FOLDING MECHANISM OF GLUTAREDOXIN 2

Samantha Gildenhuys

A thesis submitted to the Faculty of Science, University of the Witwatersrand, Johannesburg, in fulfilment of the requirements for the degree of Doctor of Philosophy.

Johannesburg, 2006

Declaration

I declare that this thesis is my own, unaided work. It is being submitted for the Degree of Doctor of Philosophy in the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination in any other University.

_____

Samantha Gildenhuys

_____ day of _____ 2006

ABSTRACT

Equilibrium unfolding, single- and double-jump kinetic studies were conducted to determine the unfolding and refolding pathway of glutaredoxin 2. Structural changes for wild-type glutaredoxin 2 were monitored by far-ultraviolet circular dichroism and intrinsic tryptophan fluorescence for equilibrium unfolding and intrinsic tryptophan fluorescence for single- and double-jump kinetics studies. Glutaredoxin 2 possesses two tryptophan residues in domain 2. In order to monitor changes in domain 1, cysteine 9 at the active site cysteines, situated in domain 1, was labelled with an extrinsic fluorophore, AEDANS, and a mutant was created (Y58W glutaredoxin 2). The AEDANS labelled protein displayed decreased alpha-helical secondary structure and conformational stability. A high degree of cooperativity and similar conformational stability was observed during the two-state transition of the urea-induced equilibrium unfolding of both the wild-type and Y58W glutaredoxin 2 proteins therefore Y58W glutaredoxin 2 could be used to assess structural changes in the local environment of domain 1 during unfolding and refolding. Two phases of unfolding, the fast and slow phase, occurred for both the wild-type and Y58W proteins. 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 isomerization of the Val48-Pro49 peptide bond to the *trans* conformation occurs during the slow phase and this isomerization is coupled to conformational unfolding of the protein. The structural separation of these phases could be represented by two structural units (unit x and unit y), these units do not represent domain 1 and 2. The units could also result in parallel refolding pathways with the folding of the x unit involving the fast and slow refolding phases and the folding of the y unit of structure is represented by the medium phase of refolding. The fast and slow phases are further separated as the fast phase represents the gross structural folding of glutaredoxin 2 for species with the Val48-Pro49 peptide bond in the native *cis* conformation. The development of the slow phase after extended unfolding delay periods during double-jump refolding studies, as well as the acceleration of the rate of the phase by the peptidyl prolyl isomerase hFKBP-12 proved that the phase involves a proline peptide bond

isomerization. This phase represents a slow isomerization coupled with conformational folding similar to the slow unfolding phase. Complex unfolding and refolding kinetics indicated the involvement of kinetic intermediates during (un)folding.

I dedicate this work to my family you have always been there for me, supported me, guided me and helped me to get to this point. Then to my friends and everyone one who has had a positive influence on me in some way, be it big or small.

"Great things are done by a series of small things brought together" --Vincent van Gogh

Acknowledgements

Professor Heini Dirr for all his support and guidance throughout my studies.

Members of the Protein Structure-Function Research Unit, University of the Witwatersrand.

The National Research Foundation, South Africa and the University of the Witwatersrand for financial support.

Table of contents

List of figures

List of tables

List of folding schemes

Abbreviations

ΔASA	change in solvent accessible surface area
[θ]	mean residue ellipticity
A_{280}	absorbance at 280 nm
AEDANS	5-[[2-[(acetyl)amino)ethyl]amino]-naphthalene-1-sulphonic acid
ANS	8-anilino-1-naphthalene sulphonate
a.u.	arbitrary units
CE	combinatorial Extension
Clic	chloride intracellular channel
C_m	urea-induced midpoint of unfolding
CRABP	cellular retinoic acid-binding protein
Da	Dalton
DCIP	2,6-dichlorophenolindophenol
Dsb	disulphide-bond isomerase
DTNB	5,5'-dithiobis(2-nitrobenzoic acid)
DTT	dithiothreitol
EDTA	ethylenediaminetetra-acetic acid
ESI MS	electrospray ionization mass spectrometry
$\Delta G(H_2O)$	change in Gibbs free energy of unfolding in the absence of denaturant
Grx	glutaredoxin
GST	glutathione S-transferase
hFKBP	human FK506-binding protein
IAEDANS	5-[[2-[(iodoacetyl)amino)ethyl]amino]-naphthalene-1-sulphonic acid
IPTG	Isopropyl-β-D-thiogalactopyranoside
k_{app}	pseudo-first order rate constant
$k(H_2O)$	apparent unfolding(k_u)/ refolding(k_f) rate in the absence of denaturant
kDa	kilodalton
k_f	rate constant for refolding (determined for conditions where rate observed is dominated by the folding rate)
k_u	rate constant for unfolding (determined for conditions where rate observed is dominated by the unfolding rate)

m_f	change in the solvent accessibility during refolding
m-value	dependence of the free energy of unfolding on denaturant concentration
m_u	change in accessible surface area during unfolding
N	native monomer
OD_{600}	optical density at 600 nm
PAPS	3'-phosphoadenylyl sulphate
PDB	protein data bank
PPI	peptidyl-prolyl isomerase
rpm	revolutions per minute
RR	ribonucleotide-diphosphate reductase
SDS-PAGE	sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SE-HPLC	size exclusion high-pressure liquid chromatography
τ	time constant, which is the inverse of the apparent rate constant
Trx	thioredoxin
U	unfolded monomer
Xaa	unspecified amino acid residue
Y58W	Tyr58 residue (wild-type) replaced with a tryptophan residue

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

1 INTRODUCTION

1.1 Protein folding: a general introduction

Proteins are synthesised according to the Central Dogma: DNA—›RNA—›protein, once synthesised these linear polypeptide chains adopt unique three-dimensional structures. The process by which these polypeptides obtain their unique structures is termed protein folding and the 'protein folding problem' refers to the question as to how these linear sequences adopt their native structures. In addressing this problem, Anfinsen (1973) demonstrated that the amino acid sequence contains all the information to specify the final fully functional 3–dimensional structure of the protein. The exact mechanism as to how the unique sequence of amino acids dictated the final structure is not fully understood. What is understood is that the time it takes for proteins to fold dictates that proteins do not attain their native structure by a random process (Levinthal, 1968) but are rather said to pass through a limited number of states along a defined pathway to the native state (Creighton, 1990; Jaenicke, 1991; Schmid, 1992).

1.1.1 Some protein folding models

In order to understand and explain how proteins fold, many folding models have been proposed. These models aim to generalize protein folding mechanisms so that one model could explain any protein's folding behaviour in the time frame at which folding occurs. One such model is the hydrophobic collapse model, this model identifies the first step in folding to be collapse of the polypeptide chain by long range hydrophobic interactions (Dill, 1985). In the hydrophobic collapse model the hydrophobic collapse occurs prior to any secondary structure formation however it has been accepted and Dill an co-workers later proposed that instead local secondary structure forms simultaneously with the collapse of the polypeptide during folding. The diffusion-collision model builds on this and proposes that small microdomains form early during folding, these microdomains

are made up of secondary structural elements (Karplus and Weaver, 1994). Folding then continues with the microdomains interacting to form multimicrodomain intermediates, these multimicrodomain elements form and disappear until the tertiary structure is formed (Karplus and Weaver, 1994). The last phase of folding is the formation of the tertiary structural elements and the correct packing of the side chains. It is also during this phase that slow rate-limiting steps such as proline isomerization occurs (Karplus and Weaver, 1994).

A model that is based on the classical nucleation model and expands on the diffusion-collision model, is the nucleation-condensation model (Fersht, 1997). In this model there is a larger more global nucleus formed, this nucleus then forms the transition state that then forms the native structure. The transition state for proteins is a high energy state cannot be characterised directly (Creighton, 1990) and is usually characterised indirectly only under certain conditions (Schmid, 1992). For the model the formation of the nucleus is coupled to the formation of secondary and tertiary structure throughout the protein hence the name nucleation-condensation (Fersht, 1997). This coupling results in a highly cooperative folding process where there is a balancing between local structure formed as well as long range tertiary structure as well (Fersht, 1997). With this model though there is minimal accumulation of intermediates (Fersht, 1997). Larger proteins form separate folding units of structure that represent individual nucleation-condensation folding units, these units then dock to form the native structure (Fersht, 1997). The larger proteins therefore are more prone to form intermediates as the individual folding units form stable structures and the rearrangement and docking of these units is the rate determining step in folding (Fersht, 1997).

Both models presented so far have not introduced the idea that a single polypeptide chain can take several routes to achieve the native state. One such model is the jigsaw-puzzle model (Harrison and Durbin, 1985). This model is based on the idea that protein folding, like solving a jigsaw puzzle, is a process

directed to an end point although the route taken to that end point can differ each. In the model it is also proposed that native-like native structures form during folding (Harrison and Durbin, 1985). This closely resembles the diffusion-collision model proposes that small microdomains form early during folding, (Karplus and Weaver, 1994). The intermediates that form during folding are not described in terms of their structural characteristics but it is said that they are a heterogeneous population that is in rapid equilibrium (Harrison and Durbin, 1985).

1.1.2 Protein folding intermediates and pathways

Protein folding models try to generalise folding mechanisms in order to create a "one fits all" model for all proteins. The formation of this model is based on the determination of protein folding pathways of several proteins. A particular protein's folding pathway is only fully understood when the starting native or denatured state as well as intermediates formed during (un)folding can be positioned in the order in which they form (Creighton, 1990; Jaenicke, 1999). The native and unfolded state as well as stable intermediates if present at equilibrium can be characterized (Creighton, 1990) although the denatured state represents an ensemble of a number of interconverting species (Jaenicke, 1999). Kinetic refolding and equilibrium unfolding of α-lactalbumin indicted that the equilibrium intermediate posses the same structure as a transient kinetic refolding intermediate (Ikeguchi *et al.,* 1986). An intermediate formed early during folding which can be detected during equilibrium unfolding studies has been termed the molten globule (Seckler and Jaenicke, 1992). Kuwajima (1989) specifies the criteria that an early intermediate must possess in order to be termed a molten globule which include a high degree of native secondary structure, an overall compactness with buried but still mobile aromatic side chains and exposed hydrophobic surfaces.

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). One such example of where the kinetic data provided more information about folding intermediates was for chymotrypsin inhibitor 2 (Jackson and Fersht, 1991). The presence of intermediates can be detected by kinetic studies but the structural details and nature of these species cannot be determined by these studies (Yon, 2001). The kinetics of folding for proteins are fitted to the sum of exponentials (Zarrine-Afsar and Davidson, 2004). When more than one exponential can be fitted to the data the presence of multiple rate limiting steps or a heterogeneous starting population of species is suspected (Creighton, 1990). A heterogeneous population of starting species often results in parallel refolding pathways as opposed to multiple kinetic phases with uniform staring species where the additional rate limiting steps such as the presence of intermediates result in sequential folding pathways. For a review on tests to distinguish parallel versus sequential folding pathways, see Wallace and Matthews (2002). Most of the work on protein folding pathways has been conducted for small single domain proteins and it was thought that the information could be transferred to larger multidomain protein by assuming that the domains folded as single units (Jaenicke, 1999).

1.2 Domains and folding

Domains were identified as compact substructures within proteins that are said to have originated from smaller nucleation sites within the protein structure that following rapid growth form these independent folding units (Wetlaufer, 1973). Many other definitions for domains also exist although they are not mutually exclusive and are based on interactions made within the domain and the domain with the rest of the polypeptide, sequence homologies, functional units, and exon expression to name a few (Jaenicke, 1999 and references therein). One relevant to the study of protein folding is where domains have been defined as having distinct equilibrium unfolding transitions during equilibrium unfolding (Rowe and Tanford, 1973).

The folding of multidomain proteins has been investigated for γII-crystallin from calf eye lens and human. For these proteins it was shown that the two homologous domains (un)fold independently (Rudolph *et al*., 1990; Wenk *et al.,* 2000). The folding of the domains of calf eye lens protein, and not the pairing of these domain, is the rate-limiting step during folding (Rudolph *et al*., 1990). The pairing or docking of the two domains of γII-crystallin from calf eye lens did not contribute at all to the kinetic mechanism of folding or the stability of the protein (Rudolph *et al*., 1990). Just as this example highlights the very independent folding of the domains of a protein there exists examples where that interdomain interactions can contribute to the folding of the domains. For type κ immunoglobulin light chain and its variable and constant fragments it was shown that the fragments or domains (un)fold independently although the interdomain interactions contributed to the refolding kinetics (Tsunenaga *et al*., 1987). Similarly interdomain contacts were said to influence to folding of the domains (which can fold independently) of yeast phosphoglycerate kinase (Missiakas *et al.,* 1992; Osváth *et al.,* 2005). These examples highlight that the folding of proteins with more than one domain is not as simple as the predicted independent folding of the domains with simultaneous pairing. This and the difficulty in correctly determining the boundaries (removing all contacts with the rest of the polypeptide) for domains can add great complexity to determining the folding of multidomain proteins (Jaenicke, 1999).

A further example that points to the complex nature of the folding of multidomain proteins is that of gene-3-protein whose two domains (un)fold independently. The larger N2 domain was less stable than the N1 domain and the first transition in the equilibrium unfolding data corresponded to the unfolding of this domain (Martin and Schmid, 2003a). It must also be mentioned that although the N1 and N2 domains unfold separately, the interdomain contacts were still stabilizing factors and the unlocking of the two domains is thermodynamically coupled to N2 domain unfolding (Martin and Schmid, 2003a). The kinetics of (un)folding of gene-3-protein also showed that the unfolding of the two domains was connected (Martin and Schmid, 2003b). Prior to the docking of the domains, unfolding of

both domains occurred fast and independently. Following docking, however the unfolding was slower and more correlated (Martin and Schmid, 2003b). The refolding indicated that domain N1 folded faster and independently of N2, with domain N2 folding in a biphasic manner (Martin and Schmid, 2003b). Although the folding and unfolding of the N1 and N2 domains could occur independently the interdomain docking or coupling play an important role in the stability and (un)folding of this gene-3-protein (Martin and Schmid, 2003a, b). Gene-3-protein has independent domain folding with a very slow docking of the folded domains (Martin and Schmid, 2003b).

The examples presented so far for the folding of multidomain proteins have demonstrated that the domains of multidomain proteins fold independently or at least as independent folding units. This however is not the case for all proteins. There are cases where there is a great deal of cross talk between domains and a high degree of cooperativity for the folding/unfolding of these proteins (Privalov and Potekhin, 1986). Proteins where the domains fold independently have optimised stability of the individual domains and then have slow pairing of these stable domains as a final step during folding. Other proteins may decrease the stability of their individual domains and use the domain interactions to stabilise the whole protein by more tightly linking the domains (Jaenicke, 1999). The kinetics of folding of the multidomain protein spectrin point to the latter mechanism, it was found that the domain domain interactions had profound effects on the folding of the proteins domains (Batey *et al*., 2006). The two phases of refolding of a two-domain spectrin unit where shown not to be due to the folding of the two individual domains, which was consistent with the equilibrium data that indicated that the folding was a cooperative all-or-nothing event (Batey *et al*., 2006).

1.3 *Cis-trans* prolyl peptide bond isomerization and protein folding

Gene-3-protein was shown to have domains N1 and N2 that fold independently of one another but the slowest step in the folding of the protein was the domain docking or association, this association reaction was accelerated by mutating the Pro213 to a glycine residue (Martin and Schmid, 2003c). The effects of this mutation indicated that the biggest rate-limiting step in the folding of this protein was the *trans* to *cis* isomerization of the Gln212-Pro213 peptide bond (Martin and Schmid, 2003c). One way that parallel folding species can arise during protein folding is due to the presence of different *cis-trans* conformations about proline peptide bonds (Schmid, 1986a; Schmid, 1992; Wedemeyer *et al*., 2002). Parallel folding channels which has been shown to occur in the folding of *Escherichia coli* thioredoxin (Georgescu *et al*., 1998; Kelley and Richards, 1987), bacteriophage T4 thioredoxin (Borden and Richards, 1990), full length α-subunit of tryptophan synthase (Bilsel *et al*., 1999), ribonuclease A (Schmid, 1986b; Houry and Scheraga 1996) and fragments of the α-subunit of tryptophan synthase (Zitzewitz and Matthews, 1999) to name a few.

It was shown by kinetic studies that proteins fold along several pathways to achieve the native state and not just one, due to the presence of different isomers of prolyl peptide bonds (Schmid, 1992). Proline is unique in that the Xaa-proline peptide bond can exist in either the *cis* or the *trans* conformation. This is unique, as the *cis* conformation, of non-proline peptide bonds, is unstable relative to the *trans* conformation. Difference in conformation of the Xaa-proline peptide bond in the unfolded state is one way in which different unfolded species can arise (Schmid, 1986a; Wedemeyer *et al*., 2002). A particular Xaa-proline peptide bond can be in either the *cis* or the *trans* conformation in the unfolded protein although the *trans* isomer is the favoured conformation (Hinderaker and Raines, 2003). This was shown to be because of steric effects but also the $n \rightarrow \pi^*$ interaction between the oxygen of the peptide bond and the subsequent carbonyl carbon in the polypeptide (Hinderaker and Raines, 2003). In the native protein, one of the

isomers will be favoured or stabilized due to structural constraints that exist for the native protein structure formation (Schmid, 1992). Ten to thirty percent of the species are present in the *cis* isomeric form in the unfolded state of the protein (Brandts *et al*., 1975; Schmid, 1992). This means that for proteins where the *cis* isomer is present in the native protein, isomerization of the other 90-70 % of the species must take place during folding (Brandt *et al*., 1975). The identification of the peptide bonds that undergo isomerization during folding is important in assessing the role certain residues play in protein folding as well as determining the folding pathways of proteins (Eyles and Gierasch, 2000; Wu and Matthews, 2002). The amino acid residue that precedes the proline residue (i.e. N-terminal residue to the proline) in the *cis* conformation affects the rate of isomerization (Reimer *et al.,* 1998), the more bulky the residue the slower isomerization (Brandts *et al*., 1975). All the folding species do not require the native isomer of the Xaa-proline bond before folding can proceed, although the incorrect isomer will decrease the rate of folding (Schmid, 1986b).

In the case of proline, both isomers are isoenergetic (Wedemeyer *et al*., 2002). The factors that result in kinetically distinct folding species include, a high energy of interconversion (20 kcal.mol^{-1}) (Brandts *et al*., 1975). One other factor is that the unfolded protein with the non-native form of the Xaa-proline peptide bond is required to undergo an isomerization reaction during folding to the native state. The isomerization reaction (time constant of 10 to 1000 seconds at 25 ºC) is slower than the conformational folding of the protein (Brandts *et al*., 1975; Schmid and Baldwin 1987; Schmid, 1992; Stewart *et al*., 1990). It involves the disruption of the partial double bond character of the peptide bond (Stewart *et al*., 1990; Wedemeyer *et al*., 2002). Prolines also play a role in folding as they decrease the conformational entropy of folding. When the Xaa-proline bond is in the *cis* conformation, it restricts the number of conformations that the unfolded polypeptide can adopt on folding to the native state.

1.3.1 *Cis*-prolines and folding nuclei

Small regions of local structure are said to form first during folding, these regions or nuclei then initiate the cooperative folding process of the entire protein as it folds to the native state. The folding proceeds from the small structural nuclei via larger structured regions until the protein is fully folded. This closely resembles the folding as described by the diffusion-collision model (Karplus and Weaver, 1994). Prolines are said play a role in forming the small structural folding nuclei. The Xaa-proline residue in the *cis* conformation, forms a tight bend thus restricting the number of conformations the polypeptide backbone can adopt as well as controlling local structural changes. Kelley and Richards (1987) demonstrated that thioredoxin *cis* Pro76 plays a role in forming these folding nuclei. Proline mutants (specifically alanine mutants) would block the formation of the *cis* conformation so destroying the folding initiation site of the local environment surrounding the proline (Dodge and Scheraga, 1996). A folding initiation site was proposed for ribonuclease A in the region of Pro93, whose preceding peptide bond adopts the *cis* isomeric form in the native state (Dodge and Scheraga, 1996).

Proline to alanine mutants of CRABPI which has three prolines all in the *trans* conformation also indicated that even *trans* isomers acted as folding initiation sites. The P105A mutant showed no change in (un)folding, it is in a loop region that is flexible and therefore accommodated the new *trans* conformation. The P85A mutant displayed kinetics where the slowest phase was lost. All the Xaa-proline bonds were in the *trans* conformation so the slowest phase of folding involved a *cis* to *trans* isomerization. The isomerization from *cis* to *trans* is the rate-limiting step in the folding of CRABPI and is therefore critical in folding. It is for this reason that Pro85 is said to be a critical folding initiation site.

1.3.2 Essential and non-essential proline residues

Prolines that form folding initiation sites would be considered essential prolines to critical proline residues. Xaa-proline bonds that are not in the native

conformation can have different effects on protein folding (Schmid and Baldwin, 1978). Certain proline residues have no effect on folding (nonessential prolines) and others slow the folding of proteins (essential prolines) while others disrupt protein folding (critical prolines) (Wedemeyer *et al*., 2002). The *cis*-proline peptide bonds are said to play an important structural role in bends and turns (Stewart *et al*., 1990). The essential and critical proline residues are normally present in positions in the polypeptide chain that will form bends and turns in the native protein. Other prolines are in positions that can accommodate either the *cis* or the *trans* isomer at that position of the native structure (Schmid and Baldwin, 1978). Prolines in flexible regions did not impact folding or their mutation to alanine or glycine whereas residues in less flexible turn regions affected (un)folding as was shown for staphylococcal nuclease A (Ikura *et al*., 1997). Mutants of prolines residues that are present in structurally constrained regions often result in destabilization of the protein and changes in the kinetics of folding. Pro40 for thioredoxin, located close to the active site, assumes *cis* conformation as it creates a sharp turn. A proline to alanine mutant of Pro40 resulted in a destabilization of mutant protein (De Lamotte-Guéry *et al*., 1997). If the mutation allows the peptide bond to maintain the conformation of the native wild-type protein then often no change is observed for the mutant. CRABP1 P85V mutant did not result in a loss of the slowest refolding phase (Burns *et al*., 2003) contradicting previous evidence that isomerization of this proline results in the slow refolding phase. The lack of correlation was explained due to the *cis*-peptide bond being maintained in the mutant (Burns *et al*., 2003). The Pro34 to serine mutant of thioredoxin (proline residue located in the active site sequence Cys32-Gly33-Pro34-Cys35) had only small effects on equilibrium and kinetics data for oxidized protein. This was said to be because the conformation of the proline peptide bond is maintained due to low flexibility of the region due to the disulphide bond (Kelley and Richards, 1987).

The evidence for essential and non-essential prolines can be found for proteins that possess more than one proline residue. During the folding of CRABP1 only the Pro105 residue was shown to have no effect on the folding kinetics of the

protein while Pro39 and Pro85 played roles as chain initiation sites (Eyles and Gierasch, 2000). The Pro39 residue of CRABP1 was shown to restrict the sampling of the conformational space, resulting in accelerated folding, as a number of non-productive folding routes are avoided (Eyles and Gierasch, 2000). Staphylococcal nuclease contains both an essential and a non-essential proline residue. Mutational studies for proline mutants of the Pro47 had little effect but studies on Pro117 mutant indicated that this proline is situated in an inflexible region, as there were changes due to the mutations (Ikura *et al*., 1997). Another protein that contains both essential and non-essential *cis* prolines is carbonic anhydrase II. Mutation of the Pro30 resulted in severely destabilised protein whereas mutation of the Pro202 residue had no effect on the refolding kinetics of the protein (Fransson *et al*., 1992).

Essential prolines have only been identified where isomerization of the proline residues from *trans* to *cis* conformation occurs. An example of a protein that has an essential proline in the *trans* conformation in the naive state where isomerization of *cis* to *trans* isomers occurs during folding is chymotrypsin inhibitor 2 (Tan *et al*., 1997). The isomerization of Pro6 resulted in slow refolding species for this protein.

1.3.3 Isomerisation within folding intermediates

Peptidyl prolyl isomerization may not result in the formation of rate-limiting intermediates in protein folding, although the isomerization plays a central role in the folding of proteins (Utiyama and Baldwin, 1986). Schmid (1986b) proposed that the rates of isomerization during protein folding can be increased or decreased by the conformation of the protein. Isomerization of prolyl peptide bonds may take place within native structural intermediates that have formed during folding (Chaffotte *et al*., 1997; Cook *et al.*, 1979; Kiefhaber *et al*., 1992b). During the folding of the α subunit of Trp synthase, it was shown that the formation of a stable intermediate provides the free energy necessary for the slow isomerization reaction of the Asp27-Pro28 peptide bond (Wu and Matthews,

2002). Borden and Richards (1990) showed that isomerization of an X-proline peptide bond occurs in a compact structural form during the folding of phage T4 thioredoxin. In the folding of antibody scFv fragment, it was shown by H/D experiments and ESI-MS that the folding of the core of one of the domains occurred prior to isomerization (Jäger and Plückthun, 2000). The domains of the antibody scFv fragment fold prior to isomerization but the isomerization results in the correct interface formation, indicating that a proline residue can be essential but does not prevent structural folding.

The formation of structure can either accelerate isomerization (Schmid, 1986b) or slow the isomerization down a great deal (Kiefhaber and Schmid, 1992). With the proline to alanine mutants of CRABPI, which has three prolines all in the *trans,* the folding of P39A is slowed compared to the wild-type. For this reason, it was stated that the isomerization forms an early structure that then restricts certain folding conformations (more specifically is a helix termination signal) resulting in faster folding than that shown by the mutant (Eyles and Gierasch, 2000).

1.4 Peptidyl prolyl *cis-trans* isomerases

It is not surprising that due to the rate limiting significance of *cis-trans* isomerization during folding that enzymes known as peptidyl prolyl *cis-trans* isomerases (EC 5.2.1.8) exists that catalyse the slow isomerization steps in protein folding (Göthel and Marahiel, 1999; Schiene and Fischer, 2000). This catalysis is necessary to prevent aggregation and proteolysis and allows functional protein to be produced (Walsh *et al*., 1992). The PPIs therefore have important roles to play in the folding and trafficking of proteins (Walsh *et al*., 1992). There are three families of PPIs (Göthel and Marahiel, 1999). The families are the cyclophilins, FKBP's and parvulins (Göthel and Marahiel, 1999; Schiene and Fischer, 2000). The cyclophilins bind cyclosporine A, the FKBP's bind FK506 and rapamycin, while the parvulins bind and are inactivated by juglone, a 5-hydroxy-1,4-naphthoquinone (Göthel and Marahiel, 1999).

Trigger factor is homologous to the FKBP proteins with three domains and a molecular mass of 48 kDa (Zarnt *et al.,* 1997) although *Escherichia coli* trigger factor shows no affinity for FK506 (Göthel and Marahiel, 1999). Trigger factor is not found in eukaryotes (Göthel and Marahiel, 1999) and is associated with the large ribosomal subunit of prokaryotic ribosomes and binds newly synthesized polypeptide chains until they are released from the ribosome (Hartl and Hartl, 2002; Schiene and Fischer, 2000). The binding of trigger factor to the polypeptide chains is based on the recognition of stretches of hydrophobic amino acids rather than proline residues (Hartl and Hartl, 2002). As well as having PPI activity, trigger factor also has the ability to hold newly synthesized polypeptides in a competent state for folding until they are released from the ribosome (Hartl and Hartl, 2002). Trigger factor therefore does not only act as a PPI but also plays a role as a chaperone in catalyzing protein folding (Liu and Zhou, 2003).

Members of the FKBP family of PPIs are found in all organisms, range in size form 12 kDa to 52 kDa and are inhibited by makrolides FK506 and rapamycin (Göthel and Marahiel, 1999). FKBP proteins are found in the cytoplasm, localized in the endoplasmic reticulum as well as membrane-localized (Walsh *et al*., 1992; Göthel and Marahiel, 1999). These PPIs have been found to be localized to the large subunit of prokaryotic ribosomes (Schiene and Fischer, 2000). FKBP-12 is an abundant, cytosolic, protein composed of a 108 amino acid residues (Göthel and Marahiel, 1999). The amino acid sequence is highly conserved between human FKBP-12 and rat FKBP-12 and between human FKBP-12 and rabbit FKBP-12 (Göthel and Marahiel, 1999). FKBP-12 protein is a peptidyl prolyl isomerase that is said to play an integral role in channel gating for the intracellular calcium release channel complex (Göthel and Marahiel, 1999).

The importance of these catalytic enzymes within the cell is to produce fully functional proteins molecules. Their use however in elucidating folding pathways is to accelerate the rate of *cis-trans* isomerization folding events. Peptidyl- prolyl *cis-trans* isomerase purified from pig kidney was shown to catalyze the slow steps

of refolding for a number of proteins (Lang *et al*., 1987). These slow steps had previously been identified as involving a *cis-trans* isomerization. By catalyzing only the steps involving isomerization they can be used to identify the isomerization step during folding.

1.5 Conserved *cis*-proline in the thioredoxin-like superfamily

The thioredoxin superfamily comprises 15 protein families. These 15 families include: the thioltransferases, protein disulphide isomerases, calsequestrin, disulphide-bond formation facilitator DsbA, N-terminal domain (domain 1) of glutathione S-transferases (GSTs), phosducin, domain 1 of endoplasmic reticulum protein ERP29, splicesomal protein U5-15kD, SH3-binding, glutamic acid-rich protein-like, Circadian oscillation regulator KaiB, C-terminal domain or domain 2 of disulphide-bond isomerase DsbC/DsbG, glutathione peroxidases, thioredoxin-like 2Fe-2S ferredoxin, arsenate reductase ArsC and thioredoxin containing proteins Txn15 (http://scop.mrc-lmb.cam.ac.uk/scop/data/scop.b.d.fh.b.html, Murzin *et al*., 1995). The thioltransferase family of proteins includes thioredoxin and small glutaredoxin proteins (http://scop.mrc-lmb.cam.ac.uk/scop/data/scop.b.d.fh.b.html, Murzin *et al*., 1995). Within the GST family, there are 16 classes. These classes include the GSTs of which there are 11 classes: the Pi, Mu, Alpha, Theta, Sigma, Omega, Zeta, Delta, Tau, Phi and Beta. Then the five other classes that are found with in the GST family include *Plasmodium falciparum* GST, the yeast prion protein (yeast prion protein 2), Grx2, GST-like domain of elongation factor 1-Gamma, and chloride intracellular channel 1 (Clic1) (http://scop.mrc-lmb.cam.ac.uk/scop/data/scop.b.d.fh.b.f.html, Murzin *et al*., 1995).

A structure based sequence alignment of members of the GST family of proteins shows the conserved proline residue (Figure 2). This is in agreement with a published alignment that indicate that there is a conserved *cis*-proline at the

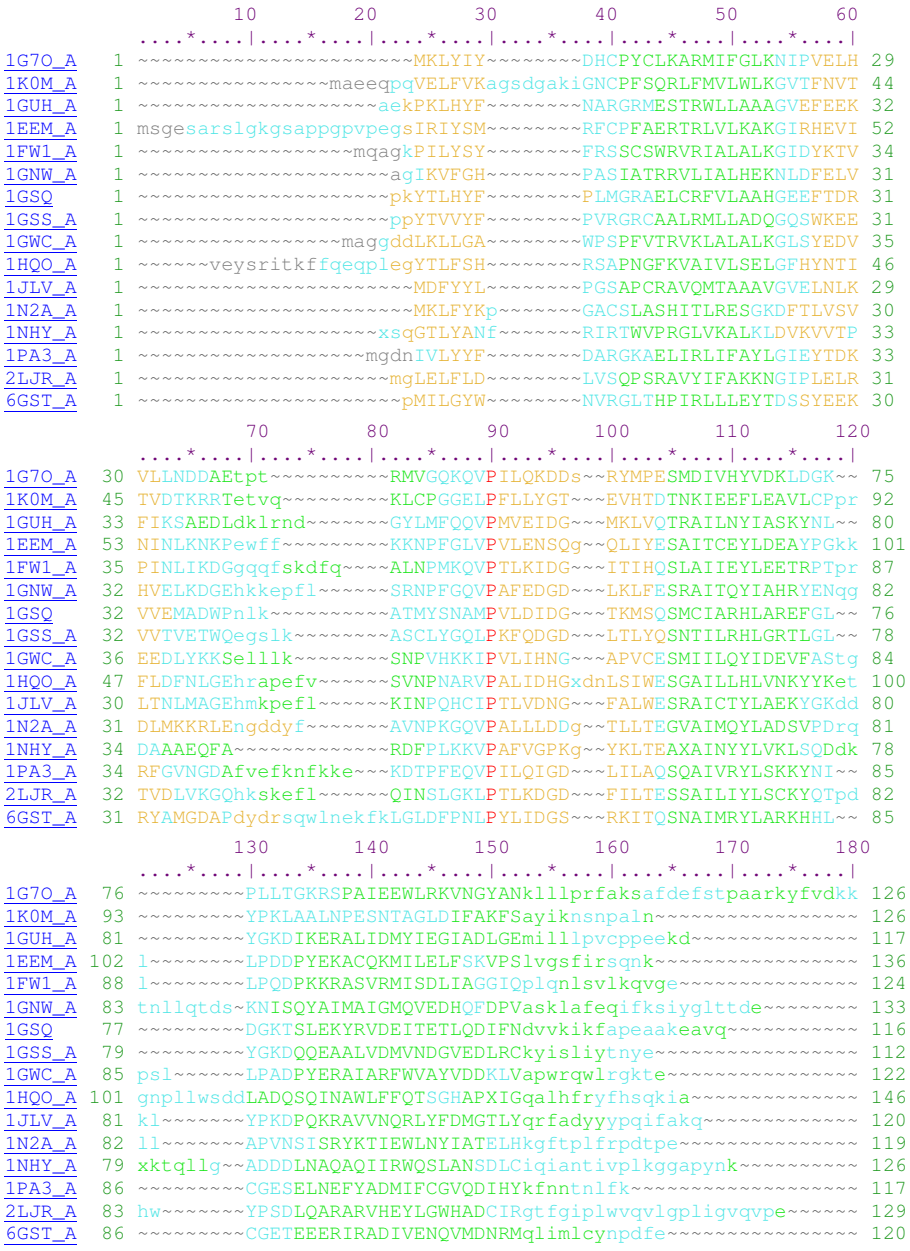


Figure 1: Structure based sequence alignment.
Alignment was conducted with sequences of GST family members (A chain unless otherwise stated), was done using Cn3D 4.1 (Wang *et al.*, 2000). The resulting alignment shows clearly the secondary structural elements of each protein, helices are in green, β-strands in orange and other structures in blue. Amino acid sequences were obtained from the NCBI (http://www.ncbi.nlm.nih.gov) the Molecular Modeling Database (MMDB)(Marchler-Bauer *et al.,* 2005) using the PDB accession codes (Table 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 proline residues are in red.

Table 1: Proteins used for the sequence alignment

The class of GST to which the proteins selected for the structure-based sequence alignment (Figure 1) are listed below as well as the reference for each of the structures. A protein from each one of the 16 classes of GST was selected for the alignment, from SCOP (http://scop.mrc-lmb.cam.ac.uk/scop/data/scop.b.d.fh.b.html, Murzin *et al*., 1995).

PDB code	GST class	Reference for structure
1G7O	Grx2	Xia *et al*., 2001
1K0M	chloride intracellular channel 1 (Clic1)	Harrop *et al*., 2001
1GUH	Alpha	Sinnig *et al*., 1993
1EEM	Omega	Board *et al.,* 2000
1FW1	Zeta	Polekhina *et al.,* 2001
1GNW	Phi	Reinemer *et al.,* 1996
1GSQ	Sigma	Ji *et al*., 1994
1GSS	Pi	Reinemer *et al.,* 1992
1GWC	Tau	Thom *et al.,* 2002
1HQO	yeast prion protein ure2p (nitrogen regulation fragment)	Umland *et al.,* 2001
1JLV	Delta	Oakley *et al.,* 2001
1N2A	Beta	Rife *et al.,* 2003
1NHY	GST-like domain of elongation factor 1-gamma	Jeppesen *et al.,* 2003
1PA3	*Plasmodium falciparum* GST	Perbandt *et al.,* 2004
2LJR	Theta	Rossjohn *et al*., 1998
6GST	Mu	Ji *et al*., 1994

N-terminus of β-strand 3, for most of the members of the thioredoxin superfamily (Nathaniel *et al*., 2003). Spliceosomal protein U5-15kD and seven of the 10 members of the glutathione peroxidase-like family did not have the conserved *cis*-proline (Nathaniel *et al*., 2003). The conserved *cis*-Proline loop falls between α2 and β3 secondary structural elements of the thioredoxin fold (Martin, 1995). GST A1-1 conserved *cis* bond is between Val54-Pro55 (Sinning *et al.,* 1993). Clic1 2,3,4,5, Omega, Pi and Beta GSTs all have the conserved *cis*-proline residue (Cromer *et al.,* 2002). The conserved *cis*-proline is required for the members of the thioredoxin superfamily that have short sequences connecting these two structural elements, and thus requiring the sharp bend the *cis*-proline residue offers (Nathaniel *et al*., 2003). *Cis*-proline peptide bonds are found primarily in bends and turns in protein structures, indicating this is their specific structural role (Stewart *et al*., 1990). Structural comparisons of the glutathione transferases indicate that, where the loop that connects the α-helix 2 and the β-strand 3 of the thioredoxin fold, there is a highly conserved *cis*-proline residue (Dirr *et al*., 1994; Sheehan *et al*., 2001). This *cis*-proline loop plays an important role in maintaining the protein's catalytically active structure (Allocati *et al*., 1999; Sheehan *et al*., 2001).

Grx2 shows low sequence identity to any other proteins as was shown by CE (Shindyalov and Bourne, 1998). The results of a search for structural neighbours Grx2 indicated that the highest sequence identity is 22.9 % (for a T4 bacteriophage glutaredoxin). Grx2 does however show structural similarity to glutathione S-transferases (GSTs) (Xia *et al*., 2001) and is said to represent the ancestral precursor of the conical GSTs (Ladner *et al.,* 2004). Cytosolic GSTs are dimeric enzymes, with each subunit possessing two domains (Dirr *et al*., 1994). The structural similarity of Grx2 and GSTs is between a single subunit of GST and the monomeric Grx2. The domain 1 of Grx2 and cytosolic GSTs both have a four stranded mixed β-sheet and three α-helices (Dirr *et al*., 1994; Sinnig *et al*., 1993; Xia *et al*., 2001). The domain 2 of Grx2 is composed of α-helices and 2 short 3_{10} helices joined by loops (Xia *et al*., 2001), and GSTs are mainly α-helical

with the helices joined by short connections (Sinnig *et al*., 1993). The overall fold of human Theta GST and Omega GST is very similar to that of Grx2 although parts of the structure are derived from different parts of the sequence (Xia *et al*., 2001). Theta GST is said to be an ancestral precursor to the dimeric GSTs (Pemble *et al.,* 1996)

1.6 Glutaredoxin 2

Glutaredoxin (Grx) was discovered as a component of a hydrogen donor system for *Escherichia coli*, which did not involve thioredoxin (Holmgren, 1976). Grx is reduced by glutathione (GSH), which in turn is reduced by NADPH and glutathione reductase. Reduced glutaredoxin is then involved in the reduction of ribonucleotide-diphosphate reductase (RR) (Holmgren, 1976) (Figure 2). Later it was shown that *Escherichia coli* has 3 glutaredoxins, namely Grx1, Grx2 and Grx3 (Aslund *et al.,* 1994).

Grx2 a cytosolic glutaredoxin with a molecular mass of 25 kDa, (Vlamis-Gardikas *et al*., 1997) but does not show activity as a hydrogen donor for ribonucleotide reductase which is atypical for glutaredoxins (Aslund *et al*., 1994). Grx2 contains the active site sequence Cys-Pro-Tyr-Cys (Vlamis-Gardikas *et al*., 1997) similar to other glutaredoxins and identical to Grx1 and Grx3 of *Escherichia coli* (Holmgren and Aslund, 1995). Grx2 has a high redox potential and oxidoreductase activity (Vlamis-Gardikas *et al.,* 1997) and catalyses the general GSH-disulphide oxidoreductase assay with β-hydroxyethyl, a small disulphide, as a substrate (Holmgren and Aslund, 1995; Vlamis-Gardikas *et al*., 1997). Grx2 also catalyses the reduction of arsenate by ArsC to arsenite, a product that is extruded from the cell (Shi *et al*., 1999). The reduction of PAPS reductase (Vlamis-Gardikas *et al*., 2002) and the cleavage and formation of glutathione protein mixed disulfide bonds (Lundstrom-Ljung *et al*., 1999) are carried out by all three glutaredoxins. Grx2 is a better catalyst than Grx1 and

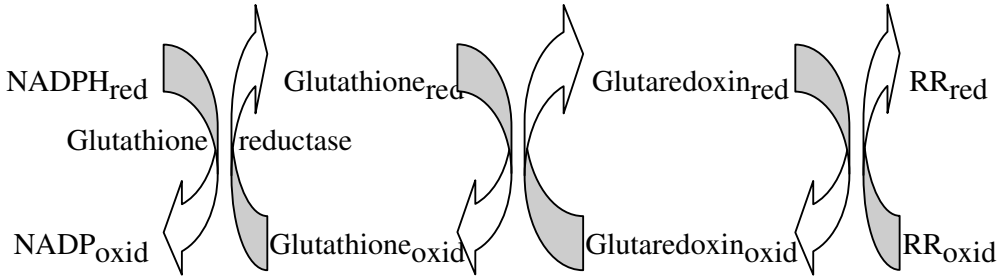


Figure 2: Schematic representation of the reduction of ribonucleotide-diphosphate reductase by glutaredoxin.
red refers to the reduced state, oxid refers to the oxidized state, and RR is ribonucleotide-diphosphate reductase.

Grx3 and protein disulfide isomerase for the reduction of glutathionylated ribonuclease A (Lundstrom-Ljung *et al*., 1999).

The Grx2 protein structure (Xia *et al*., 2001) resembles a two domain monomeric protein (Figure 3). For Grx2 to have its oxidoreductase activity, both domain 1 and domain 2 of this monomeric protein are required (Vlamis-Gardikas *et al*., 1997). Domain 1 contains residues 1 to 72 and has the typical glutaredoxin-thioredoxin fold (Xia *et al*., 2001). In Grx2 this fold 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). Domain 1 is joined to domain 2 by an 11 residue linker. Domain 2 contains residues 83 to 215 and consists of six α-helices and 2 short 3_{10} helices, these are all connected by loops. Contacts between the domains are mainly hydrophobic with some hydrogen bonds (Xia *et al*., 2001).

Grx2 has 12 proline residues, 11 of which are in the *trans* conformation. The peptide bond that is in the *cis* conformation is the Val48-Pro49 bond and this bond is located immediately before the third β strand (Xia *et al*., 2001) (Figure 3). This proline residue which is located before the third β strand is located in this same position and is in the same conformation in all 3 glutaredoxins (Xia *et al*., 2001) it represents the conserved *cis*-proline found in the thioredoxin family (see section 1.5 and Figure 1). There are two tryptophan residues in Grx2, Trp89 and Trp190 both are located in the domain 2 and two cysteine residues (Cys9 and Cys12), which are structurally close to the *cis*-proline (Figure 3).

1.7 The thioredoxin family folding and isomerization

A highly cooperative equilibrium unfolding transition was obtained with guanidine hydrochloride as denaturant and a single unfolding kinetic phase (Kelley and Stellwagen, 1984). The refolding kinetics however dispute this cooperative two-state process with three refolding phases occurring, the slowest

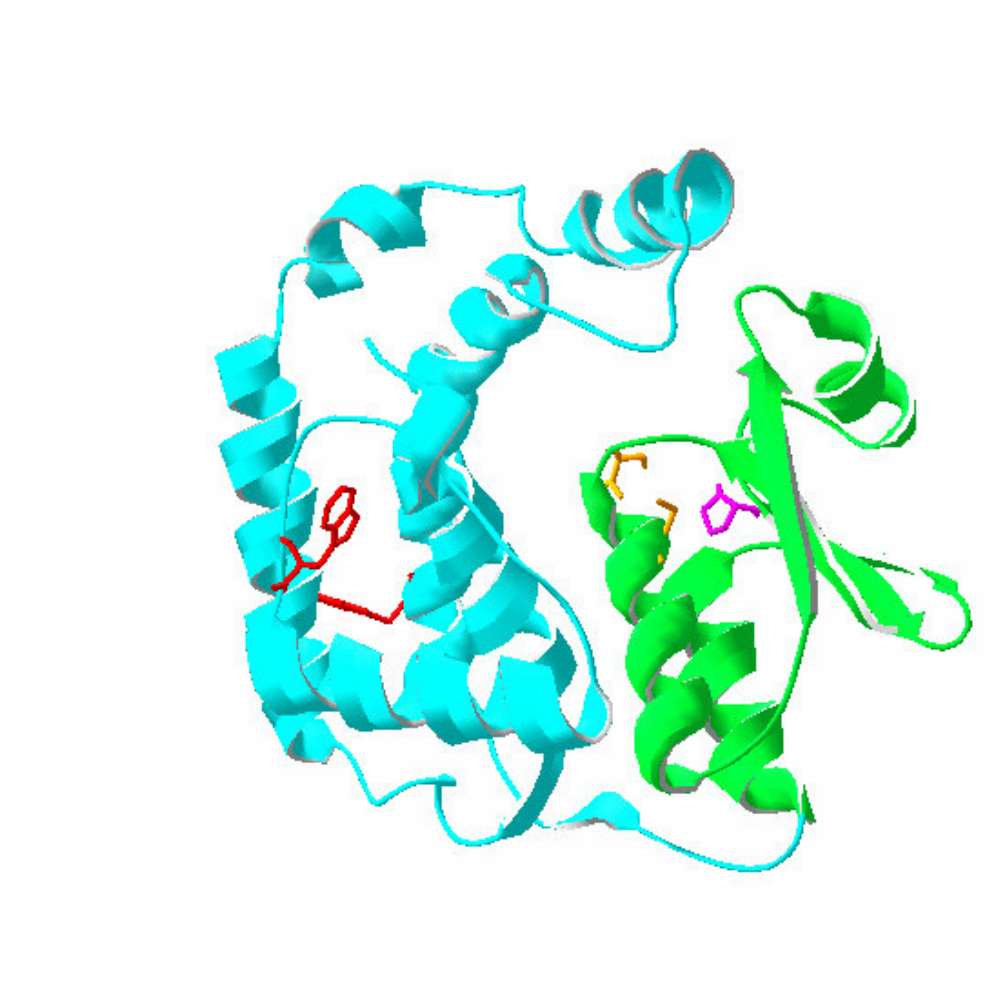


Figure 3: Structure of wild-type glutaredoxin 2
Structure was obtained from the protein data bank (1g70) and viewed with Swiss PDB viewer (ver 3.7) (Guex and Peitsch, 1997), showing the two domains, domain 1 (green) and the domain 2 (blue). The Pro49 residue, which is in the *cis* conformation, is represented in pink. The cysteine residues are yellow and the tryptophan residues are red in colour.

of which involves an isomerization (time constant of 300 to 800 sec) (Kelley and Richards, 1987; Kelley and Stellwagen, 1984). Kelly and Richards (1987) identified an intermediate for the folding of this protein and conducted mutational studies for thioredoxin. An alanine and a serine were substituted at Pro76 these mutants are destabilized and the slowest phase of refolding is abolished. The Ile75-Pro76 peptide bond is in the only one of the five X-proline peptide bonds in the *cis* conformation in the native wild-type protein (Eklund *et al*., 1984). The results indicate the Pro76 plays an important role in the folding of this protein although is not required to be in the *cis* conformation for the association of the two subdomains of thioredoxin (Chaffotte *et al*., 1997). Georgescu *et al*. (1998) identified more than the three refolding phases for *Escherichia coli* thioredoxin, a burst phase and four observable refolding phases (very rapid, rapid, medium, and slow phases). The burst phase species were heterogeneous and the slow phase was confirmed to be the isomerization phase for the Ile75-Pro76 peptide bond. A parallel folding pathway is proposed for the folding of thioredoxin, where the isomerization plays a role in forming the heterogeneous unfolded population (Georgescu *et al*., 1998). The folding of thioredoxin was further dissected and it was found that a folding nucleation site exists, this site corresponds to the β_2 and β_4 strands that zip together forming a β-sheet (Tasayco *et al*., 2000). It has been discussed how X-proline peptide bonds can undergo isomerization in compact folded intermediates (see section 1.3.3), which is said to be the case for thioredoxin (Chaffotte *et al*., 1997). Kelley *et al.* (1986) had also identified two refolding intermediates during the folding of thioredoxin that are said to form prior to isomerization.

Disulphide-bond isomerase A (DsbA), a member of the thioredoxin-like superfamily, has a conserved *cis*-proline. This Pro151 residue is involved in the stability and catalytic function of DsbA (Charbonnier *et al*., 1999). The evidence for the involvement of this residue was obtained by creating a mutant (proline to alanine) which showed a decreased stability and low activity when compared to the wild-type protein (Charbonnier *et al*., 1999).

1.7.1 Glutathione transferase family members and folding

The equilibrium unfolding of yeast prion protein 2 displayed single sigmoidal transitions, which were protein concentration independent indicting that unfolding and dimer dissociation is not linked (Perret *et al.,* 1999). An additional protein concentration dependent transition was later observed for the unfolding of this protein indicating that dimer dissociation occurs later during unfolding (Thual *et al.,* 2001). The unfolding of this protein was therefore said to involve the unfolding of the protein prior to dissociation of the dimer to form monomeric forms. A stable unfolding intermediate was also identified although the pathway for (un)folding could not be determined from these equilibrium data (Thual *et al.,* 2001). The intermediate detected by equilibrium unfolding was later identified during kinetic unfolding studies and multiple kinetic phases observed during the refolding of the protein (Galani *et al*., 2002). These multiple refolding phases are not only due to the presence of different *cis-trans* proline isomers this protein that contains 12 proline residues one of which is in the *cis* conformation in the native state but also due to the formation of dimeric intermediates of yeast prion protein 2, which were not limited by proline isomerization (Galani *et al*., 2002). During the folding of yeast prion protein 2, there is competition between folding via an on-pathway dimeric intermediate and off-pathway aggregation (Galani *et al*., 2002). The aggregation pathway is facilitated by the formation of a monomeric aggregation prone species (Galani *et al*., 2002). Isomerization does however occur late in refolding following the formation of a dimeric on-pathway intermediate.

The unfolding of human GST A1-1 was well described by a sequential unfolding pathway where the domains of the subunits dissociate followed by complete unfolding of the domains and subunits (Wallace *et al.,* 1998). The biphasic kinetics of unfolding was not expected for the two-state equilibrium unfolding transition (Wallace *et al.,* 1998). The refolding of human GST A1-1 also does not involve a simple two-state transition from unfolded monomer to folded dimer (Wallace and Dirr, 1999). During the early stages of folding, structural monomers are formed which have native secondary structure and exposed hydrophobic surfaces (Wallace and Dirr, 1999). Dimerization then occurs by parallel

intermediate and fast folding events. A *trans*-to-*cis* isomerization event occurs that limits the rate of dimerization (Wallace and Dirr, 1999). This isomerization event is associated with the intermediate phase (also an association reaction during this phase) during folding and is said to result in the parallel refolding fast and intermediate phases (Wallace and Dirr, 1999). Once dimerization has taken place, a slow folding phase takes place. It is during this phase that the GST A1-1 undergoes the final reorganization to form functional protein (Wallace and Dirr, 1999).

A P56G mutant of GST A1-1 indicated that the *cis*-proline is involved in maintaining the conformational stability as well as catalytic function of the protein (Nathaniel *et al*., 2003). A different scenario arose for glutathione transferase B1-1 from *Proteus mirabilis*. A P53S mutant had refolding and unfolding kinetics, which were indistinguishable from the wild-type protein and there was no evidence that the mutant destabilized the protein. The mutant did however result in lower antibiotic binding and showed small local structural changes in region of the Ser53 (Allocati *et al*., 1999).

1.8 Objectives

The principle out come of this study is an unfolding/refolding pathway s for the monomeric, GST homologue Grx2. The unfolding/refolding pathway for the thioredoxin, yeast prion protein 2 and GST A1-1 indicate that the folding of these related proteins does not simply involve single transition from unfolded to native protein (see section 1.7). Wild-type Grx2, AEDANS labelled Grx2 and a mutant, Y58W Grx2 were used in order to study the (un)folding of Grx2. Single-jump and double-jump kinetic studies also shed light on the unfolding/refolding pathway. The involvement of a possible isomerization of the Val48-Pro49 *cis* peptide bond was also assessed in the folding of the protein. For many of the structurally related proteins an isomerization of the conserved *cis*-proline as well as the presence of refolding intermediates influenced the folding of these proteins.

Identification of potential kinetic intermediates and the positioning of these intermediates in terms of folding reactions such as isomerization were conducted. Grx2 is composed of two domains that could affect the (un)folding of this protein. The exact effect in terms of the unfolding and refolding was to be assessed. One further question that was answered about these domains is whether they could result in separate folding pathways. If these domains do result in separate folding pathways, would a parallel or sequential folding pathway result?

2 EXPERIMENTAL PROCEDURES

2.1 Materials

Ultrapure (99.5 %) urea was purchased from BDH laboratory supplies (Poole Dorset, U.K.) and ultrapure guanidine hydrochloride was purchased from Roche Diagnostics (Mannheim, Germany). DEAE-Sepharose, IAEDANS, L-ascorbic acid, and DCIP were purchased from Sigma (St. Louis, MO USA). DTNB was purchased from Boehringer Mannheim (Ingelheim, Germany). DTT was purchased from Whitehead Scientific (Cape Town, South Africa) or Inqaba Biotech (Pretoria, South Africa). All other reagents were of analytical grade. The pET24a plasmid encoding the cDNA sequence for Grx2 was a generous gift from Dr. J. Dyson (The Scripps Institute, California, USA) (Xia *et al*., 1999). Purified hFKBP-12 was a generous gift from Dr. Jens-U Rahfeld (EMBL, Heidelberg Germany). Inqaba Biotech (Pretoria, South Africa) conducted all sequencing to confirm the identity of the plasmids.

2.2 Mutagenesis to generate the Y58W variant of Grx2

The Y58W mutant was introduced into the Grx2 coding region of the pET24a plasmid using the QuikChange™ Site-directed Mutagenesis kit (Papworth *et al*., 1996) from Stratagene (La Jolla, USA). Oligonucleotide primers for the mutagenesis, were designed in accordance with the mutagenesis kit instructions, using the web based program Primer X (http://bioinformatics.org/primerx/) as well as Gene Runner computer software package v3.05. Inqaba Biotech (Pretoria, South Africa) synthesized the primers:

Y58W forward: 5' G CAA AAA GAT GAC AGC CGC **TGG** ATG CCA GAA AGC 3'

Y58W reverse: 5' GCT TTC TGG CAT **CCA** GCG GCT GTC ATC TTT TTG C 3'.

The nucleotides in bold print are the codon for a tryptophan (TGG), originally tyrosine (TAT). The mutagenesis reaction that includes, thermocycling and post

thermocycling DNA digestion, was carried out as specified by the QuikChange™ Site-directed Mutagenesis kit from Stratagene (La Jolla, USA) (Papworth *et al*., 1996). The amount of double stranded template DNA used for the mutagenesis reaction was 20 ng and the extension time was 6 minutes at 68 ºC. Digestion of the parental DNA by the restriction enzyme DpnI, digestion was continued for a further 1 hour at 20 ºC following the specified hour long digestion at 37 °C. Plasmid encoding Y58W Grx2 was then used to transform XL1-Blue supercompetent *Escherichia coli* cells as stated in the mutagenesis kit manual. Transformed cells were grown on kanamycin containing plates from which a colony was selected. Plasmid DNA was extracted from a culture of cells grown from the selected colony, using the Flexiprep ™ Kit form Amersham Biosciences (Buckinghamshire, U.K.). The incorporation of the desired mutation in the pET24a plasmid DNA was confirmed by sequencing of the plasmid DNA, using the T7 terminator primer. The Y58W mutant plasmid DNA was then transformed into *Escherichia coli* BL21(DE3) cells to produce an expression system for the mutant protein.

2.3 Wild-type and Y58W Grx2 expression and purification

The recombinant expression system for wild-type and Y58W Grx2 is composed of the pET24a plasmid with the wild-type Grx2 gene insert or the Y58W gene insert in *Escherichia coli* strain BL21(DE3)pLysS cells or BL21(DE3) cells, respectively. Based on the expression system 30 µg/ml kanamycin was added to the growth media for both proteins and for wild-type Grx2 only, an additional 30 µg/ml chloramphenicol was added. The initial purification protocol used to obtain wild-type Grx2 involved diluting an overnight culture 50 times into 1 litre Luria-Bertani media followed by overexpression of Grx2. Overexpression was induced by addition of IPTG to cell culture (with an optical density of 0.6) to a final concentration of 1mM. Following induction the cells were grown for a further 15 hours and then harvested by centrifugation at 5000 rpm before being resuspended in 20 mM Tris-HCl buffer, pH 10, containing 2 mM magnesium chloride, 1 mM EDTA, 0.02 mg/ml DNase I and 0.02 % sodium azide. The resuspended cells

were then lysed by sonication and the lysate centrifuged at 13 000 rpm for 20 minutes. The protein was then purified using anion exchange chromatography as described previously (Aslund *et al.,* 1994; Vlamis-Gardikas *et al.,* 1997; Xia *et al.,* 2001). Bound protein was then eluted from the DEAE-Sepharose column with 50 mM Tris-HCl buffer, pH 8, containing 0.02 % sodium azide. Contaminating proteins that eluted along with wild-type Grx2 from the DEAE-Sepharose column were then removed by filtration chromatography on a G-75 Sephadex column. The column was pre-equilibrated with 50 mM sodium phosphate buffer, pH 7, containing 50 mM sodium chloride, 1 mM DTT and 0.02 % sodium azide. This purification protocol did not result in pure wild-type Grx2 and gave a low yield of 14 mg protein per litre of Luria-Bertani media, therefore it had to be optimised. Changes for the optimised protocol, which was used for the purification of wild-type and Y58W Grx2 included replacing the Luria-Bertani media with 2XYT rich media. Anion exchange chromatography was conducted using an AKTA prime system with a chart recorder (Amersham Biosciences, Vienna Austria), with this system bound wild-type Grx2 or Y58W Grx2 were eluted from the DEAE-Separose column using a pH gradient from pH 9 to pH 8. The gradient was produced by mixing 20 mM Tris-HCl, pH 9, 0.02 % sodium azide with 50 mM Tris-HCl, pH 8, 0.02 % sodium azide. No further filtration chromatography was conducted. The purified wild-type Grx2 was then dialysed and snap frozen at -70 ºC; frozen protein would then be thawed on ice and dialysed before use. Purified Y58W Grx2 once purified was dialysed and used within a week. The dialysis buffer for both proteins was a 50 mM sodium phosphate buffer, pH 7, containing 50 mM sodium chloride, 1 mM DTT and 0.02 % sodium azide. Dialysis was conducted at 4 ºC with three changes of dialysis buffer, each change lasting for 4 hours. Unless otherwise stated all experiments for wild-type and Y58W Grx2 were conducted in the dialysis buffer.

2.4 SDS-PAGE

During the purification protocols for wild-type Grx2 and Y58W Grx2 as well as for the size determination of the denatured, reduced, purified proteins, SDS-

PAGE with a discontinuous buffer system of Laemmli (1970) was used to assess the size and homogeneity of the purified proteins. SDS-PAGE gels with a 15 % separating gel and 4 % stacking gel were used. Gels were run using a Mini VE vertical electrophoresis system (Hoefer, San Francisco USA) with power pack for approximately 1 hour at 160 V to allow for complete separation of proteins. The size of the purified proteins had to be determined in relation to a set of low molecular mass standard proteins (Amersham Biosciences, Vienna Austria) run under the same denaturing, reducing conditions. The gels once run were then stained with coomasie brilliant blue stain solution for 1 hour. Stain solution is composed of 50 % (v/v) methanol, 0.05 % (w/v) coomasie brilliant blue, 10 % (v/v) acetic acid and 40 % (v/v) distilled water. Following staining, the gels were destained over night in destain solution composed of 5 % ethanol (v/v), 7 % (v/v) acetic acid and 88 % (v/v) distilled water.

2.5 SE-HPLC

The size and homogeneity of purified wild-type Grx2 was assessed using SE-HPLC. Proteins were separated using a TSKgel G 2000 SWXL size exclusion column (Amersham Biosciences, Vienna Austria) with a TSKgel SWXL guard column (Amersham Biosciences, Vienna Austria). The buffer used was a 0.1 M sodium phosphate buffer, pH 6.7, containing 0.1 M sodium sulphate and 0.05 % sodium azide. This buffer was pumped at isocratic pressure through the column using a LKB 2150 pump. Protein eluting from the column was detected using a spectraseries UV100SP absorbance detector set to record absorbance at 280 nm with a sensitivity of 0.02. The absorbance readings were recorded on a chart recorder set at a chart speed of 30 cm/hour. A set of gel filtration standard proteins (Biorad, Hercules USA) was run under the same conditions as wild-type Grx2.

2.6 Spectroscopic techniques

2.6.1 Absorbance spectroscopy

Absorbance reading were recorded using a Jasco V-550 UV-VIS spectrophotometer or a Hewlett Packard model 8452A diode array spectrophotometer. The concentration of 2-nitro-5-thiobenzoate anion, naphthalenesulphonic acid group (IAEDANS), wild-type Grx2 and Y58W Grx2, were determined using absorbance spectroscopy and the Beer-Lambert law, $A = \varepsilon cl$ where ε is the extinction coefficient, c is the concentration in M and l is the pathlength of light in centimetres. The pathlength was 1 centimetre for all cuvettes used. The extinction coefficient for the 2-nitro-5-thiobenzoate anion at 412 nm is 13 600 $M^{-1}.cm^{-1}$ (Habeeb, 1972) and for naphthalenesulphonic acid group of IAEDANS, at 338 nm is 6000 M^{-1} cm^{-1} (Hudson and Weber, 1973). Then for wild-type Grx2 the extinction coefficient at 280 nm is 21680 $M^{-1}. cm^{-1}$ (Vlamis-Garikas *et al.*, 1997). For Y58W Grx2 the extinction coefficient at 280 nm was determined to be 27670 $M^{-1}. cm^{-1}$ using the following formula described by Perkins (1986):

$$\varepsilon_{280}\ M^{-1}.\ cm^{-1} = [5550 \times (\Sigma\ \text{Trp residues})] + [1340 \times (\Sigma\ \text{Tyr residues})] + [150 \times (\Sigma\ \text{Cys residues})] \quad (1)$$

2.6.2 Circular dichroism

Far-ultraviolet circular dichroism spectra and single wavelength readings in the far-ultraviolet region, were recorded using a Jasco model J-810 spectropolarimeter. A 1 mm or 2 mm cuvette was used and readings were taken at 20 ºC with a data pitch of 0.2 nm, a response of 0.5 sec and a bandwidth of 1 nm. For the spectrum readings from 190 to 250 nm, the scanning mode was continuous with a scanning speed of 20 nm/min to 100 nm/min and an average of 10 accumulations taken per sample. When readings were taken at 222 nm wavelength only, the readings were taken over a 20 sec time period with an average of 3 accumulations for 101 data points each. Buffer contributions were subtracted for all data collected.

Data obtained was converted to mean residue ellipticity using the following formula:

$$[\theta] = 100 \text{ (signal)}/ \text{ Cnl} \qquad (2)$$

Where C is the concentration of protein in mM, n is the number of amino acid residues in the polypeptide chain and l is the pathlength in cm. The units of mean residue ellipticity are deg. cm^2. $dmol^{-1}$.

2.6.3 Fluorescence spectroscopy

Florescence spectra were recorded using a Perkin Elmer luminescence spectrometer model LS 50B. Spectra were recorded for the characterization of wild-type Grx2, Y58W Grx2 and AEDANS-labelled Grx2 as well as for urea-induced equilibrium unfolding of all proteins. The native wild-type Grx2 as well as Y58W Grx2 was excited at a wavelength 295 nm with spectra recorded from 295 nm to 400 nm. The excitation and emission slitwidths for the wild-type Grx2 were 3 nm and 3.5 nm respectively and 3 nm for both, for Y58W Grx2. AEDANS-labelled protein was excited at 338 nm and 295 nm with excitation and emission slitwidths of 3 nm. A step size of 200 nm to 300 nm was used for all samples. Excitation of the AEDANS probe occurred either directly (excitation at 338 nm) or by fluorescence resonance energy transfer (excitation at 295 nm). Spectra were therefore recorded for AEDANS labelled protein form 295 nm to 600 nm. The cuvette used for all samples had a 1 cm pathlength, spectra were recorded at 20 ºC and buffer contributions subtracted.

2.7 Cysteine modification

2.7.1 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 phosphate buffer, pH 7, containing 1 mM EDTA and 0.02 % azide. Following buffer exchange the DTNB assay was conducted by titration of a solution of 5 µM protein, 50 mM phosphate, 1 mM EDTA, 0.02 % azide, pH 7, in the absence and presence of 5.4 M guanidine hydrochloride, with 0.22 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, 1972). The concentration of the released 2-nitro-5-thiobenzoate anion was determined spectrophotometrically at 20 ºC as described previously (see section 2.6.1).

Guanidine hydrochloride used in the assay was prepared in the 50 mM phosphate buffer, pH 7, containing 1 mM EDTA and 0.02 % azide and filtered. The concentration of the guanidine hydrochloride stock solution was determined using an Atago R5000 refractometer (Tokyo, Japan). Guanidine hydrochloride was used as denaturant as it twice as efficient as urea as a denaturant (Alonso and Dill, 1991; Myers *et al*., 1995).

2.7.2 IAEDANS labelling

The reduced thiol groups of wild-type Grx2 were allowed to react with excess IAEDANS for 14 hours at 25 ºC protected from light. This resulted in the iodoacetamide group of IAEDANS reacting with the free thiol groups of the protein, and the naphthalenesulphonic acid group being covalently attached to the protein (Hudson and Weber, 1973). Following attachment any residual IAEDANS that was not bound was removed during buffer exchange on a G-25 Sephadex column. Labelled Grx2 in a 50 mM phosphate buffer, pH 7, containing 50 mM sodium chloride and 0.02 % sodium azide, was then filtered and stored at 4 ºC. Unless otherwise stated, all experiments conducted for AEDANS labelled protein were in this buffer. The concentration of the naphthalenesulphonic acid groups attached to the protein was determined spectrophotometrically as described previously (see section 2.6.1).

The absorbance of AEDANS (Hudson and Weber, 1973) overlaps with absorbance of the protein at 280 nm so therefore the concentration of protein could not be determined using the absorbance reading at 280 nm. It was for this reason that the concentration of labelled Grx2 was determined using the Bradford assay (Bradford, 1976) and bovine serum albumin. The colour change was monitored at 595 nm for a set of bovine serum albumin standards of known concentration and the labelled Grx2 sample. Concentration of the unknown labelled Grx2 was determined by comparison with the standards of known concentration.

2.8 Urea-induced equilibrium unfolding

2.8.1 Determination of reversibility of unfolding

Equilibrium between the native and the unfolded states of the proteins can only occur if the unfolding reactions are shown to be reversible. The reversibility of unfolding was determined for wild-type, Y58W and AEDANS-labelled Grx2. To determine the reversibly of unfolding, native protein was allowed to unfold in the presence of 8 M urea for 1 hour at 20 ºC before refolding was initiated by an 6 fold dilution of the unfolded protein. Refolding was allowed to proceed for 1 hour at 20 ºC prior to both native protein and refolded protein being assessed using far-ultraviolet circular dichroism (2.6.2) and intrinsic tryptophan fluorescence (2.6.3). The buffer for wild-type Grx2 is 50 mM phosphate buffer, pH 7, containing 1 mM DTT and 0.02 % sodium azide.

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

Japan) and the method of Pace (Pace, 1986). The 10 M stock solution was then frozen at -20 ºC. When the urea was to be used, it was defrosted and used the same day. Stock urea was used within a one week of being made.

2.8.2 Urea-induced equilibrium unfolding studies

Urea-induced unfolding of wild-type Grx2, Y58W Grx2 and AEDANS-labelled Grx2, was conducted at a range of urea concentrations from 0 M to 8 M urea. Unfolding was conducted at 20 ºC for 1 hour to allow equilibrium to be reached. Following 1 hour the samples for the range of urea concentrations, were monitored using far-ultraviolet circular dichroism (2.6.2) and fluorescence spectroscopy (2.6.3).

Data for all the proteins obtained using all the spectroscopic probes was analysed as two-state unfolding processes for monomeric protein. During two-state reversible unfolding there is an equilibrium reached between the unfolded species (U) and the native species (N):

$$N \rightleftharpoons U \qquad (3)$$

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

$$f_U + f_f = 1 \qquad (4)$$

where f_f 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:

$$y = y_f f_f + y_U f_U \qquad (5)$$

where y is the signal obtained for the respective spectroscopic probe, f_f represents the fraction of folded protein, f_U represents fraction unfolded protein. In addition y_f represents the y value for the folded state and can be extrapolated from linear pre-transition region of the unfolding data. The symbol y_u represents the y value

for the unfolded state and can be extrapolated from the linear post-transition region of the unfolding data. Combining equations 4 and 5, the fraction of unfolded protein can be obtained:

$$f_U = (y - y_f)/(y_U - y_f) \quad (6)$$

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

$$f_f = (y_U - y)/(y_U - y_f) \quad (7)$$

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

$$K_U = f_U / f_f \quad (8)$$

So therefore if equations 6 and 7 are substituted into 8:

$$K_U = (y - y_f) / (y_U - y) \quad (9)$$

and

$$\Delta G^\circ = - RT \ln K_{eq} \quad (10)$$

where ΔG° 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 $\Delta G(H_20)$ it is assumed that ΔG° has a linear dependence on denaturant concentration [D] for all urea concentrations (Pace, 1986). Therefore,

$$\Delta G^\circ = \Delta G(H_2O) - m [D] \quad (11)$$

where $\Delta G(H_2O)$ represents the free energy required in the absence of denaturant to destabilize the protein, m is the m-value for unfolding, and [D] is the denaturant concentration.

Combining equations 9, 10 and 11 and rearranging them gives:

$$y = [y_f + y_U * e^{-(\Delta G (H20) - m [D])/RT}]/[1 + e^{-(\Delta G (H20) - m [D])/RT}] \quad (12)$$

Data obtained was fitted to equation 12 using SigmaPlot version 8.0 (SPSS Science, Chicago, Illinois USA) and the parameters $\Delta G(H_20)$ and m were obtained.

2.9 Kinetics

All kinetic experiments were conducted using a BioSequential SX-18MV stopped-flow reaction analyzer from Applied Photophysics (Leatherhead, U.K.) The excitation pathlength is 10 mm and the emission pathlength is 2 mm. The excitation bandwidth was 2.32 nm to minimize photodecomposition. The photomultiplier voltage was set at 600 V for all experiments. The temperature of the sample handling unit was maintained at 20 ºC with a fluctuation of not more than 0.1 ºC. Analysis of all kinetic data was done using the Applied Photophysics software version 4.47 which uses the Levenberg-Marquardt algorithm (Marquardt, 1963) for non-linear least squares fitting.

2.9.1 Dead-time determination

The dead-time of the stopped-flow reaction analyzer during single mixing mode (used for single-jump studies), was determined using the reduction of 2,6-dichlorophenolindophenol (DCIP) by ascorbic acid (Tonomura *et al*., 1978). This reaction is pseudo-first order at DCIP concentrations of less than 5 % of the ascorbic acid concentration (Tonomura *et al*., 1978). DCIP was prepared in water containing 10 % propanol. Ascorbic acid was prepared by dissolving ascorbic acid in 0.02 N hydrochloric acid/0.2 M sodium chloride solution. Dilutions of the stock ascorbic acid were prepared using the 0.02 N hydrochloric acid/0.2 M sodium chloride solution. The dilutions were prepared in order to produce a range of ascorbic acid concentrations from 2 mM to 120 mM. Sodium chloride was added in order to minimize the effect of variance of ionic strength (Tonomura *et al*., 1978).

The DCIP solution was rapidly mixed with the range of ascorbic acid concentrations, in a 1 : 1 mixing ratio. The pH of the mixing solution was 2. Changes in the absorbance as the reactions occurred, was monitored at 524 nm. The resulting average of 3 traces was then fitted to a single exponential, excluding the first 2 milliseconds data, for each ascorbic acid concentration. The pseudo-

first order rate constant (k_{app}) obtained was plotted versus the ascorbic acid concentration (Figure 4 A). The plot clearly indicates that the ascorbic acid concentrations, 80 mM, 100 mM and 120 mM deviate from a linear trend (Figure 4 A). These points were therefore excluded from the plot of log of the observed absorbance versus the rate constant (Figure 4 B). The equation for the straight line fitted to this plot gives the equation:

$$\ln(\Delta A_{obs}) = \ln(\Delta A_{tot}) - k_{app} * t_d \quad (13)$$

where ΔA_{obs} is the change in the observed absorbance, ΔA_{tot} is the change in the theoretical absorbance, k_{app} is the apparent rate constant and t_d is the dead-time of the instrument. Therefore for the data obtained, the slope of the log of the observed absorbance versus the rate constant data gave a dead-time of 1.1 milliseconds for the stopped-flow instrument for non-sequential mixing. Therefore, no data within 2 milliseconds were fitted for all experiments conducted.

The dead-time for the sequential or double mixing mode of the stopped-flow reaction analyzer was given in the profiles window generated following a run, using the Applied Photophysics software version 4.47. The dead-time was displayed along with the exact drive volumes for each mixing event for that run. Data within the dead-time for each run were not included in the fitting. The average dead-time for the sequential mixing runs was 3 milliseconds.

2.9.2 Kinetics studies

All kinetic studies were conducted by monitoring changes in the intrinsic tryptophan fluorescence emission. The excitation wavelength for all experiments was 280 nm, with a cut off filter of 320 nm used to prevent excitation wavelength swamping the emission signal. It was shown prior to kinetics experiments for wild-type Grx2 and Y58W Grx2 that the proteins do not undergo any photodegradation over a period of 1000 seconds and 500 seconds, respectively.

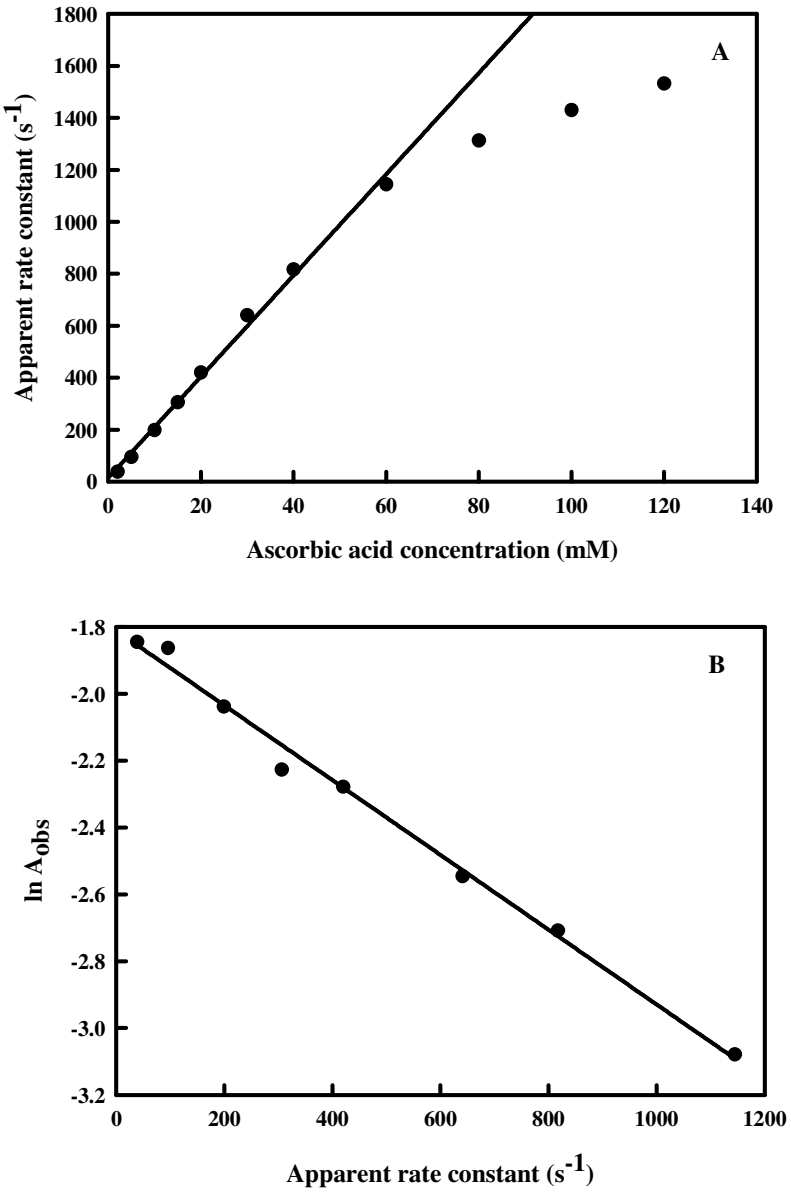


Figure 4: Dead-time determination
(A) pseudo-first order rate constant obtained (k_{app}) plotted against the ascorbic acid concentration for the reaction. Linear fit (solid line) indicating points deviating from this trend (B) log of the observed absorbance change versus the rate constant for the reaction at specific ascorbic acid concentrations. Linear fit to data indicated as a solid line.

The baseline values were obtained for 5 μM urea denatured (7.5 M urea) wild-type Grx2 and Y58W Grx2 and 5 μM native wild-type Grx2 and Y58W Grx2. Urea was prepared as stated previously (see section 2.8.1). Kinetic experiments were conducted at 20 ºC, three traces were averaged for each urea concentration or delay time of unfolding, with the final average trace analysed using the Applied Photophysics software. Data were fitted to the following general equation:

$$F_t = \Sigma F_i * \exp(-t/\tau) + y_o \quad (14)$$

where F_t is the total fluorescence amplitude, F_i is the amplitude for the phase i, 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. The fit of the data to the general equation were shown to be good by residuals plots.

2.9.2.1 Single-jump unfolding and refolding studies

To produce the unfolding reactions 30 μM native wild-type Grx2 or Y58W Grx2 was mixed in an asymmetric ratio of 1:5 with buffer containing 6 M to 9 M urea. The final unfolding conditions were 5 μM protein and urea concentrations from 5 M to 7.5 M. Refolding reactions were performed by rapidly mixing unfolded protein in a 1:5 ratio with buffer containing 0 to 3.5 M urea. Unfolded protein solution of 30 μM wild-type Grx2 or Y58W Grx2 in 6 M urea, was prepared and allowed to unfold for 1 hour at 20 ºC. The final refolding conditions following mixing were 5 μM wild-type Grx2 or Y58W Grx2 in 1 M to 4 M urea. Urea was prepared as for reversibility studies (2.8.1). Unfolding data were fitted to a double exponential function while refolding data were fitted to a triple exponential.

2.9.2.2 Kinetic double-jump refolding studies

A double-jump experiment as the name suggests involves two rapid, successive steps. The steps can be either unfolding (delay) followed by refolding or refolding (delay) followed by unfolding. The names given to these double-jump experiments are refolding double-jump (N→U→delay→N) or and unfolding

double-jump (U→N→delay→U) experiment respectively. Refolding double-jump experiments were conducted for wild-type Grx2 only. For the refolding double-jump experiment, the first jump or step is a rapid unfolding of the protein. Unfolding reactions were carried out with initial urea concentrations of 7.7 M urea and 8.4 M urea rapidly mixed with 105 µM wild-type Grx2 in a ratio of 1:2.5. The final conditions for the unfolding reactions were 30 µM unfolding wild-type Grx2 with 5.5 M urea or 6 M urea. The protein is allowed to unfold in an ageing loop for a specified delay times ranging from 10 seconds to 500 seconds before refolding was initiated by a 5 fold dilution of the unfolding reaction. The resultant refolding conditions of 5 µM wild-type Grx2 in 0.95 M residual urea or 1 M residual urea, were produced and monitored by intrinsic tryptophan fluorescence. For all conditions specified, the 5.5 M urea unfolding conditions were stated first, followed by the 6 M urea conditions. As the refolding was initiated during the double-jump experiments, the ageing loop, where the unfolding reactions take place, was flushed with 5.5 M urea or 6 M urea solution, depending on the concentration used for unfolding.

In order to perform the double-jump experiments the stopped-flow reaction analyzer was set up for sequential mixing. Refolding data were fitted to equation 14 (2.9.2) for double exponential as well as triple exponential fits depending on the delay time of unfolding. For 5.5 M double-jump experiments, refolding data for unfolding delay times shorter than 45 sec were fitted to a double exponential and for delay times greater than 45 sec to a triple exponential. Similarly for the 6 M double-jump data for unfolding delays of less than 30 seconds were fitted to a double exponential and for data of 30 sec delay times and longer a triple exponential were fitted. The fit of the data were shown to be good by residuals plots.

2.9.2.3 Kinetics of refolding in the presence of hFKBP-12

Double-jump refolding experiments using 6 M urea for unfolding as well as single-jump refolding experiments were conducted in the presence of the PPI, hFKBP-12. The double-jump refolding experiments with hFKBP-12 were conducted as stated before for 6 M urea unfolding delay experiments (see section 2.9.2.2) except 1 µM hFKBP-12 was added to the refolding buffer.

Single-jump refolding experiments in the presence of hFKBP-12 were conducted as stated previously (see section 2.9.2.1) only hFKBP-12 was added to the refolding buffer (range of concentrations form 0 µM to 3.6 µM). The final refolding conditions were 5 µM wild-type Grx2, 1 M Urea and 0 µM to 3 µM hFKBP-12. Resultant average refolding traces were fitted to a triple-exponential and the fit was shown to be the best by residuals plots.

2.10 Software for structural analysis, sequence analysis and data fitting

Unless otherwise stated all non-linear and linear least squares fitting of data were done using Sigma Plot (version 8.0) (SPSS Science, Chicago, Illinois USA). Protein illustrations were produced using Swiss PDB viewer (ver 3.7) (Guex and Peitsch, 1997). The wild-type and Y58W Grx2 DNA sequences obtained were viewed using the program Chromas version 1.45 (32 bit) (http://www.technelysium.com.au/chromas.html; Technelysium Pty. Ltd., Helensvale, Australia). The structure-based sequence alignment (Figure 1) of the GST family of structurally related proteins was done using Cn3D version 4.1 (Wang *et al*., 2000) which incorporates the structure-based sequence alignment tool SALTO (Kann *et al*., 2005) as well as the alignment refinement algorithm REFINER (Chakrabarti *et al*., 2006).

3 RESULTS

3.1 Mutagenesis to produce Y58W Grx2

Using wild-type Grx2 plasmid DNA as a template, the Y58W mutant Grx 2 was created by site-directed mutagenesis (section 2.2). Sequencing of the resultant plasmid DNA indicated that the desired mutation had been incorporated and no additional mutations were incorporated during mutagenesis (Figure 5).

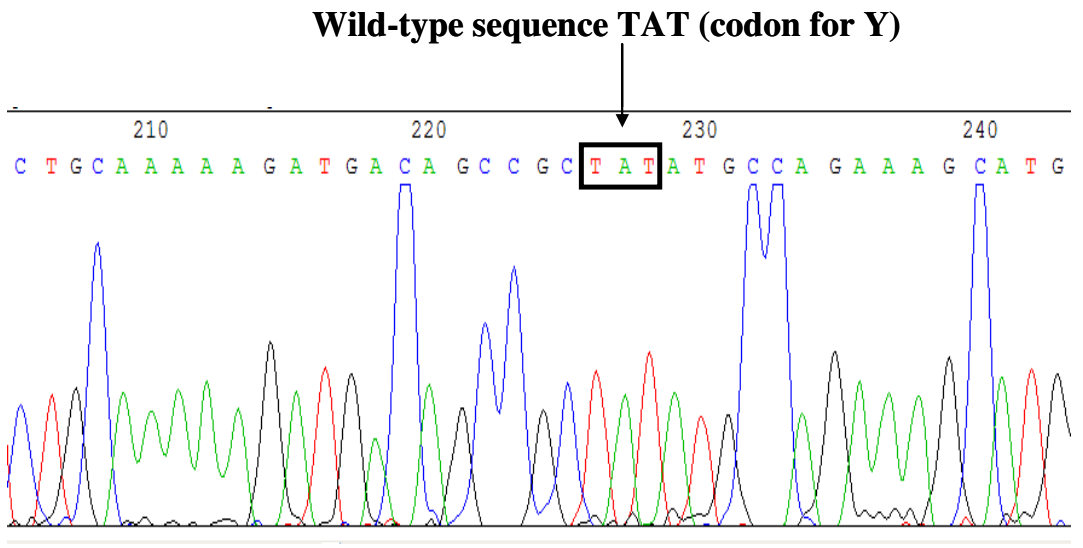
3.2 Size of wild-type and Y58W Grx2

SDS-PAGE was used to determine the size and purity of wild-type and Y58W Grx2. Both proteins were shown to be pure following purification and are 25 kDa in size (Figure 6). The size of the proteins was determined by comparing the distance they migrate in comparison to the distance the low molecular mass proteins migrate under the same denaturing, reducing conditions (Figure 6). A size of 25 kDa is consistent with the published results for wild-type Grx2 (Xia *et al*., 2001; Vlamis-Gardikas *et al*., 1997). It was confirmed that the native non-denatured size of wild-type Grx2 is 25 kDa by size exclusion high pressure liquid chromatography (Figure 8). The results also indicate that glutaredoxin 2 is monomeric, this result is in agreement with the published NMR determined structure of reduced Grx2 (Xia, *et al*., 2001).

3.3 Active site cysteine modification of wild-type Grx2

3.3.1 Thiol groups of wild-type Grx2

DTNB assays are used to determine the accessibility to solvent and redox state of cysteine residues (Habeeb, 1972). Wild-type Grx2 contains two cysteine residues (Cys9 and Cys12) (Xia, *et al*., 2001).



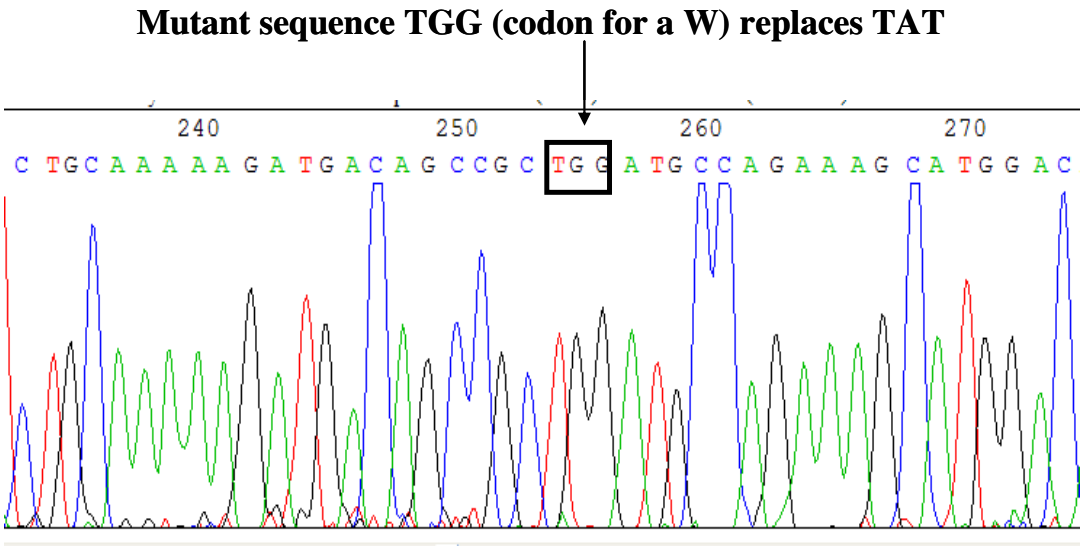


Figure 5: Y58W sequencing conformation
The sequences obtained from sequencing were viewed using the program Chromas version 1.45 (32 bit) (http://www.technelysium.com.au/chromas.html; Technelysium Pty. Ltd., Helensvale, Australia). Shown here is a small segment of the sequence. The wild-type TAT codon for a tyrosine was replaced with the TGG codon for a tryptophan. No additional mutations occurred during the mutagenesis.

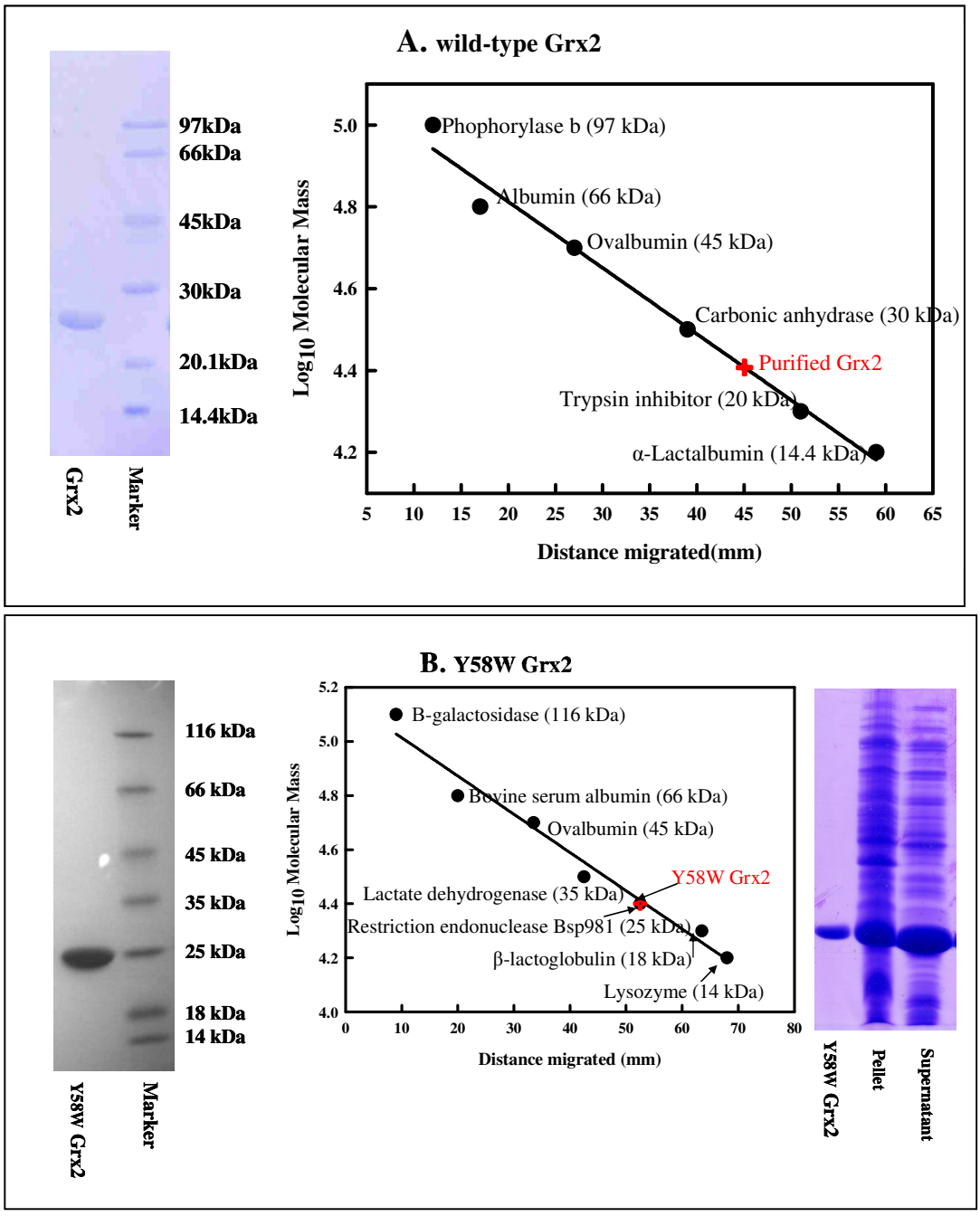


Figure 6: SDS-PAGE analysis wild-type Grx2 and Y58W Grx2
Discontinuous SDS-PAGE gels and calibration curves for (A) wild-type Grx2 and (B) Y58W Grx2. Gels are marked according to the samples that were loaded. Pellet and supernatant samples for Y58W Grx2 are shown in B. The marker lane correspond to the low molecular mass marker sample that was run, the size of the marker proteins are indicated on the gel and their names and sizes given on the calibration curves. Wild-type Grx2 and Y58W Grx2 both migrated a distance which correspond to a size of 25 kDa. For the respective calibration curves the position of both purified wild-type Grx2 and Y58W Grx2 marked on the gels (+).

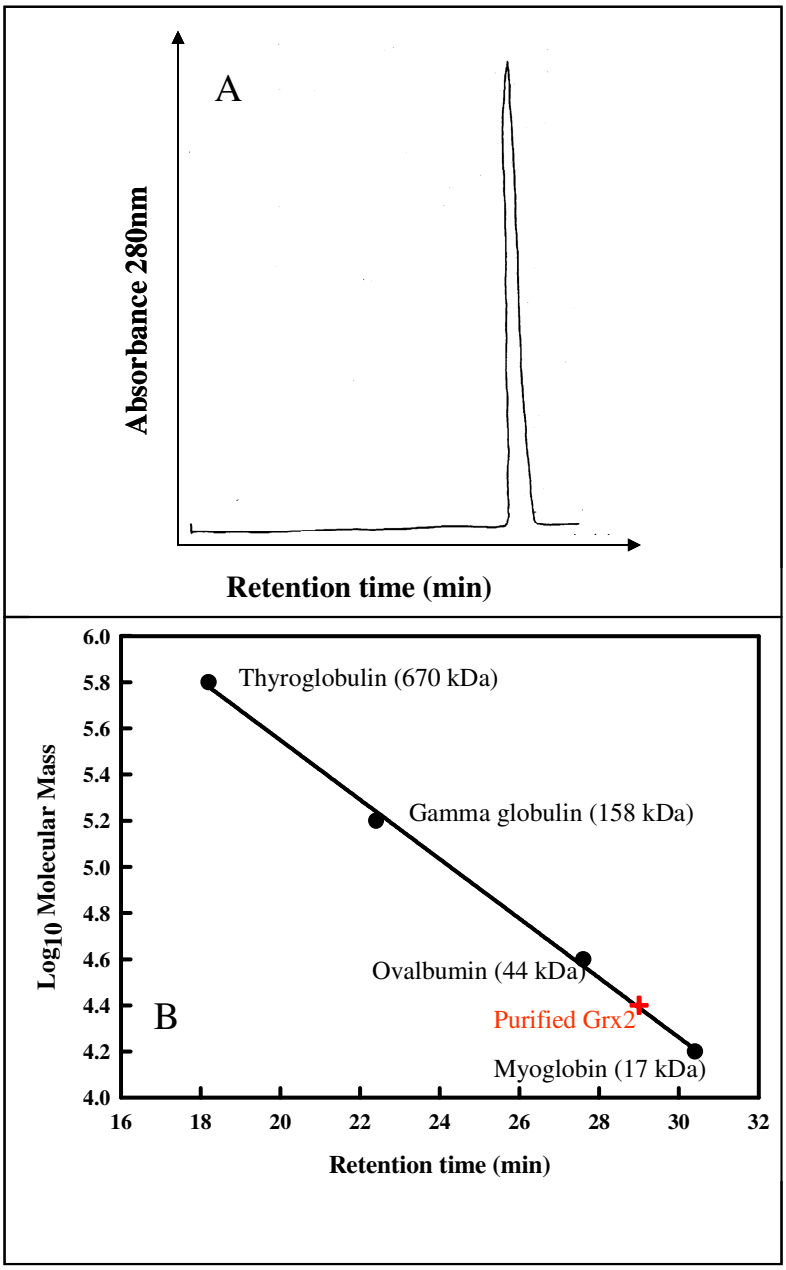


Figure 7: Size exclusion high pressure liquid chromatography
The (A) elution profile of wild-type Grx2 detected by monitoring the absorbance at 280 nm of eluent. The retention times along with the molecular masses of the standard proteins was used to construct a (B) calibration curve, from the curve it was shown that the retention time of Grx2 corresponds to a molecular mass of 25 kDa. The names and sizes of the standard proteins as well as purified wild type Grx2 are indicated.

For the DTNB assay a concentration of 5 µM native and guanidine hydrochloride denatured wild-type Grx2, bound DTNB releasing 10 µM and 8.2 µM of 2-nitro-5-thiobenzoate anion, respectively (concentration determined spectrophotometrically see section 2.6.1). The results of the DTNB assay indicated that both of the cysteine residues of wild-type Grx2 reacted with the DTNB and therefore wild-type Grx2 has two free thiol groups that are maintained in their reduced state by the presence of 1 mM DTT in the storage buffer. So although the active site is buried in the interface between the two domains the DTNB assay indicates that the active site cysteine residue are still accessible to small molecules, which is consistent with published data (Aslund, *et al*., 1994; Vlamis-Gardikas, *et al*., 1997; Xia, *et al*., 2001). The reason the concentration of released 2-nitro-5-thiobenzoate anion was lower for the denatured protein could be due to possible oxidation of the thiol groups or a loss of thiol groups that occurs when guanidine hydrochloride is in a phosphate buffer (Habeeb, 1972).

3.3.2 IAEDANS-labelling of wild-type Grx2

Both of the cysteine residues of native wild-type Grx2 were shown to be accessible to solvent (see above section 3.3.1) and in their reduced state when stored in DTT containing buffer. These residues could therefore be reduced with 1,5-I-AEDANS (iodoacetyl-aminoethyl)-5-naphthylamine-1-sulfonic acid. Once the residues had been labelled a Bradford assay (Bradford, 1976) was used to determine the concentration of the stock solution of labelled Grx2 to be 80 µM. The AEDANS concentration in the protein stock solution was 77 µM as determined spectrophotometrically (section 2.6.1). The results therefore indicated that although both cysteine residues are accessible to DTNB only one of the cysteine residues of wild-type Grx2 was labelled with AEDANS. The most likely labelled cysteine being Cys9 as it is more solvent accessible than Cys12 in the native structure of wild-type Grx2 (Xia *et al*., 2001). An alignment and structural assessment of similarities between glutaredoxins and thioredoxins indicated that the redox sulphur atoms in these proteins are accessible from one side of the molecule but not the other (Eklund *et al*., 1984). This could also explain how only

one of the cysteine residues was labelled in glutaredoxin 2 as it shares a similar active site sequence to glutaredoxin 1 and 3 (Vlamis-Gardikas *et al*., 1997) as well as catalyses some of the same reactions (Lundstrom-Ljung *et al*., 1999).

3.3.3 Secondary structural characterization

The amide bond linking amino acids in proteins acts as the most abundant chromophore absorbing circularly polarized light in the far-ultraviolet region (Woody, 1995). The circular dichroism spectra for proteins, in the far-ultraviolet region have characteristic features based on the secondary structure adopted by the proteins (Woody, 1995). Therefore it was used to assess the secondary structural nature of wild-type Grx2, AEDANS labelled Grx2 and Y58W Grx2. Proteins composed mainly of α-helices have the strongest and most characteristic spectrum with distinct troughs or minima at 208 nm and 222 nm, with a peak at 190 nm (Woody, 1995). Spectra for wild-type and Y58W Grx2 display these distinct spectra (Figure 8) indicating that the proteins are composed predominantly of α-helices and that no gross secondary structural changes had occurred to Y58W Grx2 upon substitution of a tryptophan for a tyrosine residue. The helical secondary structural nature of Grx2 is in agreement with the NMR determined structure of wild-type Grx2 (Xia *et al*., 2001) as well as the structural homologous GST proteins human GST A1-1 (Wallace and Dirr, 1999), Sigma glutathione transferase (Stevens *et al*., 1998) and Ure2 protein (Thual *et al*., 2001), which display similar far-ultraviolet circular dichroism spectrum.

The signal for the native AEDANS labelled protein was not the same as for native wild-type Grx2 and showed a loss of α-helical structure (Figure 8), this would indicate that gross structural changes occurred to the protein upon labelling. The loss in secondary structure seen for the labelled protein is almost as severe as the loss that occurs when the wild-type and Y58W Grx2 are incubated in 7 M urea (Figure 8). Far-ultraviolet circular dichroism spectra for unordered polypeptides have spectra that have weak spectra at wavelengths longer than 200 nm, the weak spectra could be either positive or negative (Woody, 1995).

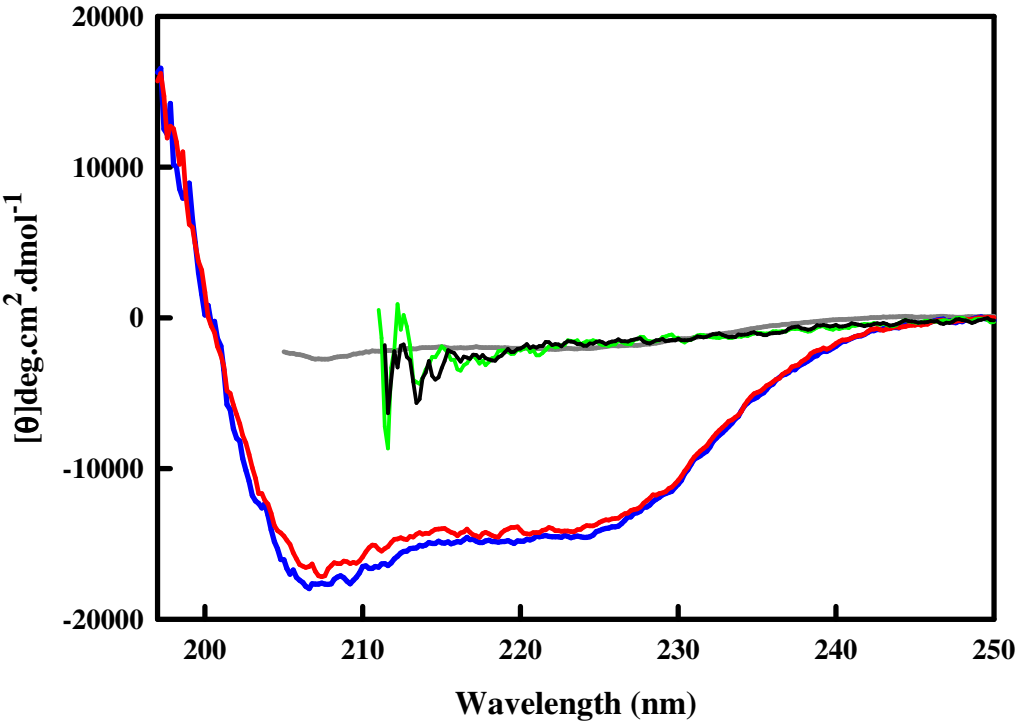


Figure 8: Far-ultraviolet circular dichroism spectra
Spectra shown for 5 µM (—) native wild-type Grx2, (—) native Y58W Grx2 and for 5µM urea denatured (7 M urea) (—) wild-type Grx2 and (—) Y58W Grx2. The spectrum for (—) 20 µM native labelled Grx2 is also included. The buffer was diluted 10 fold for the native far-ultraviolet circular dichroism spectra for wild-type and Y58W Grx2 in order to prevent the salt in the buffer causing a noisy signal.

3.3.4 Tertiary structural characterization of Grx2

Phenylalanine, tyrosine and tryptophan are the dominant fluorescent amino acids in proteins, the most dominant of the three being tryptophan which absorbs light maximally at wavelengths longer than 295 nm (Lakowicz, 1999). Glutaredoxin 2 contains two tryptophan residues (Xia *et al.,* 2001) (see section 1.6 and Figure 3), while the mutant protein (Y58W) contains three tryptophan residues. The emission spectra for native wild-type and Y58W Grx2 display fluorescence emission maxima at a wavelength of 345 nm (Figure 9). Structurally related GST proteins with buried tryptophan residues display emission maximum for the native proteins at a wavelength of 335 nm (Hornby *et al.,* 2000; Hornby *et al.,* 2002; Kaplan *et al.,* 1997; Luo *et al.,* 2002) therefore the tryptophan residue of Grx2 in comparison must be slightly exposed tryptophan residues in the native structure. The slight solvent exposure of the tryptophan residues for the wild-type protein can be confirmed by inspection of the NMR determined structure (see section 1.6) (Xia *et al.,* 2001). The emission maximum for several proteins that contain tryptophan residues red shifts from the native state emission maximum to a denatured state emission maximum of 355 nm (Teale, 1960). This characteristic red shift in the maximum wavelength as well as a decrease in the intensity occurs for wild-type and Y58W Grx2 in the presence of denaturing concentrations of urea (Figure 9). The indole ring of tryptophan residues (Lakowicz, 1999) as well as the extrinsic probe AEDANS (Hudson and Weber, 1973) are sensitive to changes in the environment as has been show by the changes in emission maxima for native and denatured proteins. It is for this reason that the intrinsic tryptophan residues as well as extrinsic AEDANS probe can be used to monitor local structural changes for Grx2 during (un)folding. The native and denatured labelled protein exhibited an emission maximum of 465 nm and 485 nm respectively, regardless of the excitation wavelength (either 295 nm or 338 nm) (Figure 10). The emission maxima of both the native and denatured labelled protein are characteristic of the AEDANS probe because the emission of the tryptophan residues overlaps with the absorbance wavelength of the AEDANS (Hudson and Weber, 1973) and energy transfer occurs (Lakowicz, 1999).

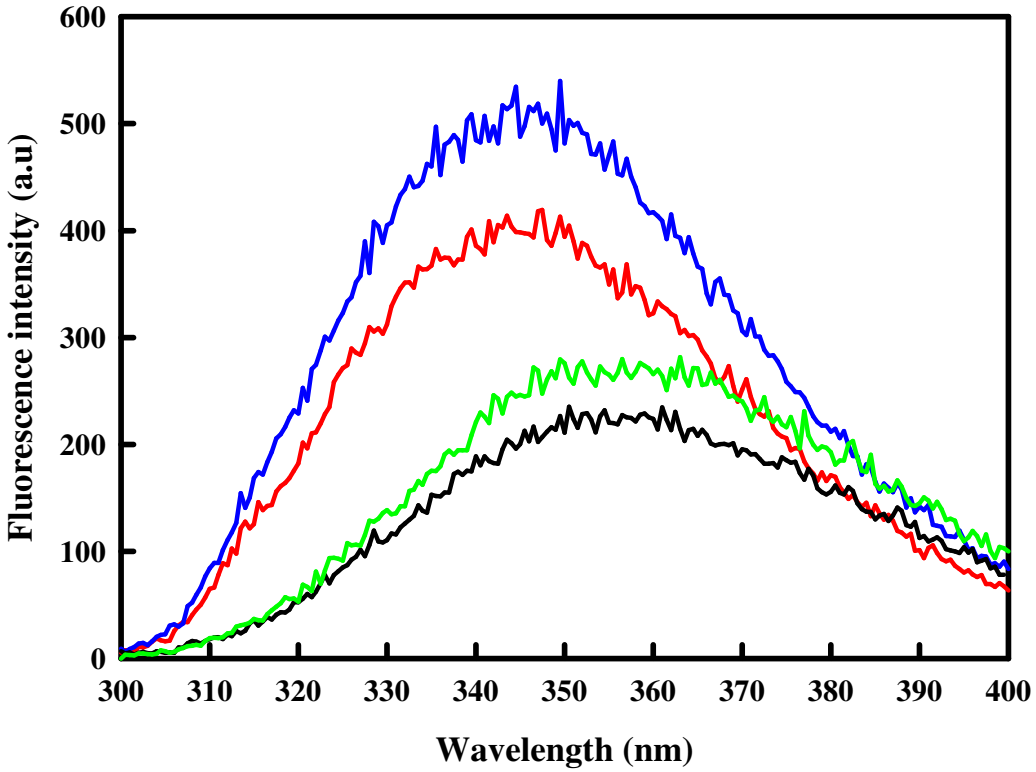


Figure 9: Fluorescence spectra for wild-type and Y58W Grx2
Emission spectra were taken of 5 µM (—) native wild-type Grx2, (—) native Y58W Grx2 and 5µM urea denatured (7 M urea), (—) wild-type Grx2 and (—) Y58W Grx2, upon excitation with light of 295 nm wavelength.

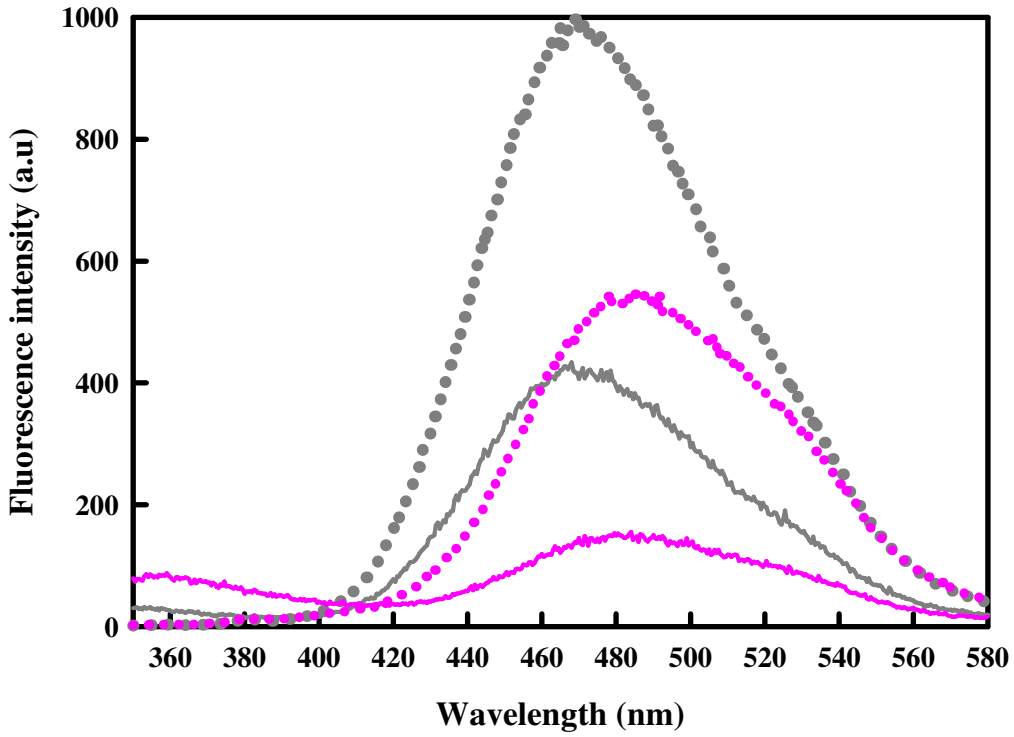


Figure 10: Fluorescence spectra AEDANS labelled Grx2
The emission spectra were recorded for 10 µM native AEDANS labelled Grx2 excited at (—) 295 nm and (····) 338 nm and for 10 µM denatured (8 M urea) AEDANS labelled Grx2 excited at (—) 295 nm and (····) 338 nm.

An increase of 20 % in the intensity of the emitted signal for the native mutant protein indicates the presence of an additional intrinsic tryptophan fluorophore for the mutant protein. There is however no changes in the position of the emission spectrum maxima for the native and denatured mutant protein, respectively.

3.4 Conformational stability

3.4.1 Reversibility of unfolding

Wild-type Grx2, Y58W Grx2 and labelled Grx2 were allowed to unfold in the presence of urea, the urea was then diluted and to allow refolding of the proteins. Then the refolded protein as well as native protein was characterized using the structural probes far-ultraviolet circular dichroism and fluorescence (Figure 11 and Figure 12 respectively). It was shown for wild-type, Y58W, and AEDANS labelled Grx2 that the far-ultraviolet circular dichroism spectra for the refolded and native proteins were almost identical (Figure 11). Both 208 nm and 222 nm are characteristic troughs in the spectra of the α-helical proteins (see section 3.3.3). For the secondary structural probe (far-ultraviolet circular dichroism) the signal to noise ratio was best at a wavelength of 222 nm in the presence of urea, it was for this reason that it was chosen and not the signal at a wavelength of 208 nm. The percentage recovery as determined using the 222 nm signal values for wild-type, Y58W and AEDANS-labelled Grx2 were calculated to be 96 %, 100 % and 100 % respectively. Although the ellipticity values at 222 nm gave a recovery of 100 % for AEDANS labelled protein, this may not be a very accurate measure as the signal is very noisy for the far-ultraviolet region due to structural changes (Figure 8).

Intrinsic tryptophan fluorescence was also used as a tertiary structural probe for native and refolded proteins (Figure 12). The change in the signal only at 345 nm is monitored for wild-type and Y58W Grx2. For the labelled protein intrinsic tryptophan fluorescence as well as extrinsic AEDANS fluorescence was used to

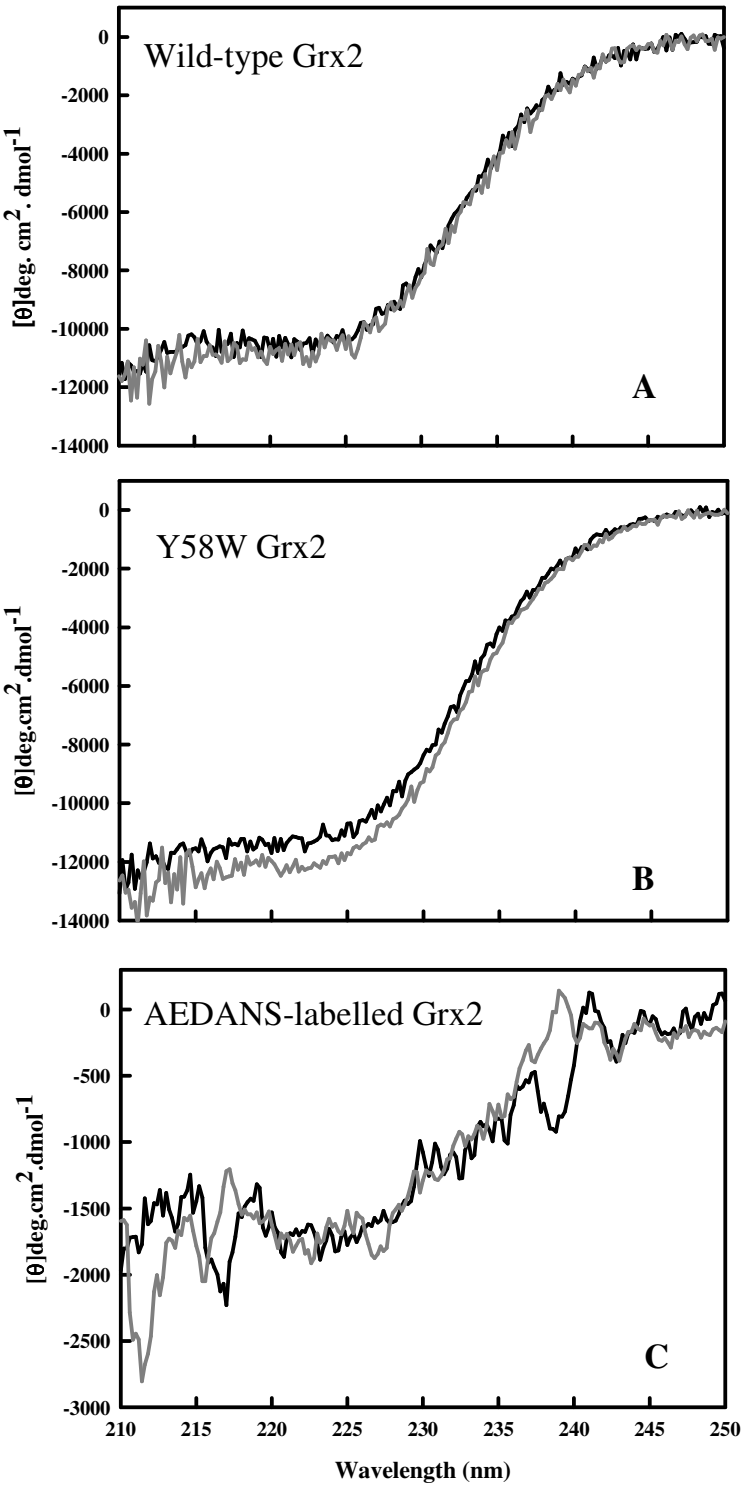


Figure 11: Reversibility of unfolding monitored by far-ultraviolet circular dichroism

The far-ultraviolet circular dichroism spectra for (A) wild-type Grx2 (B) Y58W Grx2 and (C) for AEDANS-labelled Grx2. Spectra for native proteins are represented by the black lines and spectra for the refolded proteins are represented by grey lines. The final concentration for native and refolded proteins was 5 µM.

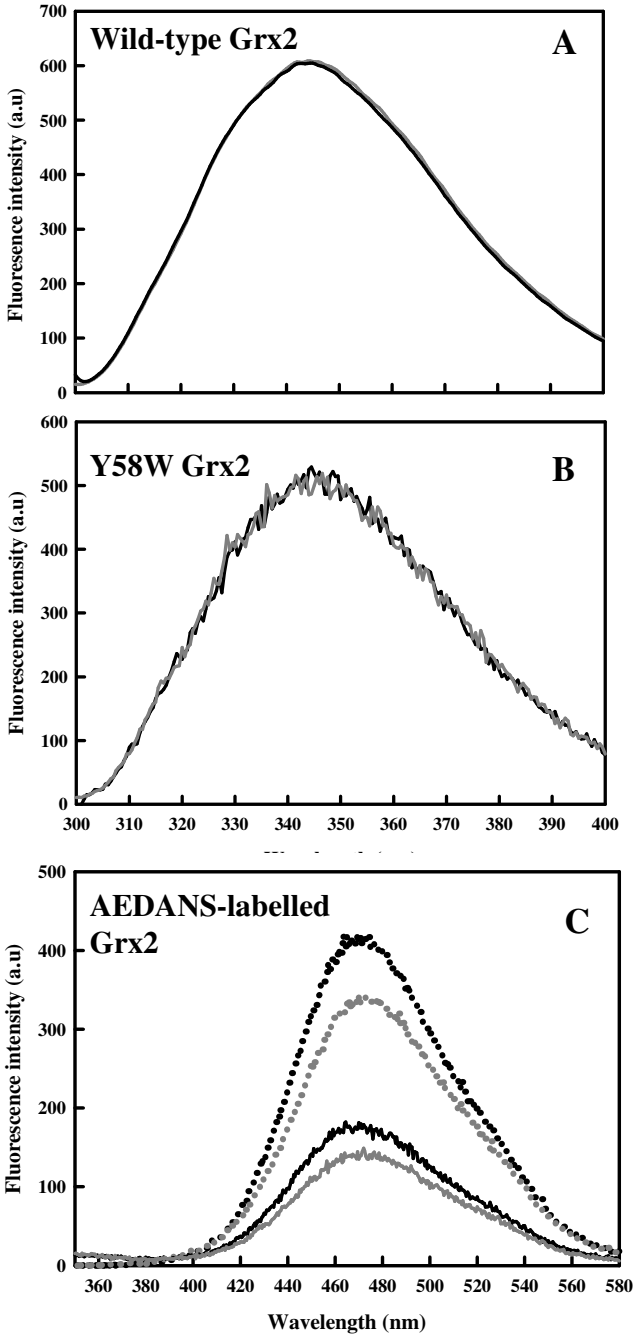


Figure 12: Reversibility of unfolding monitored by intrinsic and extrinsic fluorescence

Fluorescence emission spectra for (A) wild-type, (B) Y58W and (C) AEDANS-labelled Grx2 were recorded for native proteins (black) and refolded proteins (grey). Florescence emission spectra for all proteins was conducted for a final concentration of 5µM protein. For wild-type and Y58W Grx2 the excitation wavelength was 295 nm only. The AEDANS-labelled protein however had an additional extrinsic probe that was used. The AEDANS probe was excited with light of 338 nm wavelength, the spectra for the extrinsic probe are displayed as dotted lines.

assess refolded protein. The fluorescence measurements were taken for the labelled protein samples at 465 nm regardless of where the excitation wavelength was either 295 nm or 338 nm. Based on the values obtained from the specified florescence wavelengths for wild-type, Y58W and AEDANS-labelled Grx2 the percentage recoveries were calculated to be 100 %, 97 % and 80 % (both extrinsic and intrinsic fluorescence probes). The percentage recoveries are expected for the fluorescence spectra of the proteins. The wild-type and Y58W Grx2 native and refolded spectra in close agreement while there is a decrease in the intensity of emission maxima for the labelled protein (Figure 12).

The fact that the unfolding reaction is reversible for wild-type, Y58W and AEDANS-labelled Grx2 indicates that there are no off-pathway folding pathways during refolding that result in aggregation of protein and therefore a loss of recovery of native protein. A high refolding recovery is important for the detailed characterization of equilibrium unfolding data as well as the analysis of the kinetic data to provide information leading toward proposed mechanisms for unfolding and refolding.

3.4.2 Urea-induced equilibrium unfolding

A set of unfolding reactions for a range of urea concentrations, from 0 M urea to 8 M urea, were set up and allowed to reach equilibrium (1 hour). The resulting equilibrium reactions were monitored by the structural probes far-ultraviolet circular dichroism as well as intrinsic tryptophan fluorescence and extrinsic AEDANS fluorescence (Figure 13). The probes assess the structure of the predominating species present in the equilibrium mixture at the set urea concentration (Pace, 1986). As for the reversibility of all the proteins, for both far-ultraviolet circular dichroism and fluorescence the wavelength chosen to monitor unfolding is where there is the maximum change in the signal between the native and unfolded state (as was used in 3.4.1 for recovery percentage calculation). The data obtained were then analysed and a single sigmoidal curve

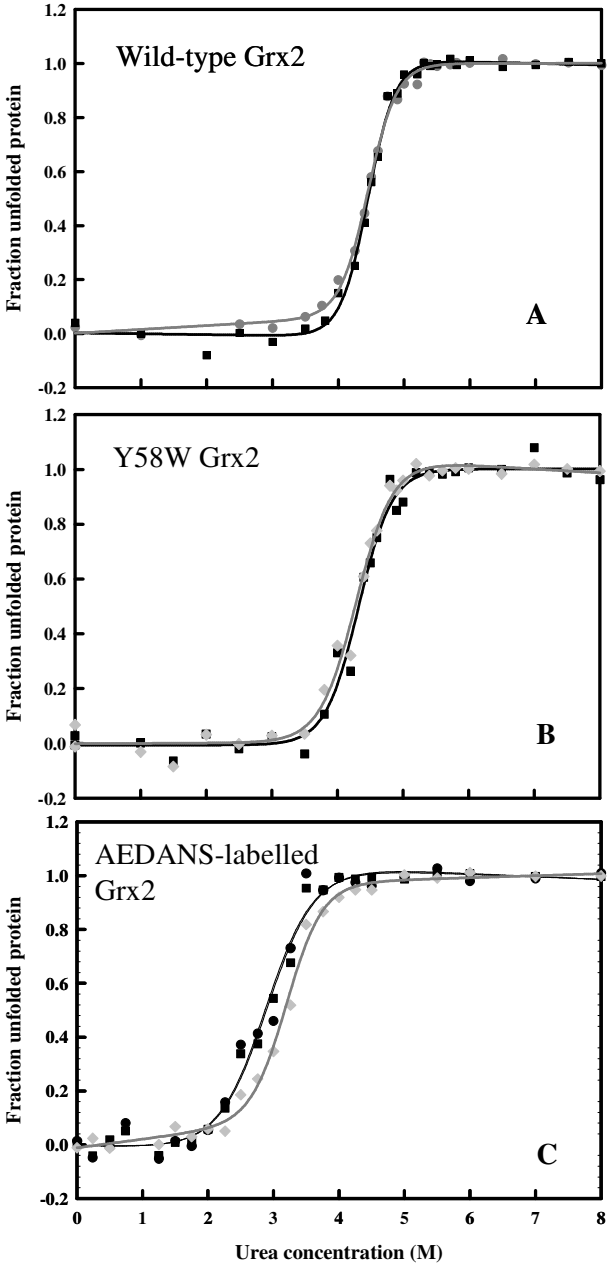


Figure 13: Urea-induced equilibrium unfolding

Unfolding reactions at equilibrium for (A) 5 µM wild-type Grx2 and (B) 5 µM Y58W Grx2 were monitored by (♦) far-ultraviolet circular dichroism ellipticity at 222 nm and (■) intrinsic tryptophan fluorescence (excitation wavelength of 295 nm and the emission monitored at 345 nm). Unfolding reactions at equilibrium for (C) 10 µM AEDANS-labelled Grx2 were monitored by fluorescence, emission at 465 nm with excitation at (■) 295 nm and (●) 338 nm, respectively. The secondary structural probe far-ultraviolet circular dichroism was also used to monitor changes for the labelled protein with the ellipticity at (♦) 222 nm used to determine the fraction of unfolded protein at each urea concentration. Data were normalised to the fraction of unfolded protein. Solid lines represent two-state sigmoidal fit to the data.

fit to the data for wild-type and Y58W Grx2 as well as AEDANS-labelled Grx2 (Figure 13). The unfolding data for both wild-type Grx2 and Y58W Grx2, shows that the predominant species of the protein is the native state at urea concentrations of 0 M to 4 M. The transition region begins at 4 M urea and ends at 5 M urea. From 5 M urea to 8 M urea the protein is unfolded. With the single sigmoidal fit being a good one, it indicates that the unfolding of wild-type and Y58W Grx2 is a two-state process with no significant population of intermediates present. The unfolding as monitored by the tertiary local structural probe intrinsic tryptophan fluorescence and the global secondary structural probe far-ultraviolet circular dichroism overlay and are in agreement. This adds further support for this conclusion that equilibrium unfolding of wild-type and Y58W Grx2 is a two-state process with no stable equilibrium intermediates accumulating during unfolding.

The unfolding is different for AEDANS-labelled Grx2 as the predominant species of the protein is the native state only up until 2 M urea. At 2 M urea, the transition region begins and this region ends at 3.5 M to 4 M urea depending on the probe. From 4 M urea to 8 M urea, labelled Grx2 is completely unfolded. The intrinsic tryptophan and the extrinsic AEDANS probes for the labelled protein agreed with one another but not with the far-ultraviolet circular dichroism data (Table 2). This indicates that during the unfolding of AEDANS-labelled protein an intermediate could be present. The non-coincidence could also reflect the less than 100 % reversibility of the unfolding reaction (see section 3.4.1). The differences detected between probes cannot be due to concentration errors because the same samples were used for both probes.

The parameters obtained from fitting the all equilibrium unfolding data for wild-type, Y58W and AEDANS-labelled Grx2 are summarized in Table 2. The $\Delta G(H_2O)$, m and C_m value obtained for the Y58W Grx2 are also in close agreement with those obtained for wild-type Grx2 whereas the values obtained for AEDANS-labelled Grx2 differ greatly. The significance of the values obtained for the labelled protein will be discussed later (4.1).

Table 2: Parameters obtained for equilibrium unfolding

The equilibrium unfolding of wild-type Grx2, Y58W Grx2 and labelled Grx2 were monitored using far-ultraviolet circular dichroism as well as fluorescence. Intrinsic tryptophan fluorescence was the tertiary structural probe used for wild-type Grx2 and Y58W Grx2. The tertiary structural probes used for labelled protein were direct excitation of the extrinsic AEDANS probe as well as fluorescence energy transfer for the tryptophan residues to AEDANS. Data obtained from all probes were fitted to a two-state model.

Wild-type glutaredoxin 2

Parameters	Far-ultraviolet circular dichroism	Intrinsic tryptophan fluorescence	Average
$\Delta G(H_2O)$ (kcal.mol^{-1})	12.0 (± 1.0)	12.5 (± 1.1)	12.3
m (kcal mol^{-1} M^{-1} urea)	2.7 (± 0.2)	2.79 (± 0.3)	2.74
C_m (M urea)	4.4	4.4	4.4

Y58W glutaredoxin 2

Parameters	Far-ultraviolet circular dichroism	Intrinsic tryptophan fluorescence	Average
$\Delta G(H_2O)$ (kcal.mol^{-1})	9.0 (± 1.2)	9.9 (± 1.1)	9.5
m (kcal mol^{-1} M^{-1} urea)	2.1 (± 0.3)	2.3 (± 0.3)	2.2
C_m (M urea)	4.3	4.3	4.3

AEDANS-labelled glutaredoxin 2

Parameters	Far-ultraviolet circular dichroism	AEDANS fluorescence	Fluorescence energy transfer	Average
$\Delta G(H_2O)$ (kcal.mol^{-1})	6.8 (± 0.9)	4.6 (± 1.1)	4.6 (± 0.7)	5.4
m (kcal.mol^{-1} .M^{-1} urea)	2.1 (± 0.3)	1.6 (± 0.4)	1.6 (± 0.2)	1.8
C_m (M urea)	3.2	2.9	2.9	3.0

3.5 Single-jump kinetic studies

Intrinsic tryptophan fluorescence was chosen as a structural probe for the changes occurring during (un)folding of wild-type and Y58W Grx2. It was chosen because there is changes in the intensity of the emission signal of the intrinsic tryptophan residues on denaturation (quenching) or renaturation (enhanced) of the proteins such that the unfolding or refolding of the protein can be monitored (see Figure 9).

The excitation wavelength chosen for the fluorescence was 280 nm, as it results in greater absorbance and excitation of the tryptophan residues. When the protein is excited with light of a wavelength of 280 nm, resonance energy transfer occurs from the tyrosine residues to the tryptophan residues and the emission spectrum characteristic of a tryptophan emission spectrum is observed (Lakowicz, 1999). In addition, the cut off filter (320 nm) eliminates any detectable contribution of the tyrosine residues (emission at 303 nm in water) (Lakowicz, 1999). The signal output for the stopped-flow reaction analyzer is in volts and therefore provides fluorescence intensity over a range of wavelengths (greater than 320 nm) not at a particular wavelength.

3.5.1 Unfolding kinetics

From the equilibrium unfolding data it can be seen that at urea concentration of 5 M to 8 M urea the predominating species is that of unfolded protein for both wild-type and Y58W Grx2 (Figure 13). Kinetic unfolding experiments were therefore conducted over this range of urea concentrations, as no refolding should occur, and the resultant traces analysed. A double exponential fit was shown by residuals plot to be the best fit for both wild-type and Y58W Grx2 (Figure 14). These two phases account for all the amplitude change during the unfolding reactions as the unfolding kinetic traces extend from the native protein baseline to the unfolded protein baseline.

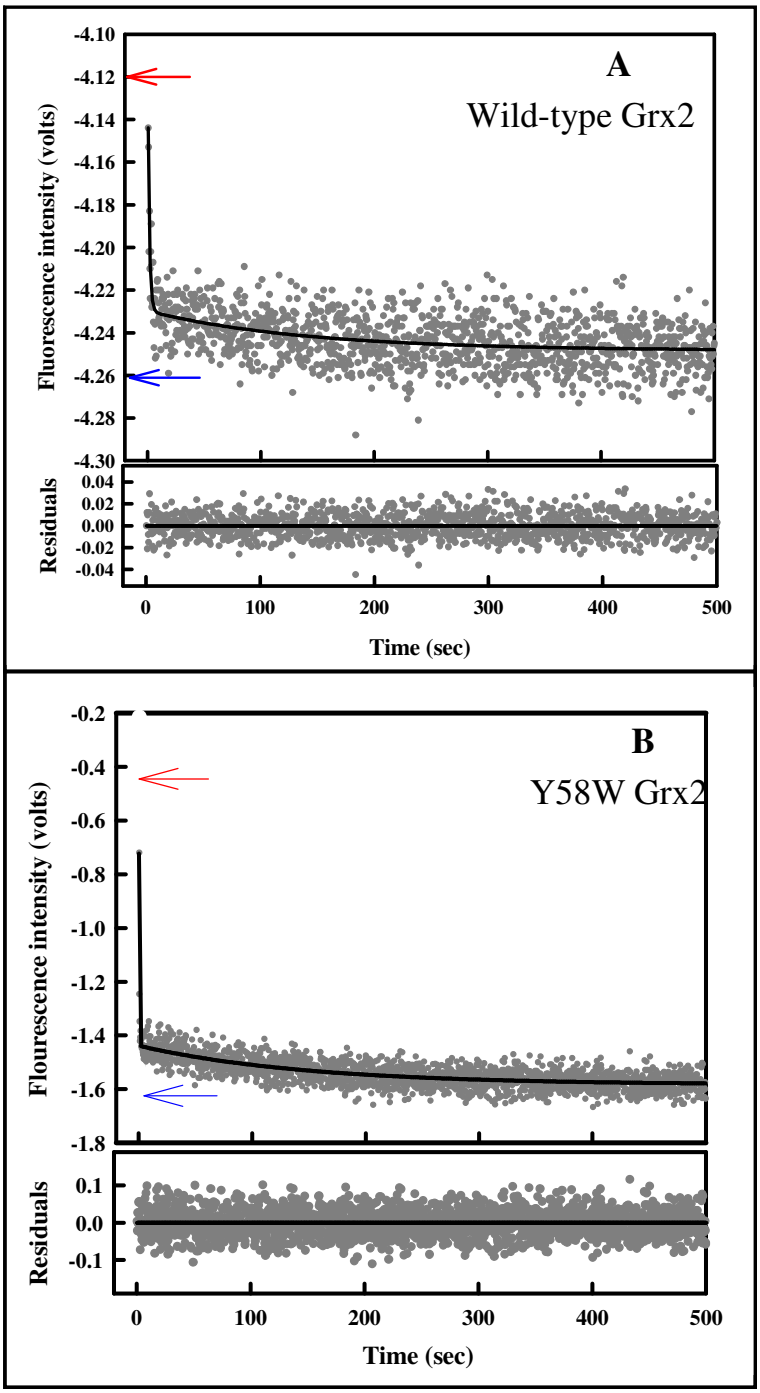


Figure 14: Unfolding kinetics traces

The average of (●) three kinetic unfolding traces (7.5 M urea) for 5 µM (A) wild-type Grx2 or (B) Y58W Grx2 in the phosphate dialysis buffer, ph 7. The excitation wavelength was set at 280 nm with the emission intensity change monitored with a 320 nm cut off filter at 20 ºC. Data were fitted to a double exponential as indicated by the solid black line. Residuals indicate the fit is good one. The red arrow indicates the position of the average value for the native baseline and the blue arrow the position of the average value for the unfolded baseline

The percentage amplitude change for each phase for the wild-type and mutant proteins was then plotted against the urea concentration for that unfolding reaction (Figure 15 A). Changes in the amplitude for wild-type Grx2, over the range of urea concentrations from 5 M to 7.5 M urea are complex, with the fast phase amplitude decreasing from 26 % at 5 M urea to 15 % at 5.25 M urea before increasing to 86 % at 5.75 M urea. From 5.75 M urea to 7.5 M urea it remains relatively constant. The amplitude changes are different for the fast phase of the mutant protein with the amplitudes increasing from 50 % at 5 M urea to 90 % at 5.6 M urea, from 5.6 M urea the amplitude remains constant. Similarly the change in the percentage amplitude of the slow phases of unfolding over the range of urea concentrations is complex for wild-type Grx2. The slow phase amplitudes increase from 74 % at 5 M urea to 85 % at 5.25 M urea before decreasing to a percentage of 14 % at 5.75 M urea. From 5.75 M urea to 7.5 M urea the amplitudes remain relatively constant. Y58W Grx2 however displays a different pattern for the changes in the amplitudes of the slow phase. The slow phase amplitudes decrease from 45 % at 5 M urea to 10 % at 5.6 M urea. From 5.6 M urea it remains roughly 10 % to a urea concentration to 7.5 M.

Wild-type Grx2 at 7.5 M urea displays a fast phase that has a time constant (inverse of the rate constant) of 1.4 sec and the slow phase has a time constant of 140 sec. Similarly, Y58W Grx2 has two phases of unfolding. At 7.5 M urea, the time constant of the fast phase is 210 milliseconds and for the slow phase is 150 seconds. The kinetic unfolding of human GST A1-1 a structurally related protein, as monitored by tryptophan fluorescence also displays two kinetic unfolding phases with comparable time ranges to the two phases of Grx2 (Wallace *et al.,* 1998). The logarithm of the rates of the fast and slow phases of unfolding for both the wild-type and Y58W Grx2 were plotted against the respective urea concentration at which they were obtained (Figure 15 B). Wild-type Grx2 displays complex urea dependence for the rates of both the fast phase and the slow phase of unfolding while Y58W Grx2 does not. For both proteins, the slow phase displays low urea dependence while the fast phases high urea dependence

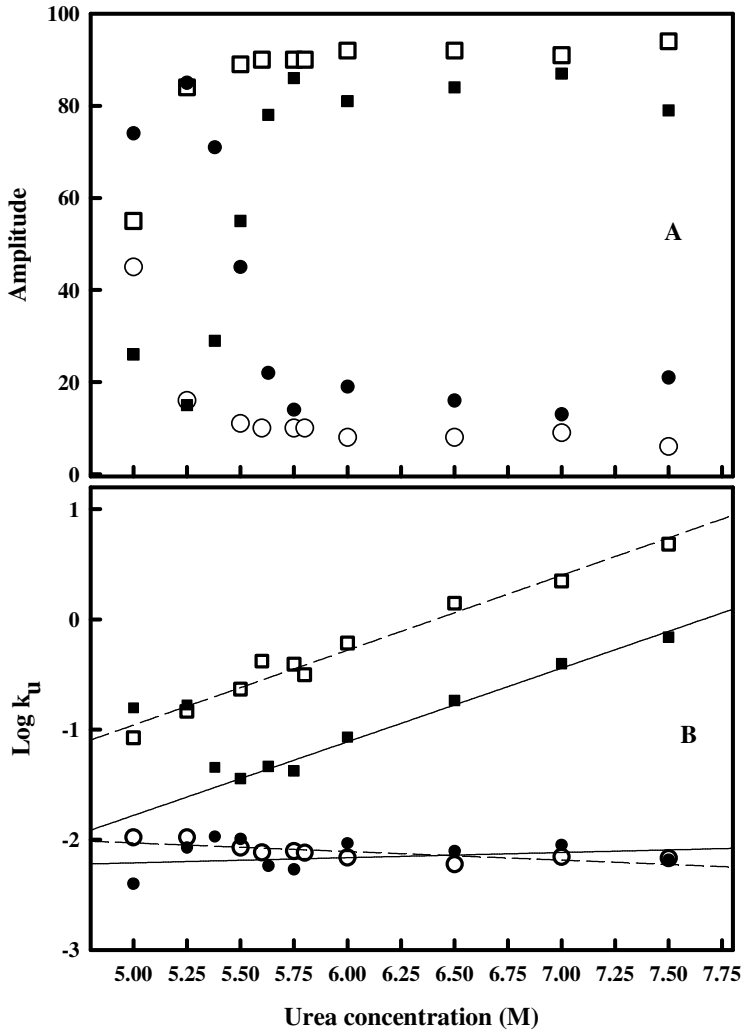


Figure 15: Affect of urea on the fast and slow phases of unfolding

The dependence of (A) amplitude and (B) log of the rates of unfolding on urea for wild-type Grx2 (closed symbols) and Y58W Grx2 (open symbols). The dependencies of the (■□) fast phase and the (○●) slow phase of unfolding were obtained from fitting a double exponential to the traces obtained for each urea concentration. The solid (wild-type Grx2) and dashed (Y58W Grx2) lines represent a linear fit to the rate data (see Table 3 for values).

Table 3: Affect of varying final urea concentration on unfolding

The logarithm of the rates of the fast and slow phases of unfolding for different final urea concentrations is displayed in **Figure 15**. Linear regression analysis was done for the entire set of data for the slow phase of the mutant protein (R^2=0.60) and the fast phase (R^2=0.97) for the mutant protein. The data set for the fast phase (R^2=0.99) and the slow phase (R^2=0.12) of unfolding for the wild-type protein was only analysed form 5.5 M urea and 5.6 M urea respectively, to 7.5 M urea. Fitting the data to a linear function results in obtaining the parameters of m_u (change in solvent accessibility during unfolding) and $k_u(H_2O)$ (apparent unfolding rate in the absence of denaturant) (Tanford, 1970), presented in this table .

Parameter	**Wild-type Grx2**		**Y58W Grx2**	
	Fast phase	Slow phase	Fast phase	Slow phase
m_u (kJ mol^{-1}/ M urea)	1.6 (± 0.077)	0.12 (± 0.16)	1.7 (± 0.10)	0.19 (± 0.055)
$k_u(H_2O)$ (ms^{-1})	7.6 x 10^{-3} (± 3.0 x 10^{-4})	3.6 (± 0.59)	0.044 (± 0.0025)	23 (± 1.9)

however, the urea dependence for the wild-type protein of both the fast phase and slow phase changes 5.5 M urea. It is also interesting to note that the wild-type data for 5 M and 5.25 M urea align with the mutant proteins data. Linear regression analysis was conducted for the complete set of data for the fast phase and the slow phase of the mutant protein and was only done for data from 5.5 M and 5.6 M urea to 7.5 M urea for the fast and slow phases of the wild-type protein respectively. The data were fitted to the following equation (Tanford, 1970):

$$\log k_u = \log k_u(H_20) + m_u[\text{Urea}] \quad (15)$$

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. 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).

3.5.2 Refolding kinetics

The equilibrium unfolding data indicates that at urea concentrations of 0 M to 4 M, the predominating species is that of native protein for both wild-type Grx2 and Y58W Grx2 (Figure 13). Kinetic refolding experiments were therefore conducted over the range of urea concentrations from 1 M to 4 M at these concentrations the unfolding reaction should not occur as these are strongly refolding conditions. Refolding was monitored by intrinsic tryptophan fluorescence and an average trace obtained from three refolding traces per urea concentration. The average trace was then fitted to a triple exponential, the fit was shown to be good by the random distribution of points on the residuals plot (Figure 16). The refolding reaction monitored and fit to a triple exponential was shown to account for the total amplitude change from unfolded protein to native protein, by the position of the baselines (Figure 16). Therefore, the folding of both wild-type Grx2 and Y58W Grx2 is composed of 3 phases: fast, medium and slow. Fitting the data gave time constants, the inverse of the rate constants as follows: for the fast phase 61 milliseconds, the medium phase 0.79 seconds and the slow phase 180 seconds,

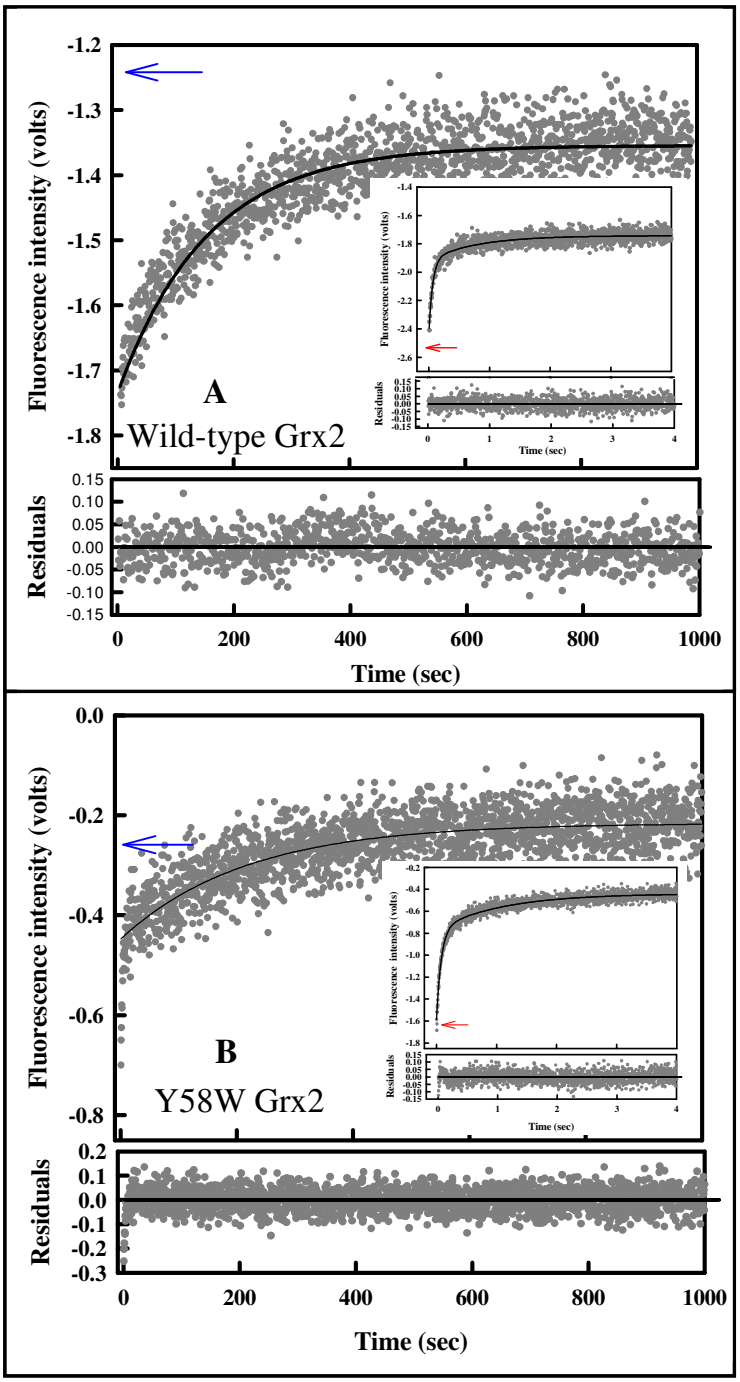


Figure 16: Refolding kinetics traces for 1M residual urea
The average of (●) 3 refolding kinetic traces for 5 µM (A) wild-type Grx2 and (B) Y58W Grx2 in buffer, pH 7 containing 1 M urea. Refolding traces were recorded for the first 4 sec (insets) with the red arrows indicating the unfolded baseline. Following the first 5 sec of refolding a slow phase of refolding occurs for the remaining 1000 sec, the native protein baseline is indicated with a blue arrow. Refolding was conducted at 20 ºC and the average data traces for three reactions were fitted to a double exponential (inset) and a single exponential (both indicated by black lines). The residuals plots indicate both fits are good.

this was all for refolding in the presence of 1 M residual urea for wild-type Grx2. The corresponding amplitude changes for wild-type Grx2 are 50 %, 17 % and 33 % for the fast, medium and slow phases respectively. The rate constants obtained for Y58W Grx2 are as follows: for the fast phase 76 milliseconds, the medium phase 1.1 seconds and the slow phase 212 seconds with corresponding amplitude changes of 57 %, 22 % and 21 % respectively.

The dependence of the amplitudes on urea concentration for all three refolding phases shows complex behaviour for both wild-type and Y58W Grx2 (Figure 17 A). The figure representing the urea dependence is obtained by plotting the amplitudes for the fast, medium and slow phases of refolding, obtained from fitting the data to a triple exponential, against the final urea concentration of refolding. The changes in the amplitude dependence of the medium phase do not resemble those of the fast and slow phases. Rather the amplitude remains constant from 1 M urea to 1.8 M urea from which point it decreases to 2.6 M urea and then again remains constant from 2.6 M urea to 4 M urea. This pattern is a complex one, but is also seen to occur for the medium phase amplitudes of Y58W Grx2 over the range of refolding urea concentrations. The slow phase amplitude, however increases for the entire range of urea concentrations although the increase becomes steeper from 1.8 M urea to 4 M urea than it is from 1 M urea to 1.8 M urea. In an opposite manner, the fast phase amplitudes decrease from 1 M urea to 4 M urea with a steeper decrease occurring from 1.8 M urea to 4 M urea. The changes in amplitude over the range of urea concentrations, for the fast and slow phase are the same for the wild-type and mutant proteins with only the percentages increasing or decreasing for the fast and slow phases respectively between 1 M and 2.5 M urea.

Amplitudes as well as rates for each phase of refolding are obtained from fitting the kinetic refolding data to a triple exponential. The log of the rates for both unfolding phases was plotted against the residual urea concentration at which refolding was conducted (Figure 17 B) giving the urea dependence of the rates.

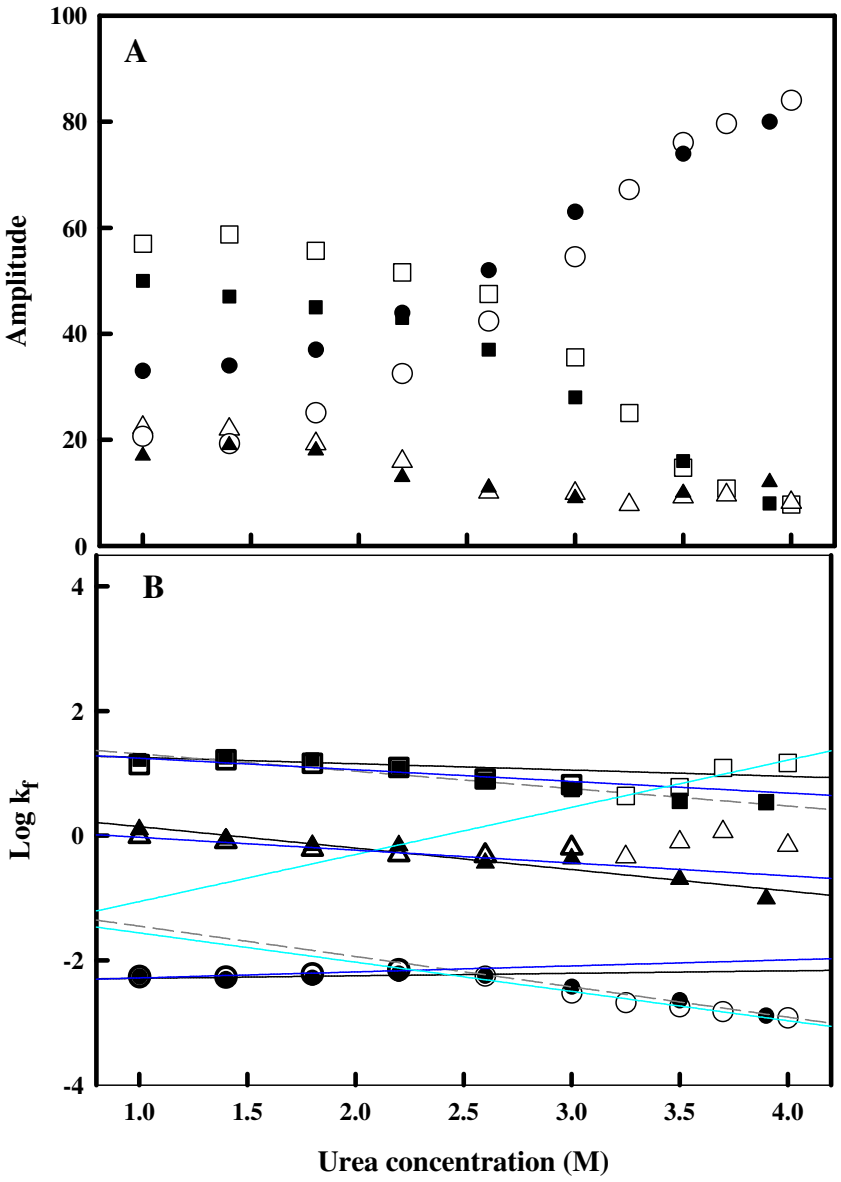


Figure 17: Effect of varying final urea concentrations on refolding
The (A) percentage amplitudes and (B) log of the rates of the (■□) fast (▲Δ) medium and (●○) slow phases of the respective refolding phases were plotted against the residual urea concentrations from 1 M to 4 M. Data for wild-type Grx2 (closed symbols) and Y58W Grx2 (open symbols) were fitted as stated for Figure 17. Linear regression analysis curves for the proteins are represented as follows: (—) wild-type Grx2 (medium phase all data and fast and slow phases from 1 M to 2.2 M urea), (---) wild-type Grx2 (from 2.2 M urea to 4 M urea for the fast and slow phases). Then for (—) Y58W Grx2 (fast phase 1-3 M urea, the medium phase 1-2.6 M urea and the slow phase 1-2.2 M urea), and (—) Y58W Grx2 from 3 M urea (fast phase) or 2.2 M urea (slow phase) to 4 M urea. See Table 4 for parameters.

Table 4: Affect of varying final urea concentration on refolding

The logarithm of the rates of the fast medium and slow phases of refolding for different final urea concentrations for both wild-type and Y58W Grx2 are displayed in **Figure 17**. Linear regression analysis was done for subsections of the data sets for each phase and each protein. For the wild-type protein, the fast phase data were separated into two sections and regression analysis done for each section: 1 M to 2.2 M urea (R^2=0.53) and 2.6 M to 4 M urea (R^2=0.94). The medium phase data was fitted as a single set (R^2=0.92) and the slow phase as for the fast phase with R^2 values of 0.32 and 0.99 respectively. The mutant protein data were also analysed in subsections for all the phases of refolding. For the fast phase: 1 M to 3 M urea (R^2=0.78) and 3.3 M urea to 4 M urea (R^2=0.92). The medium phase only a subsection of the data could be fitted form 1 M urea to 2.6 M urea (R^2=0.95). Then for the slow phase: 1 M to 2.2 M urea (R^2=0.83) and 2.6 M to 4 M urea (R^2=0.96). Fitting the data to a linear function results in obtaining the parameters of m_f (change in solvent accessibility during refolding) and $k_f(H_2O)$ (apparent refolding rate in the absence of denaturant) (Tanford, 1970), presented in this table .

Parameter	**Wild-type Grx2**					**Y58W Grx2**				
	Fast phase		Medium Phase	Slow phase		Fast phase		Medium phase	Slow phase	
Urea concentration range (M)	1-2.2	2.5-4	1-4	1-2.2	2.5-4	1-3	3.3-4	1-2.6	1-2.2	2.5-4
m_f (kJ mol^{-1}/ M urea)	0.25 (± 0.17)	0.69 (± 0.12)	0.84 (± 0.10)	0.10 (± 0.10)	1.2 (± 0.071)	0.47 (± 0.12)	1.9 (± 0.40)	0.52 (± 0.068)	0.24 (± 0.076)	1.2 (± 0.12)
$k_f(H_2O)$ (s^{-1})	23 (± 1.9)	39 (± 4.1)	3.1 (± 0.69)	4.7 x 10^{-3} (± 1.4 x 10^{-4})	0.11(± 0.011)	27 (± 2.0)	0.015 (± 4.7 x 10^{-3})	1.5 (± 0.42)	4.2 x 10^{-3} (± 8.9 x 10^{-5})	0.081 (± 0.011)

Based on the points where changes in the urea dependence occur linear regression was done for sections of the data (See Figure 17 B and Table 4). The data were fitted to the following equation (Tanford, 1970):

$$\log k_f = \log k_f(H_20) + m_f[\text{Urea}] \quad (16)$$

where k_f is the apparent first order rate constant for refolding. The m_f values (change in solvent accessibility during refolding) and the apparent rates of refolding in the absence of water ($k_f(H_2O)$) reported in Table 4. The units for the m_f 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 dependence of the rates of refolding on urea was shown to be complex for all phases of refolding for the mutant protein and for the fast and slow phases of refolding of the wild-type protein. The medium phase for the wild-type protein shows no changes in dependence over the range of urea concentrations while for the mutant protein this phase displays a change from 3.3 M urea to 4 M urea. Also the medium phase for wild-type and Y58W Grx2 (data from 1 M urea to 3.3M urea) show high urea dependence. The data for the medium phase for the mutant protein from 3.3 M urea to 4 M urea displays a complex dependence on urea in the upward direction. The fast and slow phases for wild-type Grx2 both show changes at 2.2 M urea in a downward direction although, the change for the fast phase is not as predominant as for the slow phase. This change at 2.2 M urea is also seen to occur for the slow phase of Y58W Grx2. At urea concentration of 1 M to 2.2 M urea the slow phases of refolding for both proteins are urea independent and then from 2.2 M urea to 4 M urea the urea dependence increases. The fast phase for the wild-type protein also displays a more urea independent section for 1 M to 2.2 M urea, with the data from 2.2 M urea to 4 M urea displaying a higher urea dependence. For the mutant protein, the urea dependence is different to the wild-type protein for the fast phase with a change in the urea dependence occurring at 3.3 M urea to 4 M urea in an upward direction.

3.6 Double-jump refolding kinetic studies

In order to uncouple the gross structural unfolding form the isomerization occurring during unfolding, a high denaturant concentration is required (Wallace and Matthews, 2002). At a high denaturant concentration, the slow urea independent isomerization will be kinetically separated from the rapid gross structural unfolding of the protein (Kiefhaber *et al.,* 1992a). During unfolding, the fast phase occurs rapidly at 7.5 M urea with a time constant of 61 milliseconds for wild-type Grx2. With such rapid unfolding the refolding initiated from unfolding delays within the fast phase were not feasible. The rate of the fast phase however decreases with lower urea concentrations. Therefore, to optimize the kinetic separation of gross structural unfolding from isomerization as well as capture the rapidly unfolding species, lower urea concentrations were chosen (5.5 M and 6 M). The time constants for the fast phase of unfolding are 28 sec and 12 sec at 5.5 M urea and 6 M urea respectively. For both urea concentrations, a set of delay times were carried out with the resultant kinetics of refolding monitored by intrinsic tryptophan fluorescence as for single-jump refolding (see section 3.5).

Different unfolded species created by the various initial urea concentrations as well as the different delay times of unfolding resulted in altered refolding kinetics. The final urea concentrations for the second jump or the refolding phase of the double-jump experiment were 0.9 M and 1 M urea for the 5.5 M urea and 6 M urea unfolding, respectively. Two sets of double-jump refolding experiments were therefore conducted. The first one involved the initial unfolding reaction in 5.5 M urea with varying delay times of unfolding before refolding was initiated in 0.9 M residual urea for a final protein concentration of 5 μM. The second set involved unfolding delays in 6 M urea before refolding was initiated by dilution of unfolding solution to produce refolding conditions of 1 M residual urea and 5 μM wild-type Grx2.

3.6.1 Double-jump refolding with unfolding delay in 5.5 M urea

The unfolding of wild-type Grx2 in 5.5 M urea is composed of 2 phases, the fast and slow phase. The fast phase is complete in 50 seconds at 20ºC and from 50 seconds, the slow phase of unfolding occurs. In order to refold fast and slow unfolded species the refolding reaction was initiated following 5 seconds of unfolding to 500 seconds unfolding for the double-jump experiment. Refolding of wild-type Grx2 following unfolding delays from 5 seconds to 45 seconds is only composed of two phases (Figure 18 A), a fast and medium phase. The delays of 50 seconds (Figure 18 B) and longer are composed of three phases (fast, medium and slow).

Although the refolding of equilibrium unfolded wild-type Grx2 is composed of three phases (see section 3.5.2) the slow phase amplitude is too small for proper detection and therefore cannot be fitted for refolding following unfolding delays of shorter than 50 seconds. The first time the data can be fitted to a triple exponential is for an unfolding delay of 50 seconds unfolding delay (Figure 18 B). At a 500 seconds unfolding delay time, the refolding is the same as the refolding of equilibrium unfolded wild-type Grx2 (Figure 16 ; Figure 19). The baselines indicate that the refolding of the transiently unfolded wild-type Grx2 is complete (Figure 19) after 1000 seconds following 500 seconds unfolding delay time. For many of the short unfolding delay times including 10 seconds unfolding delay, a large burst phase is observed where a large amount of the refolding amplitude change occurs in the dead-time of the instrument (Figure 18 A).

In order to identify trend in the changes that occur as the unfolding delay time increases the amplitudes (Figure 20 A) and rates of refolding (Figure 20 B) of the relative phases were plotted against the unfolding delay time for which they were obtained. The amplitudes and rates were obtained by fitting a double exponential (refolding traces for unfolding delays of 5 seconds to 45 seconds at 5.5 M urea unfolding) and a triple exponential (refolding traces for unfolding delays of 50 seconds to 500 seconds at 5.5 M urea unfolding).

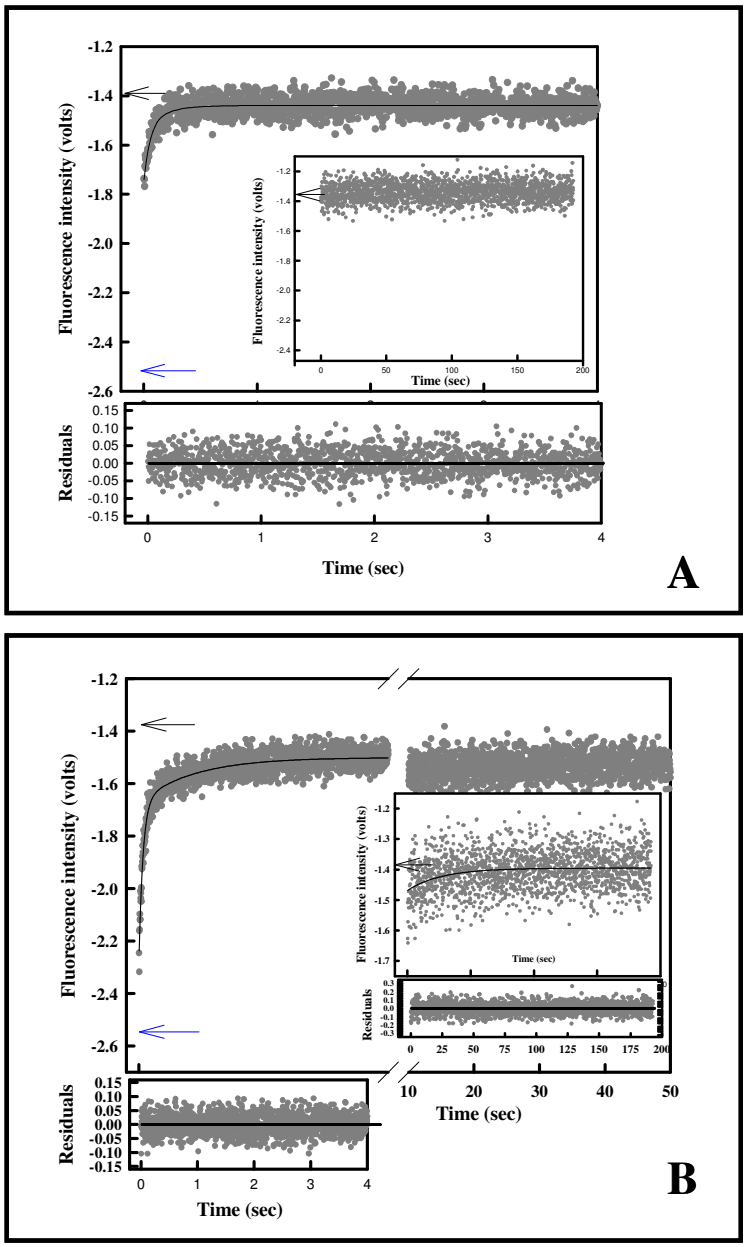


Figure 18: Double-jump refolding traces for unfolding delay in 5.5 M urea
Refolding of 5 μM wild-type Grx2 in 0.9 M residual urea following (A) 10 seconds and (B) 50 seconds unfolding delay of 30 μM wild-type Grx2 in 5.5 M urea. The refolding was monitored using intrinsic Trp fluorescence at 20 °C. For the 10 seconds and 50 seconds unfolding delay the first 4 seconds of recorded data were fitted to a double exponential. The refolding was complete by 4 seconds and no slow phase was observed for 10 seconds unfolding delay (inset) while an additional single exponential was fitted to the data for 50 seconds and longer delay times (inset). Data fitting was shown to be good by residuals plots. The arrows indicate the average baseline value for 5 μM native (black arrow) and 5 μM urea denatured (blue arrow) wild-type Grx2.

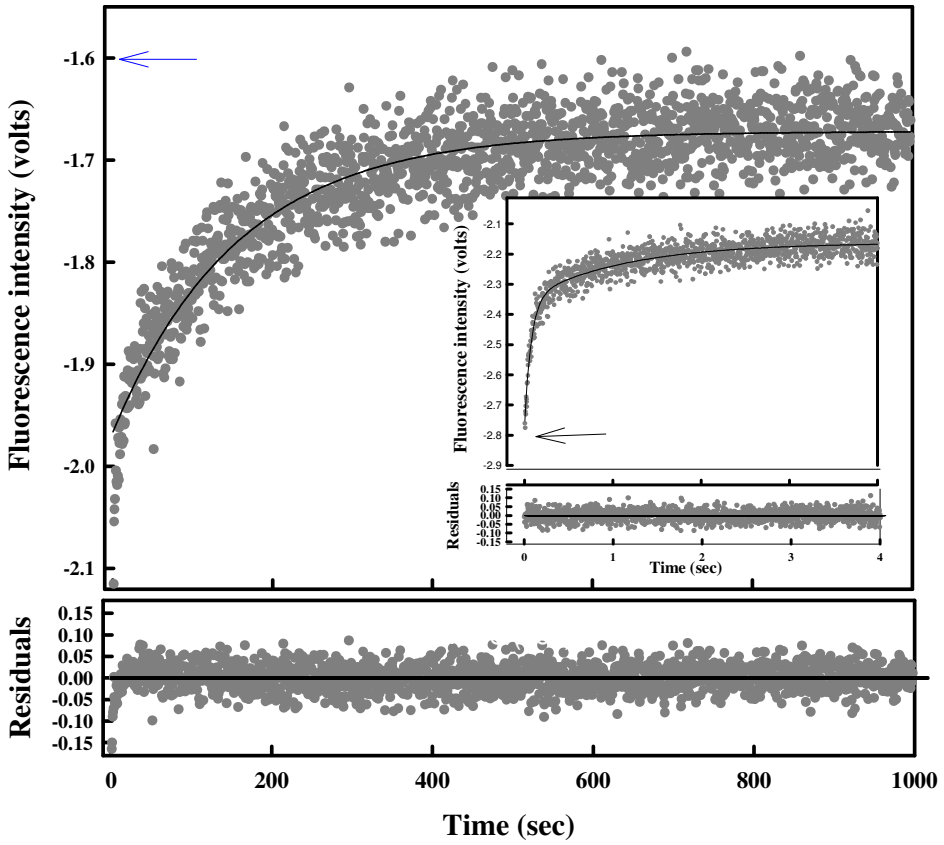


Figure 19: Refolding following 500 sec unfolding delay in 5.5 M urea
Average of three refolding trace of 5 µM of wild-type Grx2 in 0.9 M residual urea, pH 7 following unfolding delay of 500 seconds in 5.5 M urea at 20 ºC. Refolding was monitored by intrinsic Trp fluorescence and data were fitted to a tri-exponential function. This fit was shown to be a good by residuals plots. The slow phase (fitted to a single exponential function) occurs over 1000 seconds while the fast and medium phase are represented in the inset (fitted to a double exponential function) occur over 4 seconds. The average baseline value for 5 µM native (blue arrow) and denatured (black arrow) wild-type Grx2.

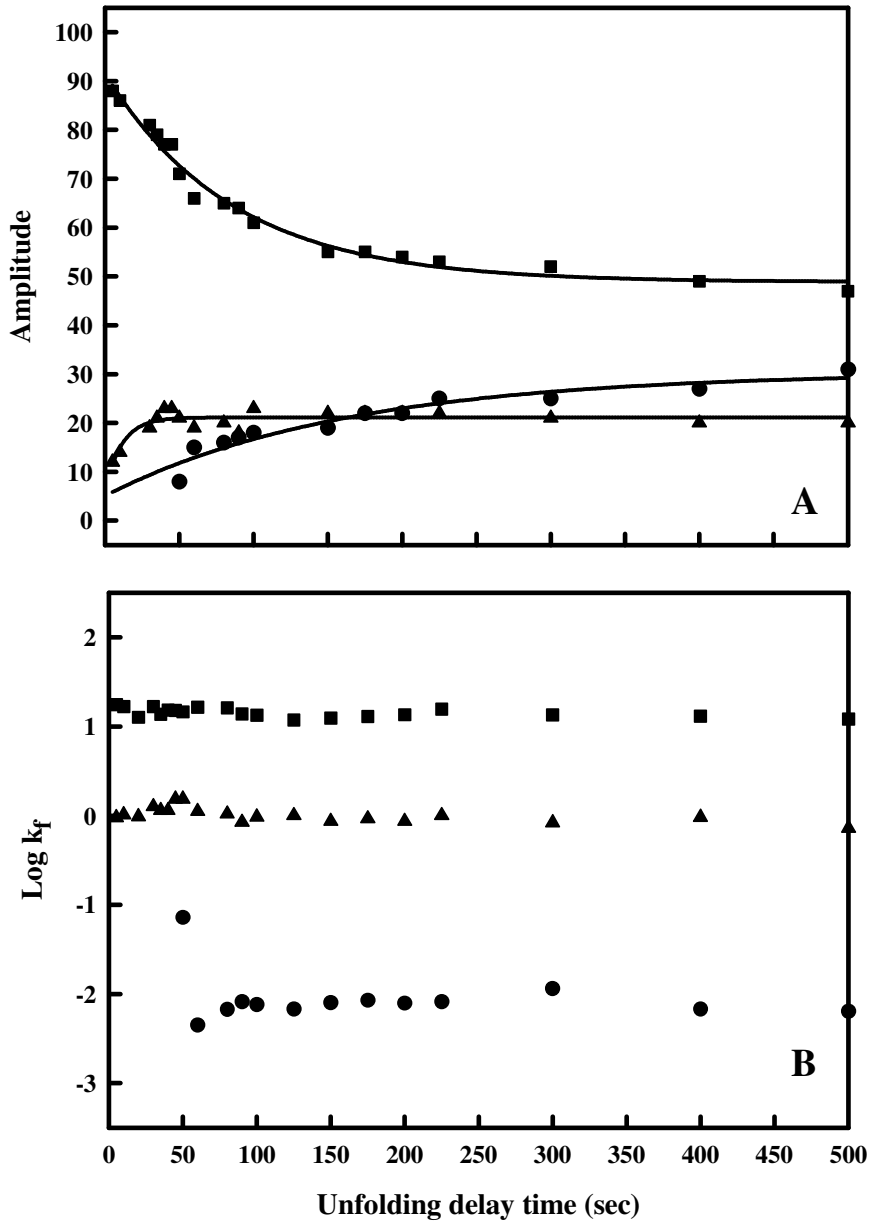


Figure 20: Amplitude and rate changes for 5.5 M urea double-jump
The (■) fast, (▲) medium and (●) slow phase, (A) amplitudes and (B) log of the rates were plotted against the unfolding delay time. Refolding traces were obtained and fitted as described previously in this section 3.6.1 (also see Figure 18 and Figure 19). The amplitudes displayed exponential changes (exponential fits indicted by solid black lines).

Overall changes in the amplitude data include a decrease in the fast phase amplitude with increasing unfolding delay time and an increase in both the slow phase and medium phase amplitude (Figure 20 A). The inability to detect the presence of a slow phase for short unfolding delay periods (5 to 45 seconds) does not support the presence of a lag in the development of the slow phase. Rather at the short unfolding delays it is expect that all the peptide bonds are in the native conformation and therefore there would be none to very few of the proteins passing through an isomerization phase to the native state. The trends in the amplitude data could be expressed mathematically and were fitted to single exponentials. Exponential decrease of the fast phase has a rate of 0.012 sec^{-1} (time constant of 85 seconds). The exponential increases for the medium and slow phases have rates of 0.0089 sec^{-1} (11 seconds time constant) and 0.061 sec^{-1}(164 seconds time constant), respectively. There is no corresponding trends for the rates of folding for all the refolding phases over the range of unfolding delay times (Figure 20 B). The only outlier for the rate data was that of the slow phase for the unfolding delay time of 50 seconds, this rate is faster because the slow phase here was fitted for a shorter time period than the other slow phases.

3.6.2 Double-jump refolding with unfolding delay in 6 M urea

A second double-jump experiment was conducted where unfolding in 6 M urea was allowed to occur for various delay times. These unfolding conditions were again selected in order to monitor the refolding of fast and slow unfolded species as well as avoid the 5.5 M urea unfolding region where changes in the unfolding occur (see section 3.5.1 and Figure 15). Similarly to the 5.5 M double-jump refolding traces obtained the refolding of transiently unfolded (6 M urea) wild-type Grx2 displayed fast and medium phases of refolding only for short delay times. For delay times of 10 and 20 seconds the slow phase amplitude was too small to be detected and fitted in the refolding traces of wild-type Grx2 (Figure 21 A). Refolding traces from 30 seconds to 50 seconds unfolding delay times showed a small slow phase and were fitted to a triple exponential (Figure 21 B).

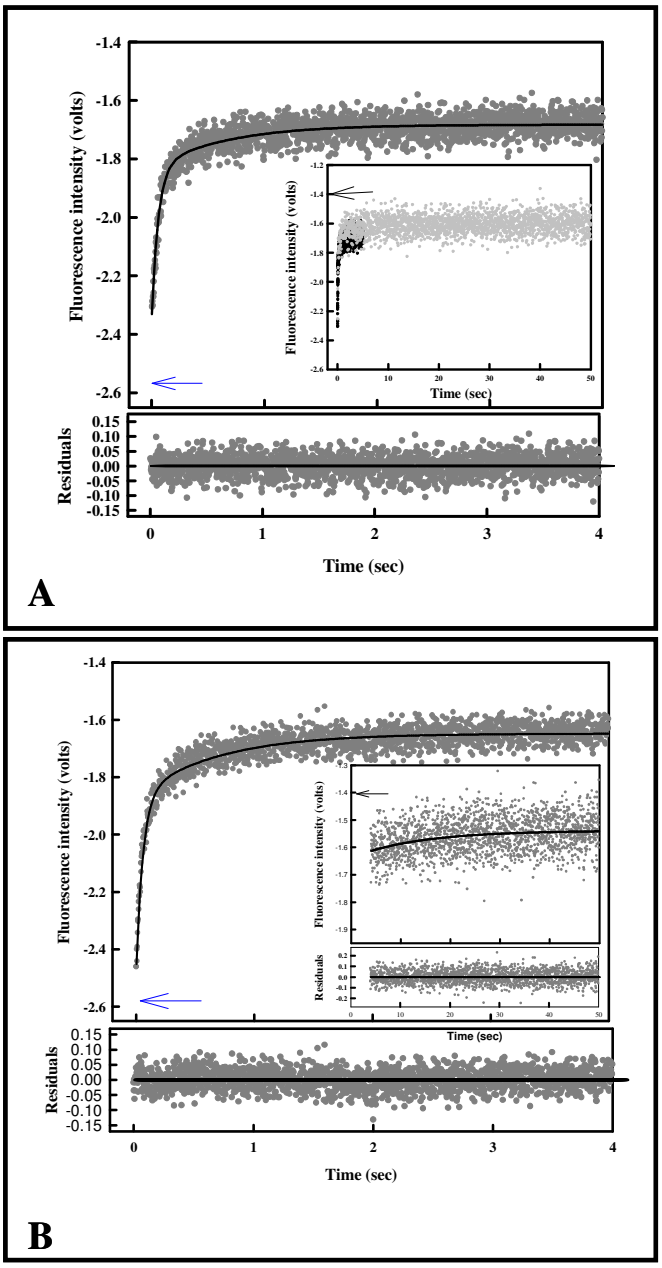


Figure 21: Double-jump refolding traces for unfolding delay in 6 M urea
Average (of three) refolding traces of 5 µM wild-type Grx2 in 1 M urea, pH 7 following unfolding delay in 6 M urea. Traces for delay times of (A) 10 seconds and (B) 30 seconds are shown. Refolding was monitored at 20 ºC, by intrinsic Trp fluorescence. The 10 and 30 second data were fitted to a double exponential for the first 4 seconds of data and for the 30 second data only an additional single exponential (inset) was fitted for 4 to 50 second refolding trace data. The exponential fits are represented by solid lines and were shown to be good by the residuals plot. In A the inset shows that refolding is complete after 4 seconds (dark grey data is 4 second data and light grey 50 second refolding trace). The arrows indicate the average baseline value for 5 µM native (black arrow) and 5 µM urea denatured (blue arrow) wild-type Grx2.

From 80 seconds unfolding delay to 500 seconds unfolding delay, the slow phase is present and fitted as for the single-jump refolding traces (Figure 22). At 500 seconds unfolding delay the slow phase has an amplitude of 30 % and a rate of 0.007 s^{-1} (time constant of 140 seconds).

The slow phase of refolding develops from 20 seconds unfolding delay and the amplitude increases exponentially with increasing unfolding delay times (Figure 23 A). As the slow phase of refolding increases with increasing unfolding delay times, the fast phase of refolding decreases exponentially (Figure 23 A). The medium phase amplitude remains independent of both the slow and fast phases and is constant from the 10 seconds unfolding delay time to the 500 seconds unfolding delay time in 6 M urea. The exponential increase of the slow phase and the decrease in the fast phase amplitude were both fitted to single exponential functions. The fast phase showed a decrease with a time constant of 76 seconds (rate of 0.013 sec^{-1}). The slow phase had an exponential development with increasing unfolding delay time with a time constant of 100 seconds (rate of 0.010 sec^{-1}). In order to determine if changes occur to the rates of the refolding phases, the log of the rates was plotted against the delay time. Due to the slow phase having a small amplitude and occurring over a shorter time period for unfolding delay times of 30 seconds to 50 seconds, the rate would appear to have changed (Figure 23 B). This change in rate is however not a true reflection that a rate change has occurred due to the differences in fitting the slow phase data. There was no effect on the rates of the fast and medium phases. The slow phase does not display a lag but similarly to the 5.5 M double-jump data no *trans* isomers have formed for unfolding delay times shorter than 30 seconds.

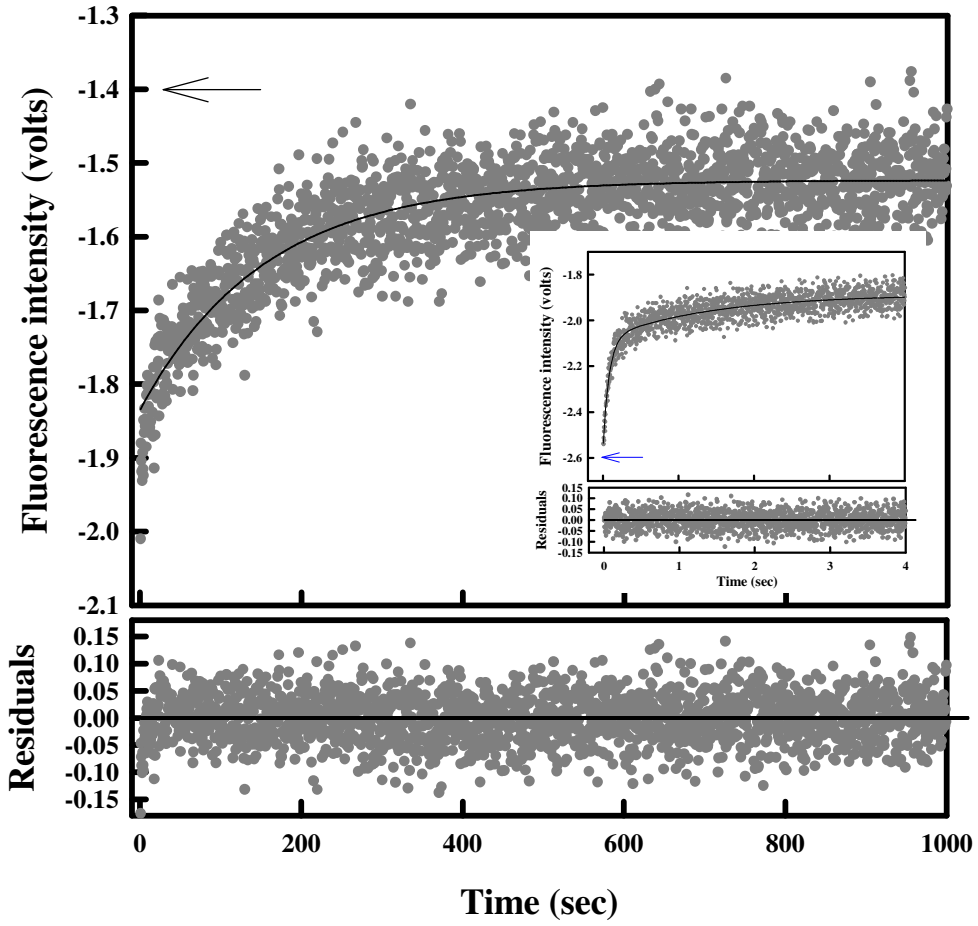


Figure 22: Refolding following 500 seconds unfolding delay in 6 M urea
Refolding trace for 5 µM Grx2 in 1 M residual urea, pH 7 recorded at 20 ºC, using intrinsic Trp fluorescence (inset first 4 seconds of refolding). The solid lines represent the exponential fits to the data (fitting show to be good by residuals plots). The slow phase fits a single exponential (main graph) while the fast and medium phases are shown in the inset fit to a double exponential. The arrows indicate the average baseline value for 5 µM native (black arrow) and 5 µM urea denatured (blue arrow) wild-type Grx2.

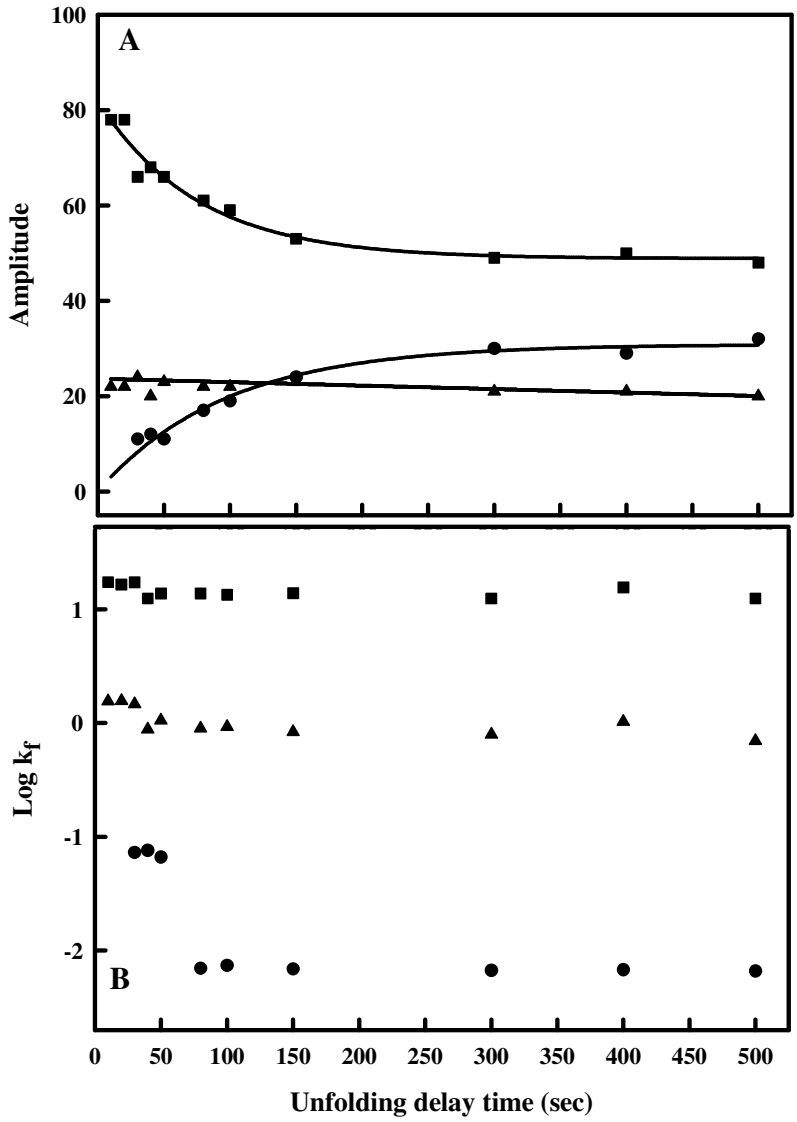


Figure 23: Amplitude and rate changes for 6 M urea double-jump
The (■) fast, (▲) medium and (●) slow phase, (A) amplitudes and (B) log of the rates were plotted against the unfolding delay time. Refolding traces were obtained and fitted as described previously in this section 3.6.2 (also seeFigure 21 and Figure 22). The fast and slow phase amplitudes displayed exponential changes (exponential fits indicted by solid black lines).The medium phase data were fitted to a linear function.

3.7 Refolding experiments in the presence of hFKBP-12

3.7.1 Double-jump refolding experiment in the presence of hFKBP-12

Refolding double-jump experiments were conducted in the presence of 1 µM hFKBP-12 a peptidyl-prolyl isomerase (PPI). Refolding double-jump experiments were conducted as described in 3.6.2 the only exception was that PPI, hFKBP-12 was introduced in the refolding buffer. In the presence of the PPI an increases in the rate of the slow phase of refolding was observed (Figure 24). The time constants for the fast and medium phases of refolding, in the absence and presence of the PPI, were the same. There was no increase in the rate or catalysis of the reactions that are taking place during these folding events, but the slow phase shows an increase in rate in the presence of PPI. The increase in the rate can be seen by a clear decrease in the time constant (the inverse of the rate) (Figure 24).

Amplitude changes observed in the absence of PPI (Figure 23 A) are the same as those observed in the presence (Figure 25) with the fast phase amplitude decreasing exponentially and the slow phase amplitude increased exponentially. The exponential increase in the slow phase indicated that the slow phase develops during unfolding with a time constant of 243 seconds (rate of 0.0041 sec^{-1}). This is greater than for the double-jump refolding experiment in the absence of PPI. The higher time constant for the development of the slow phase may not be significant, as the data set is smaller for this experiment and therefore could change the fit of the exponential to the data. The fast phase decreases exponentially with a time constant of 71 seconds this is very similar as for the decrease in the fast phase in the absence of PPI (time constant of 77 sec).

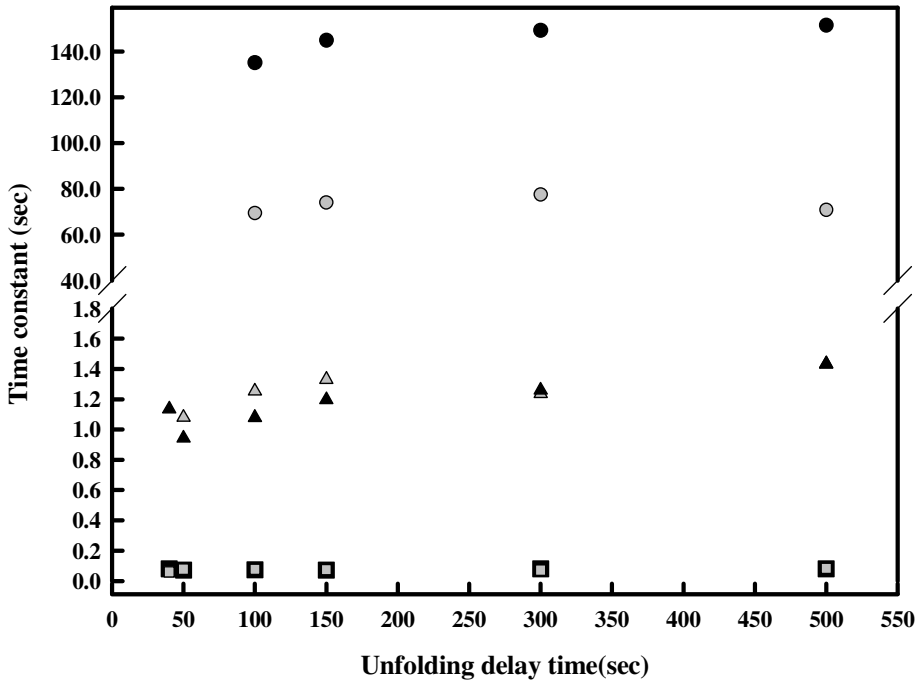


Figure 24: Time constants for double-jump refolding in the presence of hFKBP-12

The time constants obtained for the double-jump refolding in the presence of PPI for the fast (squares), medium (triangles), and slow (circles) phases. The refolding double-jump experiment was conducted as described previously (3.6.2) only 1 µM hFKBP-12 was added to the refolding buffer for each unfolding delay time. The refolding of wild-type Grx2 in the presence (black symbols) and absence (Grey symbols) of the PPIase was monitored using intrinsic tryptophan fluorescence at 20 °C (excitation 280 nm with emission monitored with a 320 nm cut off filter). The resultant refolding traces fit to a double (40 and 50 sec unfolding delays) and triple exponentials (unfolding delay times longer than 50 sec).

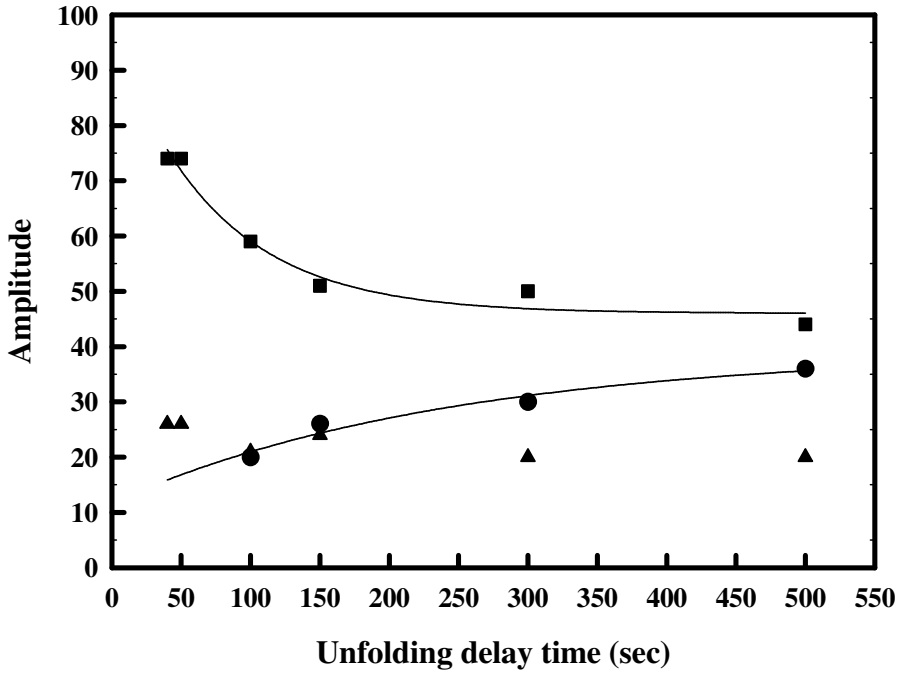


Figure 25: Amplitudes for double-jump refolding in the presence of hFKBP-12

The refolding double-jump experiment was conducted and data fitted as described previously **Figure 24**. From the data the amplitude changes for the (■) fast, (▲) medium and (●) slow phase of refolding in the presence of 1 μM hFKBP-12 are shown here. The fast phase and slow phase data were fitted to single exponential functions (solid lines).

3.7.2 Single-jump kinetic refolding in the presence of hFKBP-12

The effect of the PPI, hFKBP-12 on the refolding kinetics of wild-type Grx2 was determined by conducting single-jump refolding experiments. Wild-type Grx2 was allowed to unfold for at least an hour to allow the protein to unfold to equilibrium. Refolding of the equilibrium-unfolded protein was then initiated by dilution and monitored by intrinsic tryptophan fluorescence. PPI was added to the refolding buffer. There are no changes in the amplitudes of the respective three refolding phases (Figure 26 A). The fast phase maintains a 50 % amplitude, the medium phase a 20 % and the slow phase a 30 % amplitude. These values are consistent with those obtained for single-jump refolding (wild-type Grx2) for 1 M urea in the absence of PPI (Figure 17). Changes are however observed for the rates of the slow phase as the PPI concentration is increased. The slow phase was enhanced four fold when the final concentration of hFKBP-12 was 3 μM, there was however no affect on either the fast or the medium phases of refolding at this concentration of hFKBP-12 (Figure 26 B). A linear increase was observed with increasing hFKBP-12 concentration for the slow phase (slope of 1.1) of refolding, but no increase was observed for the fast and the medium phases (slopes of 0 and 0.1 respectively) of refolding.

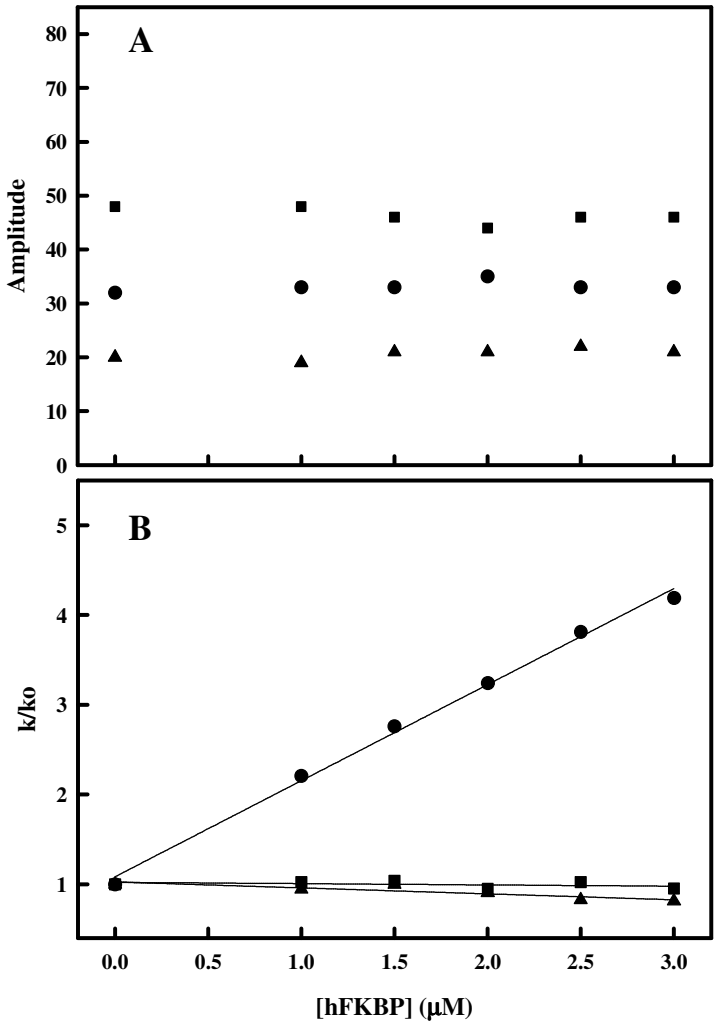


Figure 26: Refolding of wild-type Grx2 in the presence of hFKBP-12
The (A) amplitudes of the refolding phases and (B) rate of the refolding phases in the presence (k) of the PPI divided by the rate of the same phase in the absence (ko) of PPI was plotted against hFKBP-12 concentration. The (■) fast, (▲) medium and (●) slow phase data were plotted. Final refolding conditions: 5 µM wild-type Grx2 in 1 M urea with hFKBP-12 (concentration range 0 µM to 3 µM). The solid lines represent linear function fits to the data.

4 DISCUSSION

4.1 Stability of Grx2 proteins

Three Grx2 proteins were investigated in order to determine the unfolding and refolding pathways for Grx2. One of these proteins was wild-type Grx2 and then using the wild-type protein one protein was created by labelling a residue with AEDANS. The third protein was produced by the insertion of a local intrinsic probe (tryptophan residue) by mutagenesis. As the mutant protein was created to be identical to the wild-type protein in terms of stability and folding it should not show any changes in stability. Equilibrium unfolding studies were used to assess the stability of the proteins. Prior to these studies it was evident from the secondary and tertiary structural characterization of the mutant and AEDANS-labelled protein that the mutant had similar structure to the wild-type but the labelled protein had been altered. The secondary structural nature of the AEDANS-labelled protein was distinctly denatured from that of the native unlabelled wild-type protein (see far-ultraviolet circular dichroism spectra wild-type and AEDANS labelled protein Figure 8). The spectrum for the mutant protein indicates that AEDANS-labelled glutaredoxin loses the α-helical nature that was present for the native unlabelled wild-type protein just as the wild-type protein does on unfolding.

The mutant protein (Y58W) was expressed in high yield (70 mg protein per litre 2XYT media, in comparison to wild-type which yielded 80 mg protein per litre 2XYT media) as well as was found in the soluble cytoplasm fraction of whole cells (Figure 6), all indicating that the tryptophan mutant had not dramatically altered the stability of the mutant protein. The insertion of a tryptophan residue at a position that is normally occupied by a tyrosine residue should have minimal to no affect on the native protein because both residues have similar hydrophobicity values (Wolfenden *et al.,* 1979) and average buried volumes (Chothia, 1975). Also Tyr58 is situated close to the surface of the protein in a slightly solvent exposed positions as is evident from and increase in the fluorescence intensity at

345 nm (Figure 9), a wavelength that was shown for the wild-type protein to represent the emission of slightly exposed tryptophan residues.

Following the evidence that the secondary structure of AEDANS-labelled protein had been altered, the stability of the proteins was assessed by equilibrium unfolding studies. Data for wild-type Grx2, Y58W Grx2 and AEDANS-labelled Grx2 were all fitted to a single sigmoidal transition. The parameters of $\Delta G(H_2O)$ as well as the *m*-value and C_m for the transition region of unfolding are obtained from fitting the data to a sigmoidal transition (Table 2). The unfolding therefore for all three proteins was two-state, displaying a cooperative unfolding reaction.

The $\Delta G(H_2O)$ value for the unfolding of a protein in the presence of water as extrapolated from the urea-induced unfolding indicates the conformational stability of the protein (Pace, 1986). The average $\Delta G(H_2O)$ values for all the spectroscopic probes obtained from fitting the single sigmoidal curve for native unlabelled protein, mutant protein and labelled protein are 12.3 kcal.mol^{-1}, 9.5 kcal.mol^{-1} and 5.4 kcal.mol^{-1} respectively. The $\Delta G(H_2O)$ values obtained for all of the Grx2 proteins is within the range stated for monomeric proteins (6 to 14 kcal.mol^{-1}) (Neet and Timm, 1994) although the decrease in conformational stability indicates that addition of AEDANS to Grx2 destabilizes the protein while mutagenesis resulting in the substitution of Tyr58 with Trp does not significantly destabilize the protein.

As proteins unfold, there is an increased exposure of surface area. The *m*-value obtained from the sigmoidal fit gives an indication of the amount of surface area exposed during unfolding of a protein (Alonso and Dill, 1991; Myers *et al*., 1995). The change in accessible surface area for wild-type Grx2 and Y58W Grx2 was calculated to be 19088 Å^2 using the following equations of Myers *et al.* (1995):

$$(n \times 93) - 907 = \Delta ASA \qquad (17)$$

where n is the number of residues for the protein. The *m*-value for urea as denaturant can then be predicted based on the calculated change in accessible surface area:

$$(\Delta ASA \times 0.11) + 374 = m \quad (18)$$

Using this equation the predicted *m*-value for wild-type Grx2 and Y58W Grx2 is 2.5 kcal mol^{-1} M^{-1} urea (Myers *et al*., 1995). This predicted value agrees well with the experimentally obtained *m*-values for both wild-type Grx2 and Y58W Grx2 (Table 2). From these values it can be seen that the mutant exposes a similar amount of surface area as the wild-type protein and that the cooperativity of the unfolding reaction has also not been altered dramatically by the substitution of the Trp58 residue. The low *m*-value however, obtained for the unfolding of labelled Grx2 indicates that the protein has more exposed surface area in the native state than the unlabelled protein and therefore is destabilized by the modification with AEDANS. As the destabilized protein unfolds, less surface area is exposed as some is already exposed in the native folded structure. This result is consistent with the far-ultraviolet spectrum obtained for the native labelled protein that indicated that there was a loss of the proteins characteristic secondary structural elements due to the label. The cooperativity of the unfolding reaction for Y58W Grx2 has also decreased indicating that intermediates may be present during unfolding. The presence of intermediates in the unfolding pathway of the labelled protein is further confirmed by the lack of agreement of the parameters, obtained from the sigmoidal two-state fit, between the tertiary structural probes and the secondary structural probes. The lower *m*-values for all the probes used to monitor the unfolding of the labelled protein versus the two probes used to monitor the unlabelled protein indicate that there is more than one equilibrium state (Myers *et al*., 1995). This behaviour was not seen for the wild-type and Y58W Grx2 proteins, clearly indicating that the unfolding pathway as described by the equilibrium data is a two-state process with no intermediates.

The C_m value is the midpoint of the unfolding transition. The unlabelled protein has C_m value of 4.4 M urea, which is in agreement with the C_m value of the mutant protein (4.3 M). The C_m values for the labelled protein however much lower at 3.0 M urea. This indicates that lower amounts of denaturant are able to disrupt the labelled proteins structure. The C_m value for the labelled protein adds to the conclusion that the labelling of the protein has disrupted the native structure of Grx2.

The conformational stability and structure of AEDANS-labelled Grx2 is decreased substantially compared to the wild-type protein therefore the labelled protein cannot be used to assess local changes that would occur at the active site region of domain 1. The mutant protein however showed no significant changes in stability or structure in comparison to the wild-type protein with the equilibrium data displaying a single sharp transitions as for wild-type Grx2 (Figure 13). 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 rates of the slow unfolding phase as well as the three refolding phases of Y58W were not significantly altered in comparison to the wild-type protein [see section 3.5.1 (unfolding) and 3.5.2 (refolding) for rates] further indicating minimal changes in stability of Y58W Grx2 in relation to wild-type Grx2. The slight decrease in the stability of Y58W Grx2 (Table 2) is however reflected in the increased rate of the fast phase of unfolding of the mutant protein (see section 3.5.1).

Y58W Grx2 and wild-type Grx2 display two-state highly cooperative equilibrium unfolding behaviour. This two-state unfolding behaviour has been identified for homologous dimeric GST family members such as Pi (Dirr and Wallace, 1999; Erhardt and Dirr, 1995) *Schistosoma japonicum* (Kaplan *et al.,* 1997), and Alpha (Wallace *et al.,* 1998; Wallace and Dirr, 1999) GST. Interestingly there are also examples of these homologues dimeric GST's where the unfolding is multistate which can be attributed to in all cases a monomeric intermediate (Hornby *et al.,*

2000; Hornby *et al.,* 2002; Sacchetta *et al*., 1993; Stevens *et al*., 1998). There is however no evidence that the individual domains of the subunits fold independently. Rather there is a high degree of cooperativity between the domains for many of these GST proteins subunits (Erhardt and Dirr, 1995; Kaplan *et al.,* 1997; Hornby *et al.,* 2000; Hornby *et al.,* 2002; Luo *et al.,* 2002; Stevens *et al*., 1998; Wallace *et al.,* 1998). It is possible from equilibrium studies to identify domains that denature independently as is the case for human κ light chain (Rowe and Tanford, 1973). Based on the equilibrium data for Grx2 and the evidence of the structurally homologous GST proteins (Xia *et al.,* 2001) it can be concluded that the domains of Grx2 do not fold independently and display a high degree of cooperativity.

4.2 Unfolding of Grx2

The conclusion drawn from the equilibrium unfolding data is that unfolding is a highly cooperative two-state process for Grx2. Based on this we would expect to have a single kinetic phase for unfolding with no intermediates. The kinetic unfolding of wild-type Grx2 and Y58W however is composed of two phases (Figure 14) as shown by the parameters obtained from fitting the data to a double exponential decay curve, assessed by a residuals plot to be a good fit. The dependence of the unfolding amplitudes as well as the rates of each phase is complex over the range of unfolding urea concentrations for wild-type Grx2 (Figure 15). Both phases display a change in urea dependence of the rates of unfolding at 5.5 M urea and 5.6 M urea, for the fast and slow phases respectively. Changes in the urea dependence of unfolding have indicated the presence of intermediates in the unfolding of arc repressor (Jonsson *et al*., 1996). Intermediates are present in the unfolding of Ure2 a structurally related protein to Grx2 (Galani *et al*., 2002) and gene-3-protein (Martin and Schmid, 2003b). Both of these proteins display similar changes to those observed for Grx2, in the urea dependence of the rates and amplitudes of unfolding. These changes in amplitude over a range of denaturant concentration similar to those seen for wild-type Grx2 (Figure 15 A) have been observed previously where it was concluded that at least

one intermediate was present during unfolding (Zaidi *et al.,* 1997). The changes at 5.5 M urea and 5.6 M urea (for the fast and slow phases respectively) were not detected for Y58W Grx2. Therefore intermediates formed during unfolding of Grx2 could involve structural elements in the local environment of Trp89 and Trp190 and not Trp58 as Trp 89 and Trp 190 are situated in the domain 2 of the protein (Vlamis-Gardikas, *et al*., 1997; Xia, *et al*., 2001) and Trp 58 has been inserted into domain 1.

The slow phase of unfolding is urea independent (Figure 15 B) a characteristic which has been attributed to phases which involve proline peptide bond isomerizations (Kiefhaber *et al.,* 1992a; Schmid and Baldwin, 1978). This phase also possesses a time constant of 140 sec, which is within the range for isomerization reactions (Brandts *et al*., 1975). The only Xaa-proline peptide bond in Grx2 that is in the *cis* conformation is the Val48-Pro49 peptide bond (Xia *et al.,* 2001). This peptide bond will isomerise to the *trans* conformation as the *trans* conformation is favoured in the unfolded state (Brandts *et al*., 1975; Hinderaker and Raines, 2003). Many proteins however, that contain more than one *cis*-proline in the native state have simple unfolding kinetics, with only a single unfolding phase, such ribonuclease A (Dodge and Scheraga, 1996) and thioredoxin from *Escherichia coli,* a member of the same superfamily as Grx2 (Kelley and Stellwagen, 1984). Phases however, of unfolding that are isomerization phases can be detected for urea concentrations close to the transition region of unfolding and then disappear at higher urea concentrations (Brandts *et al*., 1975; Kiefhaber *et al.,* 1992a; Kiefhaber and Schmid, 1992). This behaviour is not seen for Grx2 as the slow phase maintains 20 % amplitude from 5.75 M urea to 7.5 M urea (Figure 15 A). Unfolding of staphylococcal nuclease A is biphasic with a slow isomerization phase with an amplitude of 90 % (Ikura *et al*., 1997). The slow phase of unfolding for Grx2 has an amplitude of 20 % under strong unfolding conditions. Unfolding phases that represents the major phase for low denaturant concentrations near to the transition region and then decrease with increasing denaturant concentration are said to represent *cis*-proline peptide bond isomerization events that are coupled to conformational folding events (Kiefhaber

et al., 1992a). The complex dependence of the amplitudes of the slow phase on urea concentration for wild-type Grx2 displays such behaviour so therefore this is a possible explanation for what is occurring during the slow phase of unfolding of Grx2. The complex dependence of the amplitudes and rates on urea concentration were not detected for the mutant (Y58W Grx2) and because Trp58 is situated in the local environment of the *cis*-Pro49 we would expect that changes due to the isomerization would be detected for the mutant. Alternatively the change in amplitude dependence from high amplitudes at low urea concentrations to low amplitudes at high urea concentrations for the wild-type Grx2 (Figure 15 A) could have shifted to lower urea concentrations for the mutant protein. The beginning of the increase in the slow phase amplitude with decreasing urea for the mutant protein is observed from 5.5 M to 5 M urea (Figure 15 A). The slow phase of unfolding for staphylococcal nuclease was also said to involve a proline isomerization (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). An isomerization occurs during the slow phase of unfolding and there is the presence of intermediates during unfolding of Grx2 although it is not clear whether a parallel or sequential unfolding pathway exists.

The m_u values obtained by fitting equation 15 (Tanford, 1970) can be used to indicate solvent exposure during unfolding (Doyle *et al.,* 1996). The m_u values reported in Table 3 for wild-type and Y58W Grx2 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. Using the m_u values obtained for the wild-type protein may not give an accurate description of the changes in surface area due to the complex dependence the rates show on the urea concentration and therefore the presence of intermediates (Doyle *et al.,* 1996). The m_u values for the mutant protein are still reliable and they agree well with the values for the smaller linear regions of the data that were fitted for the wild-type protein. It can therefore still be concluded that the unfolding of Grx2 involves the fast phase (a gross structural

unfolding phase) and the slow phase where only small rearrangements occur such as an isomerization with intermediates forming during unfolding, but do these phases represent a parallel or sequential unfolding pathway?

Garel and Baldwin (1973) have demonstrated that fast and slow unfolding species coexist under unfolding conditions. An initial conditions test for unfolding conducted by Dr L. A. Wallace (unpublished data) demonstrated the presence of two native-like species that would result in two parallel unfolding pathways. Double-jump refolding experiments confirmed a parallel unfolding mechanism. No lag in the development of any refolding phase (Figure 20 A, Figure 23 A) indicates that at short unfolding delays there is still the presence of all the unfolded species necessary for refolding. For a sequential unfolding pathway when short unfolding delay times are used during a refolding double-jump experiment, only part of the sequential unfolding reaction would have taken place and then the additional unfolding reaction would be required to occur prior to refolding, resulting in a lag in the development of the refolding phase. Therefore, the most likely unfolding pathway is a parallel one for Grx2.

4.3 Folding of Grx2

The single-jump refolding kinetics for wild-type and Y58W Grx2 display three phases of refolding (Figure 16). Thioredoxin, a protein closely related to Grx2 in structure, was also show to have three refolding phases fast (0.54 seconds), slow (14 seconds) and slowest (500 seconds) (Kelley and Stellwagen, 1984). The order that the Grx2 refolding phases occur in as well as the identity of these phases was investigated. The urea dependence of the rates of the three phases of refolding for wild-type and Y58W Grx2 show complex behaviour, with a “roll-overs” (Baldwin, 1996) or changes in urea dependence occurring at certain urea concentrations. The complex urea dependence of all the refolding phases (Figure 17 B) indicates the presence intermediates during folding. The detection of the folding intermediates for Grx2 are said to be early folding intermediates based on the position of the change in urea dependence at low urea concentrations (Zarrine-

Afsar and Davidson, 2004). The complete change in secondary structure formation for Grx2 occurs within the dead-time of the stopped-flow (Dr L A Wallace (unpublished data), this supports the conclusion of the formation of an early intermediates. Similar intermediates that are not due to the presence of isomerization of a peptide bond were detected for the structurally related thioredoxin protein (Georgescu *et al*., 1998). The structural similarity between Grx2 (Xia *et al.,* 2001) and this protein is only for domain 1 of Grx2 and the entire structure of thioredoxin (Holmgren *et al.,* 1975). In order to position the intermediates within a folding pathway it was necessary to identify if the folding pathway involves the parallel folding of protein along all three phases as well as what would create such heterogeneity of the unfolded state.

4.4 Slow phase in Grx2 folding is an isomerization phase

A bond isomerization requires a high energy of activation (20 kcal.mol^{-1}) and therefore is a slower reaction with a time constant of 10 to 1000 seconds at room temperature (Brandts *et al*., 1975). The slow phase of refolding for wild-type and Y58W Grx2 have time constants of 180 seconds and 212 seconds respectively. These time constants fit into the range mentioned for bond isomerization. In general the slow phase of refolding can be attributed to an isomerization of an Xaa-proline peptide bond (Brandts *et al*., 1975; Kiefhaber *et al*., 1990; Schmid and Baldwin, 1978). The isomerization for Grx2 is presumably of the Val48-Pro49 peptide bond, as it is the only Xaa-proline peptide bond in the *cis* conformation out of 12 proline residues, in the native protein (Xia *et al*., 2001). When a protein unfolds an equilibrium mixture is produced of both the *cis* and *trans* isomers. The isomers do not however have a 50:50 percent ratio rather the equilibrium is skewed (30:70 ratio *cis*:*trans*) and the *trans* conformation is favoured (Brandts *et al*., 1975; Schmid, 1992). Therefore a substantial proportion of the unfolded species are required to undergo isomerization to form the native protein. The slow phase of refolding accounts for only 30 percent of the amplitude change for the refolding of Grx2 at 1 M residual urea. This would indicate that only 30 percent of the unfolded species are in the *trans*

conformation. It has been shown that 90 to 70 percent of the unfolded species exist in the *trans* conformation (Schmid, 1992), therefore the amplitude for the isomerization phase of refolding should reflect this percentage. The percentage amplitude change does however increase to 80 % at a residual urea concentration increases to 3.9 M urea. Brandts *et al*. (1975) have demonstrated that under strong (un)folding conditions the amplitude of the isomerization phase can be greatly reduces and even absent. The amplitude of an isomerization phase has been shown to have a smaller amplitude at low denaturant concentrations and then develop an increase in amplitude as the urea concentration increases (Kiefhaber *et al.,* 1992a; Kiefhaber and Schmid, 1992; Walkenhorst *et al*., 1997).

The dependence of the slow phase amplitudes on urea is also relevant in the identification of this phase as the isomerization phase. The amplitude of slow phase of refolding increases with increasing urea concentration (Figure 17 A), this behaviour was shown to indicate a phase involved in isomerization for notch ankyrin domain (Mello *et al*., 2005). The amplitude of the slow phase for Grx2 is small at low concentrations of urea and then becomes the major phase as the urea concentration increases. This behaviour has been typically seen for phases where isomerization is coupled to conformational events (Kiefhaber *et al.,* 1992a; Kiefhaber and Schmid, 1992). The slow phase could very well represent the isomerization of the *cis*-Pro49 as well structural refolding events but the denaturant dependence of the amplitudes of the isomerization phases are also said to occur due to the presence of rapidly forming intermediates. This has been shown to occur for ribonuclease T1 (Kiefhader *et al*., 1990), thioredoxin (Kelley *et al*., 1986) and lambda Cro repressor (Satumba and Mossing, 2002). The rates of all of the refolding phases displayed changes in urea dependence at various concentrations that was said to be due to the presence of intermediates (see section 4.3). The presence of intermediates during the refolding of Grx2 would also explain how the rate of the slow phase is relatively urea independent between 1 and 2.2 M urea but then becomes highly urea dependent between 2.5 M urea and 4 M urea (Figure 17 B and for the m_f values Table 4). The m_f values obtained for all the phases of refolding indicate that for urea concentrations of 1 to 2.2 M

urea the slow phase is urea independent. Phases of refolding with rates that display urea independence are said due to this to involve proline isomerization (Kiefhaber *et al.,* 1992a).

The isomerization of the proline peptide bonds are slow steps in folding and therefore the development of the incorrect isomer during unfolding can be monitored by the double-jump refolding experiments (Brandts *et al.*, 1975). Strong unfolding conditions are used so that unfolding proceeds more rapidly than isomerization (Wallace and Matthews, 2002). As the unfolding delay is increased, the amplitude of the isomerization phase will increase as more of the non-native isomer is formed during unfolding, requiring isomerization back to the native conformation during refolding. The amplitude of the isomerization phase of refolding will increase until equilibrium is reached between the species with different conformations of the Xaa-proline peptide bond. For wild-type Grx2 the slow phase amplitude increased exponentially (for both the 5.5 M and 6 M double-jump experiments) with increasing unfolding delay time. The exponential fitted to the development of the amplitude of the phase involved in isomerization gives the time constant for the formation of the slow folding species (Schmid, 1986b; Wallace and Matthews, 2002). The time constant obtained from fitting the increase in slow phase amplitude with increased unfolding delay time is 100 seconds and 164 seconds for the 6 M and 5.5 M urea double-jump experiments, respectively. These time constants are in agreement with that of the slow phase of unfolding (time constant of 140 sec) the potential isomerization phase. The medium phase also develops exponentially for the 5.5 M double-jump experiment (Figure 20) but the time constant (11 seconds) is rather small for an isomerization event as well as there is no development of this phase during the 6 M double-jump experiment (Figure 23). Therefore, the evidence supports the conclusion that the slow phase of refolding involves the isomerization of the Val48-Pro49 peptide bond. Double-jump experiments where isomerization phases were identified by the exponential increase of the amplitude of the isomerization phase with increasing unfolding delay time were conducted for thioredoxin (Bhutani and Udgaonkar, 2001) and pectate lyase C (Kamen and Woody, 2002). The

development of the isomerization phase amplitude for refolding double-jump experiments was not the only proof the phase was the isomerization of one Xaa-Proline peptide bond in the *cis* conformation in the native protein, the enhancement of the phase by PPI was further confirmation (Kamen and Woody, 2002).

Further confirmation that the slow phase of refolding for Grx2 is the isomerization of the Val48-Pro49 peptide bond is the acceleration of this phase by the peptidyl prolyl isomerase (PPI), hFKBP-12. PPIs act as catalysts, accelerating the *cis-trans* isomerization of Xaa-proline peptide bonds (Fischer *et al*., 1984). Therefore, the phase with the enhanced rate in the presence of this catalyst must involve an isomerization event (Lang *et al*., 1987; Schreiber and Fersht, 1993). The amplitude meanwhile should not be affected at all (Lang *et al*., 1987; Schreiber and Fersht, 1993). When hFKBP-12 is added to the wild-type Grx2 refolding buffers during single-jump refolding experiments the rate of the slow phase of refolding was enhanced 4-fold with 3 μM hFKBP-12. The rate of the medium and fast phases are unchanged and none of the amplitudes were affected by increasing amounts of the catalyst (Figure 26). The PPI was also introduced into the refolding buffer in a 6 M double-jump experiment. The results of the experiment indicated that the rate of only the slow phase of refolding was enhanced (Figure 24). This data indicates that the slow phase of refolding involves the isomerization in folding of Grx2. The amplitude for the slow phase remains unaffected by the presence of the PPI, for the double-jump (Figure 26), and single-jump experiments (Figure 26 A). The same percentage of unfolded species, therefore fold via the slow phase in the presence of the hFKBP-12 as in the absence of the hFKBP-12. Therefore, there is no channelling of folding species via the slow folding phase, in the presence of hFKBP-12. In the presence of PPI, an increase in the amplitude of the slow isomerization phase of folding of ribonuclease T1 was observed (Kiefhaber *et al*., 1990). The channelling of folding along defined kinetic routes in the presence of PPI has been shown to result in a higher yield of protein (Fransson *et al*., 1992). This does not occur for Grx2.

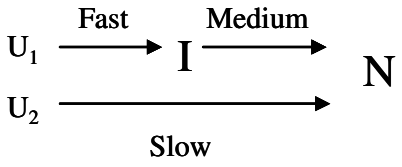
Borden and Richards (1990) concluded that the slow phase of refolding for thioredoxin from phage T4, was isomerization based on 3 observations; the denaturant independence, the time constant of the slow folding phase and the fact that slow refolding phase was dependant on dwell time or unfolding delay time during double-jump refolding experiments. Kamen and Woody (2002) concluded, based on evidence from double-jump experiments, catalysis of isomerization phases by cyclophilin (a PPI) and temperature studies, that for the folding of pectate lyase C, there are two slow phases of refolding that involve isomerization. Georgescu *et al.* (1998) also conclude that the slow phase in refolding of thioredoxin is the isomerization phase. A similar conclusion that the slow phase of refolding involves a *cis-trans* isomerization of a proline peptide bond can be made for Grx2 based on the evidence presented here. A summary of the evidence for the slow phase of refolding of Grx2 involving an isomerization are as follows:

- A time constant of 180 seconds (which is within the range of 10 to 1000 seconds at room temperature for phases said to involve Xaa-proline peptide bond isomerization (Brandts *et al*., 1975))
- The low urea dependence of the slow phase on urea concentration as well as the amplitude dependence behaviour (The amplitude of an isomerization phase has been shown to have a smaller amplitude at low denaturant concentrations and then develop an increase in amplitude as the urea concentration increases (Kiefhaber *et al.,* 1992a; Kiefhaber and Schmid, 1992; Walkenhorst *et al*., 1997)) (Figure 17 A and B)
- The development of the slow phase with a time constant of 164 sec and 100 sec with increasing unfolding delay times in 5.5 M and 6 M respectively, for double-jump refolding studies (Figure 20 A; Figure 23 A)
- Acceleration of the rate of the slow phase of refolding only during single- and double-jump refolding studies by the PPI hFKBP-12 (Figure 24; Figure 26 B)

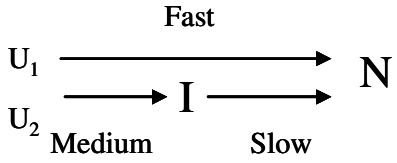
The isomerization of the conserved *cis*-proline in the thioredoxin superfamily of proteins was shown to occur for thioredoxin in a compact intermediate structure during folding and does not hinder folding (Chaffotte *et al.*, 1997; Cook *et al.*, 1979; Kiefhaber *et al.*, 1992b). It has been shown in the folding of Grx2 that burst phase species are formed that possess native secondary structure that do not bind ANS (Dr L. A Wallace, unpublished data). This would indicate that the folding of Grx2 initially forms a very native-like structure, which would mean that the isomerization of the conserved *cis*-proline peptide bond would have to occur in a structure that posses large amounts of native secondary structure in a relatively compact native structural arrangement. As well as the region where the *cis*-proline peptide bond is situated is proposed to be in a folding initiation site (to be discussed in section 4.6) which would form early in the folding of proteins. Other proteins that have been shown to fold prior to isomerization include pectate lyase C, where isomerization of a *cis*-proline peptide bond occurred in a folding intermediate and the urea dependence of the rates of the isomerization phase were shown to be urea dependent (Kamen and Woody, 2002). Although the slow phase is the isomerization phase the question still remains how does it fit into the folding pathway with the medium and fast phases of refolding.

4.5 A parallel or sequential folding pathway for Grx2?

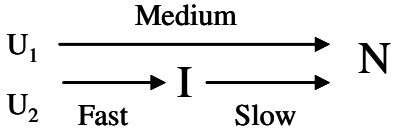
The unfolding of Grx2 is biphasic indicating the presence of potentially maximum three unfolded species (Georgescu *et al.*, 1998). While the isomerization about the Val48-Pro49 peptide bond would result in at least two unfolded species. Based on data presented here several potential folding pathways are proposed (Scheme 1, Scheme 2, Scheme 3 and Scheme 4).



Scheme 1: Proposed folding scheme 1 for Grx2

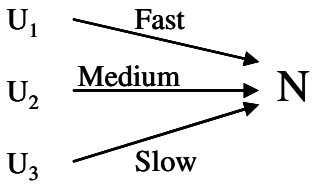


Scheme 2: Proposed folding scheme 2 for Grx2



Scheme 3: Proposed folding scheme 3 for Grx2

For Scheme 1, Scheme 2 and Scheme 3, U_1 would represent the unfolded protein with the Val48-Pro49 peptide bond in the *cis* conformation while U_2 the *trans* conformation. I is the kinetically detected intermediate and N is the native state with the Val48-Pro49 bond in the *cis* conformation.



Scheme 4: Proposed folding scheme 4 for Grx2

For Scheme 4, U_1 and U_2 would represent the unfolded protein with the Val48-Pro49 peptide bond in the *cis* conformation while U_3 the *trans* conformation. N is the native state with the Val48-Pro49 bond in the *cis* conformation

Double-jump refolding experiments allowed the folding mechanism of Grx2 to be further dissected. A parallel versus a sequential folding mechanism of folding can be determined from the double-jump refolding data. Double-jump refolding experiments monitor changes that occur as unfolding of the protein proceeds to completion over time. The different species formed throughout the unfolding reaction can be "trapped" and their presence detected by the monitoring altered refolding kinetics. The unfolding reaction can results in a group of species that can fold along alternate parallel routes to the native state, or a group of unfolded species that can only fold along a single sequential pathway to the native state (Wallace and Matthews, 2002). The double-jump refolding experiments for Grx2 showed an exponential development of the slow phase of refolding with an exponential decrease in the fast phase amplitude with increasing unfolding delay. A parallel folding pathway is proposed where unfolded species with the correct proline isomer fold during the fast phase to the native state and species with the incorrect isomer fold along the slow phase to the native state. The medium phase is not affected by the increase in unfolding delay and is constant from the shortest unfolding delay times for 6 M double-jump experiments, this indicates the phase is an independent pathway to the fast and slow pathways. The medium phase does show an exponential development for the 5.5 M double-jump experiment (Figure 20) but the development has a time constant of 11 seconds and does not appear to be linked to the fast or slow phase's changes in amplitude. The single-jump

refolding studies also point toward a parallel mechanism of folding for Grx2. Both the amplitude and rates dependence on urea shows that the fast and slow phases of refolding act in an opposite but linked way to those of the medium phase, which behaves independently of the fast and slow pathways. The rates for the fast and slow phases both show changes in urea dependence at 2.2 M urea while the medium phase does not show these changes (Figure 17 B). It has been shown that the percentage amplitude for a kinetic refolding phase represents the fraction of species that take part in that refolding phase (Hagerman and Baldwin, 1976). For Grx2 the amplitude dependence on urea shows how when the fast phase amplitude decreases the slow phase increases, therefore species if they are not folding via the fast phase they fold via the slow phase. In addition, the medium phase amplitude changes show that when folding does not occur along the fast/slow route then they fold via the medium phase. An initial conditions test conducted for the refolding of Grx2 indicated the presence three unfolded species (Dr L. A. Wallace, unpublished data). The data all supports the folding pathway proposed in Scheme 4. In order to eliminate the folding mechanisms proposed in Schemes 1, Scheme 2 and Scheme 3 an unfolding double-jump experiment would have to be conducted. Thioredoxin however, folds via a parallel folding pathway (Georgescu *et al*., 1998; Kelley and Richards, 1987). The parallel folding pathway stemming from a number of unfolded forms was determined for thioredoxin based on evidence where the amplitude of the fast phase of refolding decreased the slow phase amplitude increases (Bhutani and Udgaonkar, 2001). Other proteins that have parallel folding pathways as a result of proline isomerization include staphylococcal nuclease (Ikura *et al*., 1997; Kuwajima *et al*., 1991), ribonuclease T1 (Kiefhaber *et al*., 1990) and ribonuclease A (Dodge and Scheraga, 1996; Houry and Scheraga, 1996; Juminaga *et al*., 1998; Schultz *et al*., 1992).

A parallel mechanism of folding is the most likely mechanism of folding for Grx2. The fast and slow phases represent species that are separated by the difference in conformation about the Val48-Pro49 peptide bond. Species with the correct *cis*-peptide bond conformation fold via the fast phase of refolding and with the incorrect conformation the slow phase but there is no explanation as to

what accounts for the third phase the medium phase of refolding. What differences in the unfolded state result in the partitioning of species along the separate medium phase to the fast-slow route of folding?

4.6 Do the two domains of Grx2 play a role in (un)folding?

If the folding pathway for Grx2 involves the folding of three unfolded species of Grx2 along three parallel pathways to the native state and the slow and the fast phases, exist due to differences about the Val48-Pro49 peptide bond what results in the medium phase? Could the two domains of Grx2 (Xia, *et al*., 2001) be folding independently? Trp 89 and Trp 190 are situated in the domain 2 of the Grx2 protein (Vlamis-Gardikas, *et al*., 1997; Xia, *et al*., 2001) and therefore act as local tertiary structural probes during (un)folding of Grx2 when monitored using intrinsic tryptophan fluorescence. These residues however can only provide limited information on the changes occurring in domain 1. The labelling of one of the active site cysteine residues with an extrinsic probe AEDANS resulted in a loss of secondary structure and a decrease in stability of the protein (see section 4.1) while Y58W Grx2 showed no significant changes in structure or stability (see section 4.1) and therefore could provided information on local changes occurring in domain 1. The kinetics of unfolding for the mutant protein still displayed two phases of unfolding as for the wild-type protein, with similar urea dependencies of the rates of unfolding (see m_u values in Table 3) while the kinetics of refolding for the mutant protein also still displayed the wild-type proteins three refolding phases. Based on the data it would appear that local probes situated in the two domains of the protein are reporting the same folding behaviour for Grx2. A highly cooperative equilibrium unfolding transition and a disruption of the entire protein by a label within domain 1 of the protein would indicate that this two-domain protein behaves as one single cooperative unit. The domains of Grx2 are held together mainly by hydrophobic contacts as well as a hydrogen bond between Glu112 and His8 as well as Asn33 (Xia *et al.,* 2001). The hydrophobic nature of the interdomain contacts suggests that the two domains are not stable on their own, this was proven to be the case as neither domains has successfully been

purified in isolation (Vlamis-Gardikas *et al.,* 1997). There is also reported cooperativity between the domains of the subunits of structurally related Alpha (Wallace *et al.,* 1998), Sigma (Stevens *et al.,* 1998), Pi (Erhardt and Dirr, 1995), *Schistosomal japonicum* (Kaplan *et al.,* 1997) and Mu (Hornby *et al.,* 2000; Hornby *et al.,* 2002; Luo *et al.,* 2002) GST. It has been shown that for proteins where the individual domains (un)fold independently there will be more than one unfolding transition for the equilibrium unfolding data with more than two species taking present during the unfolding transition (Pace 1986). The equilibrium unfolding transition for gene-3-protein, a protein that possesses two domains that unfold separately, contained two distinct transitions (Martin and Schmid, 2003a). The identification of a cooperative two-state equilibrium unfolding transition does not however exclude the existence of autonomous folding units (Peng and Wu, 2000). Llinás and Marqusee (1997) discovered for T4 lysozyme that an autonomously folding unit existed even though the equilibrium unfolding data indicated that the process was a two-state one. If the domains are not responsible for the structural separation of Grx2 could there exist a separation not based on the domains?

The work conducted to identify autonomously folding units for T4 lysozyme highlighted the importance of interactions with in such a unit (Llinás and Marqusee, 1997). Tight linkage between domains 1 and 2 and cooperative equilibrium unfolding was shown for a mutant monomeric GST (Thompson *et al.,* 2006). It was also concluded based on hydrogen deuterium exchange studies, for this monomeric GST, that a stable core of interaction exist between domains I and II of the monomer (Thompson *et al.,* 2006). In order for elements from both domains to act as a single folding unit, the inter-domain contacts formed must form early in folding and be as well as play an important role. Studies involving helix 6 in human glutathione transferase P1-1, indicated that when mutations were created that disrupted contacts with helix 1 or residues of conserved motifs in the helix, the global stability of the protein was disrupted (Dragani *et al*., 1997; Stenberg *et al*., 2000). Helix 6 is said to determine the direction of the polypeptide chain during folding as well as act as a folding nucleus (Dragani *et al*., 1997;

Stenberg *et al*., 2000). The refolding of this close family member protein sheds some light on the folding of Grx2. Helix 8 of Grx2 is structurally equivalent to helix 6 in GST P1-1 and is situated between domain 1 and 2 of Grx2, it interacts with helix 1 in domain 1 (Figure 27). It could similarly act as a folding nucleation site as well as provide strong inter-domain interactions.

If helix 8 does act as a folding initiation site with its position at the domain interface it can be assumed that any structural division of the Grx2 structure will not be by the separation of the domains, instead, two structural folding units could exist: unit x and unit y. The cooperativity between domains and the local probes (Trp80 and Trp190) reporting on distant changes in the structure during folding all point to a cooperative core folding region (unit x). The kinetics of refolding however point toward a second unit (unit y) generating the parallel medium refolding phase. Helix 8 for Grx2 is said to be part of the folding x unit, as it would provide the necessary communication between the local environments of Trp 89 and Trp190 with domain 1 and more specifically the *cis*-proline (Val48-Pro49 peptide bond). This communication is seen when the wild-type protein without Trp 58 (Y58W Grx2) still reports the slow phase of refolding (Figure 16). There are also no changes in the urea dependence of the rates of refolding of the slow phase when the domain 1 Trp 58 probe is present (Figure 17).

Val48-Pro49 peptide bond occurs on the bend prior to β-strand 3 (Xia *et al.,* 2001). The slow phase reports on the isomerization of this peptide bond therefore β-strand 3 must also be part of the x unit. β-strand 3 is linked to β-strand 1 (Figure 27) as they are the equivalent β-strands to β_2 and β_4 strands of thioredoxin. β_2 and β_4 strands have been identified as a folding nucleus in *Escherichia coli* thioredoxin. It has also been shown that the β-sheets are the core of folding for thioredoxin (Bhutani *et al*., 2003). The zipping together of these strands in thioredoxin to form a β sheet is said to be a nucleation reaction for folding of the

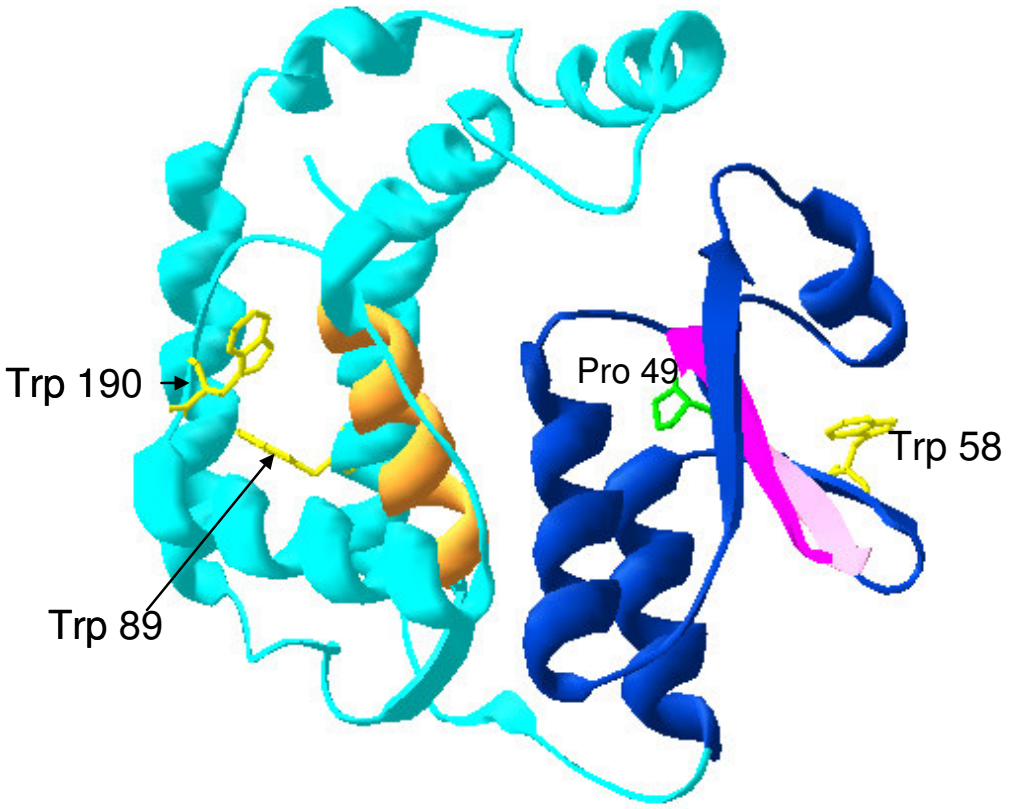


Figure 27: Structure of Y58W glutaredoxin 2
Structure of wild-type Grx2 was obtained from the protein data bank (1g70) and viewed with Swiss PDB viewer (ver 3.7) (Guex and Peitsch, 1997), showing the two domains, domain 1 (dark blue) and domain 2 (pale blue). The Pro49 residue, which is in the *cis* conformation, is represented in green and the tryptophan residues are yellow. Tyr58 in the wild-type protein was mutated using the program software to a tryptophan residue. Then β-strand 1 is purple, β-strand 3 is pink and α-helix 8 is orange.

protein (Tasayco *et al*., 2000). Therefore, β-strand 1 is also part of the x unit of folding for Grx2. It is also interesting that the protein folding core of the structurally related proteins to Grx2 occur in the region of the β-sheets in close proximity to the conserved *cis*-proline as it has been shown for ribonuclease A that often chain initiation or nucleation sites of folding exist about *cis*-proline peptide bonds (Dodge and Scheraga, 1996). The fast phase of refolding and the slow phase of refolding have been shown to be present due to different conformations about the Val48-Pro49 peptide bond, they are seen to be linked as shown by the refolding double-jump amplitude data (Figure 20 A, Figure 23 A). The slow phase represents the folding of the x unit of the Grx2 structure and because the slow phase is linked to the fast phase of refolding, the fast phase represents the folding of this x unit of the Grx2 structure too. An indication this of this comes from the amplitude changes for the single-jump refolding data (Figure 17 A). Here the mutant protein amplitudes become more pronounced for only the slow and fast phases of refolding while the medium phase remains unchanged. The increased intensity arises from the additional probe (Trp58) as can be seen from the native protein spectra for the wild-type and mutant proteins (Figure 9). So the slow and fast phases of refolding must report on the structural changes in the region of Trp58 which based on its location appears to report on the x unit of folding. This leaves the medium phase reporting on structural changes of the y unit of Grx2 during folding. If a parallel refolding pathway exists the fast and slow phases could represent the unfolding of the x and y units of Grx2 structure respectively. Then an observation that might link the fast and slow phases of unfolding to the unfolding of the x and y units respectively is that the amplitude increases for the fast phase and decreases for the slow phase of Y58W Grx2 with respect to the wild-type protein (Figure 15 A). The 5.5 M double-jump data provides further evidence because as the amplitude for the slow phase increases at 5.5 M urea unfolding (Figure 15 A) so we see changes in the medium phase of refolding (Figure 20 A). There are however no changes in the trend of the fast and slow phases of refolding with increased unfolding delay time in comparison to the 6 M data (Figure 20 A; Figure 23 A). This indicates as more molecules unfold via the slow pathway or the y unit pathway (45 % amplitude for the slow phase for

single-jump kinetic unfolding at 5.5 M urea- Figure 15 A) the medium refolding phase has to develop or the y structural elements have to rearrange to refold to the native state via the medium phase. Therefore the 5.5 M urea may stabilize the y elements in a non-native structure (unfolded) requiring a rearrangement of the y elements prior to refolding and hence the development of the medium phase during the double-jump work.

The idea that elements of secondary structure smaller than subdomains could form early folding intermediates was also shown by the renaturation of proteolytically cleaved fragments of proteins such as horse cytochrome c, which could associate forming secondary structure and tertiary contacts (Wu *et al.,* 1994). These early associated secondary structural elements were said to represent early folding intermediates. The nucleation-collision model for protein folding also indicates that secondary structural elements form and associate to form microdomains, these microdomains then associate to form mulitmicrodomains, which can represent intermediates (Karplus and Weaver, 1994). The association of the β_1 and β_3 strands in the folding of Grx2 could represent the formation of a microdomain. This microdomain may then interact via helix 8 or contacts with helix 8 to form mulitmicrodomains. Following the formation of the microdomain the structural details of which elements associate to form the x element or unit in the folding of Grx2 are not clear. Some elements of the structure have been putatively assigned as part of the x unit but the urea dependence of the refolding phases indicated that the tryptophan probes are monitoring changes for the x and y units of the Grx2 structure. The fast phase and medium phase both display changes in the dependence of the rates of refolding for Y58W Grx2 at roughly 3.3 M urea (Figure 17) which are not seen for the wild-type protein. This could point to the Trp58 probe reporting on the y unit elements while we have identified this probe as reporting on x unit folding too. The evidence presented does however clearly indicate that the folding of Grx2 does not involve the independent folding of domain 1 and domain 2. The folding of Grx2 could be separated into two regions although the structural details of exactly all of the elements of the x unit of folding and the y unit of folding cannot be defined by the scope of this study.

Further hydrogen exchange mass spectroscopy work or NMR studies would have to be conducted in order to confirm exactly which areas of the protein are nucleation sites and confirm the presence of two folding units that are not the two domains of Grx2. A similar folding of the protein spectrin via two structural elements or units that are not defined by the domains has been demonstrated (Batey *et al*., 2006). Here tryptophan reporter probes situated in each of the protein domains showed similar refolding behaviour as reported here for Grx2.

4.7 Can a final (un)folding pathway be assigned for Grx2?

The unfolding and folding of Grx2 have been discussed and it is not possible to definitively assign a final unfolding and refolding pathway for Grx2. The data presented here does not allow for the complete assignment of the parallel (un)folding pathways due to the presence of two structural folding units. The discussion however has allowed for the proposal of an idea for an (un)folding pathway. Beginning with the unfolding half of the pathway the fast and slow phases of unfolding could arise due to the presence of the x and y elements or units of structure previously identified (see section 4.6) as well as differences about the Val48-Pro49 peptide bond (see section 4.2). The slow phase of unfolding represents potentially the unfolding for the y unit where only small changes in surface area (see Table 3 for m_u values) are detected. This phase represents a coupling of the isomerization reaction and conformational unfolding (discussed in section 4.2). The unfolding of the x unit is via the fast phase and this unfolding reaction produces the largest exposure of surface area (see Table 3 for m_u values). The presence of minimum two native species is supported by the initial conditions test of Dr L A Wallace (unpublished data). It appears that the separation of these two species is based on differences in *cis* and *trans* isomers as well as a structural component. Based on the evidence presented so far the most likely unfolding pathway for Grx2 is a parallel one with the fast and slow phases occurring in parallel. Similarly parallel refolding pathways arise due the presence of a mixture of unfolded species. Two of three potential species are separated by differences in the conformation about the Val48-Pro49 peptide bond (this results

in the fast and slow pathways) (discussed in section 4.5). The fast/ slow folding species are said to be separated from the medium folding species by the structural division between the x and y folding units respectively of the Grx2 structure (discussed in section 4.6). The urea dependence is higher for the medium phase and both the regions of the fast phase in comparison to the 1 M to 2.2 M region of the slow phase. Therefore these phases represent gross structural changes rather than small rearrangements such as isomerization reactions. Again, the assignment of the pathways to a parallel mechanism involving the structural units is not a definitive one. Further experimentation would be required to eliminate any other possible mechanisms and clarify the structural details of the folding units as well as undoubtedly assign links between species formed from unfolding reactions to species required for the parallel refolding pathways. The amplitude for the slow phase of unfolding is 20 % (7.5 M urea) would have to link to the medium phase of refolding (20 % at 1 M urea). The corresponding 80 % for the fast phase refolds via the fast / slow (50 % / 30 %) refolding phases, but how can the slow unfolding isomerization phase not be linked to the slow refolding isomerization phase? Also the different (un)folding species resulting in parallel pathways are all said to arise due to a structural separation and differences in conformation about the Val48-Pro49 peptide bond but why then is there only two unfolding pathways where three are required for refolding?

Intermediates have been detected for both the unfolding (see section 4.2) and the refolding (see section 4.3) of Grx2. It has been debated if these intermediates represent on- or off- pathway structures during folding although a review by Roder and Colon (1997), concluded that they predominantly on-pathway structures which reduce the conformational space during folding. The presence of intermediates during the (un)folding of Grx2 and the fact that the slow phase of unfolding and refolding involves an isomerization coupled to structural folding (see section 4.2 and section 4.4) precludes the characterization of the transitions states for Grx2 (Schmid, 1992).

5 REFERENCES

Allocati, N., Casaone, E., Masulli, M., Ceccarelli, I., Carletti, E., Parker, M. W. and Di Ilio, C. (1999) Functional analysis of the evolutionarily conserved proline 53 residue in Proteus Mirabilis glutathione transferase B1-1. *FEBS Lett* **445**, 347-350.

Alonso, D. O. V. and Dill, K. A. (1991) Solvent denaturation and stabilization of globular proteins. *Biochemistry* **30,** 5974-5985.

Anfinsen, C. B. (1973) Principles that govern the folding of protein chains. *Science* **181**, 223-230.

Aslund, F., Ehn, B., Miranda-Vizuette, A., Pueyo, C., and Holmgren, A. (1994) Two additional glutaredoxins exist in *Escherichia coli*: Glutaredoxin 3 is a hydrogen donor for ribonucleotide reductase in a thioredoxin/glutaredoxin 1 double mutant. *Proc. Natl. Acad. Sci USA*. **91**, 9813-9817.

Batey, S., Scott, K. A. and Clarke, J. (2006) Complex folding kinetics of a multidomain protein. *Biophys. J.* **90**, 2120-2130.

Baldwin, R. L (1996) On-pathway versus off-pathway intermediates. *Folding and design*. **1**, R1-R8.

Bhutani, N. and Udgaonkar, J. B. (2001) GroEL channels the folding of thioredoxin along one kinetic route. *J. Mol. Biol.* **314**, 1167-1179.

Bilsel, O., Zitzewitz, J. A., Bowers, K. E. and Matthews, C.R. (1999) Folding mechanism of the α-subunit of tryptophan synthase, an α/β barrel protein: Global analysis highlights the interconversion of multiple native intermediate, and unfolded forms through parallel channels. *Biochemistry* **38**, 1018-1-29.

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, characterization, and crystal structure of the Omega class glutathione transferases. *J. Biol. Chem.* **275,** 24798-24806.

Borden, K. L. B., Richards, F. M. (1990) Folding kinetics of phage T4 thioredoxin. *Biochemistry* **29**, 3071-3077.

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

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* isomerizations of proline residues. *Biochemistry* **1**4, 4953-4963.

Burns, L. L., Dalessio, P. M. and Ropson, I. J. (2003) Replacement of proline with valine does not remove an apparent proline isomerization-dependent folding event in CRABP I. *Protein Sci.* **13**, 1670-1676.

Chaffotte, A. F., Li, J-H., Georgescu, R. E., Goldberg, M. E. and Tasayco, M. L. (1997) Recognition between disordered states: kinetics of the self-assembly of thioredoxin fragments. *Biochemistry* **36**, 16040-8.

Chakrabarti, S., Lanczycki, C. J., Panchenko, A. R., Przytycka, T. M., Thiessen, P. A. and Bryant, S. H. (2006) Refining multiple sequence alignments with conserved core regions. *Nucleic Acids Res.* **34**, 2598-2606.

Charbonnier, J. B., Belin, P., Moutiez, M., Strura, E. A. and Quéméneur, E. (1999) On the role of the *cis*-proline residue in the active site of DsbA. *Protein Sci.* **8**, 96-105.

Chothia, C. (1975) Structural invariants in protein folding. *Nature* **254,** 304-308.

Cook, K. H., Schmid, F. X. and Baldwin, R. L. (1979) Role of proline isomerization in folding of ribonuclease A at low temperatures. *Proc. Natl. Acad. Sci. USA* **76,** 6157-6161.

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

Cromer, B. A., Morton, C. J, Board, P. G. and Parker, M. W. (2002) From glutathione transferase to pore in a CLIC. *Eur. Biophys. J.* **31**, 356-364.

De Lamotte-Guéry, F., Pruvost, C., Minard, P., Delsuc, M.-A., Miginiac-Maslow, M., Schmitter, J.-M., Stein, M. and Decottignies, P. (1997) Structural and

functional roles of a conserved proline residue in the α2 helix of *Escherichia coli* thioredoxin. *Protein Eng.* **10**, 1425-1432.

Dill, K. A. (1985) The theory of folding and stability of globular proteins. *Biochemistry* **24**, 1501-1509.

Dirr, H., Reinemer, P. and Huber, R. (1994) X-ray crystal structures of cytosolic glutathione S-transferases. *Eur. J. Biochem.* **220**, 645-661.

Dirr, H. W. & Wallace, L. A. (1999). Role of the C-terminal helix 9 in the stability and ligandin function of class Alpha glutathione transferase A1-1. *Biochemistry* **38**, 15631-15640.

Dodge, R. W. and Scheraga, H. A. (1996) Folding and unfolding kinetics of the proline to alanine mutants of bovine pancreatic ribonuclease A. *Biochemistry* **35**, 1548-1559.

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**, 7403-7411.

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 of a buried and conserved hydrogen bond important for protein stability. *J. Biol. Chem.* **272**, 25518-23.

Eklund, H., Cambillau, C., Sjoberg, B. M., Holmgren, A., JÖrnvall, H., HÖÖg, J.O. and Branden, C. I. Conformational and functional similarities between glutaredoxin and thioredoxins. *EMBO J.* **3**,1443-1449.

Erhardt, J. and Dirr, H. (1995). Native dimer stabilizes the subunit tertiary structure of porcine class Pi glutathione S-transferase. *Eur. J. Biochem.* **230**, 614-620.

Eyles, S. J. and Gierasch, L. M. (2000) Multiple roles of prolyl residues in structure and folding. *J. Mol. Biol.* **301**, 737-747.

Fersht. A. R. (1997) Nucleation mechanisms in protein folding. *Curr. Opin. Struc. Biol.* **7,** 3-9.

Fischer, G., Bang, H. and Mech, C. (1984) Determination of enzymatic catalysis for the *cis-trans*-isomerization of peptide binding in proline containing peptides. *Biomed. Biochim. Acta*, **10**, 1101-1111.

Fransson, C., Freskgard, P., Herbertson, H., Johansson, A., Jonasson, P., Martensson, L., Svensson, M., Jonsson, B. and Carlsson, U. (1992) *Cis-trans* isomerization is rate-determining in the reactivation of denatured human carbonic anhydrase II as evidenced by proline isomerase. *FEBS lett.* **296**, 90-94.

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**, 213-227.

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.

Georgescu, R. E., Li, J., Goldberg, M. E., Tasayco, M. L. and Chaffotte, A. F. (1998) Proline isomerization-independent accumulation of an early intermediate and heterogeneity of the folding pathways of a mixed α/β protein, *Escherichia coli* thioredoxin. *Biochemistry* **37**, 10286-10297.

Goldenberg, D. P. (1992) Mutational analysis of protein folding and stability. In Protein folding (Creighton, T. E. ed.) pp. 353-403, W. H. Freeman, New York.

Gomez-Hens, A. and Perez-Bendito, D. (1991). The stopped-flow technique in analytical chemistry. *Analytica Chimica Acta.* **242**, 147-177.

Göthel, S. F. and Marahiel, M. A. (1999) Peptidyl-prolyl *cis-trans* isomerases, a superfamily of ubiquitous folding catalysts. *Cell. Mol. Life Sci.* **55**, 423-436.

Guex, N. and Peitsch, M. C. (1997) SWISS-MODEL and the swiss-pdb viewer: an environment for cooperative protein modeling. *Electrophoresis* **18**, 2714-2723.

Habeeb, A. F. S. A. (1972) Reaction of protein sulfhydryl groups with Ellman's reagent. *Methods Enzymol.* **25**, 457-464.

Hagerman, P. J. and Baldwin, R. L. (1976) A quantitative treatment of the kinetics of the folding transition of ribonuclease A. *Biochemistry* **15**, 1462-1473.

Harrison, S. C. and Durbin, R. (1985) Is there a single pathway for the folding of a polypeptide chain? *Proc. Natl. Acad. Sci. USA* **82**, 4028-4030.

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. G. (2001) Cystal structure of a soluble form of the intracellular chloride ion channel CLIC1 (NCC27) at 1.4Å resolution. *J. Biol. Chem.* **276,** 44993-45000.

Hartl, F. U. and Hartl, M. H. (2002) Molecular chaperones in the cytosol: from nascent chain to folded protein. *Science* **295**, 1852-1858.

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 *Escherichia coli* ribonucleoside-diphosphate reductase dependent upon glutathione. *Proc. Natl. Acad. Sci.* USA. **73**, 2275-2279.

Holmgren, A. and Aslund, F. (1995) Glutaredoxin. *Methods Enzymol.* **252**, 283-292.

Holmgren, A., Söderberg, B-O., Eklund, H. and Brändén, C-I. (1975) Three-dimensional structure of *Escherichia coli* thioredoxin-S_2 to 2.8 Å resolution. *Proc. Natl. Acad. Sci.* USA. **72**, 2305-2309.

Hornby, J. A. T., 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**, 14238-14247.

Hornby, J. A. T., 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**, 12336-12334.

Houry, W. A. and Scheraga, H. A. (1996) Nature of the unfolded state of ribonuclease A: effect of *cis-trans* X-pro peptide bond isomerization. *Biochemistry* **35**, 11719-11733.

Hudson, E. N and Weber, G. (1973) Synthesis and characterization of two fluorescent sulfhydryl reagents. *Biochemistry* **21**, 4154-4160.

Ikeguchi, M., Kuwajima, K., Mitani, M. and Sugai, S. (1986) Evidence for Identity between the equilibrium unfolding intermediate and a transient folding intermediate: a comparative study of the folding reactions of α-lactalbumin and lysozyme. *Biochemistry* **25,** 6965-6972.

Ikura, T., Tsurupa, G. P. and Kuwajima, K. (1997) Kinetic folding *cis*/*trans* prolyl isomerization of staphylococcal nuclease. A study by stopped-flow absorption, stopped-flow circular dichroism and molecular dynamics simulations. *Biochemistry* **36**, 6529-6528.

Jackson, S. E. and Fersht, A. R. (1991) Folding of Chymotrypsin inhibitor 2. 1. Evidence for a two-state transition. *Biochemistry* **30,** 10428-10435.

Jaenicke, R. (1991) Protein folding : local structures, domains, subunits, and assemblies. *Biochemistry* **30**, 3147-3161.

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

Jäger, M. and Plückthun, A. (2000) Direct evidence by H/D exchange and ESM-MS for transient unproductive domain interaction in the refolding of an antibody scFv fragment. *Protein Sci.* **9**, 552-563.

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 1B gamma from *Saccharomyces cerevisiae*. *J. Biol. Chem.* 278, 47190**-**47198.

Ji, X., Johnsom, 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,** 1043-1052.

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.

Juminaga, D., Wedemeyer, W. J. and Scheraga, H. A. (1998) Proline isomerization in bovine pancreatic ribonuclease A. 1. Unfolding conditions. *Biochemistry* **37**, 11614-11620.

Kamagato, K., Sawano, Y., Tanokura, M. and Kuwajima, K. (2003) Multiple parallel-pathway folding of proline-free staphylococcal nuclease. *J. Mol. Biol.* **332**, 1143-1153.

Kamen, D. E. and Woody, R. W. (2002) Folding kinetics of the protein pectate lyase C reveal fast-forming intermediates and slow proline isomerizations. *Biochemistry* **41**, 4713-4723.

Kann, M. G., Thiessen, P. A., Panchenko, A. R., Schäfer, A. A., Altschul, S.F. and Bryant, S. H. (2005) A structure-based method for protein sequence alignment. *Bioinformatics* **21**, 1451-1456.

Kaplan, W., Husler, P., Klump, H., Erhardt, J., Sluis-Cremer, N. & 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.

Karplus, M. and Weaver, D. L. (1994) Protein folding dynamics: the diffusion-collision model and experimental data. . *Protein Sci.* **3**, 650-668

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

Kelley, R. F., Wilson, J., Bryant, C. and Stellwagen, E. (1986) Effects of guanidine hydrochloride on the refolding kinetics of denatured thioredoxin. *Biochemistry* **25**, 728-732.

Kelley, R. F. and Richards, F. M. (1987) Replacement of proline-76 with alanine eliminates the slowest kinetic phase in thioredoxin folding. *Biochemistry* **26**, 6765-6774.

Kiefhaber, T., Quaas, R., Hahn, U. and Schmid, F. X. (1990) Folding of ribonuclease T1. 1. Existence of multiple unfolded states created by proline isomerization. *Biochemistry* **29**, 3053-3061.

Kiefhaber, T. and Schmid, F. X. (1992) Kinetic coupling between protein folding and prolyl isomerization. *J. Mol. Biol.* **224**, 231-240.

Kiefhaber, T., Kohler, H-H. and Schmid, F. X. (1992a) Kinetic coupling between protein folding and prolyl isomerization. *J. Mol. Biol*. **224**, 217-229.

Kiefhaber, T., Schmid, F. X., Willaert, K., Engelborghs, Y. and Chaffotte, A. (1992b) Structure of a rapidly formed intermediate in ribonuclease T1 folding. *Protein Sci.* **1**, 1162-1172.

Kuwajima, K. (1989) The molten globule state as a clue for understanding the folding and cooperativity of globular-protein structure. *Proteins.* **6**, 87-103.

Kuwajima, K., Okayama, N., Yamamoto, K., Ishihara, T. and Sugai, S. (1991) The Pro 117 to glycine mutation of staphylococcal nuclease simplifies the unfolding-folding kinetics. *FEBS lett.* **290**, 135-138.

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**, 352-361.

Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685.

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

Lang, K., Schmid, F. X. and Fischer, G. (1987) Catalysis of protein folding b prolyl isomerase. *Nature* **329**, 268-270.

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

Liu, C.-P. and Zhou, J.-M. (2004) Trigger factor-assisted folding of bovine carbonic anhydrase II. *Biochem. Biophys. Res. Communications* **313**, 509-515.

Llinás, M. and Marqusee, S. (1997) Subdomain interactions as a determinant in the folding and stability of T4 lysozyme. *Protein Sci.* **7,** 96-104.

Lundstrom-Ljung, J., Vlamis-Gardikas, A., Aslund, F. and Holmgren, A. (1999) Reactivity of glutaredoxins 1,2 and 3 from *Escherichia coli* and protein disulfide isomerase towards glutathionyl-mixed disulfides in ribonuclease A. *FEBS Lett.* **443**, 85-88.

Luo, J-K., Hornby, J. A. T., 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 μ glutathione transferase. *Protein Sci.* **11,** 2208-2217.

Marchler-Bauer, A., Anderson, J., Fedorova, N., DeWeese-Scott, C., Geer, L. Y., Hurwitz, D., Jackson, J. J., Jacobs, A., Lanczycki, C., Liebert, C., 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) MMDB: Entrez's 3D-structure database. *Nucleic Acids Res.* **33**, D192–D196.

Marquardt, D. W. (1963) An algorithm for the least squares estimation of nonlinear parameters. *J. Soc. Industr. Appl. Math*. **11**, 431-441.

Martin, J.L. (1995) Thioredoxin- a fold for all reasons. *Structure* **3,** 245-250.

Martin. A. and Schmid, F. X. (2003a) Evolutionary stabilization of the gene-3-protein of phage fd reveals the principles that govern the thermodynamic stability of two-domain proteins. *J. Mol. Biol.* **328**, 863-875.

Martin. A. and Schmid, F. X. (2003b) The folding mechanism of a two-domain protein: folding kinetics and domain docking of the gene-3-protein of phage fd. *J. Mol. Biol.* **329**, 599-610.

Martin. A. and Schmid, F. X. (2003c) A proline switch controls folding and domain interactions in the gene-3-protein of the filamentous phage fd. *J. Mol. Biol.* **331**, 1131-1140.

Mello, C. C., Bradley, C. M., Tripp, K. W. and Barrick, D. (2005) Experimental characterization of the folding kinetics of the notch ankyrin domain. *J. Mol. Biol.* **352**, 266-281.

Missiakas, D., Betton, J-M., Chaffotte, A., Minard, P. and Yon, J. M. (1992) Kinetic studies of the refolding of yeast phosphoglycerate kinase: Comparison with the isolated engineered domains. *Protein Sci.* **1**, 1485-1493.

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**, 536-540.

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**, 2138-2148.

Nathaniel, C., Wallace, L. A., Burke, J. and Dirr, H. W. (2003) The role of an evolutionary conserved *cis*-proline in the thioredoxin-like domain of human class Alpha glutathione transferase A1-1. *Biochem. J.* **372**, 241-246.

Neet, K. E. and Timm, D. E. (1994) Conformational stability of dimeric proteins: Quantitative studies by equilibrium denaturation. *Protein Sci.* **3**, 2167-2174.

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,** 2176-2185.

Osváth, S., Köhler, G. Závodszky, P. and Fidy, J. (2005) Asymmetric effect of domain interactions on the kinetics of folding in yeast phosphoglycerate kinase. *Protein Sci.* **14**, 1609-1616.

Pace, C. N. (1986) Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods Enzymol.* **131**, 266-280.

Papworth. C., Bauer, J.C., Braman, J. and Wright, D.A. (1996) Site-directed mutagenesis using double-stranded plasmid DNA templates. *Strategies* **9**, 3-4.

Pemble, S. E., Wardle, A. F. and Taylor, J. B. (1996) Glutathione S-transferase class Kappa: characterization by the cloning of rat mitochondrial GST and identification of a human homologue. *Biochem. J.* **319**, 749-754.

Peng, Z-Y. and Wu, L. C. (2000) Autonomous protein folding units. In Advances in protein chemistry (Matthews, C. R. ed.) pp. 1-47, Academic Press, USA.

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,** 1336-1342.

Perkins, S. J. (1986) Protein volumes and hydration effects. *Eur. J. Biochem*. **157**, 169-180.

Perrett, S., Freeman, S. J., Jonathan, P., Butler, G. and Fersht, A. R. (1999) Equilibrium folding properties of the yeast prion protein determination Ure2. *J. Mol. Biol.* **290**, 331-345.

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,** 1567-1576.

Privalov, P.L. and Potekhin, S. A. (1986) Scanning microcalorimetry in studying temperature-induced changes in proteins. *Methods Enzymol.* **134**, 4-51.

Reimer, U., Scherer, G., Drewello, M., Kruber, S., Schutkowiski, M. and Fischer, G. (1998) Side-chain effects on peptidyl-prolyl *cis/trans* isomerization. *J. Mol. Biol.* **279,** 449-460.

Reinemer, P., Dirr, H. W., Ladenstein, R., Huber, R., Lobelo, 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 angstroms. *J. Mol. Biol.* **227,** 214-226.

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 A resolution: structural characterization of herbicide-conjugating plant glutathione S-transferases and a novel active site architecture. *J. Mol. Biol.* **255,** 289-309.

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 *Escherichia coli*. *Proteins* **53,** 777-782.

Roder, H. and Colon, W. (1997) Kinetic role of early intermediates in protein folding. *Curr. Opin. Struc. Biol.* **7**, 15-28.

Rossjohn, J., McKinstry, W. J., Oakley, A. J., Verger, D., Flanagan, J., Chelyanayagam, 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,** 309-322.

Rowe, E. S. and Tanford, C. (1973) Equilibrium and kinetics of the denaturation of a homogeneous human immunoglobulin light chain. *Biochemistry* **12**, 4822-4827.

Rudolph, R., Siebendritt, R., Nesslaŭer, G., Sharma, A. K. Jaenicke, R. (1990) Folding of an all-β protein: Independent domain folding in γII-crystallin from calf eye lens. *Proc. Natl. Acad. Sci. USA* **87,** 4625-4629.

Sacchetta, P., Aceto, A., Bucciarelli, T., Dragani, B., Santarone, S., Allocati, N. and Di Ilio, C. (1993) Multiphasic denaturation of glutathione transferase

B1-1 by guanidinium chloride. Role of the dimeric structure on the flexibility of the active site. *Eur. J. Biochem.* **215**, 741-745.

Satumba, W.J. and Mossing, M.C. (2002) Folding and assembly of lambda Cro repressor dimers are kinetically limited by proline isomerization. *Biochemistry* **41,** 14216-14224.

Schiene, C. and Fischer, G. (2000) Enzymes that catalyze the restructuring of proteins. *Curr. Opin. Struct. Biol.* **10**, 40-45.

Schmid, F. X. (1986a) Fast-folding and slow-folding forms of unfolded proteins. *Methods Enzymol.* **131**, 70-82.

Schmid, F. X. (1986b) Proline isomerization during refolding of ribonuclease A is accelerated by the presence of folding intermediates. *FEBS Lett.* **198**, 217-220

Schmid, F. X (1992) Kinetics of unfolding and refolding of single-domain proteins. In Protein folding (Creighton, T. E., ed.), pp. 197-241, Freeman, New York.

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.

Schreiber, G. and Fersht, A. R. (1993) The refolding of *cis*- and *trans*-peptidylproplyl isomers of barstar. *Biochemistry* **32**, 11195-11203.

Schultz, D. A., Schmid, F. X. and Baldwin, R. L. (1992) *Cis* proline mutants of ribonuclease A. II. Elimination of the slow-folding forms by mutation. *Protein Sci.* **1**, 917-924.

Seckler, R. and Jaenicke, R. (1992) Protein folding and protein refolding. *FASEB J.* **6,** 2545-2552.

Sheehan, D., Meade, G., Foley, V. M. and Dowd, C. A. (2001) Structure, function and evolution of glutathione transferases: implications for classification of non-mammalian members of an ancient enzyme superfamily. *Biochem. J.* **360**, 1-16.

Shi, J., Vlamis-Gardikas, A., Aslund, F., Holmgren, A. and Rosen, B. P. (1999) Reactivity of glutaredoxins 1, 2, and 3 from *Escherichia coli* shows that glutaredoxin 2 is the primary hydrogen donor to ArsC-catalyzed arsenate reduction. *J. Biol. Chem.* **274**, 36039-36042.

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

Sinnig, I., Klewegt, G. J., Cowan, S.W., Reinemer, P., Dirr, H. W., Huber, R., Gilliland, G. L., Armstrong, R.N., Ji, X., Board, P. G., Olin, B., Mannervik, B. and Jones, T. A. (1993) Structure determining 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.

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**, 10421-10428.

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**, 15534-15541.

Stewart, D. E., Sarkar, A. and Wampler, J.E. (1990) Occurrence and role of *cis* peptide bonds in protein structures. *J. Mol. Biol.* **214**, 253-260.

Tan, Y. J., Oliveberg, M., Otzen, D. E. and Fersht, A. R. (1997) Rate of isomerisation of peptidyl-proline bonds as a probe for interactions in the physiological denatured state of chymotrypsin inhibitor 2. *J. Mol. Biol.* **269**, 611-622.

Tanford, C. (1970) Protein denaturation. *Adv. Protein Chem.* **24**, 1-95.

Tasayco, M. L., Fuchs, J., Yang, X. M., Dyalram, D. and Georgescu, R. E. (2000) Interaction between two discontiguous chain segments from the beta-sheet of *Escherichia coli* thioredoxin suggests an initiation site for folding. *Biochemistry* **39**, 10613-8.

Teale, F. W. J. (1960) The ultraviolet fluorescence of proteins in neutral solution. *Biochem. J.* **76,** 381-388.

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,** 7008-7020.

Thompson, L. C., Walters, J., Burke, J., Parsons, J. F., Armstrong, R. N. and Dirr, H. W. (2006) Double-mutation at the subunit interface of glutathione transferase rGSTM1-1 results in a stable, folded monomer. *Biochemistry* **45,** 2267-2273.

Thual , C., Bousset, L., Komar, A. A., Walter, S., Buchner, J., Cullin, C. & Melki, R. (2001). Stability, folding, dimerization, and assembly properties of the yeast prion Ure2p. *Biochemistry* **40**, 1764-1773.

Tonomura, B., Nakatani, H., Ohnishi, M., Yamaguchi, J. and Hiromi, K. (1978) Test reactions for a stopped-flow apparatus. *Anal. Biochem*. **84**, 370-383.

Tsunenaga, M., Goto, Y., Kawata, Y. and Hamaguchi, K. (1987) Unfolding and refolding of a type κ immunoglobulin light chain and its variable and constant fragments. *Biochemistry* **26**, 6044-6051.

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. USA* **98,** 1459-1464.

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

Vlamis-Gardikas, A., Aslund, F., Spyrou, G., Bergman, T. and Holmgren, A. (1997) Cloning, overexpression, and characterization of glutaredoxin 2, an atypical glutaredoxin from *Escherichia coli*. *J. Biol. Chem.* **272**, 11236-11243.

Vlamis-Gardikas, A., Potamitou, A., Zarivach, R., Hochman, A. and Holmgren, A. (2002) Characterization of *Escherichia coli* null mutants for glutaredoxin 2. *J. Biol. Chem.* **277**, 10861-10868.

Wallace, L. A. and Dirr, H. W. (1999) Folding and assembly of dimeric human glutathione transferase A1-1. *Biochemistry* **38**, 16686-16694.

Wallace, L. A. and Matthews, C. R. (2002) Sequential vs. parallel protein-folding mechanisms: experimental tests for complex folding reactions. *Biophys. Chem.* **101-102**, 113-131.

Wallace, L. A., Sluis-Cremer, N. and Dirr, H. W. (1998) Equilibrium and kinetic unfolding properties of dimeric human glutathione transferase A1-1. *Biochemistry* **37**, 5320-5328.

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

Walsch, C. T., Zydowsky, D. L. and McKeon, F. D. (1992) Cyclosporin A, the cyclophilin class of peptidylprolyl isomerases, and blockade of T cell signal transduction. *J. Biol. Chem.* **267**, 13115-13118.

Wang, Y., Geer, L. Y., Chappey, C., Kans, J. A. and Bryant, S. H. (2000) Cn3D: sequence and structure views for Entrez. *Trends Biochem. Sci.* **25**, 300-302.

Wedemeyer, W. J., Welker, E. and Scheraga, H. A. (2002) Proline *cis-trans* isomerization and protein folding. *Biochemistry* **41**, 14637-14644.

Wenk, M., Herbst, R., Hoeger, D., Kretschmar, M., Lubsen, N. H. and Jaenicke, R. (2000) Gamma S-crystallin of bovine and human eye lens: solution structure, stability and folding of the intact two-domain protein and its separate domains. *Biophys. Chem.* **86**, 95-108.

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

Wolfenden, R. V., Cullis, P. M. and Southgate, C. C. F. (1979) Water, protein folding and the genetic code. *Science* **206**, 575-577.

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

Wu, L. C., Grandori, R. and Carey, J. (1994) Autonomous subdomains in protein folding. *Protein Sci.* **3**, 369-371.

Wu, Y. and Matthews, C. R. (2002) A *cis*-prolyl peptide bond isomerization dominates the folding of the alpha subunit of Trp synthase, a TIM barrel protein. *J. Mol. Biol.* **322**, 7-13.

Xia, B., Vlamis-Gardikas, A., Holmgren, A., Wright, P. E. and Dyson, H. J. (1999) Letter to the Editor: Assignment of 1H, 13C, and 15N resonances of reduced *Escherichia coli* glutaredoxin 2. *J. Biomol. NMR.* **14**, 197-198.

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**, 907-918.

Yon, J.M. (2001) Protein folding: a perspective for biology, medicine and biotechnology. *Bras. J. Med. Biol. Res.* **34,** 419-435.

Zaidi, F. N., Nath U. and Udgaonkar, J. B. (1997) Multiple intermediates and transition states during protein folding. *Nature Struct. Biol* **4**, 1016-1024.

Zarnt, T., Tradler, T., Stoller, G., Scholz, C., Schmid, F. X. and Fischer, G. (1997) Modular structure of the trigger factor required for high activity in protein folding. *J. Mol. Biol.* **271**, 827-837.

Zarrine-Afsar, A. and Davidson, A. R. (2004) The analysis of protein folding kinetic data produced in protein engineering experiments. *Methods* **34**, 41-50.

Zitzewitz, J. A. and Matthews, C. R. (1999) molecular dissection of the folding mechanism of the α-subunit of tryptophan synthase: an amino terminal autonomous folding unit controls several rate-limiting steps in the folding of a single domain protein. *Biochemistry* **38**, 10205-10214.