from Stratagene (La Jolla, CA, USA). The incorporation of the desired mutation, and that no other mutations were generated during the mutagenesis amplification reaction, was confirmed by sequencing (Inqaba Biotech; Pretoria, South Africa) of the plasmid DNA, using the T7 terminator primer.

2.2.2 Transformation, over-expression and purification of mutant and wild-type Grx2

The plasmids containing the insert that codes for wild-type Grx2 and M17A Grx2 were used to transform *E. coli* BL21(DE3)/pLysS cells (Lucigen, Middleton, WI, USA) and *E. coli* T7 Express I^q competent cells (New England Bio-labs Inc., Ipswich, MA, USA), respectively, and both cell strains are chloramphenicol resistant. The cells were transformed using a one-step method as described by (Chung *et al.*, 1989). Competent cells were thawed on ice for 15 minutes of which 1 µl of mutant plasmid DNA (100 $ng.\mu l^{-1}$) was added and the reaction mixture incubated on ice for 30 minutes. The cells were heat-shocked at 42°C for 10 seconds in the case of the *E. coli* T7 Express I^q Competent cells and for 45 seconds in the case of the *E. coli* BL21(DE3)/pLysS cells, on a heating block, followed by a rapid transfer to ice for 2 minutes. SOC medium (2% (w/v) tryptone, 0.5% (w/v) yeast extract, 250 mM KCl, 1 M glucose, 2 M $MgCl_2$) was added to the reaction mixture followed by incubation at 37°C for 90 minutes. The cells were then plated on LB-agar plates (1% (w/v) tryptone, 0.5% (w/v) yeast extract, 1.0% (w/v) NaCl, 1.5% (w/v) agar) supplemented with the antibiotics kanamycin (30 $\mu g.ml^{-1}$) and chloramphenicol (30 $\mu g.ml^{-1}$). The plates were then incubated at 37°C for 12 - 16 hours. The transformed *E. coli* expression cells transformed with the pET24a+ plasmid containing the cDNA sequence coding for the Grx2 proteins were added to fresh, sterile 2xYT medium (1.6% (w/v) tryptone, 1.0% (w/v) yeast extract, 0.5% (w/v) NaCl) supplemented with the antibiotics. The cells were grown at 37°C with shaking at 250 rpm for 12 - 16 hours of which a 50-fold dilution was then used to inoculate into fresh, sterile 2xYT medium (1.6% (w/v) tryptone, 1.0% (w/v) yeast extract, 0.5% (w/v) NaCl) supplemented with kanamycin (30$\mu g.ml^{-1}$) and chloramphenicol (30$\mu g.ml^{-1}$). Cells were grown at 37°C with shaking at 250 rpm till an OD_{600} of ~0.6 was reached after which over-expression of the Grx2 proteins were induced by the addition of 1 mM IPTG. The cells were grown for a further 12 - 16 hours at 37°C, with shaking at 250 rpm, in order to achieve optimum protein expression. Cells were harvested via centrifugation (4200 × *g*, 25 min) and re-suspended in approximately 35ml re-suspension buffer (20 mM Tris-HCl, 1 mM EDTA, 2 mM $MgCl_2$, 0.02 % NaN_3, pH 10.0). DNAse I (10 $mg.ml^{-1}$) and lysozyme (10 $mg.ml^{-1}$) were added to the cell suspension and was rotated at 4°C for 30 minutes. The cells were then sonicated on ice for 4 cycles of 30 seconds with 40% pulse intensity using an ultrasonic liquid

processor (Misonix Inc. (model: XL-2020), Farmingdale, NY, USA). The lysed cells were centrifuged at 16 000 × *g* for 20 minutes at 4°C. Samples of whole cell extract, soluble and insoluble fractions were electrophoresed on 12% acrylamide SDS-PAGE gels (section 2.2.3).

Purification of the Grx2 proteins was performed using anion-exchange chromatography following an optimised protocol (Gildenhuys, 2006, PhD Thesis) to that of the one described previously (Vlamis-Gardikas *et al.*, 1997). The soluble fraction after centrifugation was applied onto a DEAE-Sepharose column (GE Healthcare Life Sciences, Uppsala, Sweden) pre-equilibrated with buffer A (20 mM Tris-HCl buffer, pH 10, 0.02% (w/v) NaN_3). The column was then washed with 10 column volumes of buffer A followed by 4 column volumes of wash with buffer B (20mM Tris-HCl buffer, pH 9, 0.02% (w/v) NaN_3). Bound wild-type or M17A Grx2 protein was eluted off the column using a linear pH gradient (10 column volumes) from pH 9 to pH 8. The gradient was produced by mixing buffer B with 50 mM Tris-HCl, pH 8, 0.02% (w/v) NaN_3. Anion-exchange chromatography was conducted using an ÄKTAprime system attached to a computer with PrimeView 1.0 software (GE Healthcare Life Sciences, Uppsala, Sweden). The purified wild-type Grx2 or M17A Grx2 protein was then dialysed against three changes of Grx2 storage buffer (50 mM sodium phosphate buffer, pH 7.0, containing 50 mM NaCl, 1 mM DTT and 0.02% (w/v) NaN_3), lasting 4 hours each, after which it was snap frozen and stored at -80ºC.
It should be noted that the over-expressed proteins were void of the His-tag present in the vector (Figure 2-1) as the sequence encoding the Grx2 proteins are inserted at the *Bam*H I and *Nde* I restriction sites which are upstream on the His-tag coding region and therefore not expressed (Figure 2-1). Before use, the frozen protein was thawed on ice and dialysed against Grx2 storage buffer, ensuring that the protein was in the reduced state. The Grx2 proteins were used within a week after which the Grx2 protein was placed in fresh dialysis buffer until the next dialysis. Unless otherwise stated all experiments for wild-type and M17A Grx2 were conducted in the Grx2 storage buffer.

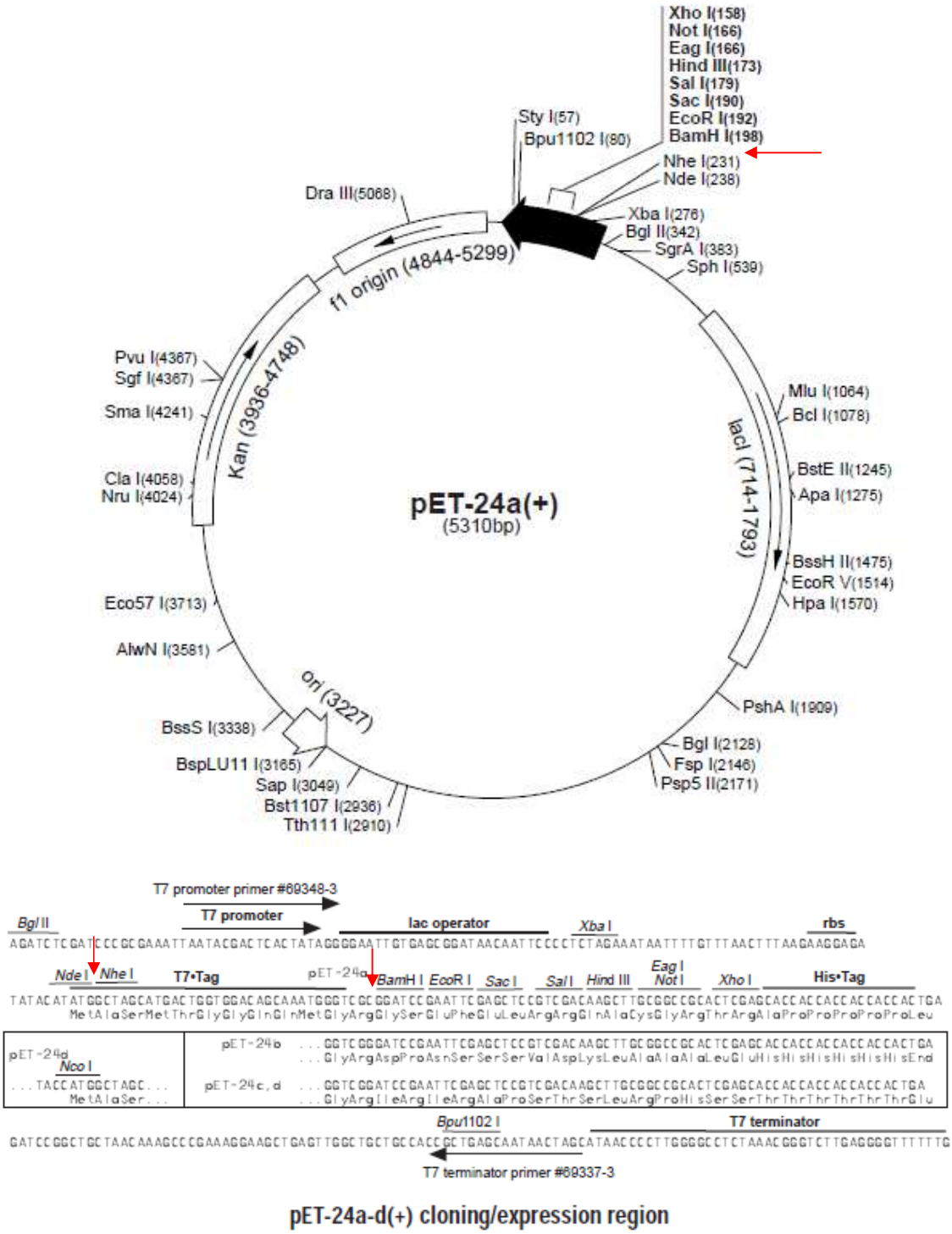


Figure 2-1. Vector map of pET24a+
The position of the insert encoding Grx2 is shown by the arrows (red). Image adapted from pET24a+ technical specification sheet (Novagen, product: TB070), Merck chemicals (Darmstadt, Germany).

2.2.3 SDS-PAGE

The solubility, homogeneity and purity of the expressed mutant proteins were assessed by separation on a 12% SDS-PAGE (Laemmli, 1970). The discontinuous gel system consisted of a 4% (w/v) acrylamide/bis-acrylamide stacking gel (0.1% (w/v) SDS, 0.05% (w/v) ammonium persulphate, 0.1% (w/v) TEMED and 0.125 M Tris-HCl buffer, pH 6.8) and a 12% (w/v) acrylamide/bis-acrylamide (w/v) separating gel (0.1% (w/v) SDS, 0.05% (w/v) ammonium persulphate, 0.1% (w/v) TEMED and 0.375 M Tris/HCl, pH 8.8). Protein samples were diluted two-fold with sample buffer (10% (w/v) glycerol, 2% (w/v) SDS, 5% (w/v) β-mercaptoethanol, 0.05% (w/v) bromophenol blue and 0.0625 M Tris-HCl buffer, pH 6.8). Samples were then boiled for 5 minutes to ensure that the proteins were denatured. The electrode buffer used contained 1% (w/v) SDS, 0.192 M glycine and 0.025 M Tris, pH 8.5. The protein samples were applied to the SDS-PAGE wells (15µl) and electrophoresed at 140 V for 2 hours using a Hoefer MiniVE electrophoresis system (Holliston, MA, USA). The molecular weight marker (Fermentas Life Sciences, (cat. no.: SM0431), Ontario, Canada) used contained a mixture of seven proteins: β-galactosidase (116 kDa), bovine serum albumin (66.2 kDa), ovalbumin (45 kDa), lactate dehydrogenase (35 kDa), restriction endonuclease Bsp98I (25 kDa), β-lactoglobulin (18.4 kDa) and lysozyme (14.4 kDa). The gels were stained in 2% (w/v) Coomassie Blue R250 staining solution containing 13.5% (v/v) glacial acetic acid and 18.75% (v/v) ethanol and destained with 40% (v/v) ethanol and 10% (v/v) glacial acetic acid until the background was clear.

2.2.4 SE - HPLC

The size and homogeneity of purified wild-type and M17A Grx2 was assessed using SE-HPLC. Proteins were separated using a TOSOH TSK gel G 2000 SWXL size exclusion column (TOSOH Corporation, Tokyo, Japan) with a TOSOH TSK gel SWXL guard column (TOSOH Corporation, Tokyo, Japan). The column was equilibrated using Grx2 storage buffer. This buffer was pumped at an isocratic pressure of 40 bar through the column using a Dionex Ultimate 3000 pump (Sunnyvale, CA, USA) at a constant flow rate of 0.4 $ml.min^{-1}$. Protein eluting from the column was detected using absorbance, fluorescence and dynamic light scattering. The absorbance detector (Spectra-Physics (model: UV100SP), Fremont, CA, USA) was set to record absorbance at 280 nm with a sensitivity of 0.02. The fluorimeter

(Jasco Inc., (model: FP 2020), Tokyo, Japan) was set have an excitation wavelength of 295 nm and emission wavelength of 345 nm, with the gain and attenuation set at 32 and 100 respectively. The sizing of particles passing through the column was detected by light scattering using a Zetasizer Nano S (Malvern Instruments Ltd., Worcestershire, UK). The instrument temperature was set to 22°C and the resolution was set to general purpose and data collected every 3 seconds using DTS software (Malvern Instruments Ltd. Worcestershire, UK).

2.2.5 Absorbance spectroscopy

The concentrations of the Grx2 proteins, ANS and DTNB were determined spectrophotometrically using a Jasco V-630 UV-VIS spectrophotometer (Jasco Inc., Tokyo, Japan) and by applying the Beer-Lambert law:

$$A = \varepsilon_{\lambda} cl \qquad (3)$$

where A is the absorbance at the respective wavelength, ε_{λ} is the molar extinction of the absorber at wavelength λ, c is the concentration of the absorbing solution and l is the path length of light through the solution (cuvette).

A molar extinction coefficient (ε_{λ}) of 21 860 $M^{-1}.cm^{-1}$ (Vlamis-Gardikas *et al.*, 1997) was used for determining the concentrations of both wild-type and M17A Grx2 proteins at 280 nm. The same extinction coefficient is applicable for the mutant protein since the amino acids involved in the mutation, i.e., Met to Ala are not chromophores. The extinction coefficients used for determining ANS concentration at 350 nm was 5000 $M^{-1}cm^{-1}$ (Weber and Young, 1964). The quantitation of free sulfhydryls, which employed the DTNB essay (section 2.2.8), at 412 nm used an extinction co-efficient of 13 600 $M^{-1}.cm^{-1}$ (Habeeb *et al.*, 1972) for the 2-nitro-5-thiobenzoate anion.

The concentrations were determined by fitting a linear regression to 10 or more points from a serial dilution. All readings were buffer corrected with the appropriate buffer used for the concentration determination.

2.2.6 Circular dichroism spectroscopy

Circular dichroism (CD) is a technique that measures the differential absorption of left- and right-handed circularly polarised light by optically active molecules. Optical activity in proteins arises from disulphide groups, aromatic side chains, and the peptide backbone (Woody, 1995). Disulphide groups and aromatic amino acids have characteristic absorption bands in the near-UV range (250 - 300 nm). In the far-UV region (180 - 250 nm), the predominant signal arises from the peptide backbone. The adoption of different secondary structures by the peptide backbone results in distinctive CD spectra (Woody, 1995). As a result, this wavelength range gives a good indication of the secondary structural content of proteins. Proteins with a high α-helical content display characteristic minima at 208 and 222 nm, as well as a stronger positive band near 190 nm (Woody, 1995). Due to the noise contribution by some buffers, hence reducing the signal to noise ratio, it is impossible to record clear spectra below 210 nm.

Far-UV CD spectra (180 - 250 nm) were recorded using 5 μM Grx2. The protein was in Grx2 storage buffer. In some cases the Grx2 storage buffer was diluted 10 fold in order to improve the signal to noise ratio. All far-UV CD spectra were recorded at 20°C and represent an average of 10 accumulations, at a scan speed of 200 $nm.min^{-1}$. The bandwidth used was 1 nm and the data pitch 0.2 nm. Measurements were obtained using on a Jasco J-810 spectropolarimeter with Spectra Manager software v1.5.00 (Jasco Inc., Tokyo, Japan) using a pathlength of 2 mm. All spectra were buffer corrected. The spectra were normalised by calculating the mean residue ellipticity [θ] $deg.cm^2.dmol^{-1}.residue^{-1}$ using the following equation (Woody, 1995):

$$[\theta] = (100.\theta)/c.n.l \qquad (4)$$

where (θ) is the ellipticity signal in mdeg, c is the protein concentration in mM, n is the number of residues in the protein chain and l is the path length in cm.

Near-UV CD spectra (250 – 350 nm) were recorded using 40 μM Grx2 at 5 °C in Grx2 storage buffer. The temperature was maintained by a Jasco PTC-423S Peltier-type temperature control system. All near-UV CD spectra were recorded at a scanning speed of 100 $nm.min^{-1}$, using a 1 cm path length cuvette. The sensitivity was set to

high (10 mdeg), data pitch was 0.05 nm, response 1 sec, bandwidth 0.5 nm, and each spectrum was the result of 10 accumulations. Low temperature CD measurements on proteins sharpen the CD bands because of lowered motility of the side-chains, and at very low temperatures increase the intensity of the signal (Strickland, 1974). Near-UV CD signal data were not converted to mean residue ellipticity since only four types of residues contribute to the signal, making averaging over all residues unjustified.

2.2.7 Fluorescence spectroscopy

Fluorescence is the emission that results from the return of an unpaired electron from the excited to the ground state (Lakowicz, 1999). The energy lost between excitation and emission, known as Stokes' shift, results in the bathochromic (red) shift of emission spectra. In proteins, the naturally occurring fluorophores are tryptophan, tyrosine and phenylalanine. Due to the small quantum yield of phenylalanine in proteins, its emission is rarely observed (Lakowicz, 1999). The fluorescence of most proteins is dominated by tryptophan, with its quantum yield being more than double that of tyrosine. In the native state, tyrosine emission is quenched by energy transfer to tryptophan, and to quenching due to nearby charged carboxyl and uncharged amino groups (Lakowicz, 1999).

The indole ring of tryptophan is highly sensitive to solvent polarity (Lakowicz, 1999). Emission spectra of this residue reflect the polarity of its surrounding environment. Therefore, tryptophan fluorescence is used to monitor tertiary structural changes in proteins. All fluorescence measurements were recorded in a quartz cuvette with a 10 mm path-length using a Perkin-Elmer luminescence spectrometer LS50B and FLwinlab v4.0 software (Waltham, MA, USA).

2.2.7.1 Intrinsic fluorescence-tryptophan fluorescence

Grx2 contains two tryptophan residues, both in domain 2, at positions 89 (Trp89) and 190 (Trp190) and eight tyrosine (Tyr) residues (5 in domain 1 and 3 in domain 2). The Trp residues were selectively excited at 295 nm. Fluorescence emission spectra were recorded using 5 μM Grx2 in the range 280 - 450 nm. The excitation and emission slit widths were 5.5 nm and 3.5 nm, respectively. The buffer used was Grx2 storage buffer. The spectra were recorded at 20°C, buffer corrected, and is an average of three accumulations at a scan speed of 200 $nm.min^{-1}$.

2.2.7.2 Extrinsic fluorescence - ANS binding

ANS is a hydrophobic dye used as an extrinsic fluorescence probe (Engelhard and Evans, 1995). It binds to hydrophobic patches in proteins. In an aqueous environment ANS fluorescence is quenched, but upon binding to a hydrophobic surface its fluorescence quantum yield increases and its maximum emission wavelength is blueshifted (Engelhard and Evans, 1995).

A stock solution (10 mM) of ANS was prepared in Grx2 storage buffer. The concentration of ANS was checked by recording the absorbance at 350 nm and using extinction coefficient of ε_{350} = 5000 $M^{-1}cm^{-1}$ (section 2.2.5). Protein (5 μM) was incubated for 60 and 90 minutes, at different concentrations of urea (0 - 8 M) with ANS added to the protein/urea mixture to a final concentration of 200 μM. The solution was incubated for at least an hour to achieve equilibrium. A series of blanks were generated, each containing 200 μM ANS with the appropriate urea concentration (0 – 8 M). The samples were excited at 390 nm and emission spectra were recorded from 390 to 600 nm. Spectra were produced from an average of three accumulations at 300 $nm.min^{-1}$ scan speed. The excitation and emission slit widths were at 5 nm. The spectra were recorded at 20°C, buffer corrected, and are an average of three accumulations at a scan speed of 300 $nm.min^{-1}$. The fluorescence emission intensities at 465 nm were extracted and plotted as a function of urea concentration.

2.2.8 DTNB assay

Wild-type Grx2 contains two cysteine residues: Cys9 and Cys12 (Xia *et al.*, 2001). The accessibility and redox state of these residues was assessed with a DTNB assay (Thannhauser *et al.*, 1984). Prior to conducting the assay wild-type Grx2 was buffer exchanged using a G-25 Sephadex column into a 50 mM sodium phosphate, pH 7.0, containing 1 mM EDTA and 0.02 % (w/v) NaN_3. Following buffer exchange the DTNB assay was conducted by titration of a solution of 5 μM protein in 50 mM sodium phosphate buffer, pH 7.0 containing 1 mM EDTA, 0.02 % (w/v) NaN_3, with 0.2 mM DTNB. The 2-nitro-5-thiobenzoate anion (absorbs maximally at a wavelength of 412 nm) is released when DTNB reacts with free thiol groups (Habeeb *et al.*, 1972). The concentration of the released 2-nitro-5-thiobenzoate anion was determined spectrophotometrically at 20ºC as described previously (section 2.2.5).

2.2.9 Thermal denaturation studies

Thermal unfolding is widely used to determine protein stability. Heat or temperature induced unfolding studies were conducted to assess the relative stabilities of wild-type and M17A Grx2.

Temperature-induced denaturation of Grx2 was determined by monitoring the ellipticity at 222 nm (section 2.2.6) over the temperature range 20°C to 80°C. The temperature unfolding profiles were recorded using 2 μM Grx2 in Grx2 storage buffer. The temperature was controlled by a Jasco PTC-423S Peltier-type temperature control system and the rate at which the temperature was increased was $1°C.min^{-1}$. The bandwidth was 1 nm and data pitch 0.2°C. Measurements were obtained using on a Jasco J-810 spectropolarimeter with Spectra Manager software v1.5.00 (Jasco Inc., Tokyo, Japan) using a pathlength of 2 mm. The temperature unfolding profiles were normalised by calculating the mean residue ellipticity $[\theta]$ $deg.cm^2.dmol^{-1}.residue^{-1}$.

2.2.10 Urea - induced equilibrium unfolding studies of wild-type and M17A Grx2

2.2.10.1 Reversibility of folding

In order to assess the conformational stability of a protein, the reversibility of the unfolding event needs to be established. It is essential to determine the degree of reversibility of the unfolding process as it will give an indication that equilibrium between the folded and unfolded states exists. The reversibility of unfolding was determined for wild-type and M17A Grx2.

The refolding of unfolded protein samples, incubated in 8 M urea for one hour at 20°C, was achieved by a 6-fold dilution of each sample reaction with Grx2 storage buffer for 1 hour at 20°C. The refolded state of the protein was assessed using far-UV CD (section 2.2.6) and intrinsic tryptophan fluorescence (section 2.2.7).

All of the urea used for experiments was prepared by the method of (Pace, 1986) using Grx2 storage buffer as the solvent. Following preparation, the pH of the stock urea solution was adjusted to pH 7, filtered and the concentration of 10 M confirmed using an Atago R5000 refractometer (Tokyo, Japan) and the method of (Pace, 1986).

2.2.10.2 Urea - induced equilibrium unfolding/refolding

Protein unfolding transitions are a convenient way of estimating the stability of a protein. The effect of an engineered mutation on the stability of a protein can be studied by comparison of the wild-type and mutant equilibrium unfolding curves. A denaturant is used to shift the equilibrium from the native to the unfolded state. The equilibrium constant (K_{eq}) can be calculated and hence the conformational stability parameters $\Delta G(H_2O)$ and *m*-value can be determined provided reversibility has been established (section 2.2.10.1).

Urea unfolding of wild-type Grx2 and M17A Grx2 was conducted over a range of urea concentrations from 0 M to 8 M urea in the absence and presence of ANS (section 2.2.7.2). Unfolding was conducted at 20ºC for 1 hour to allow equilibrium to be reached. Thereafter, the samples for the range of urea concentrations were monitored using far-UV CD (section 2.2.6) and fluorescence spectroscopy (section 2.2.7.1).

2.2.11 Data fitting

Unfolding/refolding data obtained for all the proteins were analysed according to a two-state unfolding process for a monomeric protein.

During two-state reversible unfolding there is an equilibrium reached between the native species (N) and the unfolded species (U):

$$N \rightleftharpoons U \qquad (5)$$

During a two-state unfolding transition, only the unfolded and native states are present at significant concentrations (Pace, 1986).

Therefore, for a two-state mechanism:

$$f_N + f_U = 1 \qquad (6)$$

where f_N is the fraction folded or native protein and f_U is the fraction unfolded protein. At any point during unfolding, there is a contribution to the signal from the concentration of both species:

$$y = y_N f_N + y_U f_U \quad (7)$$

where y is the signal obtained for the respective spectroscopic probe, f_N represents the fraction of folded protein, f_U represents fraction unfolded protein. In addition y_N represents the y value for the folded state and can be extrapolated from linear pre-transition region of the unfolding data and y_U represents the y value for the unfolded state and can be extrapolated from the linear post-transition region of the unfolding data (Figure 2-1). By combining equations 6 and 7, the fraction of unfolded protein can be obtained:

$$f_U = (y_N - y)/(y_N - y_U) \quad (8)$$

similarly the fraction of folded or native protein can be obtained:

$$f_N = (y - y_U)/(y_N - y_U) \quad (9)$$

The equilibrium constant for the unfolding reaction (K_{eq}) is:

$$K_{eq} = f_U / f_N \quad (10)$$

So, therefore, if equations 8 and 9 are substituted into 10:

$$K_{eq} = (y_N - y) / (y - y_U) \quad (11)$$

and

$$\Delta G° = - RT \ln K_{eq} \quad (12)$$

where $\Delta G°$ is the free energy of unfolding, R is the gas constant, T is temperature in Kelvin and K_{eq} is the equilibrium constant for a reaction. In order to determine $\Delta G(H_2O)$ it is assumed that $\Delta G°$ has a linear dependence on denaturant concentration [D] for all urea concentrations (Tanford, 1968, 1970).

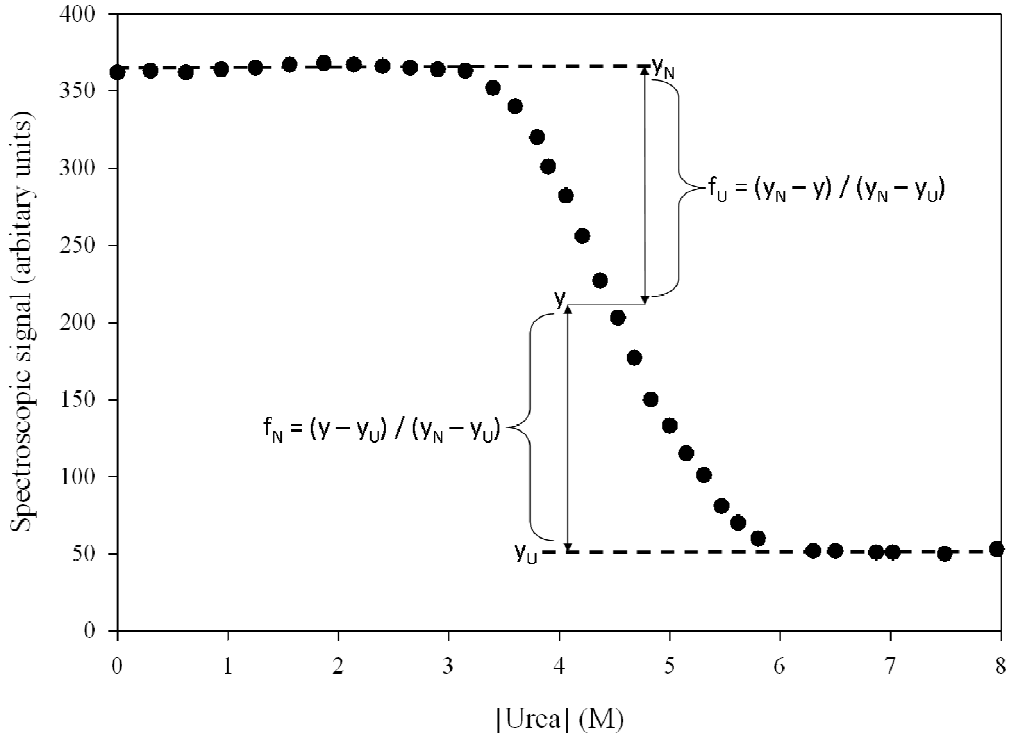


Figure 2-2. Urea denaturation curve
A two state mechanism is assumed for analysis. Image adapted from Shirley (1995).

Therefore:

$$\Delta G° = \Delta G(H_2O) - m\,[D] \qquad (13)$$

where $\Delta G(H_2O)$ represents the free energy difference between the folded and unfolded states in the absence of denaturant, m is the m-value for the dependence of free energy on denaturant concentration which is also an indicator of co-operativity and can be related to the change in solvent-accessible surface area ($\Delta SASA$) , and [D] is the denaturant concentration.

Combining equations 11, 12 and 13 thus gives:

$$y = [y_N + y_U * e^{-(\Delta G\,(H2O) - m\,[D])/RT}] / [1 + e^{-(\Delta G\,(H2O) - m\,[D])/RT}] \qquad (14)$$

The equilibrium unfolding data obtained were fitted to equation 12 using SigmaPlot version 11.0 (Systat Software Inc; Chicago, IL, USA) and the parameters $\Delta G(H_2O)$ and m were obtained.

2.2.12 Unfolding kinetics

The kinetic experiments were conducted using a Pro-data upgraded SX-18MV stopped-flow reaction analyzer from Applied Photophysics (Leatherhead, U.K.). The excitation pathlength was 10 mm and the emission pathlength was 2 mm. The excitation bandwidth was 2.3 nm to minimise photodecomposition. The photomultiplier voltage was set at 576 V for all experiments. The temperature of the sample handling unit was maintained at 20ºC with a water bath.

2.2.12.1 Kinetic studies

Kinetic studies were performed by monitoring changes in the intrinsic tryptophan fluorescence emission of Grx2. The excitation wavelength for all experiments was 280 nm, and a cut-off filter of 320 nm was used to prevent excitation wavelength swamping the emission signal. It was shown prior to kinetics experiments for wild-type Grx2 and M17A Grx2 that the proteins do not undergo any photodegradation over the period of the measurements.

2.2.12.2 Single-jump unfolding studies

The unfolding reactions (N → U) were setup by mixing 7.5 μM native wild-type or M17A Grx2 in an asymmetric ratio of 1:5 with Grx2 storage buffer containing 6 M to 9 M urea. The final unfolding conditions were 1.25 μM protein and urea concentrations from 5 M to 7.5 M.

The baseline values were obtained for 1.25 μM native and urea-denatured (in 8 M urea) wild-type Grx2 and M17A Grx2. Four traces were averaged for each kinetic experiment at each urea concentration, with the final average trace analysed using the Applied Photophysics software Data were fitted to the following general equation:

$$F_t = \Sigma F_i * \exp(-t/\tau) + y_o \qquad (15)$$

where F_t is the total fluorescence amplitude, F_i is the amplitude for the phase i at time zero, t is time, τ is the time constant (is the inverse of the apparent rate constant) and y_o is the fluorescence amplitude at infinite time.

CHAPTER 3. RESULTS

3.1 Sequence identity

The pET-24-a+ plasmid containing the open reading frames (ORF) encoding wild-type Grx2 and the mutant M17A Grx2 were sequenced. Figures 3-1 (A) and (B) show a segment of the wild-type and M17A Grx2 cDNA ORF respectively, obtained from DNA sequencing. The presence of the engineered mutation as well as the two silent mutations was confirmed and no further mutations were found to be introduced during the thermal cycling reactions.

3.2 Over-expression and purification

The recombinant expression system for wild-type Grx2 comprised of the *E. coli* BL21 (DE3)/pLysS cells transformed with the pET-24a+ plasmid containing the ORF encoding the Grx2 protein using the method described in section 2.2.2. The recombinant expression system for M17A Grx2 comprised of *E. coli* T7 Express I^q competent cells transformed with the pET-24a+ plasmid containing the ORF encoding the Grx2 protein using the method described in section 2.2.2. The expression system for M17A Grx2 initially contained the *E. coli* BL21(DE3)/pLysS system, however, the M17A Grx2 protein was found to be insoluble at growth temperatures of 37˚C and 20˚C (Figure 3-2A). *E. coli* T7 Express I^q Competent cells, an enhanced derivative to the BL21(DE3)/pLysS cells, were then transformed with pET-24a+ plasmid containing the ORF encoding the M17A Grx2 protein. The M17A Grx2 protein was then found to be soluble under the same growth conditions for wild-type Grx2 overexpression (Figure 3-2B). The fundamental difference between the strains lies with the location of the T7 RNA polymerase gene. The T7 Express I^q has thegene inserted into the lac operon on the *E. coli* chromosome and is expressed under the control of the lac promoter. Unlike the BL21(DE3)/pLysS cells that that carry the T7 RNA polymerase gene on a lysogenic prophage, the setup in the T7 Express I^q strain provides controlled induction of the polymerase and subsequently, inducible control of transcription of genes downstream of the T7 promoter thereby providing an advantage to the BL21(DE3)/pLysS strain. The soluble fractions, after sonication and centrifugation, containing wild-type and M17A Grx2 proteins were purified by means of DEAE-Sepharose anion exchange chromatography and eluted with a pH gradient (section 2.2.2). The proteins were eluted as single symmetrical peaks (Figure 3-3).

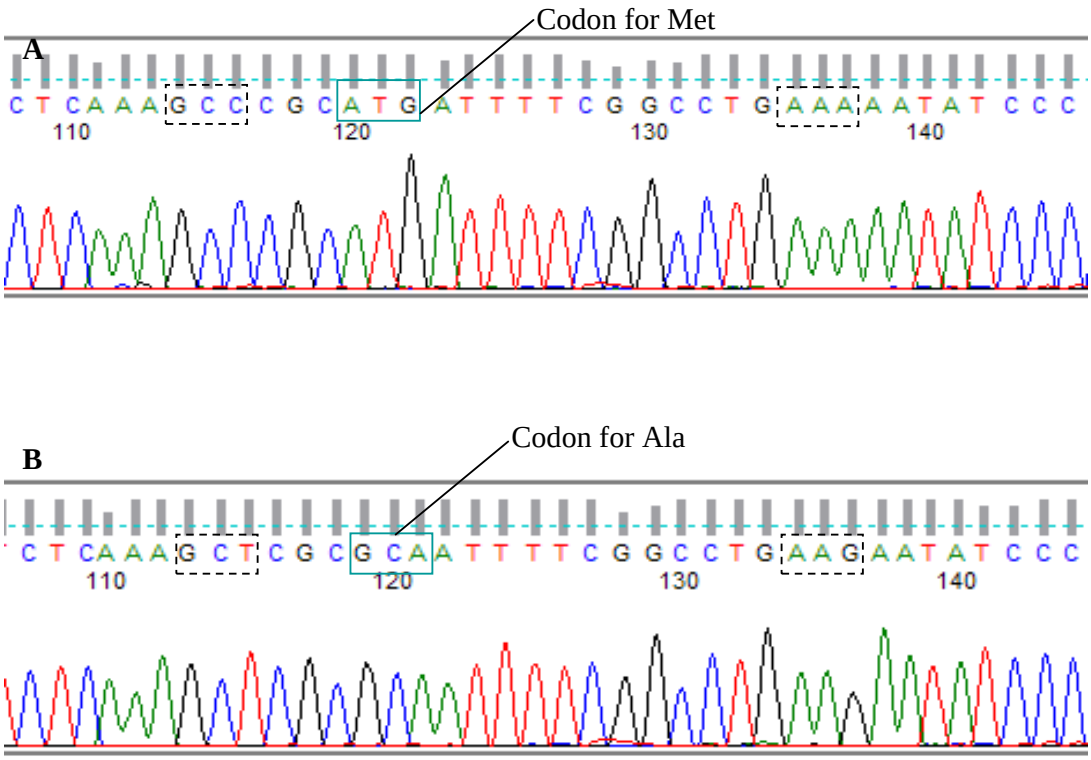


Figure 3-1. Wild-type (A) and M17A Grx2 (B) plasmid sequencing results
A selected segment of the pET24a+ plasmid sequence encoding wild-type and M17A Grx2 proteins (mutated codon boxed). The sequencing results were viewed using the program Finch TV version 1.4.0 (http://www.geospiza.com/FinchTV: Geospiza Inc.). Two silent mutations (dashed box) were incorporated in the primer.

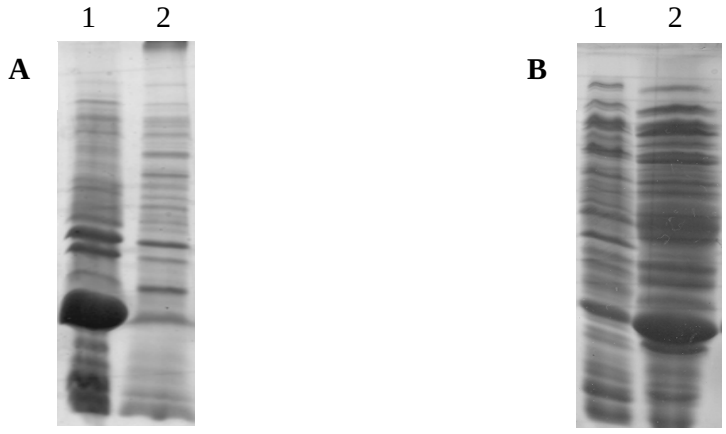


Figure 3-2. SDS-PAGE analysis of M17A Grx2 overexpression.
(A) Expression of the M17A Grx2 protein in *E. coli* BL21(DE3)/pLysS cells. Lane 1: Insoluble fraction (pellet); lane 2: soluble fraction. (B) Expression of the M17A Grx2 protein in *E. coli* T7 Express I^q competent cells. Lane 1: Insoluble fraction (pellet); lane 2: soluble fraction (pellet).

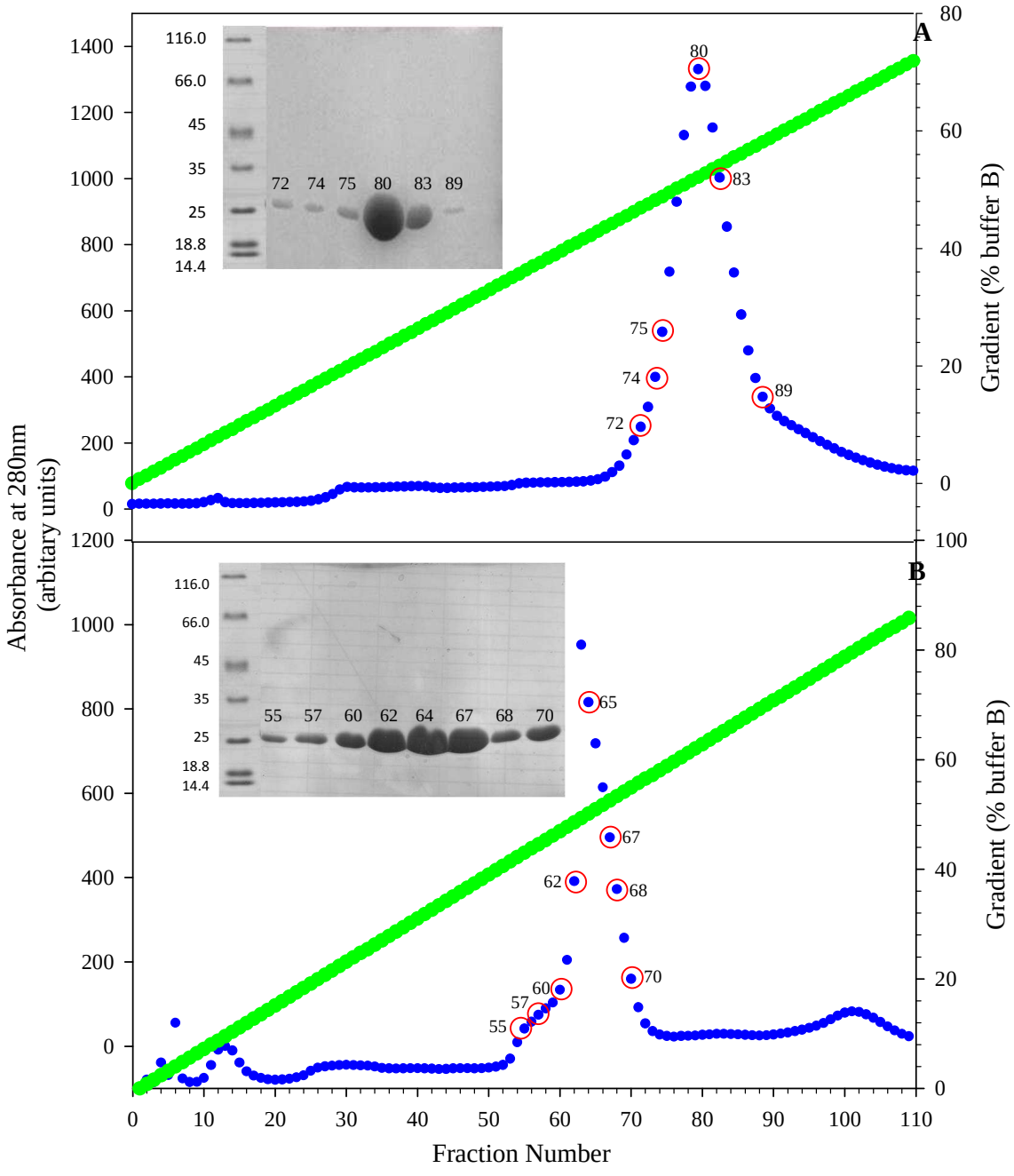


Figure 3-3 Elution profile of Grx2 proteins from purification
The A_{280} of effluent (blue) and the pH gradient (green) were recorded. (A) Elution of wild-type Grx2 and (B) M17A Grx2. SDS-PAGE analysis of selected fractions (numbered and encircled in red) of the peak (inset) with molecular weight marker (sizes in kDa) show that proteins are electrophoretically pure. The profile was obtained from the purification using DEAE ion exchange chromatography.

3.3 Size and purity determination

SDS-PAGE was used to determine the molecular mass and purity of wild-type and M17A Grx2 proteins. The purity of Grx2 was assessed by SDS-PAGE and were judged to be electrophoretically pure as seen by the single bands (Figure 3-4). One litre of culture yielded 140 mg and 100 mg for the wild-type and M17A Grx2 proteins, respectively. The Grx2 proteins were found to have a molecular mass of ~ 25 kDa (Figure 3-4). This was accomplished by comparing the distances the Grx2 proteins migrate in comparison to a set of known standards that migrate under the same denaturing, reducing conditions (Figure 4 inset). This is in agreement with the published data for wild-type Grx2 (Vlamis-Gardikas *et al*., 1997; Xia *et al.*, 2001). The proteins were also found to be homogenous using SE-HPLC under native conditions (Figure 3-5).

3.4 Quantification of free thiol groups

The DTNB assay was employed to quantify the amount of free sulfhydryls present in the protein accessible to solvent (Habeeb, 1972). Native Grx2 (5 μM) bound DTNB, releasing 10 μM of the 2-nitro- 5-thiobenzoate anion (concentration determined spectrophotometrically, section 2.2.8). The results of the DTNB assay indicated that both the cysteine residues of wild-type and M17A Grx2 reacted with the DTNB and that these two free thiol groups are maintained in their reduced state by the presence of 1 mM DTT in the storage buffer. The active site of Grx2 is buried in the interface between the two domains (Xia *et al*., 2001), hence the result of DTNB assay indicates that the mutation did not induce and major structural rearrangements at the active site and that cysteine residue are still accessible to small molecules, consistent with published data (Åslund, *et al*., 1994; Vlamis-Gardikas, *et al*., 1997; Xia, *et al*., 2001).

3.5 Structural Characterisation

3.5.1 Secondary structure characterisation

Studies of the far-UV region (typically 250 – 190 nm) can be used to assess the overall secondary structure content of the protein quantitatively. The circular dichroism spectra for proteins, in this region have characteristic features based on the secondary structure adopted by the proteins (Woody, 1995). Hence, it was used as a probe to assess the secondary structural content of the wild type and M17A Grx2

proteins. The CD spectra of the wild-type and M17A Grx2 proteins in the folded and unfolded conformational states are illustrated in Figure 3-6. There is no change in the overall secondary structure conformation of the wild-type and variant protein. Each protein exhibits minima at 222 nm and 280 nm and a peak at 190 nm, which is typical of proteins predominated by α-helices. This is consistent with the solution structure of Grx2, which reports that the protein is about 56% alpha-helical (Xia *et al*., 2001). The secondary structures of the wild-type and variant proteins were found to be completely disrupted as they unfold in the presence of 7.5 M urea as the distinctive troughs initially observed for the native conformations of each protein were no longer evident. It is evident that the mutations at the interdomain 'lock-and-key' motif interface did not induce gross conformational changes in the secondary structure.

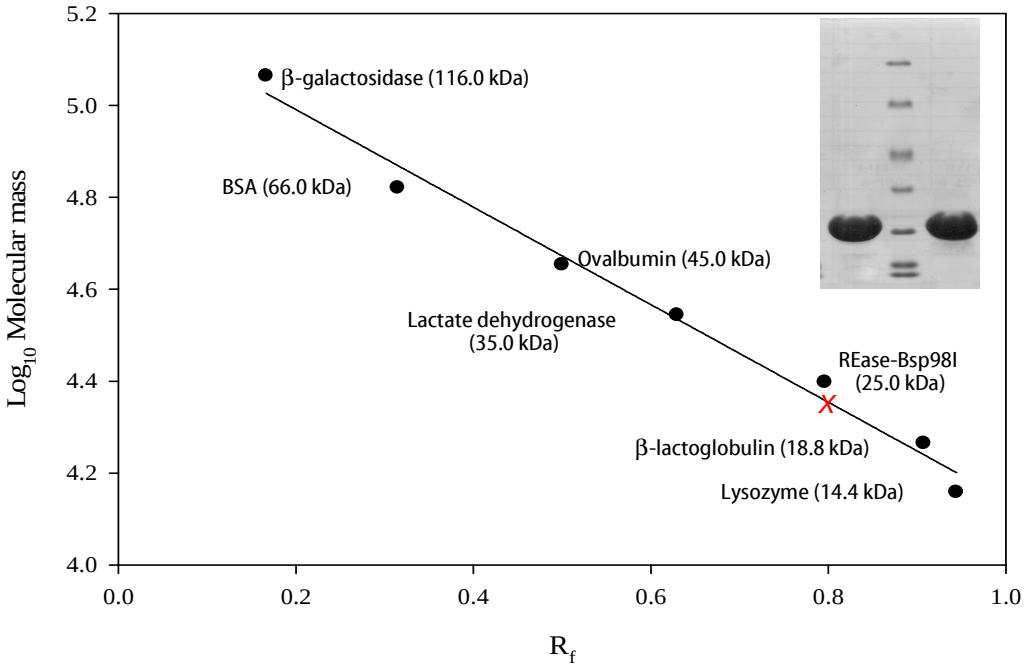


Figure 3-4. SDS-PAGE analysis of wild-type and M17A Grx2
SDS-PAGE gel (inset) and calibration curve for wild-type and M17A Grx2 proteins. The names and sizes of the marker proteins are indicated on the calibration curve. The Grx2 proteins migrated to a distance that corresponds to 25 kDa (indicated with a red cross).

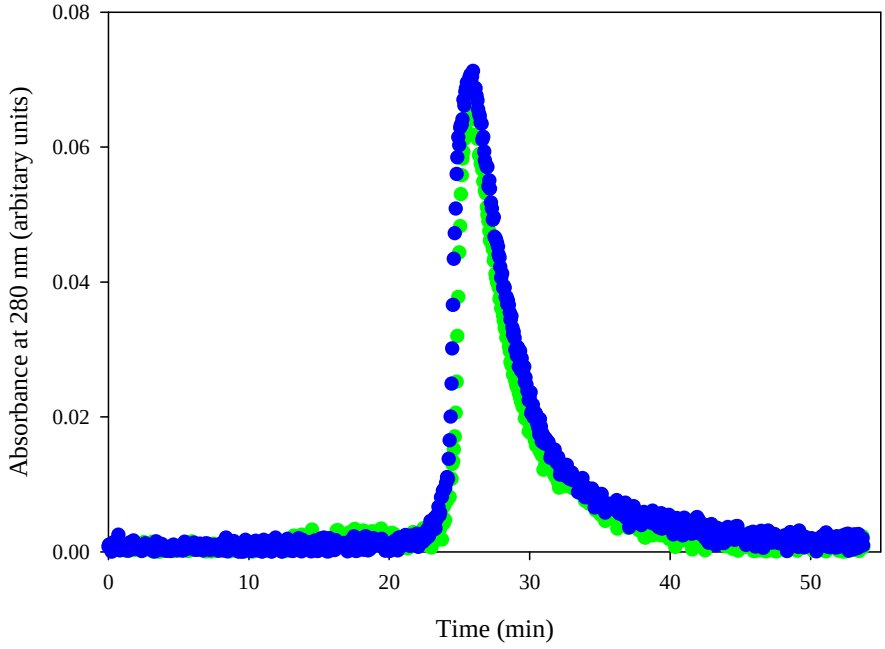


Figure 3-5. SE-HPLC of the Grx2 proteins
The elution profile for wild-type (●) and M17A (●) Grx2 proteins. Proteins were eluted at a flow rate of 0.4ml.min^{-1} at an isocratic pressure of 40bar.

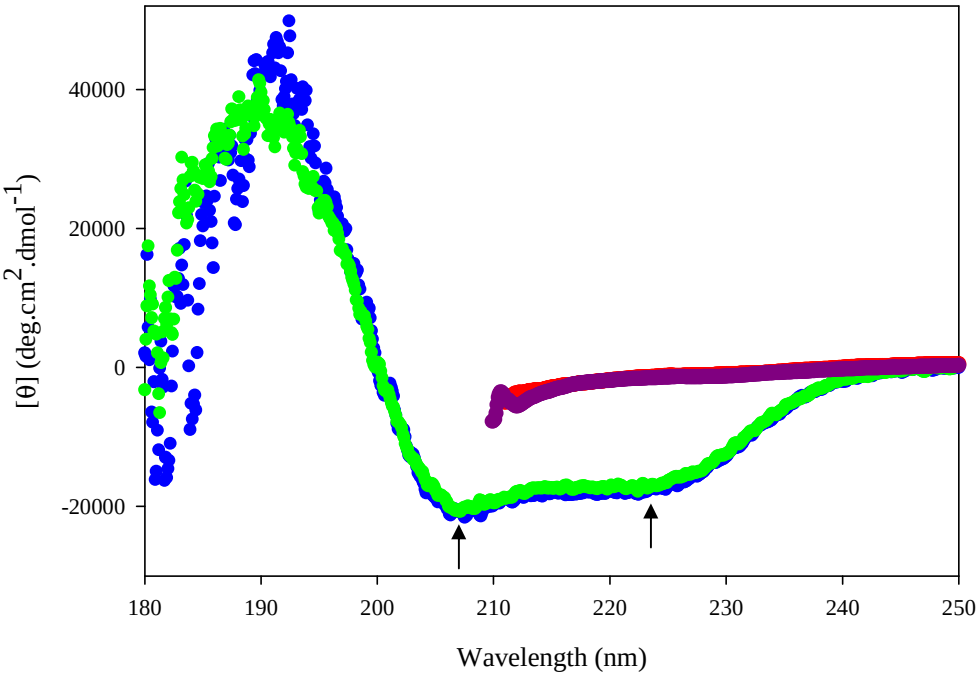


Figure 3-6. Far-UV circular dichroism spectra of Grx2
Spectra shown for the native forms of wild-type Grx2 (•) and M17A Grx2 (•) and for the unfolded (in 7.5M urea) forms for wild-type Grx2 (•) and M17A Grx2 (•). Spectra were collected using 5 μM protein in 5 mM sodium phosphate buffer, pH 7.0, containing 1 mM DTT, 0.02% NaN_3, to minimise the contribution of noise from the salt in the buffer. The minima (indicated by the arrows) are exhibited at 208 nm and 222 nm.

3.5.2 Tertiary structure characterisation

3.5.2.1 Intrinsic fluorescence

Intrinsic tryptophan fluorescence is a useful spectroscopic technique used to measure the local environment of the indole side chain of the Trp residues in the protein. The number as well as the location of this residue within a protein, provides a useful means of probing for local or global conformational changes.

E. coli Grx2 has two tryptophan residues (Trp-89 and Trp190) located in domain 2 (Figure 1-7, Xia *et al.*, 2000). Structurally related GST proteins with buried tryptophan residues display emission maximum for the native proteins at a wavelength of 335 nm (Kaplan *et al*., 1997; Hornby *et al*., 2000; Hornby *et al*., 2002; Luo *et al*., 2002). In the native state of wild-type and M17A Grx2 proteins, the indole ring of the Trp residues are slightly more solvent accessible than its dimeric GST counterparts, displaying emission maximum at 345 nm (Figure 3-7). The absence of any shift in wavelength suggests that the mutations at the interdomain 'lock-and-key' motif did not impact upon the environment of the tryptophan residues in Grx2. The emission maximum for several proteins that contain tryptophan residues displays a red shift to 355 nm when they become denatured (Teale, 1960). This characteristic red shift in the maximum wavelength as well as a decrease in the intensity occurs for wild-type and M17A Grx2 in the presence of denaturing concentrations of urea (Figure 3-7). The indole ring of tryptophan residues (Lakowicz, 1999) is sensitive to changes in the environment as has been show by the changes in emission maxima for native and denatured proteins, hence making intrinsic fluorescence a good probe to monitor local structural changes for Grx2.

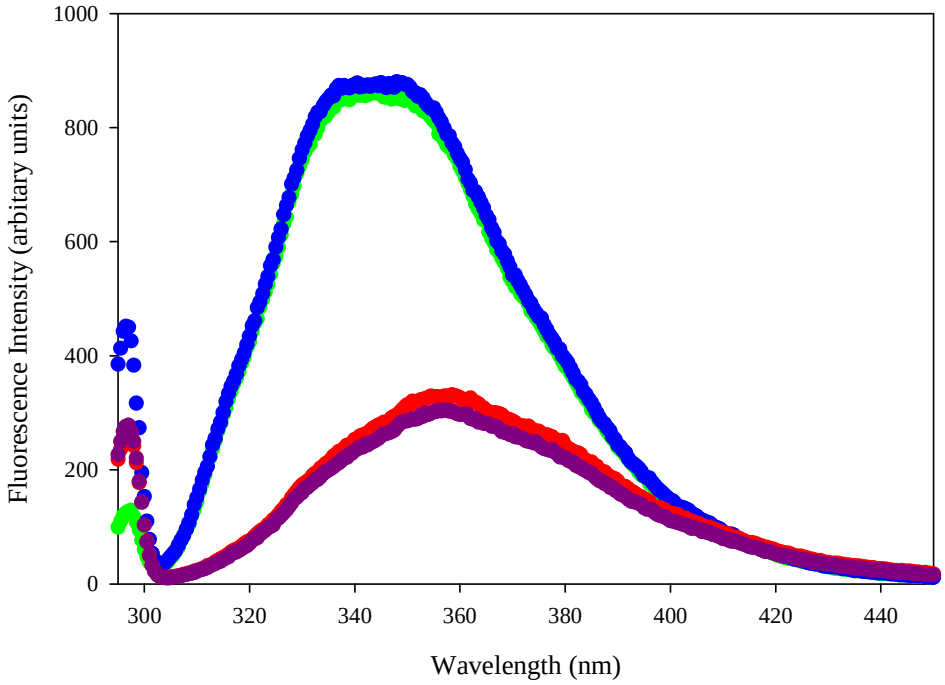


Figure 3-7. Fluorescence spectra of Grx2
Emission spectra shown for the native forms of wild-type Grx2 (●) and M17A Grx2 (●) and for the unfolded (in 7.5 M urea) forms for wild-type Grx2 (●) and M17A Grx2 (●). Spectra were collected using 5 μM protein in Grx2 storage buffer.

3.5.2.2 Near-UV circular dichroism

The near-UV CD of proteins arises from the environments of each aromatic amino acid side chain as well as possible contributions from disulphide bonds, or non-protein cofactors which might absorb in this spectral region. Near-UV CD was the second probe used to investigate the tertiary structure of Grx2 (Figure 3-8). The fine structure in these bands arises from vibronic transitions in which different vibrational levels of the excited state are involved (Strickland, 1974; Kahn, 1979). The spectra of the native proteins exhibit weak positive fine peaks between 255 nm and 270 nm corresponding to the contribution from Phe residues. The stronger negative peak between 274 nm and 282 nm are contributions from Tyr and Trp residues. The negative peak (284 nm - 288 nm) and the more pronounced positive peak (288 nm – 305 nm) is contributed by the Trp residues. Trp residues exhibit a much stronger intensity than the other aromatic amino acids hence the negative peaks observed for Tyr and Trp residues are similar in intensity due to the fact that Grx2 has 8 Tyr residues and only 2 Trp residues. This tertiary structure fingerprint reveals that this mutation induced no major structural change in the native states of Grx2.

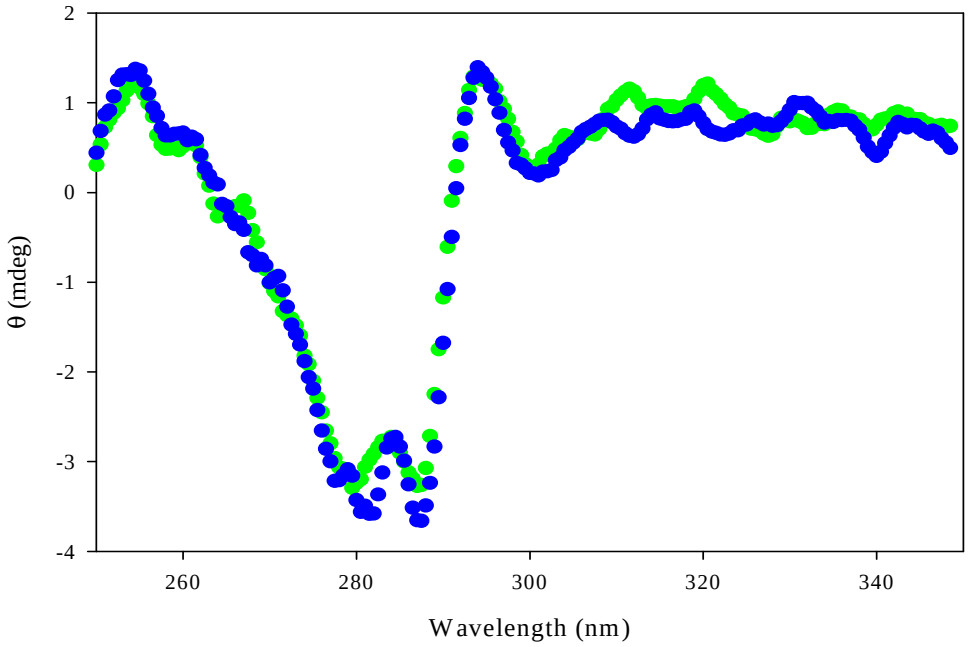


Figure 3-8. Near UV CD spectra of Grx2.
Native CD spectra of wild-type (•) and M17A (•) Grx2 proteins. Spectra were recorded using 40 µM Grx2 protein at 5ºC.

3.6. Conformational stability

3.6.1 Thermal – induced unfolding

Thermal or chemical unfolding is generally used to determine the thermodynamic stability of a protein, which is the Gibbs free energy (ΔG) difference between the folded (G_F) and the unfolded (G_U) states. In thermal unfolding experiments, protein solution is heated at a constant rate, and changes in the protein conformation or their heat effects are monitored by spectroscopy or differential scanning calorimetry (DSC), respectively. Parameters from thermal unfolding studies include the melting temperature (T_m), enthalpy ($\Delta H(T_m)$), and heat capacity increment (ΔC_p), which are used to determine protein stability function ($\Delta G(T)$) (Privalov 1979; Pace *et al.*, 1989). However, in order to obtain these parameters, the key requirement is that protein unfolding is thermodynamically reversible which is usually only valid for chemical denaturation while thermal denaturation is often irreversible. The general root cause of this irreversibility is aggregation of the heat-induced unfolded polypeptide (Benjwal *et al.*, 2005). Thermal unfolding was conducted on the wild-type and M17A Grx2 mutants using far-UV CD as a probe at 222 nm, which is proportional to the α-helical content of a protein. The proteins are shown to unfold in a co-operative manner (Figure 3-9). Wild-type Grx2, which is 56% helical largely unfolds between 58ºC and 64ºC while M17A Grx2 largely unfolds between 54ºC and 60ºC, hence displaying a distinct destabilisation by the mutation. The turbidity (dyanode voltage) recorded in CD experiments can be used as an indicator of aggregation as a protein is heat unfolded (Benjwal *et al.*, 2005). Thermal unfolding of the Grx2 proteins were shown to be irreversible as a result of aggregation shown by the increase in turbidity (Figure 3-9B) at the same temperature at which the respective proteins begin to largely unfold. Heat-induced unfolding of both wild-type and M17A Grx2 is associated with aggregation, hence irreversibility of unfolding, therefore precludes thermodynamic analysis unfolding.

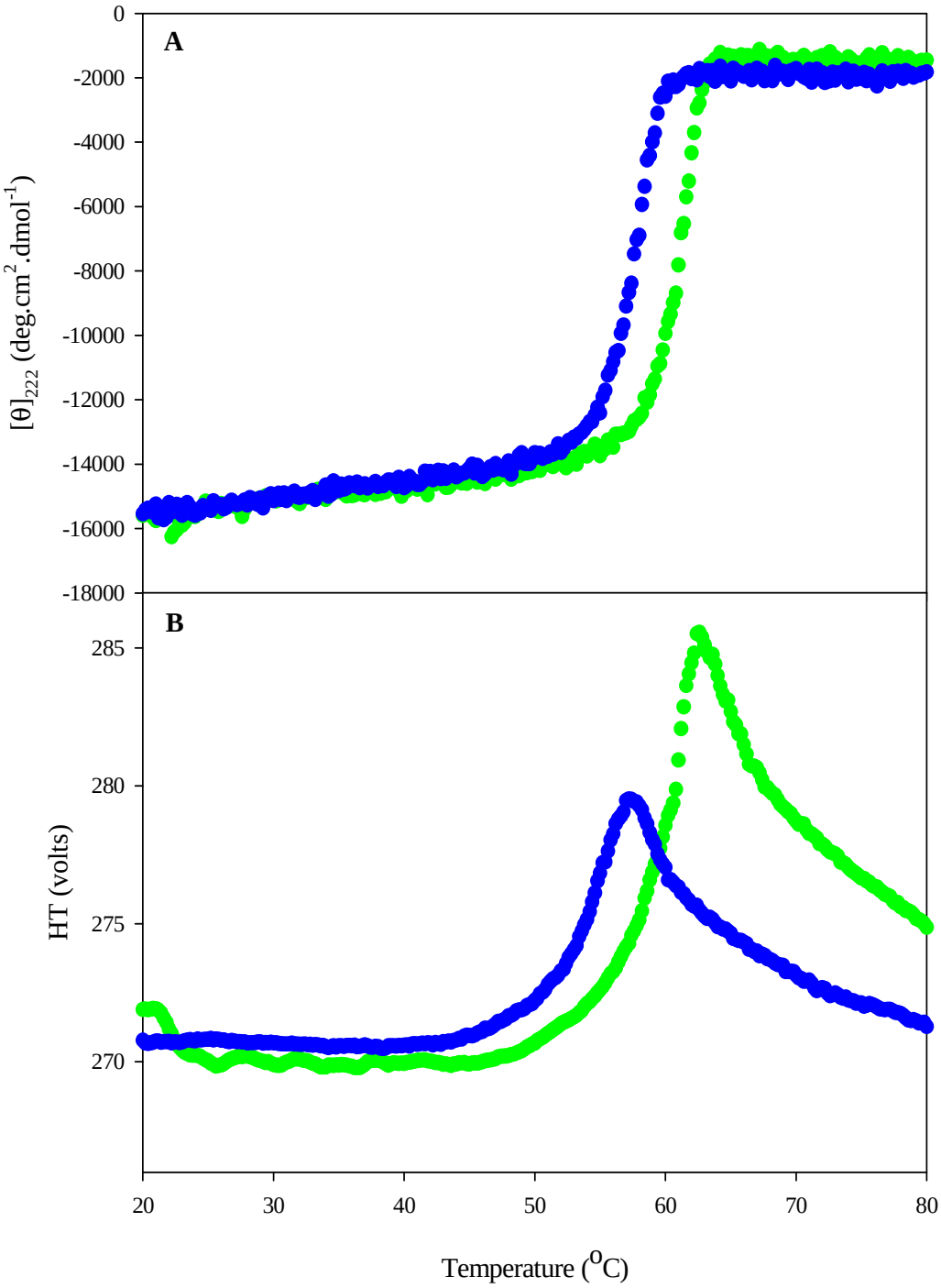


Figure 3-9. Thermal unfolding of Grx2.
(A) Heat unfolding curves of wild-type (•) and M17A (•) Grx2. (B) Turbidity (dynode voltage) plot indicating aggregation followed by precipitation of the aggregates indicated by the reduction in voltage at the temperature where the proteins are largely unfolded.

3.6.2 Reversibility of urea-induced unfolding

Equilibrium unfolding transitions can only be analysed in terms of their thermodynamic parameters if it has been established that the unfolding reaction is reversible and that the native fold can be recovered (Pace, 1986). The recovery of folded Grx2 was established using far-UV CD and intrinsic tryptophan fluorescence as secondary and tertiary probes, respectively (Figure 3-10). The proteins were unfolded in the presence of 7.5 M urea and allowed to refold in Grx2 storage buffer. The ellipticity at 222 nm was chosen as the indicator of recovery of the secondary structure as this is a signature of predominantly helical proteins. Refolded wild-type and M17A Grx2 proteins show 91% (Figure 3-10A) and 90 % (Figure 3-11A) recovery, respectively, for the secondary structure. Intrinsic tryptophan fluorescence at 345 nm (where Grx2 maximally fluoresces, Figure 3-7) shows 92% (Figure 3-10B) and 93% (Figure 3-11B) recovery for wild-type and M17A Grx2 respectively. Thus the equilibrium unfolding transitions of wild-type and M17A Grx2 can be analysed to determine the thermodynamic parameters which will indicate the conformational stability of the Grx2 proteins.

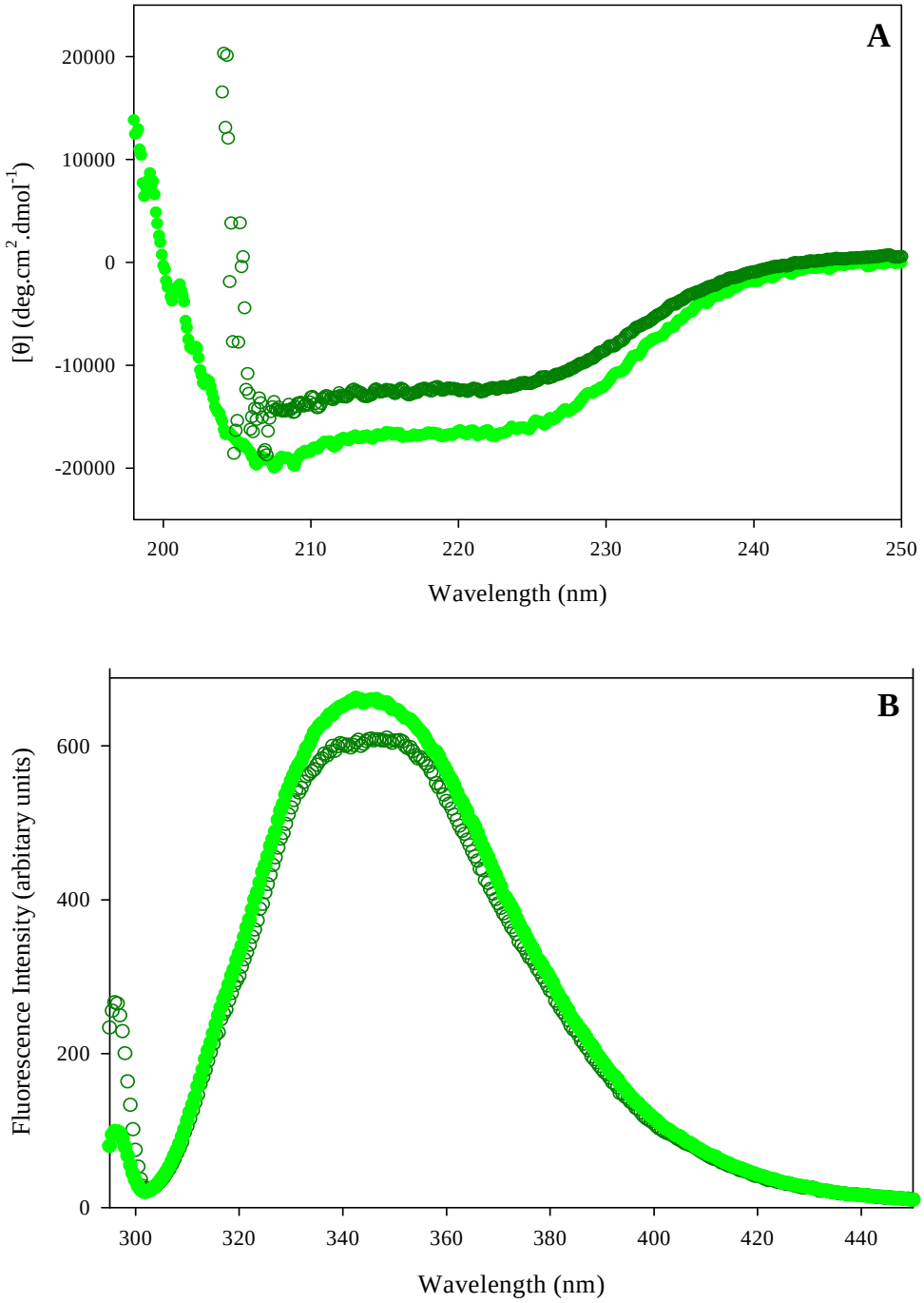


Figure 3-10. Reversibility of unfolding of wild-type Grx2 monitored by CD and fluorescence

(A) Circular Dichroism and (B) fluorescence spectra for the wild-type Grx2 protein in its native (●) and refolded (○) forms. The residual concentration of urea for the refolded form was 1.1 M.

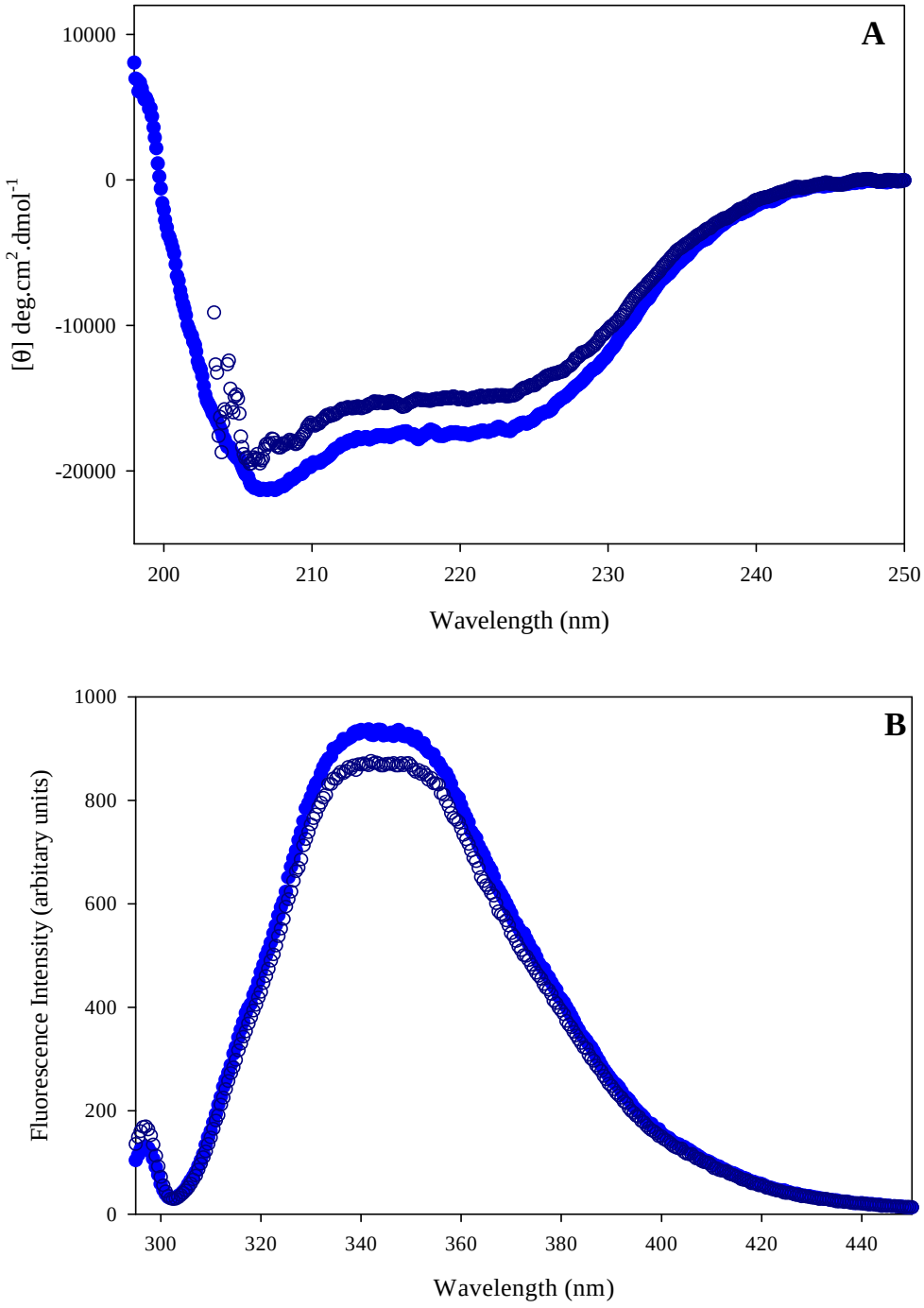


Figure 3-11. Reversibility of unfolding of M17A Grx2 monitored by CD and fluorescence

Circular Dichroism (A) and fluorescence (B) spectra for the M17A Grx2 protein in its native (•) and refolded (○) forms. The residual concentration of urea for the refolded form was 1.1 M.

3.6.3 Urea-induced equilibrium unfolding

Denaturation curves using chemical denaturants are a convenient method for estimating the conformational stability of a protein. Urea was used as the chemical denaturant and the conformational stability was determined by setting up a range (0 M – 8 M urea) of unfolding reactions and which were allowed to reach equilibrium (1 hour). The resulting urea-induced equilibrium-unfolding reactions were monitored using the structural probes far-ultraviolet circular dichroism and intrinsic tryptophan fluorescence. The probes assess the structure of the predominating species present at equilibrium at the respective urea concentration (Pace, 1986). The ellipticities at 222 nm and fluorescence emission at 345 nm for the secondary and tertiary structural probes, respectively, were plotted as a function of urea. The resulting urea-induced unfolding equilibrium curves for both structural probes (Figure 3-12) display single unfolding transitions. The data for both probes were fitted to a two-state model (N ↔ U and the parameters obtained (Table 3-1) indicate a distinct destabilisation due to the M17A mutation. The mutation has caused a decrease in the conformational stability (ΔG (H_2O)) by ~27%, as well as a shift of the midpoint of the unfolding transition C_m to a lower concentration of urea (4.5 M to 3.8 M). There is little change in the slopes (*m*-value) of the equilibrium unfolding transitions (i.e. still parallel to that of the wild-type), indicating that the co-operativity is still maintained. The parameters from the fit show that the *m*-value (slope) is decreased from 3.1 (± 0.22) $kcal.mol^{-1}M^{-1}$ to 2.7 (± 0.18) $kcal.mol^{-1}.M^{-1}$, indicating a decrease in the dependence of the free energy change of unfolding upon denaturant concentration (Pace *et al*., 1989).

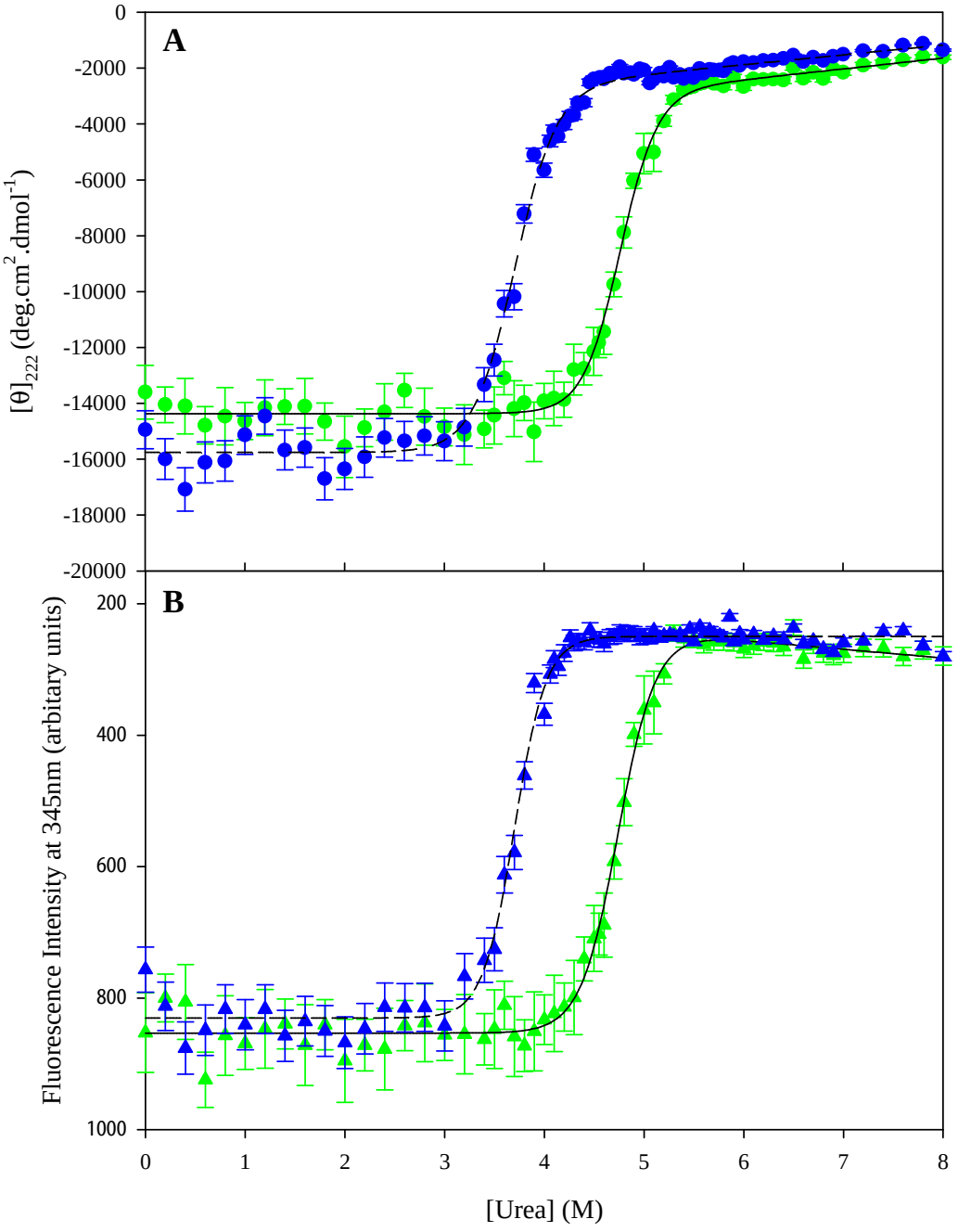


Figure 3-12. Urea-induced equilibrium unfolding of the Grx2 proteins
Unfolding curves for wild-type (green) and M17A Grx2 (blue) proteins obtained using (A) circular dichroism at 222 nm (closed circles) and (B) fluorescence intensity at 345 nm (closed triangles). The data were fitted using a two-state model represented by the solid (–) and broken lines (- -) for wild-type and M17A Grx2 respectively.

Table 3-1. Conformational stability parameters
The values were obtained from fitting the equilibrium unfolding data to a two-state model (N↔U). The parameters obtained are a result of three replicates and the numbers in parentheses represent the standard error.

Probe / *Parameter*	**Far-UV CD**		**Intrinsic Fluorescence**		***Average***	
	wild-type	***M17A***	***wild-type***	***M17A***	***wild-type***	***M17A***
$\Delta G(H_2O)$ (kcal.mol^{-1})	14.0 (± 1.05)	10.2 (±0.65)	14.1 (±1.1)	10.5 (± 0.60)	14.1	**10.3**
***m* (kcal.mol^{-1}.M^{-1} urea)**	3.1 (± 0.22)	2.7 (± 0.17)	3.1 (± 0.2)	2.7 (±0.18)	3.1	**2.7**
C_m (M urea)	**4.5**	**3.8**	**4.5**	**3.9**	**4.5**	**3.8**

3.6.4 Urea-induced equilibrium unfolding in the presence of ANS

The binding of the hydrophobic ligand, ANS, is widely used to monitor the conformational changes induced in proteins during the unfolding process (Semistonov *et al*., 1987; Semistonov *et al*., 1991). ANS is used as a probe for the presence of intermediates during protein unfolding events (Semistonov *et al*., 1991). It has been reported that ANS when bound to the molten globule state shows higher emission intensities than when bound to protein in either the folded or unfolded state (Ptitsyn, 1995). The effect of urea on the binding of ANS was assessed for both wild-type and M17A Grx2 proteins (Figure 3-13). There is negligible binding of ANS to wild-type Grx2 but significant binding to M17A Grx2. ANS was found to maximally fluoresce at 465 nm, and the intensities at this wavelength were plotted as a function of urea (Figure 3-13A). ANS binding is enhanced between 2.5 M and 6 M urea displaying a peak at 3.8 M urea, corresponding to the mid-point of the unfolding transition. ANS binding is then lost as the mutant protein is completely unfolded (5 M - 8 M urea). This enhanced binding in the unfolding transition of the protein may be indicative of the presence of an intermediate or aggregates. Light scattering can be used to detect the presence of aggregates by exciting and emitting at 340 nm. However, these data (Figure 3-13B) showed that of aggregate formation was negligible. However, the absence of a blue shift in the maximum emission wavelength in the unfolding curve (Figure 3-14) suggests that an unstable intermediate cannot be detected under the conditions used.

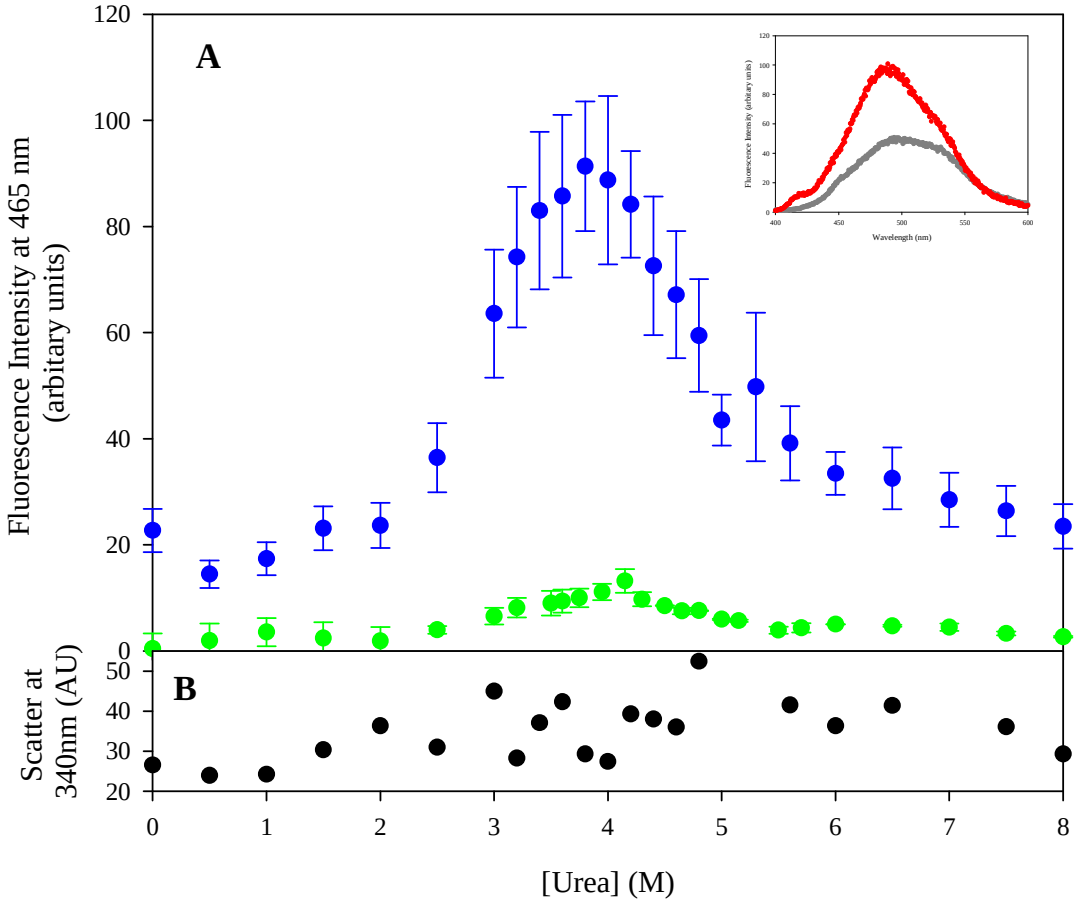


Figure 3-13. Grx2 urea-induced equilibrium unfolding in the presence of ANS
(A) Fluorescence intensity 465 nm of wild-type (•) and M17A (•) Grx2. ANS was added in excess (200 µM) of protein (5 µM). (Inset) ANS bound to M17A Grx2 at 3.8 M (•) displays an enhancement in fluorescent intensity and a hypsochromic shift in the emission spectra relative and free ANS (•). (B) Light scattering data to determine the presence of aggregates by setting the excitation and emission wavelength at 340 nm.

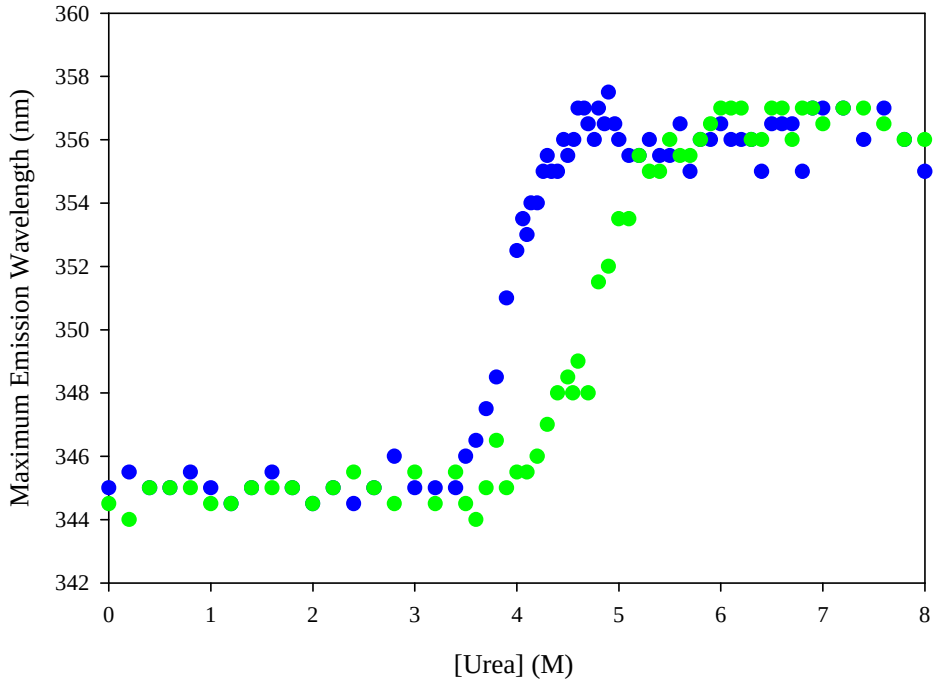


Figure 3-14. Grx2 urea-induced equilibrium unfolding expressed as maximum emission wavelength

The maximum emission wavelength at each urea concentration of the urea-induced equilibrium unfolding of wild-type (●) and M17A (●) Grx2.

3.7 Unfolding kinetics

3.7.1 Single-jump unfolding kinetics

Wild-type and M17A Grx2 are pre-dominantly unfolded in the presence of 5.2M to 8 M urea and 4.8 M to 8 M urea, respectively, as seen from the equilibrium unfolding data (section 3.6.3). These urea concentrations are therefore the choice of monitoring the kinetics of unfolding of both proteins as no refolding should occur.

The rates at which the wild-type and M17A Grx2 proteins unfold were monitored using a stopped-flow reaction analyser. Intrinsic tryptophan fluorescence was used as the structural probe to monitor the unfolding of the proteins. The kinetic traces for the unfolding reaction in 7.5 M urea for both the wild-type and M17A Grx2 proteins are illustrated in Figure 3-15. The kinetic traces were best fit to a double exponential fit as seen by the residual plots, displaying two unfolding phases for both proteins.

The fraction amplitude change for each phase for the wild-type and M17A Grx2 proteins was then plotted against the urea concentration for that unfolding reaction (Figure 3-16). The amplitude data (Figure 3-16A) for the fast phase of the wild-type Grx2 are complex. It decreases from 81% at 5.5 M to 65% at 5.75 M and then increases to 81% at 6 M, thereafter remaining constant to 7.5 M urea. The amplitude data for M17A Grx2 are less complex. The amplitude of the fast phase of unfolding increases from 55% at 4.8 M urea to 81% at 5 M urea, thereafter remaining constant to 7.5 M urea. Similar complex and less complex behaviour are seen for the slow phases of unfolding over the range of urea concentrations for the wild-type and M17A Grx2 proteins, respectively. The amplitude changes from 19% at 5.5 M urea to 35% at 5.75 M urea, then decreases to 19% thereafter and remains constant to 7.5 M for the wild-type Grx2.The amplitude data for the slow phase of unfolding for M17A Grx2 decreases from 44% at 4.8 M urea to 19% at 5.0 M urea thereafter remaining constant to 7.5 M urea.

The wild-type Grx2 protein has time constants of 2.4s and 140s for the fast and slow unfolding phases at 7.5 M urea. Similarly, the M17A Grx2 protein has time constants of 0.83s and 165s. The logarithm of the rates of the fast and slow phases of unfolding for both the wild-type and M17A Grx2 were plotted against urea concentration

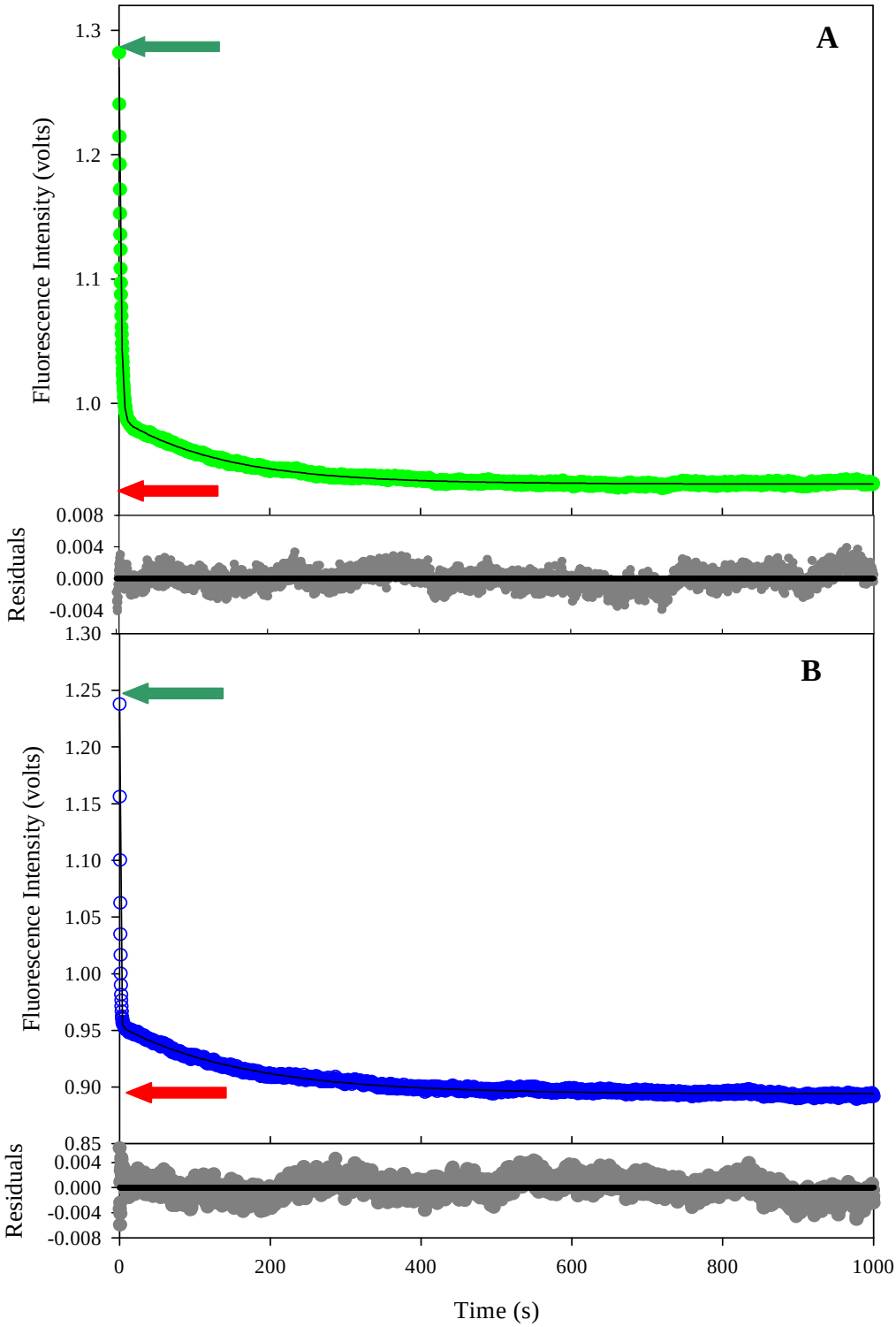


Figure 3-15. Unfolding kinetic traces for the Grx2 proteins
The kinetic traces for (A) wild-type and (B) M17A Grx2 are an average of three unfolding reactions in 7.5 M urea. The excitation was 280 nm with a cut-off filter of 320 nm. Data fitted well to a double exponential as indicated by the residuals plots. The baselines for the native forms are indicated with a green arrow and the baseline for the unfolded form is indicated by the red arrow.

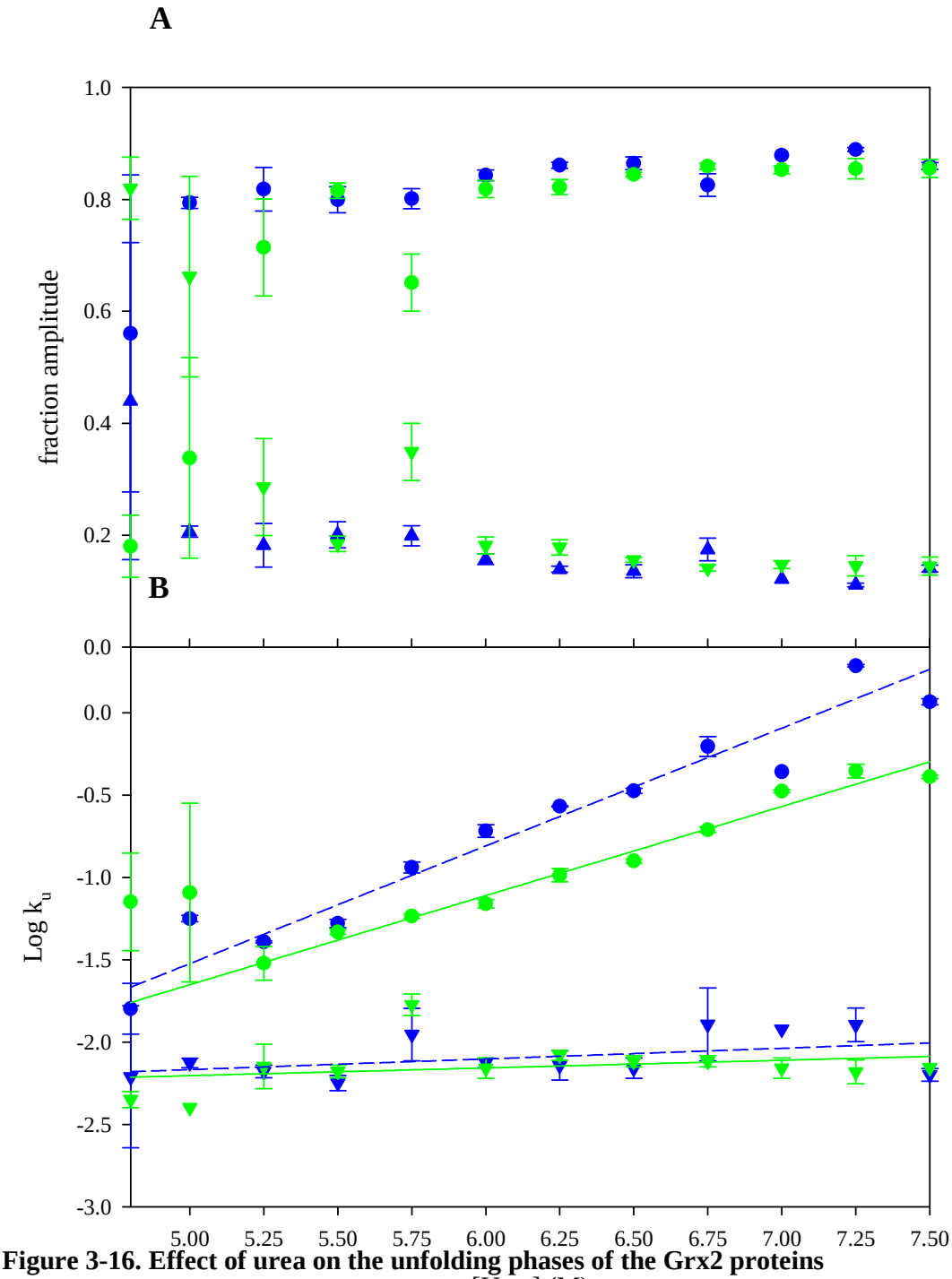


Figure 3-16. Effect of urea on the unfolding phases of the Grx2 proteins
The dependence of (A) amplitude and (B) rates on urea of the unfolding of wild-type (green) and M17A (blue) Grx2. Linear regression analysis was performed on the fast (●●) and slow (▼▼) phases of unfolding. Parameters of these fit values are tabulated in Table 3-2.

Table 3-2. Unfolding kinetic parameters

Linear regression analysis was performed on the logarithm of rates (Figure 3-16B) for both wild-type and M17A Grx2. The regression was performed on the entire data set for the M17A Grx2 (fast: $R^2 = 0.94$; slow: $R^2 = 0.1$) protein but only from 5.25 M to 7.5 M set for the wild-type Grx2 protein (fast: $R^2 = 0.97$; slow: $R^2 = 0.07$). The units for the m_u values were determined by multiplying the values by the product of the gas constant ($R = 8.31$ J k^{-1} mol^{-1}) and the temperature (298 K). The parameters obtained are a result of three replicates and the numbers in parentheses represent the standard error.

Phase / *Parameter*	***Fast***		***Slow***	
	wt Grx2	*M17A Grx2*	*wt Grx2*	***M17A Grx2***
$\boldsymbol{m_u}$ ***(kJ.mol^{-1}.M^{-1} urea)***	1.34 (± 0.18)	1.78 (± 0.033)	0.116 (± 0.03)	**0.161 (±0.035)**
$\boldsymbol{k_u}$ ***(H_2O) (s)***	**4.38×10^{-5} (± 0.1222)**	**7.94×10^{-6} (± 0.2130)**	**3.7×10^{-3} (± 0.002)**	**3.2×10^{-3} (± 0.002)**

(Figure 3-16 B). The Grx2 proteins display high and low dependences of urea for the fast and slow phases, respectively. Linear regression analysis was performed for the logarithm of fast and slow rates. The data were fitted to the following equation (Tanford, 1970):

$$\log k_u = \log k_u(H_2O) + m_u[\text{urea}] \qquad (16)$$

where k_u is the apparent first order rate constant for unfolding. The m_u values (change in solvent accessibility during unfolding) and the apparent rates of unfolding in the absence of water ($k_u(H_2O)$) reported in Table 3-2.

3.7.2 Initial conditions test

The observation of at least two kinetic phases in unfolding could be explained either by a sequential model or by a parallel channel mechanism with two stable native forms that interconvert more slowly than they unfold (Wallace and Matthews, 2002). The initial conditions test relies on the fact that the relaxation times for the reactions depend only on the final conditions while the amplitudes depend on both the final and initial conditions (Tanford, 1970; Hagerman and Baldwin, 1976).

The test for the possibility of two or more stable native conformations involves incubating samples at different starting denaturant concentrations. The Grx2 proteins were incubated in different urea concentrations ranging from 0 M to 6.5 M urea. The unfolding reactions were then initiated by rapidly transferring each sample to identical unfolding conditions (6.5 M urea) using a stopped flow reaction analyser. The resultant unfolding reaction traces best fitted (R^2 = 0.99) to a double exponential hence displaying the presence of two unfolding phases. The rates and amplitudes obtained from the fits are compared to those obtained for the unfolding of Grx2 maintained in the absence of denaturant (without prior incubation in denaturant). The amplitude data for both the fast and slow phases on unfolding for both Grx2 proteins display a sigmoidal behaviour as a function of initial urea concentration (Figure 3-17A), with its midpoints corresponding to that of its equilibrium unfolding curves (Table 3-1). The rates of these unfolding reactions of both phases are shown to be independent on the concentration of urea (Figure 3-17B).

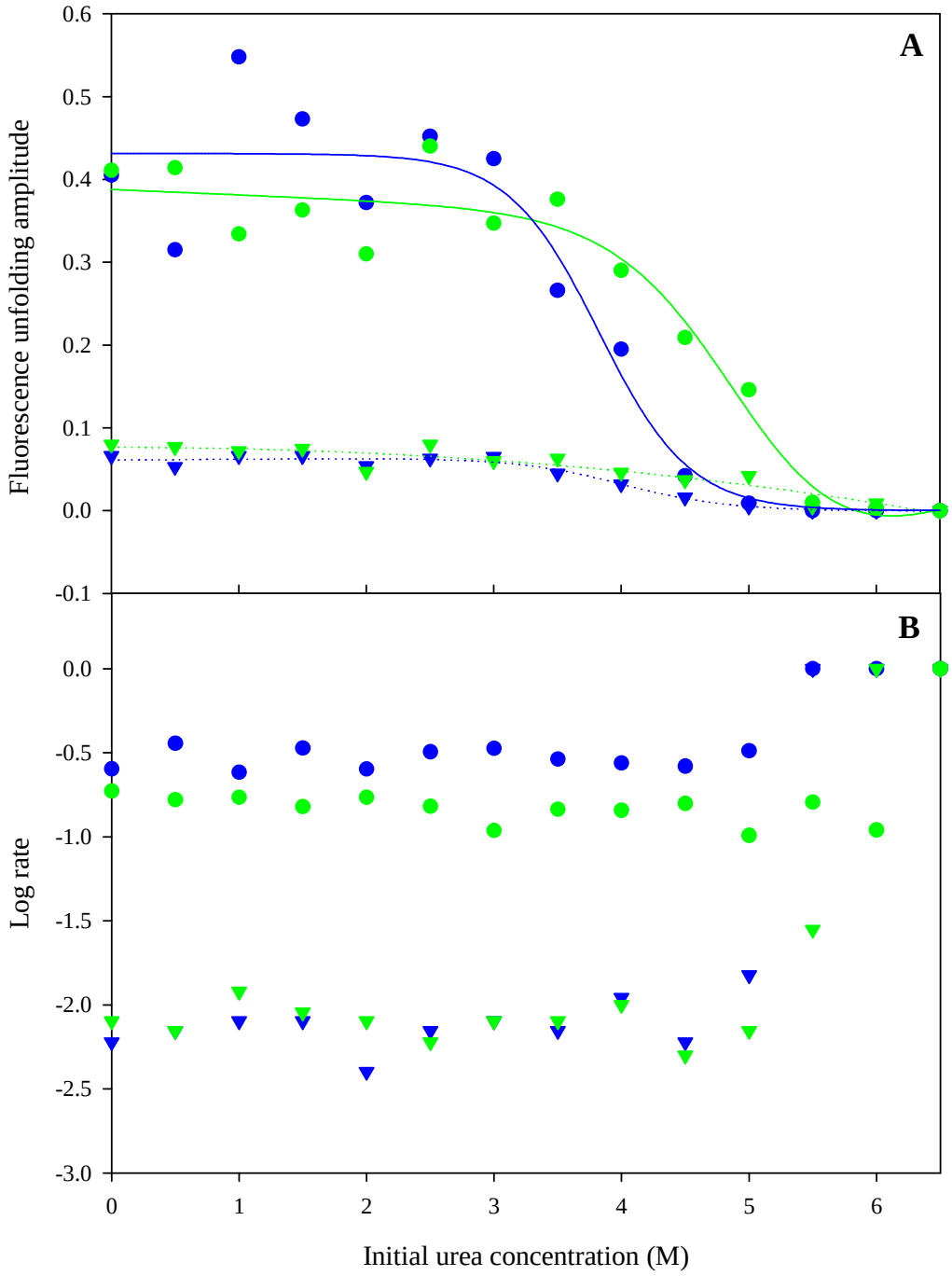


Figure 3-17. Unfolding initial conditions test
(A) The unfolding amplitudes of wild-type (green) and M17A (blue) Grx2 for the fast (closed circles) and slow (closed triangles) phases. (B) The logarithm of rates for the fast (closed circles) and slow (closed triangles) phases of the unfolding reactions. The Grx2 proteins were allowed to unfold in the initial urea concentration for 1 hour, after which the unfolding reaction (to 6.5 M urea 1.25 μM protein-final) monitored. All unfolding reactions were performed at 20ºC.

CHAPTER 4. DISCUSSION

A conserved residue in the GST family of proteins (Figure 1-3) located at the domain interface, Met17, was mutated to Ala. This was the choice of residue to eliminate the bulk of the side chain of Met that is interacting with domain two of Grx2 but maintaining the helical structure of α1, due to the high helical propensity of Ala.

4.1 Role of Met17 in the stability of Grx-2

The purification of the wild-type and mutant forms of the protein was performed under the same conditions which resulted in a yield of 140 mg and 105 mg for the wild-type and mutant, respectively, of pure protein per litre of culture. However, the expression of the M17A Grx2 protein was found to be in inclusion bodies (Figure 3-2A) in the same bacterial expression system as for wild-type Grx2, suggesting that the folding mechanism of the protein has changed as a result of the mutation and were the first signs of destabilisation. An enhanced derivative of the *E. coli* BL21(DE3)/pLysS expression system, the *E. coli* T7 express I^q, was found to render the protein soluble (Figure 3-2B) under the same growth conditions.

The mutation has not caused any major structural alterations (if any) in the native state of the protein as confirmed by the secondary (far-UV CD) (Figure 3-6) and tertiary (Trp fluorescence and near-UV CD) (Figure 3-7 and Figure 3-8) structural characterisation. The stability of the proteins was then assessed by temperature denaturation and urea-induced equilibrium unfolding studies. Temperature-induced denaturation of the Grx2 proteins was irreversible as shown for other GSTs (Kaplan *et al.*, 1997; Dragani *et al.*, 1998; Wallace and Dirr., 1999). However temperature unfolding can serve as an indicator of stability by comparing T_m values (temperature at which the protein is half folded) of variants of a protein. A decrease in T_m value is a sign of destabilisation as a result of a mutation as seen in this case where a decrease of 4ºC is observed for the T_m for M17A Grx2 (Figure 3-9).

The coincidence of the unfolding transition curves of the two structural probes used for monitoring the unfolding for both proteins shows a sigmoidal behaviour and that both the secondary structure and the local environment of the tryptophan residues unfold simultaneously. The thermodynamic parameters of unfolding were then obtained by fitting the data to a two state model. The destabilisation is further observed by a decrease in the conformational stability in the absence of denaturant ($\Delta\Delta G(H_2O)$ ~ 4 kcal.mol^{-1}) (Table 3-

1). This parameter is the free energy change of the unfolding/folding reaction in the absence of denaturant and is in agreement with the range of monomeric proteins (6 – 14 kcal.mol^{-1}) (Neet and Tim, 1994).

As proteins unfold, there is an increase in the solvent accessible surface area (ΔSASA) of the protein, which is the *m*-value (slope) of the unfolding transition (Myers *et al.*, 1995). A decrease is also observed in the *m*-value for M17A Grx2 (3.0 (±0.22) → 2.7 (±0.17) kcal.mol^{-1}.M^{-1}). The transition midpoint (C_m) is the concentration of denaturant where the population of protein is 50% (un)folded. The mutation has caused a shift in the C_m (4.4 M → 3.8 M), i.e. less denaturant is needed to unfolded the protein. The fit to two-state model shows the highly co-operative behaviour in the unfolding of wild-type and M17A Grx2. This is also seen in a mutant of Grx2 in which a tryptophan was engineered into in domain 1 to act as a spectroscopic probe to monitor any tertiary structural changes in this domain was engineered into domain 1 (Gildenhuys *et al.*, 2008) and seen amongst other classes of proteins in the GST family such as Alpha (Wallace *et al.*, 1998; Wallace and Dirr, 1999) and Pi (Erhardt and Dirr, 1995; Wallace and Dirr, 1999).

4.2 Unfolding kinetics of Grx2

Changes in the rates for folding and unfolding can also indicate a change in the stability of a protein where amino acid substitutions have been conducted (Goldenberg, 1992). The mutation caused an increase in the rate of unfolding (Table 3-2) for the fast phase but had no significant effect on the slow phase. The increase in the rate of the fast further indicates a destabilisation in the Grx2 protein.

Although the equilibrium data show that unfolding is two state with a monophasic transition, the kinetics of unfolding display more complex unfolding in which two phases, fast and slow, are observed as determined by the residual plots of the unfolding kinetic fits (Figure 3-16). The amplitudes of the phases display a change in the dependence of denaturant at 5.0 M and 4.8 M for the wild-type and M17A Grx2 protein respectively. Such changes in the unfolding dependence on urea concentration have shown to indicate intermediates in a structurally related protein, Ure2 (Galani *et al.*, 2002) as well as in the unfolding of arc repressor (Jonsson *et al.*, 1996). The unfolding reactions were monitored using fluorescence as a probe therefore the reporters are the tryptophan residues (Trp 89

and Trp 190) which are present only in domain 2, hence changes in the signal will only reveal information about the changes occurring in domain 2 alone. However, a Trp residue was engineered into domain 1 (Y58W) to serve as a reporter for any changes happening in domain 1 (Gildenhuys *et al.*, 2008), and it was found that are no change in the unfolding behaviour of the mutant protein with respect to the wild-type. The changes in unfolding monitored, therefore, can be ascribed to both domains of the protein. Similar changes are witnessed for M17A Grx2, it differs in the change of cross over which is shifted to a lower concentration of urea, consistent with the shift seen in the equilibrium unfolding curves.

The slow phase of unfolding is urea independent (Figure 3-16B) and is a characteristic which has been attributed to the isomerisation (*cis* ↔ *trans*) (Kiefhaber *et al.*, 1992a; Schmid and Baldwin, 1978) of a *cis*- proline peptide bond, Val48-Pro49 (Xia *et al.*, 2001). The slow phase on unfolding for both the wild-type and M17A Grx2 proteins show very close unfolding rate constants of 3.7 s^{-1} and 3.2 s^{-1} which is within range for isomerisation reactions (Brandts *et al.*, 1975). The only proline peptide bond in Grx2 that is in the *cis* conformation is the Val48-Pro49 peptide bond (Xia *et al.*, 2001) and will isomerise to the favoured *trans* conformation in the unfolded state (Brandts *et al.*, 1975; Hinderaker and Raines, 2003). Proteins that contain one or more *cis*-proline in the native state have simple unfolding kinetics, i.e. a single unfolding phase, such as *E. coli* thioredoxin (Kelley and Stellwagen, 1984), a member of the same superfamily as the GSTs. The slow phase of unfolding that relate to the *cis* ↔ *trans* isomerisation of the proline peptide bond, can be detected for urea concentrations close to the transition region of the unfolded state and then no longer detectable higher urea concentrations of unfolding (Brandts *et al.*, 1975; Kiefhaber *et al.*, 1992a; Kiefhaber and Schmid, 1992). The slow phase for the unfolding of both Grx2 proteins, however, is present constantly from 5.5 M and 5 M urea to 7.5 M for the wild-type and M17A Grx2 proteins, respectively, contributing to 20% of the amplitude. The complex dependence of amplitude on urea concentration near the transition maybe consistent with conformational folding events coupled to the *cis*-proline peptide bond isomerisation events (Kiefhaber *et al.*, 1992a), M17A Grx2 displays the same behaviour but the change from high to low amplitude with increasing denaturant concentration is shifted to a lower urea concentration, consistent with the lowered C_m. Similar complex dependencies were seen for the slow phase of unfolding for staphylococcal nuclease and it was also proposed to involve a proline isomerisation (Kuwajima *et al.*, 1991) but later it was shown that the curvature in the dependence of the rates of unfolding for this phase were

due to the presence of an intermediate (Walkenhorst *et al.*, 1997), hence this suggests the presence of intermediates during the unfolding of the Grx2 protein.

The m_u values reported in Table 3-2 can be used to indicate solvent exposure during unfolding (Doyle *et al.*, 1996). The *m*u values reported in Table 2 for the wild-type and M17A Grx2 proteins indicate that the fast phase involves structural rearrangements that expose large amounts of surface area while the slow phase involves structural rearrangements that expose small amounts of surface area. The mutant protein however displays a larger m_u value, indicating a larger amount of surface area exposed upon unfolding, consistent with the destabilisation as the interdomain interaction is disrupted, resulting in more solvent exposure than the wild-type under the same conditions.

The co-existing unfolding species was first demonstrated by Garel and Baldwin (1973). Two unfolding phases are observed for both the Grx2 proteins; the possibility of them occurring either sequentially or in parallel was investigated by performing initial conditions tests (Wallace and Matthews, 2002). The resulting unfolding reactions are then compared to the unfolding of the sample maintained in the absence of denaturant. This comparison provides important information: (i) equivalent rate constants insure that the system is behaving as expected, and (ii) the relative amplitudes provide a direct measure of the fraction of molecules that were present at the start of the reaction. The urea dependence of these amplitudes can serve to discriminate between sequential and parallel mechanisms as the amplitudes of the kinetic phases are proportional to the population of the species involved (Wallace and Matthews, 2002). The dependence of the rate constants and amplitudes of the unfolding reaction on the initial urea concentrations was determined by relying on the fact that rate constants depend only on the final conditions while amplitudes depend on both the initial and final conditions (Wallace and Matthews, 2002). The amplitude data (Figure 16A) indicate that the unfolding is biphasic at all initial urea concentrations while the rate constants were independent of the initial concentration of urea (Figure 3-17B). The relative amplitudes of the unfolding phases represent the fraction of molecules present at the beginning of the reaction (Hagerman and Baldwin, 1976) and the simultaneous presence of two unfolding phases indicates that the two unfolding reactions occur independently and, thus, run in parallel. Double-jump refolding experiments performed on the wild-type protein confirmed a parallel unfolding mechanism (Gildenhuys

et al., 2008). The independent parallel unfolding phases have midpoints similar to that witnessed in the equilibrium transition.

4.3 Domain architecture (Grx2 vs GSTs)

Grx2 has fewer interdomain contacts (Figure 4-1) in contrast to its monogenic homologue, Clic1 (Harrop *et al.*, 2001), which also has the topologically equivalent methionine hydrophobic contact. In addition to the hydrophobic interaction, salt-bridges and more interdomain hydrogen bonds are present at the domain interface. The mutation therefore, will have a greater impact on the soluble form of Clic1 than it does on Grx2 as more interactions are impacted on. This seen by the equilibrium unfolding studies in the case of M32A Clic1, where the unfolding transitions become biphasic as a result of the mutation and an intermediate becomes populated with the loss of the co-operative behaviour (Stoychev PhD Thesis, 2008). This is not seen in the case of Grx2, where a single unfolding transition is still observed, however there is a reduction in the *m*-value (a decrease in the dependence of the free energy change of unfolding upon denaturant concentration), suggesting the presence or accumulation of an intermediate. The use of an extrinsic probe in equilibrium unfolding, ANS, reveals that an intermediate maybe populating as a result of the mutation. With the absence bulk of the side chain, this hydrophobic interaction is now weakened as a result of cavity forming. This may be the region in which the extrinsic dye is now able to bind as a result of the enhancement of the ANS fluorescence.

The effect of the M17A destabilisation can be explained by the fact that the mutation results in the removal of the R-group ($>CH_2$-S-CH_3) of methionine that leads to a loss of hydrophobic contacts (Figure 4-1) and a decrease in $\Delta G(H_2O)$ of approximately ± 1.1 kcal.mol^{-1} per methyl (CH_2) group (Kellis, 1988; Pace, 2001). This mutation is also a cavity causing substitution which results in a less tight packing environment at the domain interface due to the loss of short-range van der Waals contacts. The loss in stability is probably greater due to the cooperative contribution of hot spot residues (Stoychev, PhD Thesis, 2008), such as Met32, toward protein stability. Wallace *et al.* (2000) showed that the mutation of this topological equivalent of this 'lock-and-key' motif at the domain interface of the dimeric homologue, GST, also causes a destabilisation, with accumulation of a dimeric intermediate. The unfolding transition of GST also becomes biphasic as a result of the loss of this hydrophobic contact.

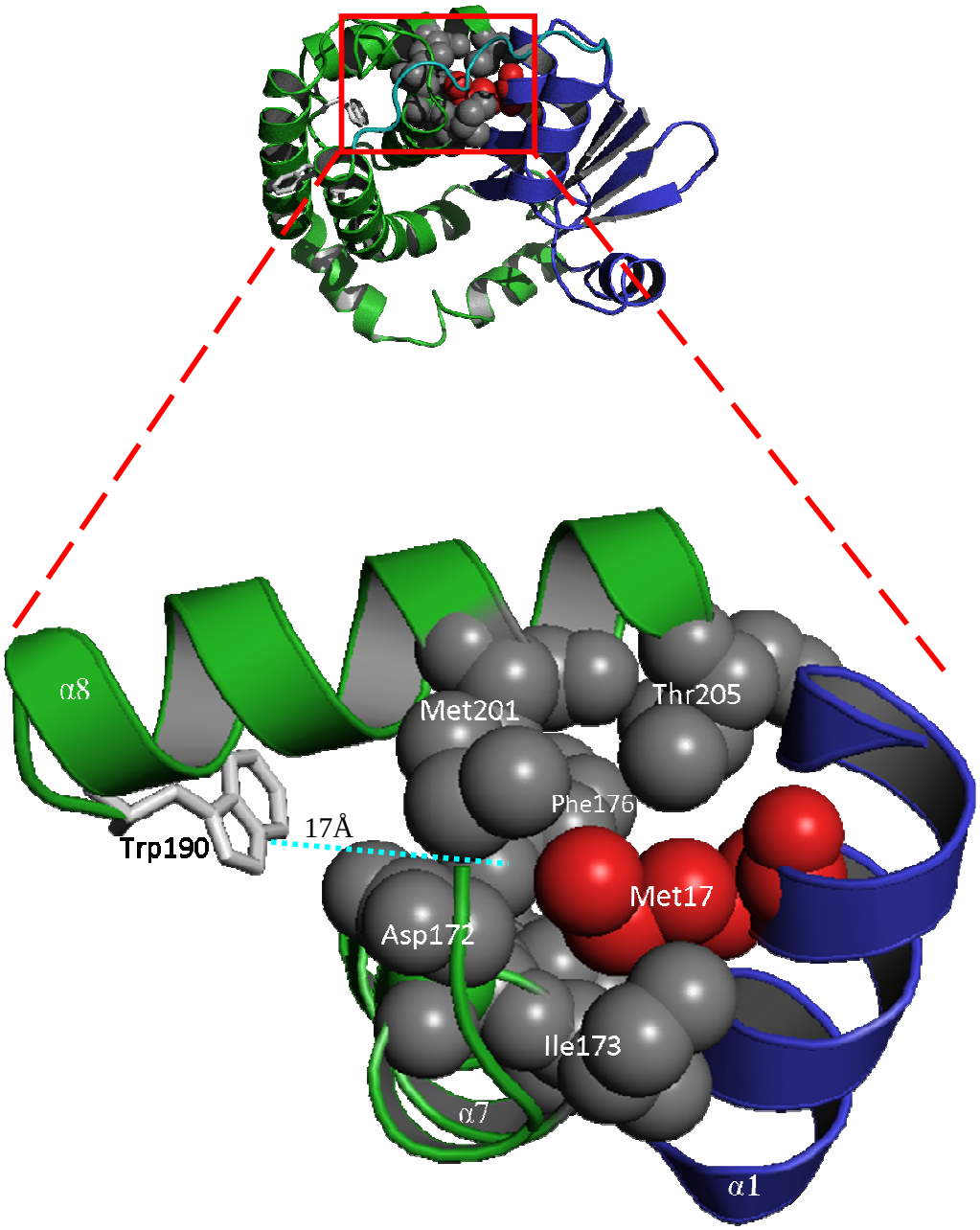


Figure 4-1. Interdomain 'lock–and-key' motif in Grx2
Key residue, Met 17 (red), located in α1 of domain 1 (blue). Residues forming the lock (grey) in domain 2 (green): Asp172, Ile173, Phe176 in α7 and Met201 and Thr 205 of α8, Residues of the 'lock-and-key' are illustrated in their van der Waals radii. The second Trp residue (Trp 190) is located at the N-terminus of α8 (17Å from the domain interface) and shown in its stick form. Image generated using PyMOL, v0.99 (DeLano Scientific).

4.4 Molten globule intermediate or hydrophobic region exposed?

The aromatic chromophore ANS, is weakly fluorescent in water, but a hypsochromic shift in its spectrum and its intensity is dramatically increased in nonpolar solvents or when it binds to nonpolar sites of proteins. Strong binding of ANS to molten globule states of proteins has been linked to the loss of tertiary structure (Semisotonov *et al.*, 1991), and has been applied to characterise transient states in protein denaturation (Das *et al.*, 1995; Guha and Bhattacharyya, 1995; Uversky *et al.*, 1996; Bismuto, 2001; Fanucchi *et al.*, 2008). There is negligible binding of ANS in the urea-induced equilibrium unfolding of wild-type Grx2 (Figure 3-13) indicating an absence of a stable intermediate. In contrast, M17A Grx2 displays significant binding of the hydrophobic dye under the same urea-induced equilibrium conditions, suggesting an intermediate formation. This is further supported by the reduced *m*-value obtain for equilibrium unfolding. This is also seen with the equivalent mutation in its monomeric counterpart, M32A Clic1 at pH 7.0 (Stoychev, PhD Thesis, 2008). As for Grx2, Met32 of Clic1 located in α1 of domain 1, but this region is the putative trans-membrane region in domain 1 for membrane insertion. The domain interface of Grx2 houses the active site of the protein (Xia, *et al.*, 2001) where it catalyses the reduction of disulphides (Holmgren and Åslund, 1995), making this helix relatively more flexible to accommodate its substrate. In catalysis, the movement of the domains often excludes water from the active site and helps position catalytic groups around the substrate. It also traps substrates and prevents the escape of reaction intermediates (Anderson *et al.*, 1979; Knowles, 1991). The loss of the 'key' from the 'lock' as result of the mutation, may expose this hydrophobic region during unfolding, thus allowing ANS to bind. However, a molten globule intermediate is questionable, as its formation is concomitant with a hypsochromic shift in the tryptophan fluorescence upon unfolding (Semistonov *et al.*, 1991) and this behaviour is absent in the unfolding of M17A Grx2 (Figure 3-14). A decrease in co-operativity (reduction in *m*-value) usually implies an intermediate formation, but resolution of spectroscopic techniques that monitor the intrinsic probes as reporters of unfolding may not be sufficient to detect these intermediates (Soulages, 1998).

4.5 Domain co-operativity in Grx2

The goal of this study was to knockout an interaction possibly involved in maintaining the co-operative folding of the domains. Equilibrium unfolding has displayed a decrease in the stability of the protein highlighting the role of the 'key' in the 'lock' of the domain interface. However, with still monophasic transitions observed, the co-operativity is still maintained between the domains, where biphasic transitions are expected in the denaturant-induced equilibrium unfolding as well as a formation of a stable intermediate (Batey *et al*, 2005). A change would also be detected in the unfolding kinetics of the protein if an intermediate were present which no change was observed other than the shift consistent with the destabilisation, therefore the kinetics data are in agreement with the equilibrium unfolding data in displaying the extensive co-operativity of the domains in Grx2.

CHAPTER 5. REFERENCES

Aloy, P., Ceulemans, H., Stark, A. and Russell, R. B. (2003). The relationship between sequence and interaction divergence in proteins. *J Mol Biol* **332**, 989-98.

Alves, C. S., Kuhnert, D. C., Sayed, Y. and Dirr, H. W. (2006). The intersubunit lock-and-key motif in human glutathione transferase A1-1: role of the key residues Met51 and Phe52 in function and dimer stability. *Biochem J* **393**, 523-8.

Anderson, M. W., Crutchley, D. J., Chaudhari, A., Wilson, A. G. and Eling, T. E. (1979). Studies on the covalent binding of an intermediate(s) in prostaglandin biosynthesis to tissue macromolecules. *Biochim Biophys Acta* **573**, 40-50.

Andujar-Sanchez, M., Clemente-Jimenez, J. M., Rodriguez-Vico, F., Las Heras-Vazquez, F. J., Jara-Perez, V. and Camara-Artigas, A. (2004). A monomer form of the glutathione S-transferase Y7F mutant from Schistosoma japonicum at acidic pH. *Biochem Biophys Res Commun* **314**, 6-10.

Anfinsen, C. B. (1973). Principles that govern the folding of protein chains. *Science* **181**, 223-30.

Armon, A., Graur, D. and Ben-Tal, N. (2001). ConSurf: an algorithmic tool for the identification of functional regions in proteins by surface mapping of phylogenetic information. *J Mol Biol* **307**, 447-63.

Armstrong, R. N. (1997). Structure, catalytic mechanism, and evolution of the glutathione transferases. *Chem Res Toxicol* **10**, 2-18.

Åslund, F., Ehn, B., Miranda-Vizuete, A., Pueyo, C. and Holmgren, A. (1994). Two additional glutaredoxins exist in *E. coli*: glutaredoxin 3 is a hydrogen donor for ribonucleotide reductase in a thioredoxin/glutaredoxin 1 double mutant. *Proc Natl Acad Sci U S A* **91**, 9813-7.

Bahadur, R. P., Chakrabarti, P., Rodier, F. and Janin, J. (2004). A dissection of specific and non-specific protein-protein interfaces. *J Mol Biol* **336**, 943-55.

Baker, E. N. and Hubbard, R. E. (1984). Hydrogen bonding in globular proteins. *Prog Biophys Mol Biol* **44**, 97-179.

Barrick, D. and Baldwin, R. L. (1993). Three-state analysis of sperm whale apomyoglobin folding. *Biochemistry* **32**, 3790-6.

Bashton, M. and Chothia, C. (2002). The geometry of domain combination in proteins. *J Mol Biol* **315**, 927-39.

Batey, S., Randles, L. G., Steward, A. and Clarke, J. (2005). Cooperative folding in a multi-domain protein. *J Mol Biol* **349**, 1045-59.

Batey, S., Scott, K. A. and Clarke, J. (2006). Complex folding kinetics of a multidomain protein. *Biophys J* **90**, 2120-30.

Bennett, M. J., Schlunegger, M. P. and Eisenberg, D. (1995). 3D domain swapping: a mechanism for oligomer assembly. *Protein Sci* **4**, 2455-68.

Benjwal, S., Verma, S., Rohm, K. H. and Gursky, O. (2006). Monitoring protein aggregation during thermal unfolding in circular dichroism experiments. *Protein Sci* **15**, 635-9.

Bismuto, E., Gratton, E. and Lamb, D. C. (2001). Dynamics of ANS binding to tuna apomyoglobin measured with fluorescence correlation spectroscopy. *Biophys J* **81**, 3510-21.

Board, P. G., Coggan, M., Chelvanayagam, G., Easteal, S., Jermiin, L. S., Schulte, G. K., Danley, D. E., Hoth, L. R., Griffor, M. C., Kamath, A. V., Rosner, M. H., Chrunyk, B. A., Perregaux, D. E., Gabel, C. A., Geoghegan, K. F. and Pandit, J. (2000). Identification, characterisation, and crystal structure of the Omega class glutathione transferases. *J Biol Chem* **275**, 24798-806.

Braman, J., Papworth, C. and Greener, A. (1996). Site-directed mutagenesis using double-stranded plasmid DNA templates. *Methods Mol Biol* **57**, 31-44.

Brandts, J. F., Halvorson, H. R. and Brennan, M. (1975). Consideration of the Possibility that the slow step in protein denaturation reactions is due to cis-trans isomerism of proline residues. *Biochemistry* **14**, 4953-63.

Brooks, L. C., Gruebelle, M., Onuchic, N. J., and Wolynes, G.P. (1998). Chemical physics of protein folding. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 11037-8.

Bychkova, V. E. and Ptitsyn, O. B. (1995). Folding intermediates are involved in genetic diseases? *FEBS Lett* **359**, 6-8.

Caffrey, D. R., Somaroo, S., Hughes, J. D., Mintseris, J. and Huang, E. S. (2004). Are protein-protein interfaces more conserved in sequence than the rest of the protein surface? *Protein Sci* **13**, 190-202.

Calamai, M., Chiti, F. and Dobson, C. M. (2005). Amyloid fibril formation can proceed from different conformations of a partially unfolded protein. *Biophys J* **89**, 4201-10.

Chaffotte, A. F., Cadieux, C., Guillou, Y. and Goldberg, M. E. (1992). A possible initial folding intermediate: the C-terminal proteolytic domain of tryptophan synthase beta chains folds in less than 4 milliseconds into a condensed state with non-native-like secondary structure. *Biochemistry* **31**, 4303-8.

Chen J, Anderson JB, DeWeese-Scott C, Fedorova ND, Geer LY, He S, Hurwitz DI, Jackson JD, Jacobs AR, Lanczycki CJ, Liebert CA, Liu C, Madej T, Marchler-Bauer A, Marchler GH, Mazumder R, Nikolskaya AN, Rao BS, Panchenko AR, Shoemaker BA, Simonyan V, Song JS, Thiessen PA, Vasudevan S, Wang Y, Yamashita RA, Yin JJ, Bryant SH. (2003). MMDB: Entrez's 3D-structure database. *Nucleic Acids Res*. **31**, 474-7.

Chung, C. T., Niemela, S. L. and Miller, R. H. (1989). One-step preparation of competent *E. coli*: transformation and storage of bacterial cells in the same solution. *Proc Natl Acad Sci U S A* **86**, 2172-5.

Cocco, M. J., Kao, Y. H., Phillips, A. T. and Lecomte, J. T. (1992). Structural comparison of apomyoglobin and metaquomyoglobin: pH titration of histidines by NMR spectroscopy. *Biochemistry* **31**, 6481-91.

Codreanu, S. G., Thompson, L. C., Hachey, D. L., Dirr, H. W. and Armstrong, R. N. (2005). Influence of the dimer interface on glutathione transferase structure and dynamics revealed by amide H/D exchange mass spectrometry. *Biochemistry* **44**, 10605-12.

Creighton, T. E. (1990). Protein folding. *Biochem J* **270**, 1-16.

Creighton, T. E. (1991). Characterising intermediates in protein folding. *Curr Biol* **1**, 8-10.

Das, B. K., Bhattacharyya, T. and Roy, S. (1995). Characterization of a urea induced molten globule intermediate state of glutaminyl-tRNA synthetase from *E. coli*. *Biochemistry* **34**, 5242-7.

Dauber, P., Hagler, A.T. (1980). Crystal packing, hydrogen bonding, and the effect of crystal forces on molecular conformation. *Accts. Chem. Res.* **13**, 105−112.

DeLano, W. L. (2002) The PyMOL Molecular Graphics System, DeLano Scientific, LLC, San Carlos, CA.

Dill, K. A. (1985). Theory for the folding and stability of globular proteins. *Biochemistry* **24**, 1501-9.

Dill, K. A. (1990). Dominant forces in protein folding. *Biochemistry* **29**, 7133-55.

Dill, K. A., Alonso, D. O. and Hutchinson, K. (1989). Thermal stabilities of globular proteins. *Biochemistry* **28**, 5439-49.

Dill, K. A. and Chan, H. S. (1997). From Levinthal to pathways to funnels. *Nat Struct Biol* **4**, 10-9.

Dill, K. A., Ozkan, S. B., Shell, M. S. and Weikl, T. R. (2008). The protein folding problem. *Annu Rev Biophys* **37**, 289-316.

Dill, K. A. and Shortle, D. (1991). Denatured states of proteins. *Annu Rev Biochem* **60**, 795-825.

Dirr, H. W. and Reinemer, P. (1991). Equilibrium unfolding of class pi glutathione S-transferase. *Biochem Biophys Res Commun* **180**, 294-300.

Dobson, C. M. and Karplus, M. (1999). The fundamentals of protein folding: bringing together theory and experiment. *Curr Opin Struct Biol* **9**, 92-101.

Dolgikh, D. A., Gilmanshin, R. I., Brazhnikov, E. V., Bychkova, V. E., Semisotnov, G. V., Venyaminov, S. and Ptitsyn, O. B. (1981). Alpha-Lactalbumin: compact state with fluctuating tertiary structure? *FEBS Lett* **136**, 311-5.

Doyle, D. F., Waldner, J. C., Parikh, S., Alcazar-Roman, L. and Pielak, G. J. (1996). Changing the transition state for protein (Un) folding. *Biochemistry* **35**, 7403-11.

Dragani, B., Stenberg, G., Melino, S., Petruzzelli, R., Mannervik, B. and Aceto, A. (1997). The conserved N-capping box in the hydrophobic core of glutathione S-transferase P1-1 is essential for refolding. Identification of a buried and conserved hydrogen bond important for protein stability. *J Biol Chem* **272**, 25518-23.

Elcock, A. H. and McCammon, J. A. (2001). Identification of protein oligomerisation states by analysis of interface conservation. *Proc Natl Acad Sci U S A* **98**, 2990-4.

Engelhard, M. and Evans, P. A. (1995). Kinetics of interaction of partially folded proteins with a hydrophobic dye: evidence that molten globule character is maximal in early folding intermediates. *Protein Sci* **4**, 1553-62.

Erhardt, J. and Dirr, H. (1995). Native dimer stabilises the subunit tertiary structure of porcine class pi glutathione S-transferase. *Eur J Biochem* **230**, 614-20.

Fanucchi, S., Adamson, R. J. and Dirr, H. W. (2008). Formation of an unfolding intermediate state of soluble chloride intracellular channel protein CLIC1 at acidic pH. *Biochemistry* **47**, 11674-81.

Finkelstein, A. V., Ivankov, D. N., Garbuzynskiy, S. O. and Galzitskaya, O. V. (2007). Understanding the folding rates and folding nuclei of globular proteins. *Curr Protein Pept Sci* **8**, 521-36.

Freire, E. (1995). Thermodynamics of partly folded intermediates in proteins. *Annu Rev Biophys Biomol Struct* **24**, 141-65.

Freire, E., Murphy, K. P., Sanchez-Ruiz, J. M., Galisteo, M. L. and Privalov, P. L. (1992). The molecular basis of cooperativity in protein folding. Thermodynamic dissection of interdomain interactions in phosphoglycerate kinase. *Biochemistry* **31**, 250-6.

Galani, D., Fersht, A. R. and Perrett, S. (2002). Folding of the yeast prion protein Ure2: kinetic evidence for folding and unfolding intermediates. *J Mol Biol* **315**, 213-27.

Garel, J. R. and Baldwin, R. L. (1973) Both the fast and slow refolding reactions of ribonuclease A. *Proc. Natl. Acad. Sci.* **70**, 3347-3351.

Geiger, A., Rahman, A. and Stillinger, F.H. (1979) Molecular dynamics study of the hydration of Lennard-Jones solutes. *J. Chem. Phys.* **70,** 263-276.

Gibrat, J. F., Madej, T. and Bryant, S. H. (1996). Surprising similarities in structure comparison. *Curr Opin Struct Biol* **6**, 377-85.

Gildenhuys, S. (2006). Folding mechanism of Gutaredoxin-2. PhD thesis.

Gildenhuys, S., Wallace, L. A. and Dirr, H. W. (2008). Stability and unfolding of reduced *E. coli* glutaredoxin 2: a monomeric structural homologue of the glutathione transferase family. *Biochemistry* **47**, 10801-8.

Goldenberg, D. P. (1992). Native and non-native intermediates in the BPTI folding pathway. *Trends Biochem Sci* **17**, 257-61.

Griko, Y. V., Freire, E., Privalov, G., van Dael, H. and Privalov, P. L. (1995). The unfolding thermodynamics of c-type lysozymes: a calorimetric study of the heat denaturation of equine lysozyme. *J Mol Biol* **252**, 447-59.

Griko, Y. V., Freire, E. and Privalov, P. L. (1994). Energetics of the alpha-lactalbumin states: a calorimetric and statistical thermodynamic study. *Biochemistry* **33**, 1889-99.

Grishin, N. V. and Phillips, M. A. (1994). The subunit interfaces of oligomeric enzymes are conserved to a similar extent to the overall protein sequences. *Protein Sci* **3**, 2455-8.

Guha, S. and Bhattacharyya, B. (1995). A partially folded intermediate during tubulin unfolding: its detection and spectroscopic characterization. *Biochemistry* **34**, 6925-31.

Habeeb, A. F., Schrohenloher, R. E. and Bennett, J. C. (1972). Studies of the component chains of human IgM by citraconylation. *Biochim Biophys Acta* **263**, 339-50.

Hagerman, P. J. and Baldwin, R. L. (1976). A quantitative treatment of the kinetics of the folding transition of ribonuclease A. *Biochemistry* **15**, 1462-73.

Hagler, A,T., Dauber, P., Lifson, S. (1979). Consistent force field studies of intermolecular forces in hydrogen bonded crystals. III. The C=O···H−O hydrogen bond and the analysis of the energetics and packing of carboxylic acids. *J. Am. Chem. Soc*. **101,** 5131−41.

Han, J. H., Batey, S., Nickson, A. A., Teichmann, S. A. and Clarke, J. (2007). The folding and evolution of multidomain proteins. *Nat Rev Mol Cell Biol* **8**, 319-30.

Han, J. H., Kerrison, N., Chothia, C. and Teichmann, S. A. (2006). Divergence of interdomain geometry in two-domain proteins. *Structure* **14**, 935-45.

Harrop, S. J., DeMaere, M. Z., Fairlie, W. D., Reztsova, T., Valenzuela, S. M., Mazzanti, M., Tonini, R., Qiu, M. R., Jankova, L., Warton, K., Bauskin, A. R., Wu, W. M., Pankhurst, S., Campbell, T. J., Breit, S. N. and Curmi, P. M. (2001). Crystal structure of a soluble form of the intracellular chloride ion channel CLIC1 (NCC27) at 1.4-A resolution. *J Biol Chem* **276**, 44993-5000.

Hartl, F. U., Hlodan, R. and Langer, T. (1994). Molecular chaperones in protein folding: the art of avoiding sticky situations. *Trends Biochem Sci* **19**, 20-5.

Henrick, K. and Thornton, J. M. (1998). PQS: a protein quaternary structure file server. *Trends Biochem Sci* **23**, 358-61.

Hinderaker, M. P. and Raines, R. T. (2003) An electronic effect on protein structure. *Protein Sci.* **12**, 1188-1194.

Holmgren, A. (1976). Hydrogen donor system for *E. coli* ribonucleoside-diphosphate reductase dependent upon glutathione. *Proc Natl Acad Sci U S A* **73**, 2275-9.

Holmgren, A. and Åslund, F. (1995). Glutaredoxin. *Methods Enzymol* **252**, 283-92.

Holmgren, A., Soderberg, B. O., Eklund, H. and Branden, C. I. (1975). Three-dimensional structure of *E. coli* thioredoxin-S2 to 2.8 Å resolution. *Proc Natl Acad Sci U S A* **72**, 2305-9.

Honig, B., Ray, A. and Levinthal, C. (1976). Conformational flexibility and protein folding: rigid structural fragments connected by flexible joints in subtilisin BPN. *Proc Natl Acad Sci U S A* **73**, 1974-8.

Hornby, J. A., Luo, J. K., Stevens, J. M., Wallace, L. A., Kaplan, W., Armstrong, R. N. and Dirr, H. W. (2000). Equilibrium folding of dimeric class mu glutathione transferases involves a stable monomeric intermediate. *Biochemistry* **39**, 12336-44.

Hornby, J. A., Codreanu, S. G., Armstrong, R. N. and Dirr, H. W. (2002). Molecular recognition at the dimer interface of a class mu glutathione transferase: role of a hydrophobic interaction motif in dimer stability and protein function. *Biochemistry* **41**, 14238-47.

Horwick, A. (2002). Protein aggregation in disease: a role for folding intermediates forming specific multimeric interactions. *J Clin Invest* **11,** 1221 – 32.

Jaenicke, R. (1999). Stability and folding of domain proteins. *Prog Biophys Mol Biol* **71**, 155-241.

Jeppesen, M. G., Ortiz, P., Shepard, W., Kinzy, T. G., Nyborg, J. and Andersen, G. R. (2003). The crystal structure of the glutathione S-transferase-like domain of elongation factor 1Bgamma from Saccharomyces cerevisiae. *J Biol Chem* **278**, 47190-8.

Ji, X., Johnson, W. W., Sesay, M. A., Dickert, L., Prasad, S. M., Ammon, H. L., Armstrong, R. N. and Gilliland, G. L. (1994). Structure and function of the xenobiotic substrate binding site of a glutathione S-transferase as revealed by X-ray crystallographic analysis of product complexes with the diastereomers of 9-(S-glutathionyl)-10-hydroxy-9,10-dihydrophenanthrene. *Biochemistry* **33**, 1043-52.

Ji, X., von Rosenvinge, E. C., Johnson, W. W., Tomarev, S. I., Piatigorsky, J., Armstrong, R. N. and Gilliland, G. L. (1995). Three-dimensional structure, catalytic properties, and evolution of a sigma class glutathione transferase from squid, a progenitor of the lens S-crystallins of cephalopods. *Biochemistry* **34**, 5317-28.

Jones, C.M., Henry, E.R., Hu, Y., Chan, C,K., Luck, S.D., Bhuyan, A., Roder, H., Hofrichter, J., and Eaton, W,A, (1993), Fast events in protein folding initiated by nanosecond laser photolysis. *Proc. Natl. Acad. Sci. USA.* **90,** 11860−64.

Jones, S. and Thornton, J. M. (1995). Protein-protein interactions: a review of protein dimer structures. *Prog Biophys Mol Biol* **63**, 31-65.

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

Jones, S., Marin, A. and Thornton, J. M. (2000). Protein domain interfaces: characterisation and comparison with oligomeric protein interfaces. *Protein Eng* **13**, 77-82.

Jonsson, T., Waldburger, C. D. and Sauer, R. T. (1996) Nonlinear free energy relationship in Arc repressor unfolding imply the existence of unstable, native-like folding intermediates. *Biochemistry* **35**, 4795-4802

Kaplan, W., Husler, P., Klump, H., Erhardt, J., Sluis-Cremer, N. and Dirr, H. (1997). Conformational stability of pGEX-expressed Schistosoma japonicum glutathione S-transferase: a detoxification enzyme and fusion-protein affinity tag. *Protein Sci* **6**, 399-406.

Kahn, P. C. (1979). The interpretation of near-ultraviolet circular dichroism. *Methods Enzymol* **61**, 339-78.

Kelley, R. F. and Stellwagen, E. (1984) Conformational transitions of thioredoxin in guanidine hydrochloride. *Biochemistry* **23**, 5095-5102.

Kellis, J. T., Jr., Nyberg, K., Sali, D. and Fersht, A. R. (1988). Contribution of hydrophobic interactions to protein stability. *Nature* **333**, 784-6.

Keskin, O., Ma, B. and Nussinov, R. (2005). Hot regions in protein--protein interactions: the organisation and contribution of structurally conserved hot spot residues. *J Mol Biol* **345**, 1281-94.

Kiefhaber, T. and Schmid, F. X. (1992). Kinetic coupling between protein folding and prolyl isomerization. II. Folding of ribonuclease A and ribonuclease T1. *J Mol Biol* **224**, 231-40.

Knowles, J. R. (1991). To build an enzyme. *Philos Trans R Soc Lond B Biol Sci* **332**, 115-21.

Kosloff, M., Han, G. W., Krishna, S. S., Schwarzenbacher, R., Fasnacht, M., Elsliger, M. A., Abdubek, P., Agarwalla, S., Ambing, E., Astakhova, T., Axelrod, H. L., Canaves, J. M., Carlton, D., Chiu, H. J., Clayton, T., DiDonato, M., Duan, L., Feuerhelm, J., Grittini, C., Grzechnik, S. K., Hale, J., Hampton, E., Haugen, J., Jaroszewski, L., Jin, K. K., Johnson, H., Klock, H. E., Knuth, M. W., Koesema, E., Kreusch, A., Kuhn, P., Levin, I., McMullan, D., Miller, M. D., Morse, A. T., Moy, K., Nigoghossian, E., Okach, L., Oommachen, S., Page, R., Paulsen, J., Quijano, K., Reyes, R., Rife, C. L., Sims, E., Spraggon, G., Sridhar, V., Stevens, R. C., van den Bedem, H., Velasquez, J., White, A., Wolf, G., Xu, Q., Hodgson, K. O., Wooley, J., Deacon, A. M., Godzik, A., Lesley, S. A. and Wilson, I. A. (2006). Comparative structural analysis of a novel glutathioneS-transferase (ATU5508) from Agrobacterium tumefaciens at 2.0 Å resolution. *Proteins* **65**, 527-37.

Kumar, S. and Nussinov, R. (1999). Salt bridge stability in monomeric proteins. *J Mol Biol* **293**, 1241-55.

Kuwajima, K. (1977). A folding model of alpha-lactalbumin deduced from the three-state denaturation mechanism. *J Mol Biol* **114**, 241-58.

Kuwajima, K., Okayama, N., Yamamoto, K., Ishihara, T. and Sugai, S. (1991). The Pro117 to glycine mutation of staphylococcal nuclease simplifies the unfolding-folding kinetics. *FEBS Lett* **290**, 135-8.

Lakowicz, J. R. (1999) Principles of fluorescence spectroscopy, Protein fluorescence, pp 445-465. Plenum Publishers, USA.

Ladner, J. E., Parsons, J. F., Rife, C. L., Gilliland, G. L. and Armstrong, R. N. (2004). Parallel evolutionary pathways for glutathione transferases: structure and mechanism of the mitochondrial class kappa enzyme rGSTK1-1. *Biochemistry* **43**, 352-61.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-5.

Laurent, T. C., Moore, E. C. and Reichard, P. (1964). Enzymatic Synthesis of Deoxyribonucleotides. Iv. Isolation and Characterisation of Thioredoxin, the Hydrogen Donor from *E. coli* B. *J Biol Chem* **239**, 3436-44.

Lee, B. and Richards, F. M. (1971). The interpretation of protein structures: estimation of static accessibility. *J Mol Biol* **55**, 379-400.

Levinthal, C. (1968). Are there pathways for protein folding? *J Chim. Phys.* **65**, 44-45.

Lillig, C. H., Prior, A., Schwenn, J. D., Åslund, F., Ritz, D., Vlamis-Gardikas, A. and Holmgren, A. (1999). New thioredoxins and glutaredoxins as electron donors of 3'-phosphoadenylylsulfate reductase. *J Biol Chem* **274**, 7695-8.

Lo Conte, L., Chothia, C. and Janin, J. (1999). The atomic structure of protein-protein recognition sites. *J Mol Biol* **285**, 2177-98.

Luo, J. K., Hornby, J. A., Wallace, L. A., Chen, J., Armstrong, R. N. and Dirr, H. W. (2002). Impact of domain interchange on conformational stability and equilibrium folding of chimeric class micro glutathione transferases. *Protein Sci* **11**, 2208-17.

Lyon, R. P. and Atkins, W. M. (2002). Kinetic characterisation of native and cysteine 112-modified glutathione S-transferase A1-1: reassessment of nonsubstrate ligand binding. *Biochemistry* **41**, 10920-7.

Marchler-Bauer, A., Anderson, J. B., Cherukuri, P. F., DeWeese-Scott, C., Geer, L. Y., Gwadz, M., He, S., Hurwitz, D. I., Jackson, J. D., Ke, Z., Lanczycki, C. J., Liebert, C. A., Liu, C., Lu, F., Marchler, G. H., Mullokandov, M., Shoemaker, B. A., Simonyan, V., Song, J. S., Thiessen, P. A., Yamashita, R. A., Yin, J. J., Zhang, D. and Bryant, S. H. (2005). CDD: a Conserved Domain Database for protein classification. *Nucleic Acids Res* **33**, D192-6.

Murphy, K. P. (1999). Predicting binding energetics from structure: looking beyond DeltaG degrees. *Med Res Rev* **19**, 333-9.

Murzin, A. G., Brenner, S. E., Hubbard, T. and Chothia, C. (1995). SCOP: a structural classification of proteins database for the investigation of sequences and structures. *J Mol Biol* **247**, 536-40.

Myers, J. K., Pace, C. N. and Scholtz, J. M. (1995). Denaturant m values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding. *Protein Sci* **4**, 2138-48.

Nelson, M. and McClelland, M. (1992). Use of DNA methyltransferase/endonuclease enzyme combinations for megabase mapping of chromosomes. Methods Enzymol 216, 279-303.

Nölting, B., Golbik, R., Fersht, A.R. (1995). Submillisecond events in protein folding. *Proc. Natl. Acad. Sci. USA* **92,**10668−10672.

Nooren, I. M. and Thornton, J. M. (2003). Structural characterisation and functional significance of transient protein-protein interactions. *J Mol Biol* **325**, 991-1018.

Oakley, A. J., Harnnoi, T., Udomsinprasert, R., Jirajaroenrat, K., Ketterman, A. J. and Wilce, M. C. (2001). The crystal structures of glutathione S-transferases isozymes 1-3 and 1-4 from Anopheles dirus species B. *Protein Sci* **10**, 2176-85.

Ohgushi, M. and Wada, A. (1983). 'Molten-globule state': a compact form of globular proteins with mobile side-chains. *FEBS Lett* **164**, 21-4.

Pace, C. N. (1986). Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods Enzymol* **131**, 266-80.

Pace, C. N. (1990). Conformational stability of globular proteins. *Trends Biochem Sci* **15**, 14-7.

Pace, C. N. (2001). Polar group burial contributes more to protein stability than nonpolar group burial. *Biochemistry* **40**, 310-3.

Pace, C. N., Shirley, B. A., McNutt, M. and Gajiwala, K. (1996). Forces contributing to the conformational stability of proteins. *Faseb J* **10**, 75-83.

Perbandt, M., Burmeister, C., Walter, R. D., Betzel, C. and Liebau, E. (2004). Native and inhibited structure of a Mu class-related glutathione S-transferase from Plasmodium falciparum. *J Biol Chem* **279**, 1336-42.

Polekhina, G., Board, P. G., Blackburn, A. C. and Parker, M. W. (2001). Crystal structure of maleylacetoacetate isomerase/glutathione transferase zeta reveals the molecular basis for its remarkable catalytic promiscuity. *Biochemistry* **40**, 1567-76.

Ponstingl, H., Henrick, K. and Thornton, J. M. (2000). Discriminating between homodimeric and monomeric proteins in the crystalline state. *Proteins* **41**, 47-57.

Privalov, P. L. (1979). Stability of proteins: small globular proteins. *Adv Protein Chem* **33**, 167-241.

Privalov, P. L. (1996). Intermediate states in protein folding. *J Mol Biol* **258**, 707-25.

Privalov, P. L. and Makhatadze, G. I. (1993). Contribution of hydration to protein folding thermodynamics. II. The entropy and Gibbs energy of hydration. *J Mol Biol* **232**, 660-79.

Ptitsyn, O. B. (1973). [Stages in the mechanism of self-organisation of protein molecules]. *Dokl Akad Nauk SSSR* **210**, 1213-5.

Ptitsyn, O.B. (1992). The Molten Globule State. In *Protein folding* (ed. T.E. Creighton), pp 243-300. New York: WH. Freeman and Co.

Ptitsyn, O.B. (1995) Structures of folding intermediates. *Curr. Opin. Str. Biol.* **5,** 74-78.

Ptitsyn, O. B., Bychkova, V. E. and Uversky, V. N. (1995). Kinetic and equilibrium folding intermediates. *Philos Trans R Soc Lond B Biol Sci* **348**, 35-41.

Pupko, T., Bell, R. E., Mayrose, I., Glaser, F. and Ben-Tal, N. (2002). Rate4Site: an algorithmic tool for the identification of functional regions in proteins by surface mapping of evolutionary determinants within their homologues. *Bioinformatics* **18 Suppl 1**, S71-7.

Ratnaparkhi, G. S. and Varadarajan, R. (2000). Thermodynamic and structural studies of cavity formation in proteins suggest that loss of packing interactions rather than the hydrophobic effect dominates the observed energetics. *Biochemistry* **39**, 12365-74.

Reichard, P. (1993). The anaerobic ribonucleotide reductase from *E. coli*. *J Biol Chem* **268**, 8383-6.

Reinemer, P., Dirr, H. W., Ladenstein, R., Huber, R., Lo Bello, M., Federici, G. and Parker, M. W. (1992). Three-dimensional structure of class pi glutathione S-transferase from human placenta in complex with S-hexylglutathione at 2.8 Å resolution. *J Mol Biol* **227**, 214-26.

Reinemer, P., Prade, L., Hof, P., Neuefeind, T., Huber, R., Zettl, R., Palme, K., Schell, J., Koelln, I., Bartunik, H. D. and Bieseler, B. (1996). Three-dimensional structure of glutathione S-transferase from Arabidopsis thaliana at 2.2 Å resolution: structural characterisation of herbicide-conjugating plant glutathione S-transferases and a novel active site architecture. *J Mol Biol* **255**, 289-309.

Richardson, J. S. (1981). The anatomy and taxonomy of protein structure. *Adv Protein Chem* **34**, 167-339.

Rife, C. L., Parsons, J. F., Xiao, G., Gilliland, G. L. and Armstrong, R. N. (2003). Conserved structural elements in glutathione transferase homologues encoded in the genome of *E. coli*. *Proteins* **53**, 777-82.

Rossjohn, J., McKinstry, W. J., Oakley, A. J., Verger, D., Flanagan, J., Chelvanayagam, G., Tan, K. L., Board, P. G. and Parker, M. W. (1998). Human theta class glutathione transferase: the crystal structure reveals a sulfate-binding pocket within a buried active site. *Structure* **6**, 309-22.

Sayed, Y., Wallace, L. A. and Dirr, H. W. (2000). The hydrophobic lock-and-key intersubunit motif of glutathione transferase A1-1: implications for catalysis, ligandin function and stability. *FEBS Lett* **465**, 169-72.

Schmid, F. X. and Baldwin, R. L. (1978) Acid catalysis of the formation of the slow-folding species of RNase A: Evidence that the reaction is proline isomerization. *Proc. Natl. Acad. Sci* USA **75**, 4764-4768.

Semisotnov, G. V., Rodionova, N. A., Razgulyaev, O. I., Uversky, V. N., Gripas, A. F. and Gilmanshin, R. I. (1991). Study of the "molten globule" intermediate state in protein folding by a hydrophobic fluorescent probe. *Biopolymers* **31**, 119-28.

Shastry, M. C. and Roder, H. (1998). Evidence for barrier-limited protein folding kinetics on the microsecond time scale. *Nat Struct Biol* **5**, 385-92.

Shirley, B.A. (1995). Urea and guanidine hydrochloride denaturation curves in Methods in *Molecular Biology: Protein Stability and Folding: Theory and Practice* (Shirley, B.A., Ed.). Vol. 40. pp 177-190. Humana Press Inc. Totowa, New Jersey.

Shortle, D. (1993). Denatured states of proteins and their roles in folding and stability. *Curr. Opin. Struct. Biol.* **3**, 66–74.

Shortle, D. (1995). Staphylococcal nuclease: a showcase of *m*-value effects. *Adv.Protein Chem.* **46,** 217–278.

Shortle, D. (1996). The denatured state (the other half of the folding equation) and its role in protein stability. *FASEB J.* **10**, 27-34.

Sigurskjold, B. W., Altman, E. and Bundle, D. R. (1991). Sensitive titration microcalorimetric study of the binding of Salmonella O-antigenic oligosaccharides by a monoclonal antibody. *Eur J Biochem* **197**, 239-46.

Sinning, I., Kleywegt, G. J., Cowan, S. W., Reinemer, P., Dirr, H. W., Huber, R., Gilliland, G. L., Armstrong, R. N., Ji, X., Board, P. G. and *et al*. (1993). Structure determination and refinement of human alpha class glutathione transferase A1-1, and a comparison with the Mu and Pi class enzymes. *J Mol Biol* **232**, 192-212.

Soulages, J. L. (1998). Chemical denaturation: potential impact of undetected intermediates in the free energy of unfolding and m-values obtained from a two-state assumption. *Biophys J* **75**, 484-92.

Stenberg, G., Dragani, B., Cocco, R., Mannervik, B. and Aceto, A. (2000). A conserved "hydrophobic staple motif" plays a crucial role in the refolding of human glutathione transferase P1-1. *J Biol Chem* **275**, 10421-8.

Stevens, J. M., Armstrong, R. N. and Dirr, H. W. (2000). Electrostatic interactions affecting the active site of class sigma glutathione S-transferase. *Biochem J* **347 Pt 1**, 193-7.

Stevens, J. M., Hornby, J. A., Armstrong, R. N. and Dirr, H. W. (1998). Class sigma glutathione transferase unfolds via a dimeric and a monomeric intermediate: impact of subunit interface on conformational stability in the superfamily. *Biochemistry* **37**, 15534-41.

Stickle, D. F., Presta, L. G., Dill, K. A. and Rose, G. D. (1992). Hydrogen bonding in globular proteins. *J Mol Biol* **226**, 1143-59.

Stillinger, F. H. (1980). Water Revisited. *Science* **209**, 451-457.

Stites, W. E. (1997). Proteinminus signProtein Interactions: Interface Structure, Binding Thermodynamics, and Mutational Analysis. *Chem Rev* **97**, 1233-1250.

Stoychev, S. (2008). The role of the domain interface in the stability, folding and function of Clic1. PhD thesis.

Strickland, E.H. (1974) Aromatic contributions to circular dichroism spectra of proteins. *Crit. Rev. Biochem.* **2,** 113-75.

Tanford, C. (1968). Protein denaturation. *Adv Protein Chem* **23**, 121-282.

Tanford, C. (1970). Protein denaturation. C. Theoretical models for the mechanism of denaturation. *Adv Protein Chem* **24**, 1-95.

Teale, F. W. J. (1960). The ultraviolet fluorescence of proteins in neutral solution. *Biochem J* **76,** 381-8.

Teichmann, S. A., Chothia, C. and Gerstein, M. (1999). Advances in structural genomics. *Curr Opin Struct Biol* **9**, 390-9.

Thannhauser, T. W., Konishi, Y. and Scheraga, H. A. (1984). Sensitive quantitative analysis of disulfide bonds in polypeptides and proteins. *Anal Biochem* **138**, 181-8.

Thom, R., Cummins, I., Dixon, D. P., Edwards, R., Cole, D. J. and Lapthorn, A. J. (2002). Structure of a tau class glutathione S-transferase from wheat active in herbicide detoxification. *Biochemistry* **41**, 7008-20.

Tsai, C. J., Lin, S. L., Wolfson, H. J. and Nussinov, R. (1996). Protein-protein interfaces: architectures and interactions in protein-protein interfaces and in protein cores. Their similarities and differences. *Crit Rev Biochem Mol Biol* **31**, 127-52.

Umland, T. C., Taylor, K. L., Rhee, S., Wickner, R. B. and Davies, D. R. (2001). The crystal structure of the nitrogen regulation fragment of the yeast prion protein Ure2p. *Proc Natl Acad Sci U S A* **98**, 1459-64.

Utiyama, H. and Baldwin, R. L. (1986). Kinetic mechanisms of protein folding. *Methods Enzymol* **131**, 51-70.

Uversky, V. N., Semisotnov, G. V., Pain, R. H. and Ptitsyn, O. B. (1992). 'All-or-none' mechanism of the molten globule unfolding. *FEBS Lett* **314**, 89-92.

Uversky, V. N. and Ptitsyn, O. B. (1996). All-or-none solvent-induced transitions between native, molten globule and unfolded states in globular proteins. *Fold Des* **1**, 117-22.

van Dael, H., Haezebrouck, P., Morozova, L., Arico-Muandel, C. and Dobson, C.M. (1993). Partially folded states of equine lysozyme. Structural characterisation and significance for protein folding. *Biochemistry*. **32**, 1186–94.

Vajda, S. and Camacho, C. J. (2004). Protein-protein docking: is the glass half-full or half-empty? *Trends Biotechnol* **22**, 110-6.

Valdar, W. S. and Thornton, J. M. (2001). Protein-protein interfaces: analysis of amino acid conservation in homodimers. *Proteins* **42**, 108-24.

Vlamis-Gardikas, A., Åslund, F., Spyrou, G., Bergman, T. and Holmgren, A. (1997). Cloning, overexpression, and characterisation of glutaredoxin 2, an atypical glutaredoxin from *Escherichia coli*. *J Biol Chem* **272**, 11236-43.

Vreuls, C., Filee, P., Van Melckebeke, H., Aerts, T., De Deyn, P., Llabres, G., Matagne, A., Simorre, J. P., Frere, J. M. and Joris, B. (2004). Guanidinium chloride denaturation of the dimeric Bacillus licheniformis BlaI repressor highlights an independent domain unfolding pathway. *Biochem J* **384**, 179-90.

Walkenhorst, W. F., Green, S. M. and Roder, H. (1997) Kinetic evidence for folding and unfolding intermediates in staphylococcal nuclease. *Biochemistry* **36,** 5795-5805.

Wallace, L. A., Blatch, G. L. and Dirr, H. W. (1998). A topologically conserved aliphatic residue in alpha-helix 6 stabilises the hydrophobic core in domain II of glutathione transferases and is a structural determinant for the unfolding pathway. *Biochem J* **336 (Pt 2)**, 413-8.

Wallace, L. A. and Dirr, H. W. (1999). Folding and assembly of dimeric human glutathione transferase A1-1. *Biochemistry* **38**, 16686-9

Wallace, L. A., Burke, J. and Dirr, H. W. (2000). Domain-domain interface packing at conserved Trp-20 in class alpha glutathione transferase impacts on protein stability. *Biochim Biophys Acta* **1478**, 325-32.

Wallace, L. A. and Matthews, C. R. (2002). Sequential vs. parallel protein-folding mechanisms: experimental tests for complex folding reactions. *Biophys Chem* **101-102**, 113-31.

Weber, G. and Young, L. B. (1964). Fragmentation of Bovine Serum Albumin by Pepsin. I. the Origin of the Acid Expansion of the Albumin Molecule. *J Biol Chem* **239**, 1415-23.

Wetlaufer, D. B. (1973). Nucleation, rapid folding, and globular intrachain regions in proteins. *Proc Natl Acad Sci U S A* **70**, 697-701.

Wilce, M. C. and Parker, M. W. (1994). Structure and function of glutathione S-transferases. *Biochim Biophys Acta* **1205**, 1-18.

Winter, C., Henschel, A., Kim, W. K. and Schroeder, M. (2006). SCOPPI: a structural classification of protein-protein interfaces. *Nucleic Acids Res* **34**, D310-4.

Wodak, S. J. and Janin, J. (1981). Location of structural domains in protein. *Biochemistry* **20**, 6544-52.

Wolynes, P. G., Onuchic, J. N. and Thirumalai, D. (1995). Navigating the folding routes. *Science* **267**, 1619-20.

Wongsantichon, J. and Ketterman, A. J. (2006). An intersubunit lock-and-key 'clasp' motif in the dimer interface of Delta class glutathione transferase. *Biochem J* **394**, 135-44.

Woody, R. W. (1995). Circular dichroism. *Methods Enzymol* **246**, 34-71.

Xia, B., Chung, J., Vlamis-Gardikas, A., Holmgren, A., Wright, P. E. and Dyson, H. J. (1999). Assignment of 1H, 13C, and 15N resonances of reduced *Escherichia coli* glutaredoxin 2. *J Biomol NMR* **14**, 197-8.

Xia, B., Vlamis-Gardikas, A., Holmgren, A., Wright, P. E. and Dyson, H. J. (2001). Solution structure of *Escherichia coli* glutaredoxin-2 shows similarity to mammalian glutathione-S-transferases. *J Mol Biol* **310**, 907-18.

Yamada, T., Komoto, J., Watanabe, K., Ohmiya, Y. and Takusagawa, F. (2005). Crystal structure and possible catalytic mechanism of microsomal prostaglandin E synthase type 2 (mPGES-2). *J Mol Biol* **348**, 1163-76.

Yassin, Z., Ortiz-Salmeron, E., Garcia-Maroto, F., Baron, C. and Garcia-Fuentes, L. (2004). Implications of the ligandin binding site on the binding of non-substrate ligands to *Schistosoma japonicum*-glutathione transferase. *Biochim Biophys Acta* **1698**, 227-37.

Yon, J. M. (2001). Protein folding: a perspective for biology, medicine and biotechnology. *Braz J Med Biol Res* **34**, 419-35.