THE UNFOLDING AND REFOLDING OF HUMAN GLUTATHIONE TRANSFERASE A1-1

Louise Annette Wallace

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

Johannesburg, February 1998.

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 for any degree or examination in any other University.

L Wallace

Louise A. Wallace

16th day of February 1998

To those who have supported me:

Jim, Jackie, Julie, John, Jenni, Clyde, Nan, Vicki, Don, Hans.

"...Two roads diverged in a wood, and I –
I took the one less travelled by,
And that has made all the difference."

From *The Road Not Taken*
By Robert Frost (1874-1963)

The results presented in this thesis have been submitted for publication:

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**, 15, 5320-5328

Wallace,L.A., Blatch,G.L. and Dirr,H.W. (1998) The hydrophobic core in domain II of glutathione transferases is stabilised by a topologically conserved aliphatic residue in alpha helix 6. *Manuscript submitted.*

Wallace,L.A. and Dirr, H.W. (1998) The folding of homodimeric Human Glutathione Transferase A1-1. *Manuscript in preparation*

ABSTRACT

The thermodynamic stability and the properties of the unfolding/refolding pathways of homodimeric human glutathione transferase A1-1 (hGST A1-1) were investigated. The conformational stability, assessed by urea- and temperature-induced denaturation studies, was consistent with a folded dimer/unfolded monomer transition with no stable intermediates. The high energy of stabilisation and the highly co-operative transition implies that the subunit-subunit interactions are necessary to maintain the three-dimensional state of the individual subunits. The stopped-flow-unfolding pathway, monitored using Trp fluorescence, was biphasic with a fast and slow unfolding event. Urea-dependence and thermodynamic activation parameters suggest that the transition state for each phase is well structured and is closely related to the native protein in terms of solvent exposure. The unfolding pathways monitored by energy transfer or direct excitation of AEDANS covalently linked to Cys111 in hGST A1-1 were monophasic with urea and temperature properties similar to those observed for the slow unfolding phase (described above). A two-step sequential unfolding mechanism involving the partial dissociation of the two structurally distinct domains per subunit followed by complete domain and subunit unfolding is proposed.

The crystal structures of all cytosolic glutathione transferases show that the alpha helices 5, 6 and 7 pack tightly against each other to form the hydrophobic core of domain II. Leu164 in class alpha glutathione transferase is a topologically conserved residue in the alpha helix 6. The replacement of Leu164 with alanine did not impact on the functional or gross structural properties of hGST A1-1. The urea-induced equilibrium and kinetic unfolding pathways were similar to those observed for the wild-type protein. The free energy change of unfolding was equivalent to the energetic cost of deleting three methylene groups. Furthermore, the decreased co-operativity of the unfolding transition is consistent with a decrease in co-operativity of the forces that maintain the native state of hGST A1-1. The biphasic kinetic unfolding pathway indicated that the fast phase was destabilised to a greater extent than the slow unfolding phase. Urea-dependence and thermodynamic activation parameters for the

slow unfolding phase indicated an increase in solvent exposure (~ 5%) for the mutant transition state relative to the wild-type protein. The fractional value of ϕ_U for the transition state of the slow unfolding phase suggests a partial loss of van der Waals interactions. This indicates that the hydrophobic core in domain II was weakened by the Leu-to-Ala substitution and that disruption of this region is partly rate-limiting during the unfolding pathway for the glutathione transferases.

The refolding pathway, monitored using Trp fluorescence, displayed a number of complexities depending on the final urea concentration. At urea concentrations between 2.5M and 4.25M the reaction was biphasic, whilst at urea concentrations less than 2.5M the reaction was triphasic. The phases are referred to as fast, intermediate and slow refolding events. The rate-limiting step for the major fast and intermediate events was bimolecular and consistent with an association reaction involving two disordered or partially folded monomers. Urea-dependence studies and thermodynamic activation parameters suggested that the transition state which involves a partially folded dimer or dimeric intermediate was largely solvent accessible. The reaction profile, therefore, has a narrow activation barrier with only a limited part of the reaction occurring at the transition state level. Long- and middle-range electrostatic forces are postulated to play a role in stabilising this initial dimeric state. A sequential and parallel pathway with evidence for on-pathway intermediates is proposed.

ACKNOWLEDGEMENTS

Prof. Heini Dirr for his continuous enthusiasm and encouragement for this research and for the opportunity and privilege of working in his lab.

Nicolas Sluis-Cremer for his helpful discussions and encouragement

Dr Greg Blatch for his help with the construction of the L164A hGST A1-1

Prof. H. Klump and Dr P. Hüsler, University of Cape Town, for the use of their Differential Scanning Calorimeter and for the analysis of the results.

The members of the Protein Structure-Function Research Programme and the Department of Biochemistry, University of the Witwatersrand.

The Foundation for Research and Development, South Africa and the University of the Witwatersrand for financial assistance.

TABLE OF CONTENTS

ABBREVIATIONS

AEDANS	5-[[2-[(acetyl)amino)ethyl]amino]-naphthalene-1-sulphonic acid
ANS	8-anilino-1-naphthalene sulphonate
bp	base pairs
C_m	urea-induced midpoint of unfolding
CDNB	1-chloro-2,4-dinitrobenzene
ΔC_{pN-U}	change in heat capacity for complete unfolding
$\Delta C_p^{\ddagger}$	change in heat capacity between native (N)/unfolded (U) states and transition state
DSC	differential scanning calorimetry
E_A	activation energy
EA	ethacrynic acid
$\Delta\Delta G^{\ddagger}_U$	change in activation energy upon unfolding as a consequence of a mutation
$\Delta\Delta G^{\ddagger}_{N-U}$	difference in free energy of unfolding for wild-type and mutant
$\Delta G(H_20)$	Gibbs free energy of unfolding in the absence of denaturant
$\Delta G^{\ddagger}(T_o)$	Gibbs free energy of activation at standard temperature
G-Site	glutathione binding site
GSH	glutathione
GST	glutathione transferase
ΔH_{cal}	calorimetric enthalpy
$\Delta H_{v,H}$	van't Hoff enthalpy

$\Delta H^{\ddagger}(T_o)$	activation enthalpy at standard temperature
H-Site	hydrophobic-electrophilic substrate binding site
hGST A1-1	human class alpha glutathione transferase with two type one subunits
I_2	dimeric intermediate
IAEDANS	5-[[2-[(iodoacetyl)amino)ethyl]amino]-naphthalene-1-sulphonic acid
L-site	non-substrate ligand binding site
$k(H_20)$	apparent unfolding (u)/refolding (f) rate in the absence of denaturant
k_{app}	apparent unfolding(u)/refolding(f) rate constant
k_{cat}/K_m	catalytic efficiency
k_f	refolding second-order rate constant
K_m	Michaelis-Menten constant
L164A	replacement of wild-type leucine (L) with alanine (A) at position 164
m	susceptibility of protein to denaturant i.e., co-operativity of unfolding transition
N_2	native folded dimer
N_2^*	native-like dimeric intermediate
NATA	N-acetyl-L-tryptophanamide
ϕ_U	structure of the transition state probed by mutation
PCR	polymerase chain reaction
$\Delta S^{\ddagger}(T_o)$	activation entropy at standard temperature
SASA	solvent accessible surface area
SEC-HPLC	size exclusion high performance liquid chromatography

τ	time constant
T_o	standard temperature (298K)
T_m	temperature-induced midpoint of unfolding
TGGE	thermal gradient gel electrophoresis
U	unfolded monomer
V_m	maximum velocity

The IUPAC-IUBMB three and one letter codes for amino acids are used.

Enzyme: Glutathione transferase (E.C. 2.5.1.18)

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1
PROTEIN FOLDING

1.1 Biosynthesis of proteins *in vivo*

The biosynthesis of proteins *in vivo* requires the specification of the amino acid (one-dimensional) sequence of a molecule that can spontaneously assume its native tertiary (three-dimensional) state as it is synthesised in the cell. The unique three-dimensional state is essential for the biological function of the protein and determines the role of the protein within the cell.

The general mechanism of protein biosynthesis has been well established. The process is highly complex but follows the central dogma of DNA→RNA→protein in the presence of various factors. The real importance of the gene is obscured by the lack of detail available on how information contained within the primary amino acid sequence is translated into the corresponding tertiary structure. The transfer of information from the gene level to the final functional product is known to be mediated by two codes, the "genetic code" and the "folding code" (Jaenicke, 1984). The "genetic code" which translates the mRNA into a linear polypeptide has been deciphered and is a degenerate triplet code with 64 combinations of 4 bases (A,C,G and T). Evidence for a protein "folding code" was provided by the spontaneous (co-translational or post-translational) formation of the native functional three-dimensional state of protein *in vivo* (Anfinsen, 1973). The "folding code" simplistically refers to the transition from the linear (one-dimensional) array of amino acids to the three-dimensional fold for a protein. The one-dimensional to the three-dimensional transition, however, encompasses a large number of hierarchical levels that result in the structural organisation of the polypeptide chain.

Linderstøm-Lang and Bernal first recognised the presence of hierarchical levels (see Jaenicke and Ruldolph, 1986). There are to date six known levels of organisation. The primary one-dimensional structure refers simply to the sequence of covalently joined

specific amino acids in a polypeptide chain. The secondary level defines the regular rearrangement of the backbone into alpha helices and beta sheets and the joining together of these elements, to form for example coiled coils or beta meanders is known as the super-secondary level. The tertiary level of organisation refers to the three-dimensional fold of the secondary elements into a specific compact state often composed of structurally or functionally spatially separated entities (referred to as domains). The domain structure, of many proteins, is important for facilitating folding and for the specific function of the protein (Chothia, 1984). The quaternary level of organisation is the regular arrangement and assembly of several polypeptide chains comprised of the above structural organisation. The quaternary level of organisation is only found in some proteins.

1.2 Historical perspectives of protein folding

The process of protein folding and the rules of how the one-dimensional state codes for the three-dimensional state has been extensively studied (for reviews see Dill, 1990; Ruddon and Bedows, 1997). The story of protein folding starts long before the pioneering work of Anfinsen and his co-workers. The first milestone noted was in 1911 when Chick and Martin discovered the denaturation of proteins *in vitro* and distinguished this event from the process of aggregation. The next ten years saw an effort by researchers to understand the denaturation process and it was proposed to be either the result of hydrolysis of the peptide bond or the result of dehydration of the protein. In 1929 and 1931, Hsien Wu postulated that protein denaturation was an unfolding process and that the three-dimensional structure of the native protein was formed by the regular repetition of repeated folding units held together by non-covalent interactions (this paper is reprinted by Edsall, 1995). The reversibility of the unfolding process and information on the native-like properties of the refolded protein was provided by studies on haemoglobin, bovine serum albumin and other proteins. In the 1950's studies, by Eisenberg and Schwert and by Schellman, demonstrated that denaturation was thermodynamically reversible and that the

process of denaturation resulted in a conformational change of the native state of a protein.

The pioneering work of Anfinsen dealt with a more detailed and specific study on the process of reversibility. Using bovine pancreatic ribonuclease (RNase), he demonstrated that *in vitro* the reoxidation of RNase to a native biologically active form of the enzyme occurred in the absence of any co-factors or helper proteins (Anfinsen, 1973). This led to the still valid conclusion that "no special genetic information beyond that contained with the primary amino acid sequence is required for the proper folding of the molecule". Anfinsen's experiments also led to the proposal of the "thermodynamic hypothesis" that states that the three-dimensional state of the protein represents the lowest Gibbs free energy state (see section 1.4).

At the same time researchers began to realise that whilst the folding of proteins *in vitro* was occurring spontaneously it was occurring on a time scale too slow to work inside a cell. In 1968, Levinthal computed, albeit crudely, how long it would take for a protein to fold if there were no intermediates along the pathway and if all conformations about the polypeptide backbone could be sampled before the native state was reached. By considering a polypeptide with 100 amino acid residues, two rotable bond (ϕ and ψ) and three possible conformations around each rotable bond, it was calculated that there would be 3^{200} possible conformations of the backbone. By assuming that the rate of bond rotation was $10^{13}sec^{-1}$ then the time (t) it would take to sample each polypeptide conformation was $t = 3^{2n}/(2^n \times 10^{13}) = 1.3 \times 10^{80}$sec or 4×10^{72} years (Wetlaufer, 1973). The conclusion was, therefore, that the folding of a polypeptide chain would take longer than the lifetime of an organism and quite possibly the lifetime of the universe! It is, therefore, clear that protein folding *in vivo* follows a specific pathway and that the cell must have special factors to increase the efficiency of the process.

It was the pioneering work of Anfinsen's "the thermodynamic hypothesis" and Levinthal's paradox that formed the basis of the protein-folding problem that is still under investigation today.

1.3 The protein folding problem

The protein folding problem i.e., the pursuit to understand how the amino acid sequence specifies the three-dimensional structure of a protein has attracted the interest of a large number of researchers, biophysicists and biochemists alike. This problem, driven by the human genome project and by the development of experimental methods such as protein engineering, stopped-flow kinetics and NMR, remains one of the 20^{th} century's greatest scientific challenges.

The elucidation of the rules, by which the primary structure of a protein determines its secondary, tertiary and quaternary structures can be divided into two parts (Fersht, 1994). Firstly, we need to know how to predict the structure of a protein that is known to occur in nature from its amino acid sequence. Knowing this will enable the prediction of the structure for another protein that displays amino acid similarity since a similar structure can be assigned. Once this is achieved, the goal would be to construct a three-dimensional model of a protein (novel or unnatural) based entirely on its amino acid sequence. Secondly, the problem would require the elucidation of the sequence of events that occur in order to allow the protein to assume its unique native conformation. Both parts of the problem are not entirely independent.

The set of rules that govern the formation of the adopted conformation of the protein will enable the design of novel proteins, the rational alteration of existing proteins and enable the precise nature of protein structure-function relationships such as ligand binding and conformational changes as the result of protein-protein interactions to be understood.

1.4 Thermodynamic versus kinetic control of protein folding

Two hypotheses to explain the principles that govern the folding process and hence the formation of the final product have been proposed: the thermodynamic hypothesis proposed by the work of Anfinsen (1973) and the kinetic hypothesis based on the calculations of Levinthal (1968).

The "thermodynamic hypothesis" states that the "three-dimensional state of the native protein in its normal physiological milieu (pH, solvent, temperature....) is the one in which the Gibbs free energy of the whole system is lowest"(Anfinsen, 1973). In this view, the final product of protein folding is constrained in its native state by a balance of stabilising and destabilising interactions (Jaenicke, 1995). A consequence of this view is that the kinetic pathway results in the most stable product and hence energy minimisation models could predict the product formed using its amino acids sequence (Baldwin and Eisenberg, 1987). The proposition of the thermodynamic control of protein folding by Anfinsen has not been seriously disputed. An argument for this view has been provided using equilibrium and kinetic unfolding/refolding experiments. If a true equilibrium is reached then the same transition curve or the same apparent rate constant is obtained whether the experiment begins with native or unfolded protein (Baldwin and Eisenberg, 1987; Kim and Baldwin, 1990) indicating that the most stable native state is thermodynamically accessible.

The kinetic hypothesis for protein folding assumes that the final product is formed via the fastest route possible and it, therefore, follows that the final product formed may not be the most thermodynamically stable state (Baldwin and Eisenberg, 1987; Matthews, 1993). The potential role of kinetic factors in controlling the protein folding pathway has been addressed and illustrated using bovine pancreatic trypsin inhibitor, the alpha-lytic protease, serpins and dihydrofolate reductase (Matthews, 1993). For example, the protease inhibitors of the serpin family refold to a metastable but active form, which is slowly converted to form a more stable but inactive state. The kinetic control of protein folding is supported by the mutation of residues that do

not stabilise the folded state but do alter the folding pathway because if the thermodynamic view were correct then such mutants would not exist (Baldwin and Eisenberg, 1987).

1.5 The protein folding pathway

The complete folding pathway of a protein requires the characterisation of the initial and final protein states; any stable intermediate(s) along the pathway and the properties of the transition states which link them.

1.6 The folded state

The folded native state of a protein is regarded as a highly ordered and co-operative macroscopic state that is generally regarded as a single conformation. The uniqueness of the native state is defined by the position and state of each atom and group of atoms that are in turn determined by the state and position of other atoms. The folded state of a protein i.e., the specific three-dimensional structure that allows it to function in a particular biological role refers to the hierarchical levels of structural organisation (see section 1.1).

The three-dimensional structure of all proteins is the result of the rotation of the linear polypeptide about single bonds. The free rotation is only about two bonds adjacent to the alpha carbon of each amino acid, namely the N-C_α, where the angle of rotation is designated ϕ, and the C_α-C bond, where the angle of rotation is designated ψ. The third peptide bond (designated ω) forms the backbone of the polypeptide chain and is constrained to be planar, in either the *cis* (0°) or *trans* (180°). With few exceptions, such as the case for the amino acid preceding proline, the peptide bond is constrained to assume a *trans* configuration. Free rotations about the other two bonds are sterically limited to three small regions in the Ramachandran plot (Ramachandran and Sosisekharan, 1968). Ramachandran plots enable any amino acid residue based on its ϕ and ψ angles to be described by a point on the map. The plot describes which conformations are sterically feasible ("allowed") and which are sterically excluded

("nonallowed") according to their van der Waals radii. Generally, for a given protein most points fall in the allowed areas with the exception of glycine for which there is a greater range of ϕ and ψ angles available because of the absence of a beta carbon. The Ramachandran plot also illustrates the positions of various regular secondary structures such as alpha helices and beta sheets. The angles of these structures fall within sterically "allowed" areas. Other arrangements of the chain such as the coil conformations although more difficult to describe are no less ordered than the alpha or beta structures.

The prediction of the tertiary structural elements has proved to be more difficult because of the assignment of side chain interactions. Side chain interactions are typically formed between residues that are far apart from one another in the linear sequence.

The next order of structural organisation is typically seen with larger polypeptide chains (molecular masses between 40-100kDa) and involves the specific interactions of individual units. A multi-subunit protein may be composed of either identical or non-identical polypeptide chains that are closely associated to form an interface that closely resembles the compact core of a single protein unit. The overall quaternary structure of a protein may exhibit a wide range of symmetries depending on the geometry of the interactions.

The native state is viewed as being conformationally flexible and its existence depends on a delicate balance between stabilising and destabilising forces. In general, for globular proteins the free energy of stabilisation is marginal, between 10-30kcal/mol (Seckler and Jaenicke, 1992; Neet and Timm, 1994). This energy has been described to be equivalent to either a few hydrogen bonds or a few ion pairs or hydrophobic patches although it is known that all of these interactions are necessary to maintain the native three-dimensional state (Jaenicke, 1991). The marginal stability of a protein has been described to be the result of evolution since proteins, in general,

have multiple functions ranging from the first which is folding to the last which involves protein degradation (Jaenicke, 1991).

1.6.1 Forces that stabilise the folded state

In general, the marginal stability of proteins implies that the maintenance of its unique three-dimensional state is dependent on the co-operative interactions of a large number of non-covalent forces. Over the years, there has been much speculation on which force is the most dominant and acts as the driving force for the stabilisation of a protein.

An understanding of the forces that stabilise proteins has enabled development of techniques to modify a protein to increase its stability. Over the past few years a number of proteins have been stabilised to the level found typically in extremophiles (Lee and Vasmatzis, 1997). The focus has been on the stabilisation of the intrinsic properties of a protein through mutational studies. One mode of stabilisation was achieved through "small-to-large" mutations within the hydrophobic core of a protein. This "size-swap" mutation could either distort the polypeptide due to a reorganisation of the packing or it may improve the packing. The role of buried charges has been investigated and it has been found that two oppositely charged residues buried together stabilise a protein by 3-5kcal/mol. For example, a buried charge triad mutant of the Arc repressor (Waldburger *et al.*, 1995) was found to be more stable than the wild type protein. Buried charges have been extensively studied using salt bridges (Dao-pin *et al.*, 1991), parallel coiled-coil helices and parallel helices (Yu *et al.*, 1996). The introduction of disulphide bonds or the cross-linking of any groups far apart in a sequence but close together in the native three-dimensional conformation have also been shown to stabilise the native conformation of a protein (Pace *et al.*, 1988).

Presently, the hydrophobic interaction and the hydrogen bond are thought to be the most dominant forces (Pace, 1992; Myers and Pace, 1996). The major interactions

that maintain the native fold have been described qualitatively rather than quantitatively (for review, see Dill, 1990). The non-covalent interactions, which are found typically in proteins, are the van der Waals interactions between charged groups (electrostatic), polar bonds (hydrogen bonds) and apolar groups (hydrophobic interactions).

1.6.1.1 Hydrogen bonding and van der Waals interactions

The self organisation and tight packing of the interior of a protein is determined by van der Waals interactions that arise from transient dipoles that non-bonded atoms induce in each other (Murphy, 1995). These interactions are ubiquitous in nature and are short ranged. The hydrogen bond is a non-covalent interaction that is formed as the result of partial charges between a hydrogen bond donor (e.g., a hydroxyl or amino group) and a hydrogen bond acceptor (e.g., an oxygen or a nitrogen). The hydrogen bond is, therefore, essentially a dipole-dipole interaction although it is often described as an electrostatic interaction between two partial dipoles (Murphy, 1995).

Mirsky and Pauling (1936) were the first to describe the importance of the hydrogen bond in protein stability (Dill, 1990). It was natural to assume that the hydrogen bond and van der Waals forces played a role in protein stability for a number of reasons. Firstly, the amino acids that comprise the polypeptide chain are dipolar and therefore are able to hydrogen bond. Secondly, the secondary structure features (the alpha helices and beta sheets) are the result of hydrogen bonding and thirdly the interactions between the polymer and solvent were short ranged indicating that the chains were self-attractive (Dill, 1990). In addition, the stabilising effects of the hydrogen bond may be the consequence of its slightly ionic nature, which would facilitate the internal packing of the protein. The hydrogen bond although important has a minor stabilising role because of competing interactions between protein groups and water. Mutagenesis studies have suggested that hydrogen bonding contributes 1 to 2kcal/mol per hydrogen bond to the stability of a protein (Pace, 1992; Myers and Pace, 1996). The importance of the hydrogen bond as a stabilising force has been supported by

thermal stability studies, where a positive correlation between the number of hydrogen bonds and salt bridges and thermostability was found (Vogt and Argos, 1997).

1.6.1.2 Hydrophobic interactions

Kauzmann (1959) first identified the role of hydrophobic interactions in stabilising the folded native state. Kauzmann's arguments were based on the relative partitioning of small apolar molecules between an apolar solvent, such as methane, and water. Using this as a model he proposed that apolar amino acids residues such as Ile, Val, Leu, Trp, Tyr and Phe would tend to be largely buried away from the solvent whilst the charged, polar side chains would be orientated to ensure contact with the solvent. This distribution of amino acids whilst made long before the elucidation of the first three-dimensional structure turned out to be an accurate prediction. Crystallographic studies have indicated that the interior of a protein is tightly packed and that such tight packing is essential for stabilisation of the proteins native conformation (Richards, 1977; Chothia, 1975). It is this tight packing that maximises van der Waals interactions and results in hydrophobic stabilisation.

The role of hydrophobic interactions in stabilising the native state has been extensively studied and have, in general, been concluded to be the major driving force in protein folding (for review see Rose *et al.*, 1985). The burial of hydrophobic residues has been proposed to have two general effects in terms of protein stability. The first effect, known as the hydrophobic effect, is the stabilisation of the native protein as a result of shielding from the aqueous environment. The magnitude of this effect has been compared with hydrophobicity scales of amino acids which were derived from the partitioning of small apolar molecules between water (representative of the unfolded state) and an apolar solvent or vapour phase (representative of the folded state). These scales indicate that the burial of a hydrophobic side chain results in an increase in the stability of a protein (Matsumura *et al.*, 1988). In general, this effect contributes 1.0-1.6kcal/mol to the overall stability of a protein (Sandberg and

Terwilliger, 1992). The second effect has been referred to as the "packing effect" and is the result of differences between the interior of a protein and an apolar liquid. In this effect, a protein is stabilised by 0.8-2.9kcal/mol as a result of the combination of close packing, the distortion of the proteins packing and the effect of the polarity of the interior of a protein. It is the combination of both effects that contributes to the stabilising effect of the hydrophobic interactions (Sandberg and Terwilliger, 1992).

The measurement of solvation effects has shown that the origin of hydrophobic effects is largely entropic at room temperature. The unfavourable entropy of solvation has been ascribed to the decrease in entropy of water surrounding the apolar groups. The solubility of apolar groups is lower in water. Therefore, the free energy of transfer to water is positive. At room temperature, the lower solubility in water could be explained by the ordering of water around the apolar groups, which would in turn lower the entropy of the system and increase the free energy ($\Delta G = \Delta H - T\Delta S$).

1.6.1.3 Electrostatic interactions

Electrostatic interactions simply occur between charges on protein groups such as amino- and carboxyl-termini and ionizable side chains. Barlow and Thornton (1983) have extensively studied Coulombic interactions (or ion pairs) and have defined an ion pair as two charged groups within 4Å of each other. Most ion pairs occur between distant residues in the sequence and between different segments of secondary structure elements and therefore exert their stabilising effect on the tertiary structure of a protein. Electrostatic interactions have been found to affect protein stability in two ways: firstly, through classical effects which result from non-specific repulsions that are formed when a protein is charged and secondly, through specific effects as the result of ion pairing. Ion pairing has been found to occur on the surface of a protein as well as in the protein core. Generally, a protein has only five ion pairs per 150 residues only one of which is buried (Barlow and Thornton, 1983). The energetic cost of burying an ion pair is large and therefore if a charge is buried it is essential for the specific role of the protein.

X-ray crystallography has provided evidence for ion pairing occurring on the surface of proteins and mutagenesis studies have shown that variations in ion pairing can influence the stability of a protein. However, evidence against it as a dominant force has been provided for a number of reasons: (i) The stability of a protein is not significantly affected by pH or salt at or near its isoelectric point, (ii) the model composed shows that the change in volume on formation of an ion pair (-14ml/mol per ion pair) contrasts with experimental studies which show a partial increase in the volume, (iii) mutagenesis studies of a large number of ion pairs in proteins has little effect on the stability of a protein and (iv) the conservation of ion pairs in evolutionary linked protein families is low (Barlow and Thornton, 1983).

1.6.1.4 Configurational entropy

Another important determinant of protein stability is configurational entropy, which tends to destabilise the native protein structure. The gain in entropy relates to the increased degrees of freedom of the side chains and backbone available to the unfolded protein relative to the native state. Thus the unfolded state has greater entropy and the conformational entropy is a destabilising component of the free energy. The destabilising effect is, therefore, due to the loss of configurational entropy of the folded protein and this is compensated for by the many small stabilising interactions (Baldwin and Eisenberg, 1987).

1.6.2 Major determinants of the protein fold

The inspection of the native state of proteins using information collected on protein structures using the protein data bank (PDB) (Bernstein *et al.*, 1977) indicated five basic but major determinates of a protein fold (Thornton *et al.*, 1995 and references therein).

First, it was apparent that all proteins exhibited a tightly bound hydrophobic core (Richards, 1977; Hubbard, 1994). The entropic tendency to exclude water has three

consequences. The tertiary and quaternary structures of globular proteins have a tightly packed interior with apolar side chains in the centre and polar and ionizable side chains at the surface, the "inner" core is unsolvated (about one water molecule is buried per 27 residues) with the average water sized cavity constituting only 1% of the total volume of the protein and the association of proteins leads to maximum stabilisation and complementation such that the quaternary structure has a unique fit which is highly specific. Second, the protein in the native state adopts a fold so that the torsional angles ϕ and ψ are confined to low energy conformations and therefore the interactions even within the proteins interior are strong enough to provide constraints on the torsional freedom. Third, the hydrogen bond potential for nearly all-main chain groups (buried or exposed) is satisfied indicating its contribution to stability. The satisfaction of the hydrogen bond potential is the driving force for the formation of secondary structural elements of globular proteins. Fourth, the interactions between side chains are specific except for apolar-apolar interactions, therefore, indicating that the side chain interaction between oppositely charged residues, polar residues and aromatic side chains play a role in protein stabilisation. Fifth, the protein structure is on average (approximately 60% of all residues) dominated by secondary structural elements that interact to fulfil the requirement for close packing.

Therefore, based on theoretical and structural evidence, the forces that maintain the three-dimensional state of a protein are balanced between those that satisfy the local conformation preference and those that fulfil the global requirement to bury an apolar side chain whilst satisfying the potential hydrogen bond acceptors and donors (Thornton *et al.*, 1995).

1.7 The unfolded state

The unfolded state of proteins, unlike the native state, is not well defined by a set of co-ordinates and therefore it is described as an ensemble of conformations. The unfolded state is described by Dill and Shortle (1991) as a specific subset of

denatured states for which the conformations are open and solvent exposed with little or no residual structure i.e., it is the state that is formed under strongly denaturing conditions. The extent of unfolding of the denatured state and its importance in influencing the (re) folding pathway has been extensively studied (Dill and Shortle, 1991). Little is known about the unfolded state but it is essential to characterise it in order to complete the folding pathway.

Tanford was the first to study the denatured state and he demonstrated that the denatured protein in 6M guandinium chloride has a hydrodynamic radii that was close to that expected for random coils based on theoretical estimates (Tanford 1968; 1970). However, several recent studies have shown that even in high concentrations of denaturant significant amounts of residual structure can be detected (Shortle, 1996). For example, the denatured state has been characterised for ribonuclease A (Sosnick and Trewhella, 1992) using X-ray scattering and for the 434-Repressor (Neri *et al.*, 1992) and barnase (Arcus *et al.*, 1995) using NMR. The residual structure and heterogeneous nature of the denatured state has important implications when studying refolding pathways.

1.7.1 Denaturing agents

Denaturants, often termed as chaotropes, are agents that result in the disruption of the balance of forces between the protein interactions with itself and its interactions with the environment (Schiffer and Dötsch, 1996). The disruption of these forces is the result of environmental changes such as temperature extremes, pH extremes or solvent. The most widely used denaturants are urea and guandinium-chloride because of the persistence of structure found in the acid and temperature denatured states (Dill and Shortle, 1991; Matthews, 1993). A combination of denaturants such as urea of guandinium-chloride with pH and temperature has been shown to have an additive effect (Schiffer and Dötsch, 1996).

1.7.1.1 Solvents

The action and details of the chemistry on how these denaturants work is poorly understood. Recently, computational methods such as molecular dynamics simulations and experimental studies have implicated the role of water in the initial stages of unfolding (Schiffer and Dötsch, 1996). Simply, the denaturants are thought to have two mechanisms of action. First, there is an initial decrease of water-water interactions that result in an increase in the interaction between the water and the protein that may initiate the unfolding of the protein. Second, the exposed interactions are further destabilised by direct interactions with the denaturant. The interactions involve the weak binding of the denaturant or adsorption of the denaturant to the apolar surfaces. A large number of solvents act as chaotropes, for example guanidinium chloride, urea, NaSCN, SDS, GuSCN, and ethanol.

In the presence of concentrated urea (8-9M) or guandinium-chloride (4M), the ordered structure of water is non-existent (Collins and Washabaugh, 1985; Pace, 1992). As a result of the interactions between denaturant and water being weaker than the water-water interactions the relative interactions of water and denaturant have increased. Therefore, for proteins, the water competes with internal hydrogen bonds and this initiates unfolding. The exposed apolar surfaces are then acted on by the denaturant. Urea has been shown using NMR to interact directly with aliphatic side chains (Dötsch *et al.*, 1995) and its interaction with aromatic side chains has been shown computationally (Tirado-Rives and Jorgensen, 1993). Therefore the denaturant acts on the exposed surface in the denatured state and promotes the dissociation and unfolding of the protein (Shortle, 1996).

1.7.1.2 pH

Denaturation as a result of extremes in pH have been implicated to act on the denatured state (Shortle, 1996). Proteins in the native state are composed of basic and acidic residues and therefore possess a net charge in an acidic or basic solution which is related to its characteristic isoelectric point. In the native state, the basic and acidic

amino acid groups have pKa' s which have values altered by the environment in which they are placed. In the denatured state, these groups are exposed and therefore the pKa's of the basic and amino acids are typically similar to those expected for free amino acids. This implies that additional protons must bind to the acid denatured state and protons must be lost from the basic denatured state (Shortle, 1996). Therefore, in order for a protein to refold, protons must be lost (or retrieved) depending on whether the pH is low (or high) and the energetic cost of doing this is higher than the free energy necessary to stabilise the folded conformation of a protein (Shortle, 1996).

1.7.1.3 Temperature

Heat (or cold) are the most complex denaturing agents which are thought, albeit subjectively, to act on the denatured state of a protein (Shortle, 1996). Thermal denaturation as the result of high temperature results in an increase in the vibrational freedom of all bonds. The increase in entropy results in a decrease in the free energy ($-T\Delta S$) of the system. At low temperatures, near 0°C, some proteins denature (Schellman, 1987). In this case denaturation is the result of a decrease in the enthalpy of the system and is the result of the ordering of water (in a clathrate-like manner) around exposed hydrophobic chains (Shortle, 1996).

1.8 The intermediate state(s)

The efficiency of protein folding (Levinthal's paradox) suggests that the process proceeds through intermediate states and this raises the possibility of the observation of partially or incompletely folded states. These states would enable the folding pathway to proceed through a definite sequence of events with decreasing Gibbs free energies (Privalov, 1996). The nature of intermediates and their actual role in speeding up a folding process (i.e., their ability to act as "folding domains") is largely debated. The structural features of intermediates can be inferred from the analysis of the transition states (section 1.9.2).

The concept of "folding domains" is feasible for large proteins since they generally fold through definite steps that correspond to the folding of individual domains. However, for smal proteins and for some oligomeric proteins the concept is problematic because the rate of folding is fast and therefore, suggestive of folding without any domain organisation. Intermediates have been postulated to be important in directing the folding process (Kim and Baldwin, 1990) but recent evidence using kinetic models (Fersht, 1995; Baldwin, 1996) has suggested that significantly populated intermediates may slow the folding process. The reasoning for this is based on the view of the protein folding pathway as an energy landscape where the folding rates are a function of the "roughness" of the energy function on the path between the native and unfolded state (Burton *et al.*, 1996). Therefore, the folding path would be slowed down by any crenulations (i.e., by transient intermediates) in the energy surface.

For some proteins, for example, alpha lactalbumin, apomyoglobin, DnaK and staphylococcal nuclease, an equilibrium intermediate that is neither native nor completely unfolded has been observed. The state was called a " molten globule" (Kuwajima, 1977) and was initially described to be a collapsed state of the polypeptide chain without fixed, specific long-range interactions. A number of experimentalists characterised the molten globule and found it to be a compact species closer to the native state than to the unfolded state with a substantial amount of secondary structure and little or no tertiary structure (Privalov, 1996). The state was compact but had a significant amount of hydrophobic surface exposed when compared with the native state. Since the first report of the molten globule, a large number of studies have described the appearance of the molten globule and there is confusion of the actual meaning of molten globule. The term may be used to describe an unstructured collapsed specie with a hydrophobic core or to describe a structured species with a fixed specific secondary structure (Fink, 1995). To resolve this issue two types of molten globules were proposed, the "true" molten globule which is structured and the "collapsed" molten globule which is unstructured (Baldwin, 1991).

There is evidence for two types of compact denatured states: the compact intermediates (the molten globule) which represent a global free minimum and the compact form of the unfolded state (Palleros *et al.*, 1993). To experimentally distinguish between the two is difficult and therefore theoretical models of each denatured state have been proposed (Ptitsyn, 1992; Finklestein and Shakhnovich, 1989). It is, therefore, questionable whether the molten globule is a true thermodynamic state or a substate of the unfolded state.

Privalov (1996) addresses the concept of the molten globule as a thermodynamic state in detail. The criterion for a specific thermodynamic state is the presence of a phase transition between that state and the next state. If the transition results in a large change in thermodynamic parameters, such as enthalpy and entropy, then the state is regarded as a first-order phase transition. The consideration of a protein as a macroscopic system means that the different states (folded/unfolded/intermediate) would be separated from each other by a first-order phase transition. For small proteins, the co-operative transition between the folded and unfolded states implies that the two-state transition should be regarded as a first-order phase transition. In all reported cases, the molten globule state is separated from the native state by a highly co-operative two-state transition but the order of the transition is not clear. However, analysis by statistical methods has suggested that the transition from unfolded state to compact molten globule would be a second order phase transition which would be accompanied by a change in heat capacity and compressibility (Privalov, 1996 and references therein).

Recently, the idea of the molten globule has been criticised for a number of reasons (Privalov, 1996). Firstly, the definition that this state is a dynamic highly mobile conformation is disputed by the fact that the molten globule and unfolded state have different retention times in chromatographic experiments meaning that the rate of interconversion between the two states is slow. And secondly, the description of this

state as molten (i.e., as being liquefied) is conflicting with its definition as a conformation with fixed long-range interactions.

1.9 Experimental methods to monitor protein stability and the unfolding/refolding pathway

The molecular mechanisms of protein folding have been studied in detail for small monomeric proteins (molecular masses between 30 and 40kDa). These studies have provided information on how the amino acid determines the secondary and tertiary structures of a protein (Kim and Baldwin, 1990; Matthews, 1993). However, the use in applying small monomeric proteins as model systems to understand the folding mechanism of dimeric/oligomeric proteins is ambiguous. The folding mechanisms of larger proteins (40-90kDa) addresses the issue of domain and subunit association reactions that result in the formation of the quaternary structure (Jaenicke, 1987). In order to gain a general folding mechanism it is essential that folding and unfolding studies be performed on different structural classes of proteins with varying degrees of oligomeric states.

Conformational stability (equilibrium) studies have been extensively used to provide information on the structure, stability and folding of the initial and final states of the folding pathway and any stable intermediates (Neet and Timm, 1994). Kinetic studies of the unfolding and refolding properties of proteins have provided information on the pathway and mechanism of protein folding and unfolding and the transition states. The combination of these two techniques with "protein engineering" (mutagenesis studies) allows specific interactions that control the pathways to be mapped out at the level of individual residues (Matouschek *et al.*, 1989; Fersht *et al.*, 1992).

1.9.1 Conformational stability studies

The stability of a protein is defined as the decrease in Gibbs free energy of a structure-less polypeptide chain when it folds to give a native protein molecule in water ($\Delta G(H_20)$). The stability of a protein can be determined from the unfolding

transition curve induced by a denaturant (such as pH, temperature or solvent). The denaturation of many small single domain proteins and some multi-domain proteins is a reversible process. For a schematic representation of a sigmoidal denaturant-induced unfolding transition fitted to a two-state approximation see, for example, figure 6. The two-state denaturation model has been observed for monomeric proteins (Kim and Baldwin, 1990) and some dimeric proteins (for example, the Arc repressor (Bowie and Sauer, 1989) and the *E.coli. Trp* aporepressor (Gittelman and Matthews, 1990)) and is the consequence of the co-operative nature of a number of non-covalent interactions, which result in the marginal stability of a protein (section 1.6.1).

The minimal hypothesis for a two-state transition assumes that at any point in the unfolding transition that the individual residues of the protein exist in only one of two states: the native or unfolded state. In other words, for monomeric proteins and some dimeric proteins the native monomer (or intermediate) does not exist at significant concentrations at equilibrium (Neet and Timm, 1994). It is important to note that the two-state approximation does not rule out the presence of a high-energy intermediate (which would correspond to a local minimum on the reaction co-ordinate diagram)(Wolynes *et al.*, 1995). If the energy of an intermediate were sufficiently high then the state would not be populated to a significant extent and therefore would not have any effect on the experimental observations.

Larger proteins generally have equilibrium unfolding transitions, which are more complex. Generally, these proteins which display non-co-operative multi-state unfolding are composed of individual domains which are able to fold and unfold independently (Neet and Timm, 1994). A general three-state model for oligomeric proteins would involve the native dimer, the unfolded monomer and folded monomer or monomeric intermediate. The presence of a stable equilibrium intermediate increases the complexity of the unfolding/refolding pathway. Some examples of oligomeric proteins unfold via an intermediate are the β2-subunit of tryptophan synthetase (Zetina and Goldberg, 1980), the lambda repressor (Banik *et al.*, 1992) and

superoxide dismutase (Mei *et al.*, 1992). The intermediate may be either dimeric or monomeric and may not have spectral properties similar to those found for the folded or unfolded protein states.

A sigmoidal transition for equilibrium unfolding implies that the interactions within the native protein are co-operative. This co-operativity may be either two-state (i.e., a first-order phase transition) or one-state (i.e., a higher-order phase transition) (Chan *et al.*, 1995)(figure 1). The only way to distinguish between the two requires the observation of the species populated at the midpoint of the unfolding transition. For a two-state transition, there are two identifiable states in contrast to only one for the one-state transition. Therefore, the two-state transition implies an energy surface with two minima separated by a barrier and the one-state transition implies a single energy minimum with no barrier between the native and unfolded state (Dill and Shortle, 1991; Chan *et al.*, 1995; Fersht, 1995).

Three models have been used to describe the co-operativity of non-covalent interaction within the native protein (Chan *et al.*, 1995). Model 1, the "helix-coil" model where co-operativity is described to be the result of interactions between neighbouring atoms in the sequence (Schellman, 1958; Zimm and Bragg, 1959; Poland and Scheraga, 1970). Model 2, the "side-chain packing" or "jigsaw-puzzle" model where co-operativity is described to be the result of the complementary fits of side chains (Shakhnovich and Finklestein, 1989) and model 3, the "hydrophobic collapse" model where co-operativity is the result of the assembly of apolar residues into a compact core (Dill, 1995). Only the third model has been supported by experimental evidence.

The analysis of the equilibrium unfolding transition curve is essential for a number of reasons (Utiyama and Baldwin, 1986). First, it enables the reversibility of the unfolding/refolding transition to be assessed. Second, it characterises the final and initial states of the protein-folding pathway. Third, it establishes the presence/absence

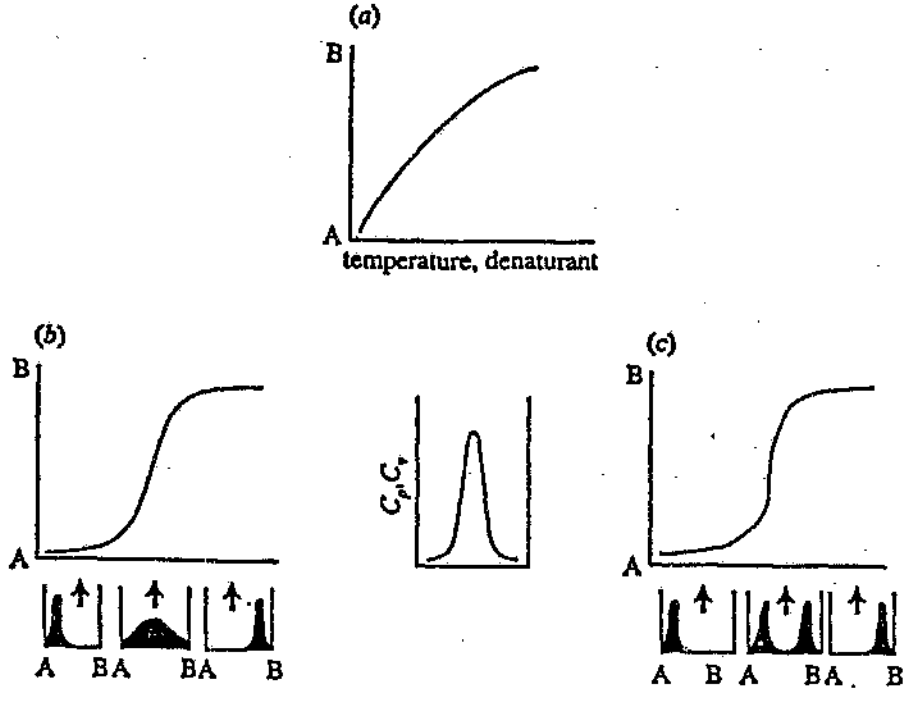


Figure 1. Models of co-operativity: two states, A and B, such as native and denatured states, can change populations with increasing denaturant concentration. (a) Gradual change - no co-operativity. (b) Co-operative transition of the one-state type. (c) Co-operative transition of the two-state type. Both one-state and two-state transitions can have sigmoidal behaviour and heat absorption (a peak in the C_p plot); they cannot be distinguished on these bases or from the steepness of the sigmoidal curve. The main distinction is whether there is one broad peak involving populations of intermediates near the denaturant midpoint (one-state) or whether there are two populated states and less intermediate population (small plots at the bottom of (b) and (c)). (Taken from Chan *et al.*, 1995).

of an equilibrium intermediate. Fourth, it enables the design of kinetic unfolding/refolding experiments and fifth, it enables the consistency between kinetic and equilibrium experiments to be tested.

1.9.2 Unfolding/refolding kinetic studies

The kinetics of protein unfolding/refolding are used to characterise the folding mechanism of a protein. In addition, the elucidation of the kinetic mechanism for protein unfolding/refolding enables any transient intermediates and the transition state between them to be characterised.

The details on how a protein unfolds/folds have increased due to the large amount of structural information available from fast reaction measurements such as stopped-flow kinetics (interfaced with fluorescence, circular dichroism or NMR); fluorescence life time measurements (Elöve *et al.*, 1992); mass spectrometry (Fersht, 1993; 1994) and "protein engineering" studies. These techniques have enabled the observation of early unfolding/refolding events (millisecond to microsecond time range).

In stopped-flow kinetic studies, the protein is unfolded by changing conditions to favour the unfolded state and vice versa for the folding of the protein. The unfolding/refolding pathway measured depends on the probe and the experimental conditions used. The experiments are designed using the equilibrium unfolding transition curve and enable the different aspects of the pathway to be characterised (Kiefhaber, 1995).

1.9.2.1 Unfolding kinetics

The kinetics of unfolding for most proteins, single and multi-domain, is a simple monophasic process. This is consistent with an "all-or-none" process of unfolding for a homogeneous population of native protein and is suggestive of a single transition state for unfolding or a set of energetically similar transition states (Kim and Baldwin, 1990). Complexities in kinetic unfolding reactions are rare but have been observed

for some proteins; for example, RNase (Houry *et al.*, 1996; Houry and Scheraga, 1996), dihydrofolate reductase (DHFR)(Touchette *et al.*, 1986), Staphylococcal nuclease (Walkenhorst *et al.*, 1997) and *Rhodobacter capsulatus* cytochrome c_2 (Sauder *et al.*, 1996). The complexities may be the result of either multiple states of the native protein or the result of kinetic coupling between peptide bond isomerisation and the actual folding reaction (Kiefhaber, 1995) or the result of formation of a structured intermediate (Walkenhorst *et al.*, 1997).

1.9.2.2 Refolding kinetics

The kinetics of refolding are generally more complex with at least two kinetic phases. The complexities may be the result of a heterogeneous mixture of unfolded protein which interconvert slowly as the result of different isomers about the peptide bond or as a result of disulphide bond formation. In addition, complexities may be the result of the population and accumulation of refolding intermediates. Generally, all molecules within the population have the same probability of folding via a pathway that is governed by a rate-limiting step. For some small proteins (for example, the chymotrypsin inhibitor, CI2 (Jackson and Fersht, 1991) and for a few oligomeric proteins (for example the Arc repressor (Milla and Sauer, 1994)) studied, a single major refolding phase was observed indicating that all the conformations interconvert rapidly and fold via a single pathway. For oligomeric proteins, the single refolding phase is indicative of tight coupling between the folding and the association reactions (Milla and Sauer, 1994). Generally, however, the folding of oligomeric proteins requires the co-ordination of secondary, tertiary and quaternary structures, which results in a complex refolding pathway.

1.9.2.3 The transition state for unfolding/refolding

Kinetics of unfolding/refolding is the only method available for analysing the structures of the transition states using the transition state theory. The transition states for protein unfolding/refolding have been defined as the intermediate species with the greatest free energy (Matouschek *et al.*, 1989). It is the most transient specie and

therefore it can only be characterised indirectly. The indirect conventional methods involve the characterisation of the transition state by monitoring the temperature and urea-dependence of the apparent unfolding/refolding rate constants (Tanford, 1970; Chen *et al.*, 1992; Schindler *et al.*, 1995). Recently, the "protein engineering" method has been used to map the interactions of the transition state at the level of the amino acid residues (Matouschek *et al.*, 1989). The two parameters that are used to characterise the structure of the transition state are $\alpha^{\ddagger}$ and ϕ. The $\alpha^{\ddagger}$ parameter monitors the solvent accessibility of the transition state based on its susceptibility to urea and the ϕ parameter probes the structure of the transition state by mutation.

1.10 Theoretical studies of protein folding

The large amount of interest in the protein folding pathway by physicists and mathematicians has led to an increase in a wide range of theoretical approaches to the problem (for review, see Karplus and Šali, 1995). The theoretical approach varies from the simple classification of single amino acids by their tendency to form a specific type of secondary structure to the classification of how secondary structural elements combine to form the tertiary folding patterns (Chothia, 1984).

The limitations of computing power necessitate that only simplified models are used to simulate the unfolding/refolding process. The simplifications omit details of the proteins structure (i.e., only the backbone is included) and any interactions that maintain the native state. In addition, simulations of the folding process are restricted because of the random coil-like nature of the unfolded state and hence the large number of configurations that the unfolded state can occupy.

The mechanism of protein folding has been investigated using molecular simulation dynamics (MD) and lattice models (Karplus and Šali, 1995). MD simulations provide information on the initial stages of the folding process using an all-atom model and have provided information on the structure and solvation of the partially unfolded state. However, because these simulations are approximate it is necessary to validate

them with experimental studies. The lattice models are used to provide information on the protein-folding pathway. In the lattice heteropolymer simulations the protein is represented as a string of beads and provides information on the final stages of protein folding (i.e., up to the stage of side chain packing). The lattice models have enabled the proposal of a folding mechanism for small proteins and have resolved the Levinthal paradox. The main conclusions from these simulations were that a global energy minimum is necessary for the rapid folding of a protein into its stable native state and that the folding mechanism starts with a rapid event that reduces the number of conformations that need to be searched and from this collapsed state folding proceeds to the native state through a number of transition states. Although the two types of simulations have provided information on the unfolding/folding mechanisms, it is necessary to assess their validity by means of experimental methods.

1.11 Models of protein folding

Several models have been proposed to describe the pathways of protein folding. These vary from the model that predicts a single pathway to one that predicts a series of pathways.

The "jigsaw-puzzle" model predicts that folding is a series of sequential folding events i.e., each protein molecule folds via a unique pathway (Harrison and Durbin, 1985). In this model, the individual molecules are conformationally diverse until the native state of the protein is reached. The "nucleation-growth" now known preferably as the "nucleation-condensation" model is in contrast to the above model (Fersht, 1995). This model proposes the formation of a stable specific substate that acts as a nucleus. The nucleation site directs the rest of the molecule to fold around it. This model, was first proposed by Wetlaufer in 1973, and is presently the most popular. In this model it is proposed that the native structure of the protein is formed early in the folding process and has been used to describe the early kinetic events in folding (Abkevich *et al.*, 1994; Fersht, 1995) for the chymotrypsin inhibitor, CI2 (Itzhaki *et*

al., 1995), the acyl-CoA binding protein (Kragelund *et al.*, 1995) and the MYL mutant of the Arc repressor (Waldburger *et al.*, 1996).

1.12 Protein folding funnels

Levinthal's hypothesis (section 1.2) led to the search for protein folding pathways with populated intermediate states. The observation of intermediates along the folding pathway was postulated to be important for directing the polypeptide chain towards its final, folded conformation (Kim and Baldwin, 1990). The role of intermediates along a protein folding pathway is largely debatable (section 1.8) since it is possible that the intermediate may act as a trap rather than speed up the folding process (Fersht, 1995; Baldwin, 1996).

The idea of the folding pathway has since been replaced by the view of protein folding kinetics as an "energy landscape" and "folding funnels" (for review see Dill and Chan, 1997). In this view, the folding process is described by a number of parallel folding events which result from an ensemble of conformations. Each ensemble, therefore, folds asynchronously to a distinctive conformation which ultimately channels into the same native structure.

The "energy landscape" is a function of the internal free energy of each configuration (which includes the sum of hydrogen bonds, ion pairs, torsional angles, hydrophobic and solvation free energies etc.) versus the degrees of freedom of each configuration (for example, the dihedral bond angles, ϕ and ψ).

The energy landscape is described by many features with peaks (hills) which correspond to high-energy conformations (for example, conformations that result in the burial of hydrophobic surfaces or unfavourable dihedral angles) and troughs (valleys) which correspond to more energetically favourable conformations. Therefore, the folding rates are a function of the "ruggedness" of the energy function on the path between native and denatured states (Burton *et al.*, 1996 and references

therein) and any bumps slow down the folding process. An idealised protein-folding funnel is smooth and, in this case, the kinetics of protein folding is fast and best described as a two-state folding mechanism. In general, however, folding funnels are more complex with bumpy landscapes that describe multi-state kinetics. The "ruggedness" of the path acts as a trap for the accumulation of misfolded intermediates that must undergo conformational rearrangements before continuing along the path towards the native state.

The modelling of protein folding kinetics using "energy landscapes" and "folding funnels" is in its elementary stages and it represents an idealised concept to aid the understanding of the protein-folding kinetics.

1.13 Protein folding *in vivo* versus *in vitro*

The folding process of a protein in the cell is likely for obvious reasons to differ from that monitored *in vitro*. Apart from differences in solvent, pH, ionic strength and temperature, the initial event of the refolding pathway is artificial (Radford and Dobson, 1995). *In vivo* the process of folding begins from a nascent polypeptide chain whilst *in vitro* the initiation of folding is from a protein that was denatured.

Peptide synthesis from the C terminus to the N terminus has been shown to result in native protein indicating that the vectorial nature of protein synthesis is not important. In addition, the rate of refolding of protein under suitable conditions is similar to the folding rate *in vivo* and the final product after refolding has been found to be similar if not identical to the native state (Rainer and Jaenicke, 1986).

There have been some similarities as well as differences observed between protein folding *in vivo* and *in vitro* (Ruddon and Bedows, 1997 and references therein). One main difference is the rate and efficiency of folding. *In vitro* the folding rate is not as high as that observed *in vivo* and may be explained by competing pathways that exist between correct folding and misfolding and aggregation. *In vitro* studies require the

accurate control of protein concentration in order to minimise and prevent aggregation. The efficiency of the folding process *in vitro* can be facilitated by protein catalysts such as the protein disulphide isomerase (PDI) and the prolyl peptide isomerases (PPI). In addition, proteins called molecular chaperones (DnaK/DnaJ for bacterial proteins and their eukaryotic analogues, Hsc70 family and BiP) have been found to increase the yield of native protein by preventing "off-pathway" folding events.

The role of chaperones in folding, assembly and intracellular translocation of proteins has been reviewed (Hartl, 1996). Generally, chaperones act sequentially in the folding pathway by binding folding intermediates and passing them onto the next chaperone or chaperone complex in the cascade eventually releasing a native protein. The binding of the chaperones to the protein involves the interaction with hydrophobic residues exposed on the surface of the unfolded protein and its release often requires ATP hydrolysis (Ruddon and Bedow, 1997).

It is important to note that although foldases and chaperones are needed to assist protein folding events in the living cell they do not negate the fact that the folding of a polypeptide chain occurs spontaneously, i.e., Anfinsen's hypothesis.

The interest in mimicking folding events *in vivo* has its application in treating human diseases th[illegible]he result of protein folding defects (Ruddon *et al.*, 1996), for example, cy[illegible]brosis, alpha 1-antitrypsin deficiency, Alzheimer's disease, neurodegenerative diseases such as Huntington's chorea and prion diseases such as bovine spongiform encephalopathy (BSE; mad cow disease) and Creutzfeld-Jakob disease. It is, therefore, essential that *in vitro* folding studies be supplemented with studies of intracellular folding pathways.

CHAPTER 2
GLUTATHIONE TRANSFERASES

2.1 A supergene family of detoxification enzymes

Cytosolic glutathione transferases (GSTs, E.C. 2.5.1.18) are a supergene family of multi-functional proteins found in all vertebrates, plants, insects, yeast and aerobic bacteria. As a group of proteins, they are involved in the biotransformation of a number of reactive endogenous and toxic xenobiotics. They function primarily as phase II detoxification enzymes (Jakoby and Ziegler, 1990) and catalyse the nucleophilic addition of reduced glutathione (GSH) to numerous chemically diverse compounds, such as alkyl- and arylhalides, lactones, epoxides, quinones, esters and activated alkenes (Mannervik and Danielson, 1988). As a consequence of this reaction, a less reactive but more polar glutathionyl-S-conjugate is formed which can be exported from animal cells by a membrane-mediated ATP-dependent pump system (Ishikawa, 1992), catabolized via the mercapturate pathway and excreted (Boyland and Chasseaud, 1969). In plants, the absence of an effective excretion pathway means that the conjugate is internally compartmentalised or stored (Sandermann, 1992).

In addition, certain glutathione transferases are able to catalyse selenium-independent peroxidase activity (Ketterer *et al.*, 1987) with organic peroxides as the substrates. The catalysis of the glutathione-dependent isomerisation of 3-ketosteroids (Benson *et al.*, 1977) and the biosynthesis of peptide leukotrienes (Tsuchida *et al.*, 1987) are two other important physiological functions of glutathione transferases.

In addition to the catalytic diversity of the GSTs, they are also able to function as ligand-binding proteins ("ligandins"). The ligand binding function of GSTs facilitates the intracellular storage and transport of a variety of hydrophobic non-substrate compounds and hormones, metabolites and drugs (Listowski, 1993). The role of GSTs in the transport and catalysis of xenobiotic compounds has resulted in the proteins being responsible for the acquisition of resistance towards a variety of

xenobiotic compounds such as carcinogens, therapeutic agents and pesticides (Hayes and Wolf, 1988; Tsuchida and Sato, 1992).

The glutathione transferase family is represented by a large number of enzymes that, according to sequence homologies, are classified into one of seven species-independent gene classes. Six are the cytosolic GSTs referred to as alpha, mu, pi (Mannervik *et al.*, 1985), sigma (Buetler and Eaton, 1992), theta (Meyer *et al.*, 1991) and kappa (Pemble *et al.*, 1996) and the seventh is the class of microsomal enzymes (DeJong *et al.*, 1988). The microsomal GSTs are distinct from their soluble counterparts. Recently, a new class called zeta has been identified by sequence alignments and phylogenetic analyses in a range of species from plants and humans (Board *et al.*, 1997).

The various isoenzyme classes evolved along a divergent pathway (figure 2). A comparison of the cDNA and gene sequence of class alpha/mu/pi with that of the class theta/sigma from eukaryotes and prokaryotes suggested that the alpha/mu/pi class enzymes originated from the duplication of the theta gene (Pemble and Taylor, 1992; Pemble *et al.*, 1996; Armstrong, 1997). The highly conserved 3' non-coding sequences of the mu and theta genes also suggest that the class mu diverged from the theta precursor before the alpha and pi gene. The class theta enzyme resembles most closely the ancestral bacterial proteins and the progenitor of the theta class may be the kappa class, which encodes the mitochondrial enzyme (Pemble *et al.*, 1996). The class sigma GST from the squid digestive gland is proposed to have diverged from the theta or an early alpha/mu/pi precursor for the purpose of the production of S-crystallins (refractory proteins) in the lens of the cephalopod eye (Tomarev *et al.*, 1993).

Examinations of the three-dimensional structures of gene classes have also suggested an evolutionary distinction between the theta/sigma and the alpha/mu/pi enzyme classes. The most significant structural differences involve a residue at the active site

and the dimer interface (Armstrong, 1997). The precursor class theta enzyme has an essential serine at the active site (Board *et al.*, 1995) as opposed to the highly conserved tyrosine residue found in all other gene classes (see section 2.3). The structural distinction at the dimer interface relates to a more hydrophilic interface in the class theta/sigma as opposed to a more complex hydrophobic "lock-and-key" interface displayed in the class alpha/mu/pi. The serine/tyrosine switch supports the theta class as the ancestral precursor of the sigma/alpha/mu/pi precursor which then evolved to the alpha/mu/pi precursor through alteration of the dimer interface (Armstrong, 1997) (figure 2).

2.2 The three-dimensional structure of cytosolic glutathione transferases

All cytosolic GSTs are roughly globular proteins with dimensions of about 62Å x 51Å x 46Å. All cytosolic GSTs have dimeric quaternary structures ($M_r \approx 50\,000$) and exist as either stable homo- or heterodimeric structures. The heterodimeric isoenzyme forms exist (in the alpha and mu class only) as a result of the expression of multiple genes and intraclass subunit hybridisation. A variety of three-dimensional structures for the mammalian GST isoenzymes have been elucidated (for review see Dirr *et al.*, 1994b; Wilce and Parker, 1994). Three-dimensional structures of isoenzymes from the alpha (Sinning *et al.*, 1993; Cameron *et al.*, 1995), mu (Ji *et al.*, 1992; Raghunathan *et al.*, 1994), pi (Reinemer *et al.*, 1991; 1992; Dirr *et al.*, 1994a; Garcia-Sáez *et al.*, 1994; Oakley *et al.*, 1997), sigma (Ji *et al.*, 1995; 1996) and theta (Wilce *et al.*, 1995; Reinemer *et al.*, 1996) have been determined. In addition, a three-dimensional representative of the helminth *Schistosoma japonicum* GST (Lim *et al.*, 1994; McTigue *et al.*, 1995) has also been elucidated. The structural information provided by these three-dimensional structures has provided a wealth of information on the tertiary and quaternary architecture of these proteins and has enabled an understanding on the structure and function of the different gene classes.

The various three-dimensional representatives for the cytosolic GSTs share a similar fold and topology (Dirr *et al.*, 1994b; Wilce and Parker, 1994). Within each class,

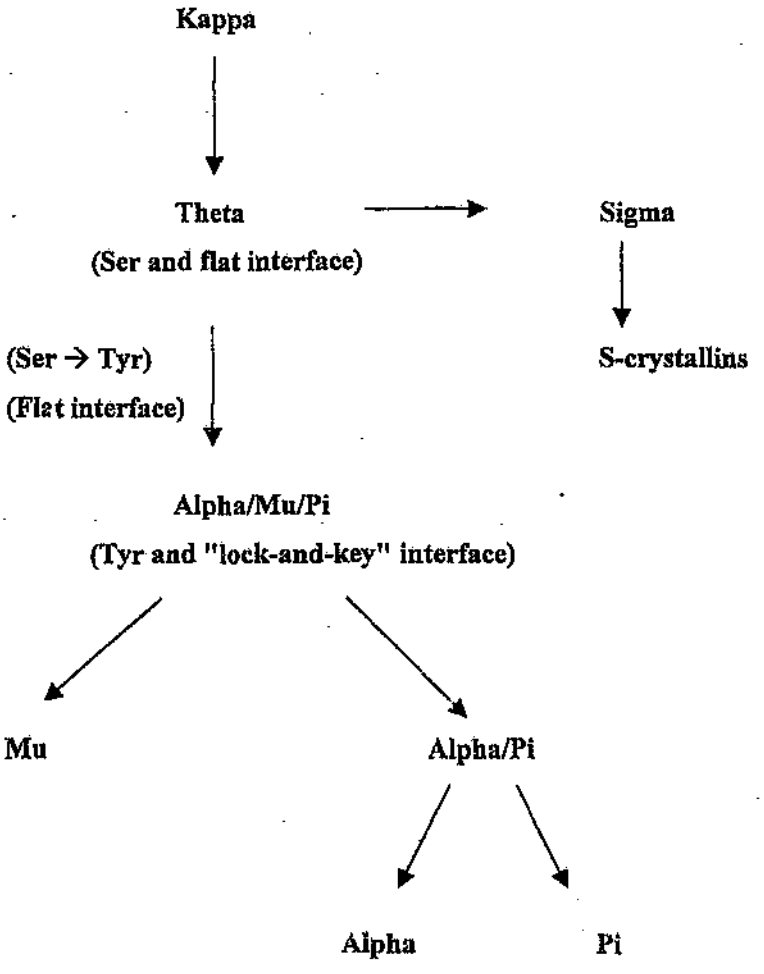


Figure 2. The proposed evolutionary scheme for the supergene family of cytosolic glutathione transferases (adapted from Armstrong , 1997)

there is a high sequence homology (> 75% identity) while between the classes the homology is much lower (20-30% identity)(Sinning *et al.*, 1993). The pair-wise sequence identities are alpha-mu 20%; pi-mu 20%; pi-mu 30%; alpha-pi 32%; theta-alpha 20%; theta-mu 13%; theta-pi 16%; theta-sigma 14%. A sequence alignment for a representative from each class is indicated in figure 3 and a comparison of the secondary structure elements for representative classes of GSTs is shown in table 1.

Despite the conservation of the overall fold, each class displays unique features about the active site and at the C-terminus (Wilce and Parker, 1994). For example, in the mu class there is an insertion of residues 36-42 in domain I (Ji *et al.*, 1992; Raghunathan *et al.*, 1994). These extra residues form the so-called "mu-loop" which serves as a lid to the active site shielding it from the bulk of the solvent. In the alpha class GST, an additional helix at the C-terminus (alpha helix 9) and a short beta strand at the end of the C-terminal segment are unique to this gene class (Sinning *et al.*, 1993) (see section 2.3.2.2).

2.2.1 Subunit structures

The homodimeric class alpha glutathione transferase A1-1 (hGST A1-1)[1] consists of 221 amino acid residues per subunit. Four three-dimensional structures of human class alpha GST are available in the PDB (Bernstein *et al.*, 1977). The class alpha structures available and their code accession numbers from the Brookhaven Protein Data Bank are: (1) in complex with S-benzylglutathione (pdb1guh.ent); (2) in complex with ethacrynic acid (pdb1gse.ent); (3) in complex with glutathione-ethacrynic acid (pdb1gsf.ent) and (4) apoenzyme form (pdb1gsd.ent).

1 The nomenclature used throughout this thesis is that suggested by Mannervik *et al.*, 1992. For example, the acronym hGST A1-1 denotes the human (h) class alpha (A) glutathione transferase (GST), which is composed of two identical type 1 subunits (1-1).

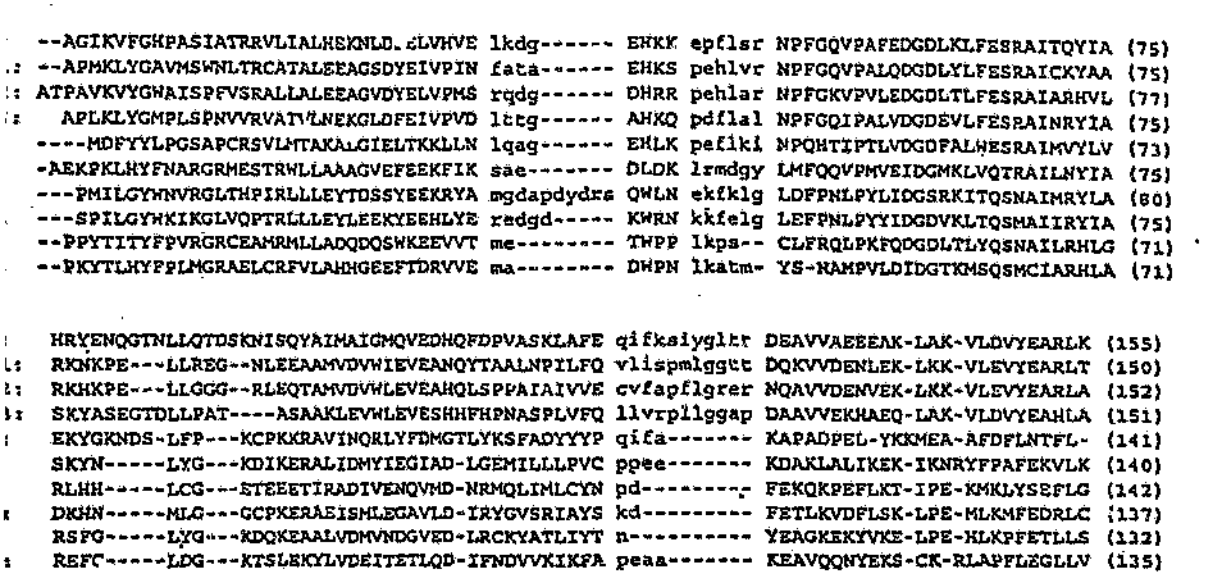


Figure 3. Sequence alignment of class theta *A. thaliana* (At; Reinemer *et al.*, 1996), class theta *L. cuprina* (Lc; Wilce *et al.*, 1995), class alpha (A; Sinning *et al.*, 1993), class mu (M; Ji *et al.*, 1992), *S. japonicum* (Sj; Lim *et al.*, 1994), class pi (P; Reinemer, 1991) and class sigma (Sq; Ji *et al.*, 1995) GSTs. Lower-case characters are used to identify sequences in regions with low structural similarity (Taken from Reinemer *et al.*, 1996).

Table 1. The secondary structure elements for representative classes of GSTs. Adapted from Reinemer *et al.* 1996.

Domain	structure element	Residue	Corresponding element in GST class or isoenzyme					
			alpha[a]	mu[b]	pi[c]	*S.japonicum*[d]	sigma[e]	*L.cuprina*[f]
I	β1	3-7	β1(5-8)	β1(2-7)	β1(3-7)	β1(2-6)	β1(3-7)	β1(2-5)
	α1	12-23	α1(17-26)	α1(14-22)	αA(15-23)	α1(14-22)	α1(15-22)	α1(10-21)
	β2	29-32	β2(28-35)	β2(27-32)	β2(29-32)	β2(27-32)	β2(29-32)	β2(27-30)
	$3_{10}'$	35-37						
	$3_{10}''$	39-41	α2(38-47)	α2/3_{10}(43-46,49-51)	αB/3_{10}(38-44)	α2(38-43)	α2(38-41)	α2(41-44)
	$3_{10}'''$	45-47						
	β3	56-59	β3(57-60)	β3(61-64)	β3(52-55)	β3(56-59)	β3(52-55)	β3(54-57)
	β4	62-65	β4(63-66)	β4(67-70)	β4(58-61)	β4(62-65)	β4(58-61)	β4(60-63)
	α2	67-77	α3(68-79)	α3(72-82)	αC(63-74)	α3(67-77)	α3(63-74)	α3(65-76)
II	α3'	93-108						
	α3''	111-121	α4(86-111)	α4(90-114)	αD(81-107)	α4(85-110)	α4(81-104)	α4(86-111)
	α3'''	123-126						
	α4	133-154	α5'(114-129) α5''(132-143)	α5a(119-127) α5b(130-141)	αE(109-132)	α5(113-136)	α5(112-135)	α5(124-141)
	α5'	167-170						
	α5''	173-179	α6(155-160)	α6(154-169)	αF(148-163)	α6(149-164)	α6(151-166)	α6(154-169)
	α6	185-189	--	---	3_{10}(167-170)	---	---	---
	α7	192-202	α7(179-190)	α7(187-189)	αG(172-182)	α7(173-184)	α7(175-186)	α7(177-189)
	α8	205-208	α8(192-198)	α8(191-197)	αH(185-192)	α8(186-192)	α8(188-196)	---

All structural analyses were calculated according to Kabsch and Sander (1983) and taken from GST structure analyses of (a) alpha (Sinning *et al.*, 1993), (b) mu (Ji *et al.*, 1992), (c) pi (Reinemer *et al.*, 1991, Dirr *et al.*, 1994a), (d) *S.japonicum* (Lim *et al.*, 1994), (e) squid (Ji *et al.*, 1995), (f) *L.cuprina* (Wilce *et al.*, 1995).

The individual subunits are composed of two spatially distinct domains. The smaller N-terminal alpha/beta domain, domain I (81 residues), has a βαβαββα folding topology. The overall topology of this domain resembles that observed for the bacteriophage T4 glutaredoxin (Eklund *et al.*, 1992) and related folds have been found in the *E.coli* Thioredoxin (Holmgren *et al.*, 1975) and the glutathione peroxidase (Epp *et al.*, 1983). The larger all alpha-helical C-terminal domain, domain II (140 residues) is composed of six amphipathic helices (figure 4). These helices pack tightly together to form a conse hydrophobic core. In the class alpha GST, domain II has an additional three-res eta sheet near the C-terminal segment and a larger alpha helix 5. In addition, the class alpha GST has a unique alpha helix 9 formed by the folding of fifteen residues of the C-terminal segment of the polypeptide chain. Domain I and II associate to form a subunit with interactions between the two domains being hydrophobic and hydrophilic.

2.2.2 Subunit-subunit interactions

The dimer of cytosolic GSTs is formed by an almost perfect two-fold axis (C2 symmetry)(figure 4). As a result of dimerization, 14% of the initial water accessible surface area becomes buried at the subunit interface (Dirr *et al.*, 1994b). All cytosolic GSTs have distinct quaternary structures and the molecular recognition at the subunit is class specific because dimerization involves only subunits from the same gene class. The interactions formed at the subunit interface are primarily between domain I (alpha helix 2) of the one subunit and domain II (alpha helices 3 and 4) of the other subunit. The interaction between the two subunits results in the formation of a solvent accessible V-shaped crevice (see section 2.2.3.3).

There are two typical types of subunit-subunit interactions observed. At the ends of the interface, the interactions are predominantly hydrophobic and have been described as a "lock-and-key" type interaction. This interaction is observed in the alpha/mu/pi/*S.japonicum* structures and is stabilised by the wedging of a hydrophobic

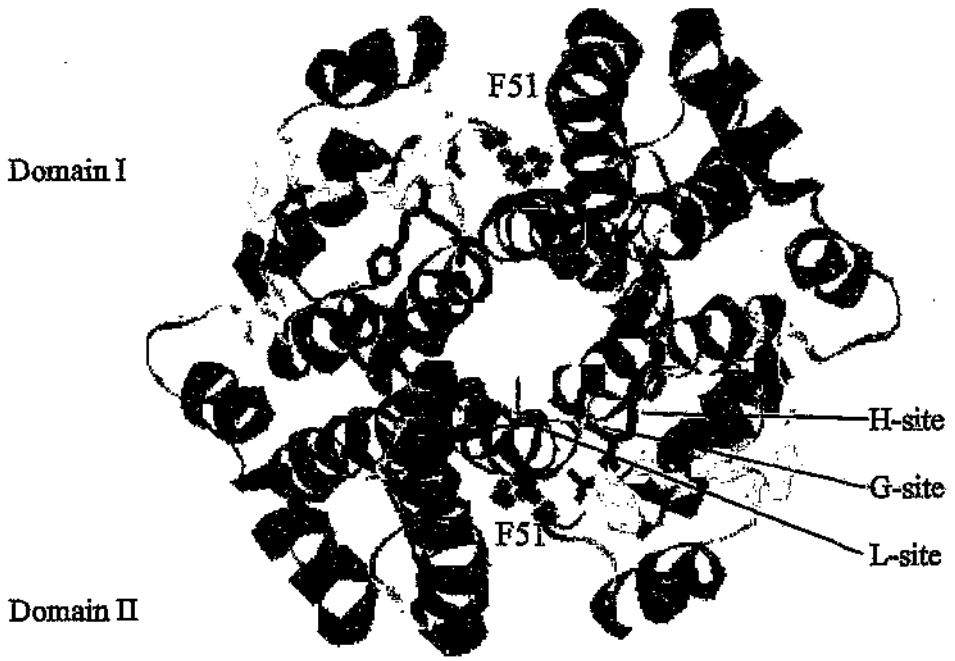


Figure 4. Cartoon representative of the homodimeric hGST A1-1 in complex with S-benzylglutathione (Sinning *et al.*, 1993). The view is parallel to the crystallographic two-fold axis. The H-site is occupied by the benzyl moiety, the G-site by glutathione and the L-site is unoccupied. The hydrophobic side chain of Phe51, which is a part of the "lock-and-key" type interaction, is shown. The figure was generated using Rasmol (Sayle, 1994).

side chain (Phe51, alpha; Phe56, mu; Phe47, pi; Phe51, *S.japonicum*) from one monomer into a hydrophobic pocket on the other side of the interface formed by the side chains of five residues belonging to the alpha helices 4 and 5 (Met93, Gly97, Ala134, Phe135 and Val138, alpha; Ile98, Gln102, Leu136, Tyr137 and Phe140, mu; Met89, Gly93, Pro126, Phe127 and Leu130, pi; Met93, Ala97, Met131, Phe132 and Arg135, *S. japonicum*). This interaction is absent in the theta and sigma classes with the result that the dimer interface is more hydrophilic. The absence of this interaction is compensated for by an increase in electrostatic interactions and is of evolutionary importance (section 2.1) and may impact on the stability (section 2.5). At the middle of the interface, there is a conserved intermolecular contact that results from the stacking of symmetrically equivalent arginine guanidino groups (Arg68, sigma; Arg69, alpha; Arg77, mu; Arg68, pi; Arg72, *S. japonicum*). The arginine residues are not conserved in the aligned sequences but a structural alignment indicates that they occupy topologically equivalent positions. The charges of the guanidino group are mitigated by the close proximity of a glutamate or asparate side chain. This type of interface has not been observed in the class theta structure (Wilce *et al.*, 1995). Recently, a mixed-charged cluster has been identified at the dimer interface of the mu class GST (Zhu and Karlin, 1996). The mixed-charged cluster which involves inter- and intrachain contacts may play an essential role in stabilising the quaternary structure of GSTs and may facilitate the folding process.

2.2.3 The active site

Cytosolic GSTs have two active sites per dimer, which behave independently of each other (Danielson and Mannervik, 1985). Each active site is composed of a specific binding site for reduced glutathione (GSH) (the G-site) and a less specific hydrophobic-electrophilic substrate-binding site (the H-site). The less specific H-site enables the accommodation of a variety of toxic agents. In all mammalian GSTs, the conserved core of domain I forms the structural framework for the active site.

2.2.3.1 The glutathione binding site (the G-site)

The G-site, as mentioned above, is specific for glutathione with the molecular recognition of the tripeptide (γ-glutamyl-cysteinyl-glycine) involving primarily polar interactions from domain I of one subunit and either no (class theta), one (class sigma/pi/mu/*S. japonicum*) or two (class alpha) amino acid residues from domain II of the other subunit (Dirr *et al.*, 1994b). The interactions with the tripeptide for the different gene classes are indicated in table 2 and the interactions between the tripeptide and protein moieties of the class alpha GST are indicated in figure 5. For class alpha GST, the G-site is made up of residues from domain I (the loop between the beta sheet1 and the alpha helix 1, the carboxyl-end of alpha helix 2, the end of the loop before the beta sheet 3 and the turn between the beta sheet 4 and the alpha helix 4) and residues Asp100 and Arg130 (in the middle of the alpha helices 4 and 5) from domain II of the other subunit. The G-site is lined with polar residues. The Pro55 residue in the *cis*-configuration maintains a functional active site (Wang *et al.*, 1993a). The tripeptide binds to the G-site in an extended conformation. The γ-glutamyl moiety of glutathione plays an important role in the glutathione binding (Adang *et al.*, 1989; 1990) and interacts with the enzyme in a similar manner (although the specific interactions differ) for all the gene classes (Dirr *et al.*, 1994b; Ji *et al.*, 1995). The γ-glutamyl arm points downward towards the dimer interface i.e., into the base of the pocket that is lined with polar residues. The interactions between the cysteinyl and glycine residues of glutathione are more variable between the gene classes. In the class mu GST, the conformation of glutathione is different which leads to an interaction of the carbonyl oxygen of the cysteinyl residue with the indole side chain of Trp7 (Ji *et al.*, 1992). A similar binding mode has been observed with the *S. japonicum* GST (Lim *et al.*, 1994). In the class sigma GST, the carbonyl oxygen atom interacts with the backbone of Met50 resulting in a different conformation (Ji *et al.*, 1995). This binding mode is similar to that observed for the class alpha and pi GSTs. At the glycine end of glutathione, the interactions are even more different. The glycyl moiety points away from the core of the N-terminal domain of the enzyme towards the solvent. Mutagenesis studies with a truncated tripeptide or with an analogue (L-

Table 2. Interactions for the different glutathione transferase gene classes with glutathione (adapted from Reinemer *et al.*, 1996)

Glutathione moieties		Alpha[a]	Mu[b]	Pi[c]	Sj26[d]	Sigma[e]	Theta (*L.c*)[f]	Theta (*A.t*)[g]
γ-Glu	NH_3^+	Glu66	Gln71	Gln62	Gln66	Gln62	Glu64	Glu66
		Asp100*	Asp105*	Asp96*	Asp100*	Asp96*		
	COO^-	Thr67	Ser72	Arg13	Thr67	Gln62	Ser65	Glu66"
				Ser63		Ser63	Arg66	Ser67
								Arg68
	C=O	-----	-----	Gln49	-----	-----	-----	-----
Cys	NH	Val55	Leu59	Leu50	Leu54	Met50	Ile52	Val54
	C=O	Val55	Trp7	Leu50	Trp7	Met50	Ile52	Val54
Gly	NH	-----	Asn58	-------	Asn53	----	----	------
	COO^-	Arg44	Arg42	Trp38	Lys44	Trp38	His38	His40"
		Arg130*	Trp45	Lys42		Lys42	His50	Lys41
			Lys49	Gln49		Asn48		Gln5[illegible]

(a) Sinning *et al.*, 1993; (b) Ji *et al.*, 1992; Raghunathan *et al.*, 1994 (c) Reinemer *et al.*, 1991; 1992; Dirr *et al.*, 1994a; Garcia-Sáez *et al.*, 1994; (d) Lim *et al.*, 1994; McTigue *et al.*, 1995 (e) Ji *et al.*, 1995 (f) Wilce *et al.*, 1995 (g) Reinemer *et al.*, 1992. * Interactions from the neighbouring subunit
"Interaction mediated by water

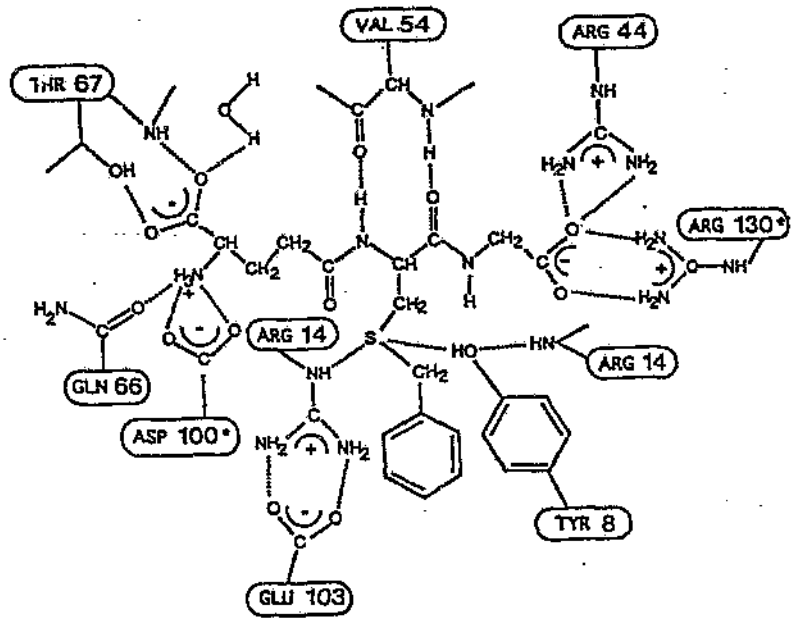


Figure 5. A schematic representation of the residues of hGST A1-1 which interact via H-bonds and salt linkages with the ligand, S-benzylglutathione. The two residues marked with an asterisk are from the two-fold related molecule. Adapted from Sinning *et al.* 1993.

Glu-L-Cys) have indicated this moiety to be the least restrictive (Adang *et al.*, 1989; 1990; Widersten *et al.*, 1996). In the class alpha, the removal of the glycyl moiety from the tripeptide resulted in a greater affinity for the electrophilic substrate suggesting that this moiety may influence the topology of the active site and assist the binding of the second substrate of the enzyme (Widersten *et al.*, 1996).

2.2.3.2 The hydrophobic binding site (the H-site)

The co-crystallisation of many representative GST gene classes with a substrate-product analogue such as S-hexylglutathione (Reinemer *et al.*, 1992, 1996; Wilce *et al.*, 1995) or S-benzylglutathione (Sinning *et al.*, 1993) and S-(-p-nitrobenzyl)glutathione (Garcia-Sáez *et al.*, 1994), transition state analogues (S-2,4-dinitrophenyl) glutathione (Ji *et al.*, 1993; 1995) as well as those crystallised in the presence of ethacrynic acid and the glutathione-ethacrynic acid complex (Cameron *et al.*, 1995; Oakley *et al.*, 1997), has enabled the identification of the hydrophobic-electrophilic-substrate (xenobiotic) binding site. In these three-dimensional representatives the tripeptide, glutathione, adopts its usual extended conformation and the moiety covalently attached to the sulphur atom is located in the H-site (figure 4).

The molecular details on the structure of the H-site varies between the gene classes, however, despite this there are clusters of apolar side chains that provide a highly hydrophobic protein surface that in the absence of xenobiotic substrates or product is accessible to solvent. The different topologies of the H-site enable the accommodation of a wide variety of hydrophobic-electrophilic substrates of different size and polarity. The H-site is formed by the residues of the C-terminal segment of the polypeptide chain, the loop between the beta sheet 1 and the alpha helix 1 and the C-terminal end of the alpha helix 4. The C-terminal region of the polypeptide chain in the class alpha GST has through photoaffinity labelling studies (Hoesch and Boyer, 1989), mutagenesis studies (Board and Mannervik, 1991) and crystallographic evidence (Sinning *et al.*, 1993; Cameron *et al.*, 1995) been shown to be an integral part of the H-site. The C-terminus folds into an additional alpha helix 9 and this

region displays conformational flexibility in the absence/presence of ligand (Cameron *et al.*, 1995; Atkins *et al.*, 1997). In the presence of ligand, the alpha helix 9 forms a "lid" over the H-site, blocking the active site from solvent i.e., rendering it more hydrophobic. It is not essential for the binding of glutathione but deletion of this segment does impair catalytic activity (Board and Mannervik, 1991).

The crystallisation of human class alpha GST with ethacrynic acid (EA) and the ethacrynic acid-GSH (EA-GSH) complex (Cameron *et al.*, 1995) indicates that EA binds to the H-site in a similar position to the benzyl- or hexyl-moieties of S-benzylglutathione or S-hexylglutathione (Reinemer *et al.*, 1992; Sinning *et al.*, 1993), respectively. In class alpha, the dichlorophenoxy ring of EA makes hydrophobic contacts with the alpha carbon of Gly13 and the side chains of residues Phe9, Leu106, Leu107, Val110, Met207 and Phe221. The ethacrynic acid assumes an orientation at the H-site so that the carboxylate group forms hydrogen bonds with Tyr8 and possibly Arg14 (Cameron *et al.*, 1995). A comparison of the EA and the EA-GSH complexes show that the position that EA occupies varies. In the unconjugated complex, the ethacrynic acid has rotated moving the whole molecule closer to the G-site. The position of the unconjugated ethacrynic acid, regardless of its orientation means that the substrate is bound in a non-productive manner. It has to move in order for the glutathione to bind. A similar mode of ethacrynic acid binding has been observed for the class pi (Oakley *et al.*, 1997).

2.2.3.3 The non-substrate ligand binding site (the L-site)

The dimerization of the individual subunits of all GSTs results in the formation of a cleft at the subunit interface. The cleft (figure 4) found in the centre of the dimer is open to the active site and to solvent and has similar features for all GST classes. However, there are distinct differences in the topology of the cleft that may explain the wide range of specificity's and affinities that exist for the different anionic ligands (Danielson and Mannervik, 1985). This site has been implicated as the binding or

transport site for the anionic hydrophobic ligands, 8-anilino-1-naphthalene sulphonate (ANS) and sulphobromophthalein (BSP).

There is very little structural data available about the location and nature of this site. Some detail on the location has been provided by the crystal structure of the *Schistosoma japonicum* enzyme in complex with the leading anti-schistosomal drug, praziquantel (McTigue *et al.*, 1995) and the crystal structure of the class sigma GST in complex with S-(3-iodobenzyl)glutathione (Ji *et al.*, 1996). In the *Schistosoma japonicum* GST structure, one molecule of the drug was found bound per protein dimer, in the dimer groove that adjoins the two catalytic sites (McTigue *et al.*, 1995). In the cephalopod class sigma GST, the dimeric enzyme binds three molecules of S-(3-iodobenzyl)glutathione (Ji *et al.*, 1996). One molecule is associated with each active site and the other is bound as the cleft between the two subunits.

In mammalian GSTs, the dimer interface near the two-fold axis was proposed using fluorescence resonance energy transfer, as the binding site for the hydrophobic anionic ligand, ANS (Sluis-Cremer *et al.*, 1996). In addition, the binding of aflatoxin B1 has also been implicated to be at this site (Sluis-Cremer and Dirr, unpublished results). It is not clear at this stage if this is the only binding site. The non-substrate ligand binding site has also been identified using 3β-(Iodoacetoxy)dehydroisoandrosterone (3β-IDA) as an affinity label (Barycki and Colman, 1997).

2.3 The catalytic mechanism

The catalytic mechanism of GSTs involves the nucleophilic addition of glutathione to a wide variety of electrophilic compounds (most commonly the aryl chloride, 1-chloro-2,4-dinitrobenzene). The addition of the electrophile to the nucleophilic species in the active site (i.e., the highly reactive thiolate anion), occurs sequentially involving a ternary complex of the enzyme, glutathione and the electrophile

(Armstrong, 1991). The sequential catalytic mechanism may be represented schematically as follows:

$$E + GSH \leftrightarrow E{\bullet}GS^{-} + H^{+} \overset{R\text{-}X}{\leftrightarrow} E{\bullet}GS^{-}{\bullet}R\text{-}X \leftrightarrow E{\bullet}GSR{\bullet}X^{-} \leftrightarrow E + GSR + X^{-}$$

where E is the enzyme (i.e., GST), R-X is the hydrophobic-electrophilic substrate and GSH is the nucleophilic physiological substrate. Glutathione and substrate are added in the catalytic reaction in a random order. However, because of the high concentration of GSH in the cell (1-10mM) it is likely that the addition of glutathione will be first.

A central aspect of the catalytic mechanism involves the interaction of the protein with the sulphur of the tripeptide and an evolutionary conserved tyrosine residue at the G-site. This tyrosine residue is located at or near the C-terminus of the beta strand 1 and is followed by the conserved loop connecting this strand and the alpha helix 1. This loop forms an essential component of the active site and for the reactivity of the tyrosine side chain (Dirr *et al.*, 1994b). In classes sigma/alpha/mu/pi/ *S.japonicum*, the hydroxyl group of the tyrosine residue (Tyr7, sigma; Tyr8, alpha; Tyr6, mu; Tyr7, pi; Tyr6, *S.japonicum*) acts as a hydrogen bond donor to the sulphur atom of GSH. In doing this, the pKa of the thiol in the enzyme-GSH complex is lowered and therefore it exists as an ionised form at physiological pH.

The role of this tyrosine residue in the catalytic mechanism has been exploited by site-directed mutagenesis studies in the class mu (Liu *et al.*, 1992), the class alpha (Stenberg *et al.*, 1991b; Wang *et al.*, 1992a; Dietze *et al.*, 1996) and the class pi (Manoharan *et al.*, 1992; Kolm *et al.*, 1992; Kong *et al.*, 1992). In all these studies, the enzymatic activity was reduced although there was little change in the affinity for the tripeptide. The tyrosine residue is implicated to play a role in increasing the nucleophilicity of the thiol group of glutathione. The hydroxyl group of the tyrosine residue acts as a hydrogen bond donor promoting the formation of the thiolate anion

by decreasing the pKa of the glutathione, which provides the electrophilic stabilisation of the thiolate anion.

Another role of the tyrosine residue has been suggested using the class alpha and pi enzymes (Atkins *et al.*, 1993; Karshikoff *et al.*, 1993; Meyer *et al.*, 1993). In this role, the tyrosine residue acts as a base i.e., a proton donor. In class alpha and pi, the pKa of the tyrosyl hydroxyl group is extremely low and the tyrosinate anion acts as a general base and therefore removes a proton from the GSH and creates the thiolate anion. A recent study in which all fourteen tyrosyl residues of the mu class GST were replaced with 3-fluorotyrosine (Parsons and Armstrong, 1996) indicated that the hydroxyl group of the tyrosyl residue stabilises the thiolate anion of bound GSH prior to nucleophilic attack. This study, therefore, argues against the general base catalytic model. In addition, theoretical studies using molecular dynamics simulations and quantum mechanics have modelled the reaction mechanism of the class pi GST and suggest that the tyrosine residue in the unionised form contributes to the catalytic process of GSTs (Orozco *et al.*, 1997).

In the class theta (*L. cuprina*) enzyme the tyrosine residue (Tyr5), in the same position as the evolutionary conserved tyrosine, has through mutagenesis studies been ruled out of playing a role in the catalytic mechanism. The tyrosine residue is too far from the active site (13.9Å from the sulphur atom) (Wilce *et al.*, 1995). The hydroxyl group of a serine residue (Ser9) has through mutagenesis studies (Board *et al.*, 1995) been shown to be essential in the catalytic mechanism of the theta class GSTs. The serine residue is 3.9Å away from the sulphur atom of glutathione and its hydroxyl group superimposes with that of the evolutionary conserved tyrosine residue. In addition, another residue (Tyr113) may influence the reactivity of glutathione by hydrogen bonding through a water molecule at the active site (Board *et al.*, 1995). Similarly, in the *A. thaliana* enzyme (Reinemer *et al.*, 1996) the structurally corresponding residue is a glycine (Gly7) which obviously has no role in the catalysis but a serine residue (Ser11) may take over the role of the tyrosine.

2.4 Mutagenesis studies of the class alpha glutathione transferase

Site-directed mutagenesis studies have been used to study the contribution of amino acids to the catalytic mechanism, substrate binding and stability of the human class alpha GST (hGST A1-1) and the rat class alpha GST (rat YaYa). A summary of the residues mutated, their location within the polypeptide chain, their possible function and the result of the mutation are tabulated (table 3). Mutagenesis studies have primarily focussed on those residues thought to be involved in the catalytic mechanism.

2.5 Conformational stability of the glutathione transferases

The equilibrium unfolding and dissociation of the class pi (Dirr and Reinemer, 1991; Erhardt and Dirr, 1995) and the *Schistosoma japonicum* (Kaplan *et al.*, 1997) GSTs are consistent with a highly co-operative two-state transition involving the native dimer and the unfolded monomer. The energy of stabilisation of the native three-dimensional state is approximately 25kcal/mol. In contrast, conformational stability studies of the class mu and class sigma (unpublished results) and the theta (Sacchetta *et al.*, 1993) GSTs suggest a multi-state equilibrium unfolding pathway. The lower conformational stability for the class theta and sigma GSTs may be explained in part by the differences observed at the subunit-subunit interface (see section 2.2.2).

The conformational stability of the class pi GST has also been investigated in the presence of G-site ligands (Erhardt and Dirr, 1996) and have indicated that glutathione destabilises the stability of the enzyme in contrast the glutathione analogues, S-hexylglutathione and glutathione sulphonate, have a stabilising effect on the enzyme. The conformational stability of the class alpha, pi and *S.japonicum* GST has been determined in the presence of non-substrate ligands (Ohlmeyer; MSc dissertation, 1997; Kaplan; PhD thesis, 1997) and G-site ligands (Parsons; MSc dissertation, 1997). The non-substrate ligands (ANS, BSP, bilirubin and the drug, praziquantel) have little or no significant effect on the conformational stability.

2.6 Objectives

A pursuit to understand the relationship between the intrinsically complex structure of the homodimeric human class alpha glutathione transferase and its biological function in cellular detoxification and transport of non-substrate ligands formed the basis of this study. The purpose was to characterise its complete folding/unfolding pathway using equilibrium and kinetic unfolding/refolding studies. In addition, kinetic unfolding/refolding studies were performed to identify and characterise the transition states between the native and unfolded states. The contribution of a topologically conserved hydrophobic residue, in the domain II of the glutathione transferases (Leu164, alpha; Leu157, pi; Leu163, mu; Leu160, sigma; Leu158, *S. japonicum*; Val163, theta), to the overall stability was characterised, using the class alpha GST, by means of equilibrium and kinetic unfolding studies. This study will facilitate the understanding of the molecular mechanism of subunit folding and the assembly of dimers for the glutathione transferases.

Table 3. Site-directed mutagenesis studies for the class alpha glutathione transferase

Residue	Location; possible function; effect of mutation	Reference
His7	β1; possible catalytic role; conserved only in the α class; mutant less stable, ligand binding unchanged, does not affect catalytic activity, nonessential role in catalysis	Wang *et al.*, 1991; 1992b; Björnestedt *et al.*, 1995
Tyr8	β1-active site; conserved in all classes except theta; catalytic role deprotonates GSH and stabilises the thiolate anion; mutation hydroxyl group abolished activity; binding of GSH unaltered	Stenberg *et al.*, 1991b; Wang *et al.*, 1992a; Dietze *et al.*, 1996
Asn10	between β1 and α1-only in the alpha class; possible role in catalysis; no change in catalytic parameters	Björnestedt *et al.*, 1995
Ala11	between β1 and α1-only in the alpha class; possible role in catalysis; no change in catalytic parameters	Björnestedt *et al.*, 1995
Arg12	β1-α1 link, conserved $\alpha/\mu/\pi$, stabilises the active site loop, interacts with domain II; mutation decreased yield, specific activity and affinity for GSH	Stenberg *et al.*, 1991a; Wang *et al.*, 1993a
Arg14	between β1-α1, conserved α/π, amide group is within H-bonding distance to the sulphur atom of GSH - possible functional role, structurally NH backbone interacts with OH of Tyr8, $-CH_2-$ groups located in the H- site, guandinium group salt bridge with Glu103; decreased catalytic activity	Wang *et al.*, 1993a; Björnestedt *et al.*, 1995

Cys17	α1-rat alpha class only; role in catalysis; not essential for catalysis	Wang *et al.*, 1992b
Arg19	α1-conserved α/μ/π; buried with salt bridge to Glu31; mutation to Ala -decreased specific activity; decreased affinity for GSH; increases susceptibility to inhibitor (S-hexylGSH; BSP); mutation to Lys; Ile no effect	Stenberg *et al.*, 1991a; Wang *et al*, 1993a
Trp20	α1-present only in alpha class; possible role in catalysis, GSH/S-hexylGSH/heme binding; mutation to Phe no effect on catalysis or ligand binding	Wang *et al.*, 1992b; Atkins *et al.*, 1993
Lys44	α2-rat α class/μ/π; possible role in GSH binding; mutation to Arg/Ile indicated it is not an essential residue for rat alpha class	Wang *et al.*, 1993a
Gln53	between α2 and β3-conserved in α/μ/π; role in GSH binding; increased affinity for GSH	Wang *et al.*, 1993a
Pro55	precedes β3-conserved all classes; *cis* configuration, role at GSH binding site to maintain functional site; mutation abolished catalytic activity, decreased affinity to S-hexylGSH; less stable	Wang *et al.*, 1993a
A[illegible]	between β3 and β4-not conserved; possible role catalysis; mutation no effect on catalytic activity	Wang *et al.*, 1992a
Gln66	bend between β4 and α3-conserved in all classes except θ; carboxyl group H-bonds with amide group of γ-glutamyl of GSH; Interacts with Asp100 of domain II via H-bond; decreased affinity for GSH and catalytic activity	Wang *et al.*, 1993a

Thr67	α3-beginning of TR(SN)AIL sequence, conservatively replaced with Ser (other classes), interacts with the α-carboxyl group of glutamyl residue of GSH; binding for electrophiles increased, possible role in stabilisation of catalytic intermediate	Widersten *et al.*, 1996
Arg68	α3-part of the TRAIL sequence in α class; possible role in GSH binding, stabilisation of the subunit-subunit interface ; decreased conjugating activity, increased sensitivity towards BSP/S-hexylGSH	Stenberg *et al.*, 1991a, Wang *et al.*, 1993a
Asp92	α4-domain II; no role in catalysis or ligand binding	Wang *et al.*, 1992a
Asp100	α4-domain II; involved in GSH binding via a salt bridge with the amino group of γ-glutamyl arm of GSH. Conserved in all classes except θ; negative charge on Asp essential for binding GSH, no role in heme binding	Wang *et al.*, 1992a
Cys111	between α4 and α5, alpha class only; not essential for catalysis or the conformation of the protein	Wang *et al.*, 1992b
Tyr131	α5-alpha class only; interacts with water molecule in active site; marginal increase in activity, σ-complex between GSH and TNB less stable	Widersten *et al.*, 1996
Asp156	α6-part of hydrophobic core, hydrogen bonds to the amide nitrogen of Ser153/Leu146/ as well as the side chain of Ser153; possible structural role; mutation to Asn abolished enzyme activity, further role not pursued	Wang *et al.*, 1992a
Met208	β5-hydrophobic residue which precedes the C-terminal helix; in alpha class it interacts with the substrate occupying the H-site; randomly mutated, M208W increases stability of the Meisenheimer complex, M208Y and M208W decreased affinity for S-(2,4-	Widersten *et al.*, 1994

del209-221	dinitrobenzyl)GSH; role in aromatic substitution reactions by interacting with the delocalized negative charge of the aromatic ring in the transition state complex α9-unique to α class(C- terminus), extra helix shows conformational flexibility, in presence of ligand forms a part of the H-site; truncation diminished catalytic activity, but no effect on GSH binding	Board and Mannervik, 1991

CHAPTER 3
EXPERIMENTAL PROCEDURFS

3.1 Materials

Ultrapure urea was purchased from ICN Biomedicals, Inc. ANS, NATA, N-acetyl-L-tyrosinamide, IAEDANS, DCIP, L-ascorbic acid, CDNB and spectroscopic grade sucrose were purchased from Sigma. GSH and glutathione sulphonate were purchased from Boehringer Mannheim. S-Hexylglutathione was prepared according to the method of Vince *et al.* (1971). The restriction enzymes (Sac I, Pst I, Pvu I) and the DNA sequencing kit (Sequenase v2.0) were purchased from Amersham Life Science International. The ExSite mutagenesis kit was purchased from Stratagene. All other reagents were of an analytical grade. The plasmid, pKHA1 was a kind gift from B.Mannervik (Stenberg *et al.*, 1992).

3.2 Construction of L164A hGST A1-1

3.2.1 Oligonucleotide primer design

Oligonucleotide primers for PCR mutagenesis were designed using the computer package Gene Runner, v3.04. The primers were designed according to the nucleotide sequence encoding the human glutathione transferase A1-1 from human Hepatoma HepG2 cells (Stenberg *et al.*, 1992). The primers were designed according to the manufacturer's recommendations. The mutagenesis oligonucleotide primers used to create the L164A hGST A1-1 have the following sequences: L164AFP 5'-TAC TAC GTC GAG GAG CTT G-3' and L164ARP 5'-*TGC* GAG CTC CAC CAG ATG AAT GTC-3'. Both primers annealed to different strands of the template and the 5' end of the L164AFP primer was phosphorylated. In primer L164ARP, the underlined nucleotides represent the translationally silent mutation, which incorporated a unique restriction site, Sac I, for screening for the mutant plasmid and the italicised nucleotides represent the mutation that generates the Leu to Ala substitution.

3.2.2 PCR mutagenesis

Site-directed mutagenesis was performed by combining inverse PCR with long PCR to introduce the mutation directly into an intact plasmid template (Weiner *et al.*, 1994; ExSite mutagenesis kit). The kit enabled site directed mutagenesis to be performed by PCR using double stranded DNA. The construct pKHA1 (Stenberg *et al.*, 1992) encoding the hGST A1-1 was used as the template DNA. PCR reactions were performed in a total volume of 25µl. The reaction mixture contained 0.7pmol of template DNA, 15pmol of each mutagenesis primer, 1.04mM of each dNTP in buffer. The buffer was 200mM Tris-HCl pH 8.75, containing 100mM KCl, 100mM $(NH_4)_2SO_4$, 20mM $MgSO_4$, 1%(v/v) Triton X-100 and 1mg/ml BSA. After a 4 minute hot start at 94°C *Taq* DNA polymerase (5U/µl) and *Taq* extender PCR additive (5U/µl) were added. The PCR product was generated through twelve amplification cycles of 100 seconds at 94°C to denature the DNA, 160 seconds at 52°C to anneal the primers and 150 seconds at 72°C for DNA extension. The PCR product was analysed on a 1% (w/v) agarose gel. The linear PCR reaction mixture was treated using DpnI (10U/µl) to digest parental template DNA and the linear PCR product was polished using Pfu DNA polymerase (4U/µl). The ligated PCR product was transformed into *E.coli* XLI-Blue supercompetent cells.

The transformed cells were screened for mutant DNA using restriction analysis. An overnight culture of a selected colony was grown in 1.5mL LB media (10g tryptone, 10g NaCl and 5g yeast extract) at 37°C. The culture was centrifuged (13 000rpm, 5 minutes) and the pellet was resuspended in 100µl ice-cold 25mM Tris-HCl, pH 8.0, containing 10mM EDTA and 50mM glucose. The suspension, after incubation at room temperature for 5 minutes, was treated with 200µl 0.2M NaOH and 1% SDS to denature the DNA. The denatured DNA was incubated on ice and treated with 150µl ice-cold 5M potassium acetate. The suspension was centrifuged at 14 000rpm for 10 minutes at 4°C. The supernatant was removed, treated with an equal volume of isopropanol and centrifuged at 14 000rpm for 15 minutes at 4°C. The pellet was

washed twice with 70% ice-cold ethanol and air dried. The DNA was resuspended in 20µl of sterile milli-Q water. 1µg of the DNA was digested using Sac I (12U/µl) at 37°C for at least 1 hour. The restricted DNA was analysed on a 1% (w/v) agarose gel containing ethidium bromide and the size of the fragments was compared with standard DNA markers.

3.2.3 Sequencing of mutant DNA using the dideoxy-chain termination method

The entire cDNA encoding the L164A hGST A1-1 was sequenced to ensure that no other mutations were introduced into the nucleotide sequence during PCR. DNA sequencing was performed by the dideoxy-chain termination method using a modified T7 DNA polymerase (Sequenase version 2.0; Sanger *et al.*, 1997; Tabor and Richardson, 1987). Briefly, the purified double stranded mutant DNA was denatured by alkaline denaturation: 2M NaOH was added to 2pmol of plasmid DNA. The DNA was precipitated using 3M sodium acetate, pH 5.2, and washed using 70% ethanol. The pellet was air dried and annealed to primers designed for sequencing. The primers designed for sequencing were SEQ1: 5' CCC GTA TGT CCA CCT GAG 3' and L164NO3: 5' TTT CCC ATA GAG GTT GTA 3'. The primer L164ARP (section 3.2.1) was also used to sequence the mutant DNA. 0.3-5pmol of plasmid DNA was annealed to 0.5pmol primer in the 5X Sequenase buffer (200mM Tris-HCl, pH 7.5, containing 100mM $MgCl_2$ and 250mM NaCl) at 37°C for 30 minutes. To each tube containing the DNA and annealed primer, 5µl of a labelling mix containing 0.1M DTT, a 1:10 dilution of dGTP labelling mix, enzyme dilution buffer (10mM Tris-HCl, pH 7.5, containing 5mM DTT and 0.5mg/ml BSA), Sequenase (3.25U/reaction) and [α-^{35}S]dATP(5µCi/reaction) was prepared. The mix was incubated at 15-18°C for 4 minutes. The sequencing reactions were set with four tubes (A, C, G, or T) per template to be sequenced. The labelling reaction was terminated with the appropriate termination mix e.g., ddGTP for the G nucleotide etc. by incubation at 45°C for 5 minutes. Finally, the reaction was stopped by addition of a stop solution (95% formamide, 20mM EDTA, 0.05% bromophenol blue and 0.05% xylene cyanol FF).

The reaction products, heated to 80°C prior to loading, were resolved on a 6% denaturing polyacrylamide gel for 2hours (short run), 4hours (medium run) and 6 hours (long run). The gel was dried and exposed to X-ray film for 72hours. The autoradiogram was developed, fixed in acetic acid and rinsed with water before analysis.

3.3 Expression and protein purification

Recombinant hGST A1-1 was expressed in *E.coli* JM 103 cells and the L164A hGST A1-1 was expressed in *E.coli* XL1 blue cells at 37°C. The wild type and mutant enzymes were purified using S-hexylglutathione affinity chromatography (Stenberg *et al.*, 1992). The protein was eluted from the affinity matrix using either 1mM S-hexylglutathione in 20mM Tris-HCl, pH 7.8, containing 200mM NaCl and 1mM EDTA or 50mM glycine-NaOH, pH 10 (Cameron *et al.*, 1995). The elution of the protein using a high pH was used to prepare protein in the apoenzyme (ligand-free) form (Cameron *et al.*, 1995). Wild type and mutant enzyme were stored in 20mM sodium phosphate buffer, pH 6.5, containing 0.1M NaCl and 1mM EDTA.

The homogeneity of the protein was assessed using SDS-PAGE (Laemmli, 1970) and SEC-HPLC. Wild type and mutant protein electrophoresed as a single band, on a 15% polyacrylamide gel, with an apparent molecular mass of 27kDa and eluted as a single symmetrical peak of molecular mass approximately 55kDa using SEC-HPLC.

The protein concentration of water-soluble globular proteins can be calculated using the average values of the extinction coefficient of tryptophan ($5550 M^{-1}cm^{-1}$), tyrosine ($1340 M^{-1}cm^{-1}$) and cysteine ($150 M^{-1}cm^{-1}$) (Mach *et al.*, 1995). Protein concentrations calculated in this manner have typically less than 2% error (Mach *et al.*, 1995). The protein concentration for dimeric wild-type and mutant enzyme were estimated using a molar extinction coefficient of 38 $200 M^{-1}cm^{-1}$ at 280nm calculated by the method described by Perkins (1986):

$$\varepsilon = 5550\Sigma Trp + 1340\Sigma Tyr + 150\Sigma Cys \quad (1)$$

where the coefficient (ε) for the specific residue is multiplied by the number of amino acids in the protein sequence. The concentration of the protein was determined using the absorbance measurement at 280nm and the Beer-Lambert law, $C=A_{280}/\varepsilon l$, where l is the pathlength of the light through the cuvette (cm).

3.4 Alkylation of hGST A1-1

The single cysteine residue (Cys111) of wild-type hGST A1-1 was chemically modified using 5-[[2-[(iodoacetyl)amino)ethyl]amino]-naphthalene-1-sulphonic acid (IAEDANS), a fluorescent sulphydryl reagent that reacts readily with thiol compounds and sulphydryl groups (Hudson and Weber, 1973). The reagent contains an iodoacetamide group that reacts readily with the sulphydryl group and in doing this the naphthalene sulphonic acid group is added that has fluorescence spectroscopic properties. The chemical modification of hGST A1-1 using the fluorescent properties of the AEDANS group provided an additional spectroscopic probe for studying the unfolding properties of the protein. hGST A1-1 was alkylated overnight with a 100-fold molar excess of IAEDANS, at room temperature in 50mM Tris-HCl, pH 7.5. Unreacted IAEDANS was separated from modified protein using gel filtration (Sephadex-G25). The stoichiometry of labelling was determined spectrophotometrically at 338nm using an extinction coefficient of $6000M^{-1}cm^{-1}$ for AEDANS (Hudson and Weber, 1973). The alkylated enzyme was assessed using specific activity, fluorescence spectroscopy and SEC-HPLC and found to be identical to native unmodified enzyme.

3.5 Steady-state enzyme kinetic properties

3.5.1 General CDNB-conjugating assay

The steady-state enzyme kinetics for wild type and mutant enzyme were performed spectrophotometrically at 340nm in 0.1M potassium phosphate, 1mM EDTA, pH 6.5 (Habig and Jakoby, 1981). The standard assay contained a final concentration of

1mM reduced glutathione (GSH) and 1mM 1-chloro-2,4-dinitrobenzene (CDNB) in 3%(v/v) ethanol. All reactions were followed for 60 seconds at room temperature (typically 21°C) and the standard error between assays was about 10%. The assays were performed on a Hewlett Packard model 8452A-diode array spectrophotometer interfaced with a Vectra CS computer. The specific activities were determined using identical protein concentrations (determined as described above) and the extinction coefficient of $9600M^{-1}cm^{-1}$ was used to determine the amount of product (1-(S-glutathionyl)-2,4-dinitrobenzene) formed. The non-enzymatic rate was subtracted from the assay-containing enzyme.

Enzyme activity measurements of hGST A1-1, in the presence of denaturant (0-8M urea) were assessed as described above. The final concentration of protein was 3nM and the residual denaturant (up to 27mM urea) had no effect on the activity of the enzyme. Conditions for the assay were chosen such that linear progress curves were observed. Less than 10% reactivation of denatured protein was observed during the assay.

3.5.2 Kinetic parameters (V_m, K_m, k_{cat}/K_m) for wild-type and mutant

The enzymes were assayed in the presence of 0-6mM GSH and 1.6mM CDNB to determine the Michaelis-Menten constant (K_m) towards GSH. Similarly, the Michaelis-Menten constant (K_m) towards CDNB was determined in the presence of 0-1.6mM CDNB and 2mM GSH. The solubility limit of CDNB (in ethanol) in aqueous solutions was assessed by monitoring the absorbance at 600nm and was estimated to be between 1.6mM and 2mM. The non-enzymatic background without enzyme was subtracted from the assay-containing enzyme. The catalytic efficiency (k_{cat}/K_m) values for GSH and CDNB were calculated from the slope of a linear plot of velocity versus substrate concentration at final concentrations between 0.1-1mM GSH and 0.05-0.15mM CDNB:

$$v = (k_{cat}/K_m)[E]_t[S] \qquad (2)$$

3.5.3 Inhibition of the CDNB-conjugating activity

The inhibition of the CDNB-conjugating activity for wild-type and mutant were assessed using the standard enzymatic assay conditions described above. The glutathione analogues, S-hexylglutathione (0-70μM) and glutathione sulphonate (0-100μM) and 8-anilino-1-naphthalene sulphonate (ANS)(0-100μM) were added to 7nM of enzyme. The activity of the enzyme in the presence of the inhibitors was measured by the addition of 1mM re ,d GSH and 1mM CDNB.

3.6 Thermal-inactivation

1μM of wild-type and mutant enzyme in 20mM sodium phosphate buffer, pH 6.5, containing 0.1M NaCl and 1mM EDTA were incubated for 30 minutes at the required temperature (10°-80°C). The remaining activity was assayed as described in section 3.5.1. At each temperature, aggregation was monitored using fluorescence spectroscopy where the wavelengths of excitation and emission were identical. The thermal inactivation was irreversible even after prolonged incubation (24 hours) at room temperature.

3.7. Spectroscopic studies

The spectroscopic properties of 5μM wild type and 5μM mutant protein were studied. The intrinsic fluorescence emission spectra of tryptophan alone (excitation at 295nm) and tyrosine and tryptophan (excitation at 280nm) were measured with a Hitachi model 850-fluorescence spectrofluorimeter (see section 3.9.1.1). Second-derivative analyses of the UV spectra (250nm to 320nm) were used to compare the extent of tyrosine exposure for each protein (see section 3.9.1.2). A comparison of the second-derivative spectra for wild type and mutant protein enabled the differences in the extent of tyrosine exposure to be determined (Ragone *et al.*, 1984).

3.8 Ligand binding assays

The affinity of wild type and mutant hGST A1-1 for 8-anilino-1-naphthalene sulphonate (ANS) was determined by measuring the quenching of the intrinsic

tryptophan fluorescence. All fluorescence measurements were performed as detailed in section 3.9.1.1. 1μM of enzyme in sodium phosphate buffer, pH 6.5, was titrated with ANS to a final concentration of 40μM. The final dilution factor did not exceed 10% of the initial volume.

The correction for inner-filter effects was performed according to Birdsall *et al.* (1983). The absorbencies of the ligand were measured at the excitation (A_{ex}; 295nm) and emission (A_{em}; 325nm) wavelengths and the observed fluorescence (F_{obs}) was corrected according to the equation:

$$F_{corr} = F_{obs} \times 10^{(Aex+Aem/2)} \quad (3)$$

The correction for the inner-filter effect can only be used when $A_{ex}+A_{em}$ is less than 0.2. The quenching data and the dissociation constant (K_d) was determined according to the equation:

$$1/\Delta F = K_d/\{\Delta F_{max}[L]\} + 1/\Delta F_{max} \quad (4)$$

where ΔF is the corrected fluorescence signal, ΔF_{max}, the maximum fluorescence, [L] is the ligand concentration and K_d the dissociation constant. The plot yields a linear line with the intercept of the abscissa equal to $-1/K_d$.

3.9 Equilibrium urea-induced unfolding studies.

The conformational stability of wild-type (modified or unmodified) hGST A1-1 and mutant protein were investigated using urea-induced denaturation. The assessment of the degree of reversibility of the urea-induced denaturation process was essential. Reversibility of the unfolding reaction means that the native and unfolded protein state are in equilibrium with each other in the folding transition zone and that the refolding reaction results in 100% recovery of native protein.

The equilibrium unfolding and reversibility study was performed at room temperature in 20mM sodium phosphate 1mM EDTA and 100mM NaCl buffer, pH 6.5. Native protein (0.1μM-5μM), AEDANS-modified protein (0.8μM) and L164A mutant protein (0.1μM-3μM) were unfolded in 0-8M urea. The 10M-urea stock solution was prepared fresh daily, as described by Pace *et al.* (1989). Each point of the denaturation curve was determined using a separate solution of varying urea concentration. The urea concentration of each solution was determined by varying the amount of urea and buffer added. The protein sample was incubated in the different urea concentrations for at least 1-hour. The reversibility of urea-induced unfolding was initiated by a 10-fold dilution of wild-type (10μM in 8M urea), AEDANS-modified (7.5μM in 7M urea) and mutant (4.5μM in 7M urea) enzyme and assessed by enzyme activity and fluorescence spectroscopy.

3.9.1 Physical techniques used to monitor the unfolding transition

A large number of techniques such as biological activity measurements, immunochemical techniques and spectroscopic techniques such as fluorescence, circular dichroism and nuclear magnetic resonance are used to distinguish the native and denatured states from one another (Shirley, 1995). Fluorescence has its application in monitoring the thermodynamics and kinetics of protein unfolding/refolding reactions. Fluorescence is the phenomenon whereby a molecule after absorbing visible or ultraviolet radiation emits radiation at a longer wavelength (Stokes' shift). Absorption of photons elevates electrons in these molecules to a higher energy state in less than 10^{-15} seconds. After absorption, the energy is lost rapidly by collision (as heat) resulting in the energy of the excited molecules falling rapidly to that of the minimum vibrational energy in the lowest excited state. The period of less than 10^{-8} seconds gives rise to a fluorescence peak with a Stokes' shift (Lacowicz, 1983). The fluorophore can be either intrinsic (tryptophan, tyrosine or a co-factor) or extrinsic (AEDANS, fluorescein or dansyl). In proteins, the tryptophan residues are particularly valuable probes since the indole ring is sensitive to

environmental changes and they can be used as probes of secondary and tertiary changes.

3.9.1.1 Steady-state fluorescence

All fluorescence measurements were made using a Hitachi model 850-fluorescence spectrofluorimeter. The excitation and emission bandwidths were set to 5nm. Tryptophan was selectively excited at 295nm and emission monitored at 325nm (folded protein) and 355nm (unfolded protein). Emission spectra were collected from 300nm-400nm at a scan rate of 60nm/min. Monitoring the intensity, when the excitation and emission wavelengths were set at 295nm assessed the extent of Rayleigh scattering (due to aggregation) during every unfolding study. The limit of the protein concentration, to minimise aggregation, was assessed to be 5μM and 3μM for the wild type and mutant enzyme, respectively.

Anisotropy decays, to assess the structural integrity of the enzyme, were measured using Hitachi polarisation attachments. Anisotropy (A) was calculated using:

$$A = (I_{VV}-GI_{VH})/(I_{VV}+2GI_{VH}) \text{ and } G = I_{HV}/I_{HH} \quad (5)$$

G is the correction factor (typically 1.07 in this study) for the differing efficiencies that the excitation and emission monochromators have for the horizontal and vertically polarised light. I_{VV}, I_{VH}, I_{VH} and I_{HV} refer to the fluorescence intensities measured when the polarisers were positioned horizontally (H) or vertically (V) (Lacowicz, 1983). Anisotropy is, therefore, the difference between parallel and perpendicular emission intensity with respect to the total intensity when parallel polarised light is used (Royer, 1995). Excitation at 295nm and emission at 355nm were used to monitor anisotropy decays.

The binding of the hydrophobic ligand, ANS, was monitored using steady-state fluorescence measurements. The ligand, ANS was added in saturation (200μM) to

wild type and mutant protein pre-equilibrated in 0-8M urea. Fluorescence measurements were as described above. In addition, the effect of ANS on the conformational stability was assessed by pre-incubation of the wild-type enzyme with saturating concentrations of ANS for at least 1hour prior to the addition of urea.

3.9.1.2 Second-derivative spectroscopy

The application of second-derivative spectroscopy to zero-order absorption spectra enables the spectral features from tyrosine residues to be monitored during an unfolding transition (Ragone *et al.*, 1984). The ratio between two second derivative peak-to-peak distances is used to evaluate the average polarity of the tyrosine residues in proteins (e.g., see figure 23). The ratio (r) is the peak-to-peak distance between the maximum at 287nm and the minimum at 283nm (a) and the peak-to-peak distance between the maximum at 295nm and the minimum at 290.5nm (b). The position of the maxima and minima may change marginally. The ratio (r) is a/b. The degree of tyrosine exposure,α, was calculated using:

$$\alpha = (r_n - r_u / r_u - r_a) \qquad (6)$$

where r_a is the ratio of a mixture of aromatic amino acids in ethylene glycol i.e., in a solvent with the same characteristics of the interior of the protein matrix. r_n and r_u are the ratio for the native and unfolded protein (in 8M urea), respectively.

Second-derivative analyses of [illegible] traviolet spectra (240nm-320nm), for 5μM wild-type hGST A1-1 in 0-8M [illegible]ea were monitored using a GBC UV-visible (911A) spectrophotometer (scan speed 250nm/min and slit width 1nm). Three-to-five spectra were collected and averaged.

3.9.2 Analysis of the equilibrium unfolding transition

The single sigmoidal transition i.e., the absence of any thermodynamically stable intermediates, for wild-type and mutant enzyme enabled the satisfactory analysis of

the unfolding curves according to the two-state assumption (Pace *et al.*, 1989). Any measured property (y_{OBS}) of wild type and mutant enzyme (i.e., Trp fluorescence; anisotropy decays; ANS binding; tyrosine exposure; enzyme activity) was converted to the fraction of protein populated in the unfolded form (f_U).

In the two-state assumption, only the native (N) state and the unfolded (U) state are present in significant concentrations at the midpoint of the unfolding transition. For the two-transition, therefore, $f_N + f_U = 1$ and y_{OBS} which is the measured property is equal to $y_N f_N + y_U f_U$ where f_N and f_U are the fraction unfolded native or unfolded protein state and y_N and y_U are the measured properties of the native and unfolded state, respectively. y_N and y_U are estimated from the linear extrapolation of the data in the pre- (y_N) and post-(y_U) transition baselines (figure 6). Therefore,

$$f_U = (y_N - y_{OBS}/y_N - y_U) \qquad (7)$$

where y_{OBS} is the observed property (i.e., the ratio of the Trp fluorescence intensity of unfolded to folded protein; the anisotropy value (A); the percentage solvent exposure of Tyr or the percentage residual CDNB-conjugating activity).

For the two-state model for dimeric proteins (Pace *et al.*, 1989; Bowie and Sauer, 1989), the equilibrium constant, K_U, was calculated using each point in the transition region of the denaturation curve:

$$N_2 \overset{K_U}{\leftrightarrow} 2U$$

$$K_U = [U]^2/[N_2] = 2P_t[f_U^2/(1-f_U)] \qquad (8)$$

where P_t is the protein concentration of the monomer, f_U is the fraction-unfolded protein. The free energy change was calculated according to the linear free energy model (Schellman, 1978; Pace *et al.*, 1986). The linear dependence between Gibbs

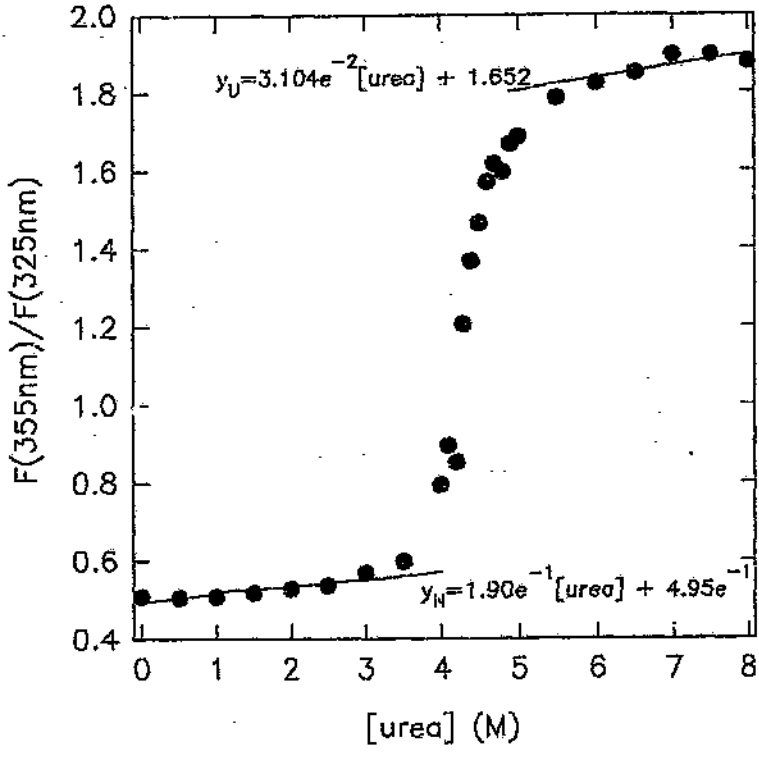


Figure 6. Urea-induced unfolding transition for 1μM hGST A1-1 prior to normalisation according to the method of Pace *et al.* (1989). The fluorescence intensity of folded protein (325nm) and unfolded protein (355nm) were monitored by excitation at 295nm. Linear regressions through the pre- and post-transition regions were fitted and the data was converted to the fraction of unfolded protein (f_U).

free energy change ($\Delta G = -RT\ln K_U$) on unfolding and the urea concentrations was calculated using:

$$\Delta G = \Delta G(H_20) - m\,[\text{denaturant}] \qquad (9)$$

where $\Delta G(H_20)$ is the Gibbs free energy in the absence of denaturant (i.e., an estimate of the conformational stability of the protein) and m is the susceptibility of the protein to urea (the m-value) and indicates the change in solvent exposure of the protein as it unfolds. The m-value is a measure of the co-operativity of the unfolding transition.

3.9.3 Differential scanning calorimetry (DSC)

DSC is the only technique available for directly accessing the thermodynamics associated with a thermally induced transition (Privalov and Potekhin, 1986). Differential scanning microcalorimetric measurements were determined using an adiabatic differential microcalorimeter model-4 (DASM-4; Mashpriborintork, Moscow). 20μM hGST A1-1 in 20mM sodium phosphate buffer, pH 6.5, containing 0.1M NaCl and 1mM EDTA was heated at a rate of 1°C/min from 20°- 90°C. In a similar experiment, 20μM hGST A1-1 in the presence of saturating concentrations of ANS (200μM) was heated as above. For each sample, three calorimetric scans were performed: a buffer scan (which was subtracted from the sample scan), a second scan of the protein and a third scan (i.e., a re- n of the sample scan) to assess the reversibility of the thermal denaturation. For hGST A1-1, the third scan indicated a signal identical to that for buffer indicating that the thermal unfolding process is irreversible. Therefore, based on this the DSC profile was not used for the evaluation of the thermodynamic parameters such as the change in entropy and the heat capacity change, it was, however, integrated to obtain the calorimetric enthalpy change (ΔH_{cal}) for comparison with the van't Hoff enthalpy change (ΔH_{vH}). The DSC profile was fitted to a two-state reaction scheme:

$$N_2 \overset{K}{\leftrightarrow} 2U$$

where K is the equilibrium constant defined as a function of the enthalpy change (ΔH), entropy change (ΔS) and the partial specific heat capacity change (ΔC_p). The equilibrium constant can then be expressed in terms of the Gibbs-Helmholtz free energy as follows:

$$K = \exp(-\Delta G/RT) \quad (10)$$

$$\Delta G = \Delta H + \Delta C_p(T-T_R) - T\{\Delta S + \Delta C_p \ln (T/T_R) - R\ln(2C_t)\} \quad (11)$$

where ΔH and ΔS are the enthalpy and entropy changes at some reference temperature (T_R), The calorimetric enthalpy (ΔH_{cal}) was calculated by integrating the ΔC_p versus temperature over the transition interval, taking into account the protein concentration (C_t). ΔC_p is the apparent heat capacity change from the initial to the final state. The equilibrium constant, K differs from the conventional equilibrium constant because the molar concentration term C_t has been included. The unfolding of hGST A1-1 was monitored by absorbance at 280nm. The van't Hoff enthalpy (ΔH_{vH}) was determined from a plot ln K versus 1/T (Pace *et al.*, 1989):

$$\delta(\ln K)/\delta(1/T) = -\Delta H_{v,H}/R \quad (12)$$

where R is the universal gas constant (1.978cal/mol/K).

3.3.4 Urea SEC-HPLC

SEC-HPLC was performed, in the presence of 0-8M urea (Corbett and Roche. 1984), at room temperature (typically 21°C) using a BioSep SEC-S3000 column (Phenomenex) (300mm x 7.80mm). The hydrodynamic volume of 5μM hGST A1-1

that was pre-equilibrated in 0-8M urea was assessed from the elution profile (at 280nm) using a flow rate of 0.5ml/min. The column was pre-equilibrated with the appropriate urea concentration prior to loading the sample.

3.9.5 Thermal-gradient gel electrophoresis (TGGE)

TGGE was performed as described by Birmes *et al.* (1990) and the buffer selected was 0.023M Na_2HPO_4, 0.132M NaH_2PO_4, pH 6.0. The selection of this buffer was based on the pI of the protein (i.e., 8.9) and on the fact that this buffer gives a small change in pH with a temperature increase of 50°C (ΔpH/50°C is +0.10) (Birmes *et al.*, 1990). A slab gel (300mm x 300mm) composed of 12.5% polyacrylamide, 0.29% bisacrylamide, 2mM DTT, 100mM NaCl, lmM EDTA in the sodium phosphate buffer, pH 6.0, was placed vertically against an aluminium gradient block insulated with a thin layer of Teflon sheeting. 150µg of protein (wild-type protein) in 1ml of the sodium phosphate buffer, pH 6.0, containing 10% glycerol and methylene green, was applied uniformly across the top of the gel. Electrophoresis was performed at 50V for 30minutes for smooth application into the gel and 120V for 3hours to allow the sample to enter the gel. The electrophoresis was then stopped and a linear temperature gradient from 20°-80°C formed for at least 30minutes by passing water at 5°C through one extreme of the block and water at 85°C through the other extreme of the block. Once the gradient was formed the electrophoresis was continued whilst maintaining the temperature gradient for 6hours at 120V.The gel was stained in 0.25% Coomassie brilliant blue R-250.

3.10 Urea-induced unfolding /refolding kinetic studies

Stopped-flow kinetics facilitates the study of the kinetic mechanism of the unfolding and refolding pathway for a protein and the characterisation of all of the species on the pathway. Stopped-flow experiments enable the detection of early events, in the millisecond time range, of the pathway. The kinetic experiments designed to monitor the unfolding/refolding reactions were based on the urea-induced equilibrium unfolding transition for hGST A1-1 (figure 6). The urea-induced unfolding transition

can be divided into three different regions of protein stability: the native baseline, the transition region and the unfolded baseline. Therefore, various aspects of the pathway can be monitored and characterised by changing the experimental conditions.

3.10.1 Fluorescence-detected stopped-flow kinetics

Fluorescence detection of the unfolding and refolding of hGST A1-1 were monitored using a stopped flow-mixing device (SX-18MV) from Applied Photophysics (UK). The excitation path length was 10mm and the emission path length was 2mm. An excitation bandwidth of 2.32nm was employed. For the unfolding of unmodified hGST A1-1 and L164A hGST A1-1, fluorescence was measured by excitation of Tyr and Trp at 280nm or Trp alone at 295nm and the emission was monitored using a 320nm cut-off filter. For the unfolding of AEDANS-modified protein, fluorescence was measured by direct excitation of AEDANS at 340nm or energy transfer from Trp to AEDANS at 295nm and emission was monitored using a 400nm cut-off filter. For the unfolding of hGST A1-1 in the presence of ANS, fluorescence was measured by excitation of Tyr/Trp at 280nm and emission was monitored using a 400nm cut-off filter. For the refolding of hGST A1-1, fluorescence was measured by excitation of Tyr and Trp at 280nm and the emission was monitored using a 320nm cut-off filter. The temperature in the stopped-flow cell was regulated within 0.1°C of the required temperature using a thermostated water bath.

The kinetics of unfolding and refolding in the equilibrium transition i.e., at final urea concentrations between 3.5M and 4.5M urea, were followed using manual mixing methods which have a dead time of 8-10s. The reactions were stirred continuously using a magnetic stirring mixer at approximately 25°C (i.e., the temperature inside the fluorimeter). Fluorescence methods were made using a Hitachi model 850-fluorescence spectrofluorimeter with excitation at 280nm and emission at 325nm. The excitation and emission bandwidths were set to 5nm each. The reactions were scanned at 10mm/min, for the appropriate time range.

In all unfolding and refolding experiments, to ensure the absence of mixing artifacts various control experiments were performed. Identical "unfolding" and "refolding" experiments were performed using 2μM N-acetyl-L-tryptophanamide (NATA) and 20μM N-acetyl-L-tyrosinamide instead of 1μM hGST A1-1. In addition, if the phase was a mixing artifact or the result of aggregation it would be observed at any wavelength as a consequence of a change in refractive index. Absorbance at 340nm was used to monitor aggregation in this study. Finally, all control runs (buffer only; buffer/urea and buffer/protein) were performed for every kinetic experiment. The effect of the slit width on the photodegradation of tryptophan residues was assessed. A slit width of 0.5mm results in little or no decrease in the tryptophan emission i.e., there is a minimal amount of Trp photodegradation. The effect of residual S-hexylglutathione on the unfolding and refolding kinetics was also investigated by purification of the protein in the absence of ligand (Cameron *et al.*, 1995).

3.10.1.1 Dead-time for the stopped-flow instrument

The reaction between 2,6-dichlorophenolindophenol (DCIP) and L-ascorbic acid (Tonomura *et al.*, 1978) was used to determine the dead-time of the instrument. L-ascorbic acid reduces DCIP and the reaction is a pseudo-first order reaction when the concentration of L-ascorbic acid was in excess. The disappearance of the colour of DCIP was followed at 524nm. 500μM of DCIP (in 10% propanol containing 0.2M NaCl) was mixed (1:1 symmetric mixing) with 2mM - 120mM of ascorbic acid in 0.02M HCl and 0.2M NaCl. The final pH of the mixture was 2.0 and all solutions contained 0.2M NaCl to minimise the effect of changes in the ionic strength.

The plot of the pseudo-first order rate constant (k_{app}) versus ascorbic acid concentration was linear up to 40-50mM ascorbic acid and the upper limit of the rate constant is $1200s^{-1}$. The mixing time of the instrument is 0.5ms and the dead time was calculated to be 1.5ms, therefore, no data within 2ms was fitted (figure 7).

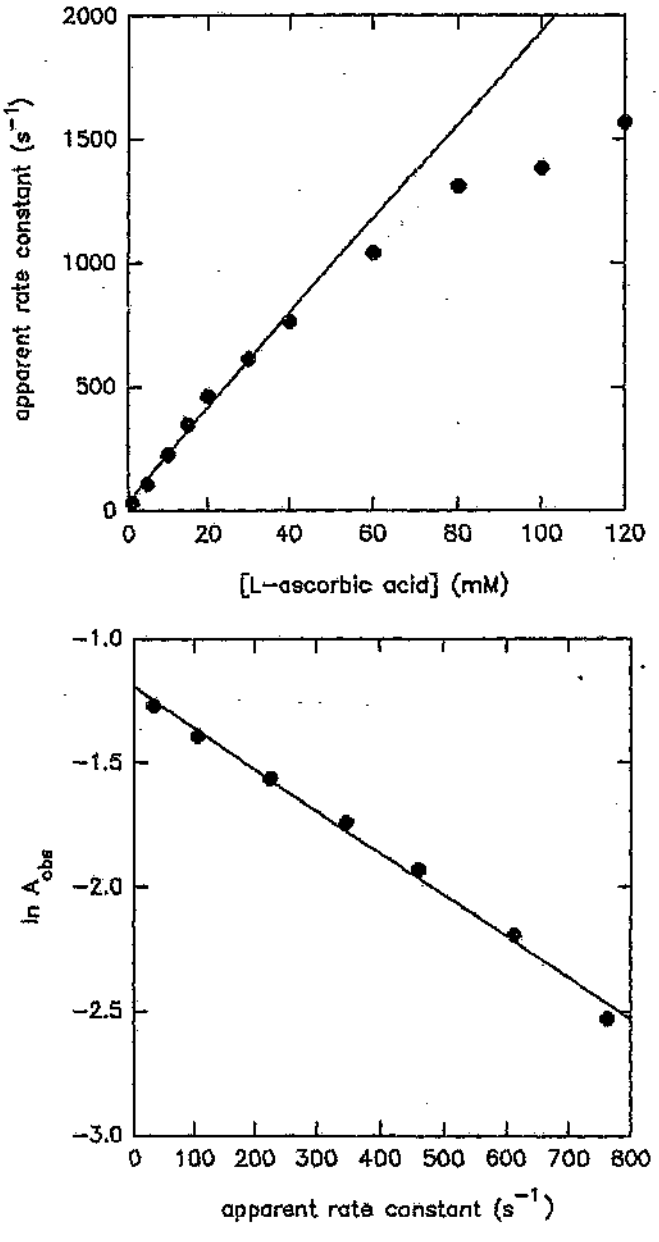


Figure 7. Dead time determination for the SX-18MV stopped-flow apparatus (Applied Photophysics, UK).

3.10.2 Kinetic data analysis

All kinetic data obtained using the stopped-flow instrument was fitted using the Applied Photophysics software, version 4.24. The program utilises the algorithm of Levenberg-Marquardt (Marquardt, 1963) for non-linear least squares fitting. The kinetic data obtained using the manual mixing method were fitted using the iteration program in Sigma Plot (v 5.0) (Jandel Corporation). Three to four kinetic runs were averaged per experiment and standard deviations calculated for at least 3 separate experiments. Kinetic analysis of the time dependent change of fluorescence signal requires consideration of the final urea concentration with reference to the equilibrium transition.

For urea unfolding jumps ending at urea concentrations above the equilibrium transition zone (final urea concentrations greater than 5M), the unfolding reactions of the fast and slow phases were fitted using a single exponential function. Each phase was fitted individually to a single exponential function (by selecting a fitting range).

For urea refolding jumps ending at urea concentrations below the equilibrium transition zone (final urea concentrations less than 4M), the refolding reactions of the fast and intermediate phases were fitted to a bi-exponential function. The slow refolding phase was fitted to a single exponential function (by selecting a fitting range).

$$F_t = \Sigma F_i \exp(-t/\tau) + F_\infty \qquad (13)$$

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

For urea unfolding and refolding jumps ending with urea concentrations within the equilibrium transition zone (final urea concentrations 3.5M to 5.8M), both unfolding

and refolding reactions must be considered since in this region the model relaxes to equilibrium. Data in this region was analysed using the relaxation expression described by Bernasconi, 1976.

For fluorescence measurements:

$$F_t = \Sigma F_i(\exp(-t/\tau))/(1+q'F_i(1-\exp(-t/\tau))) + F_\infty \qquad (14)$$

where F_t, F_i, F_∞ and τ have the same meaning as in equation 13. For fluorescence measurements, q' relates the fluorescence intensity change to the change in protein concentration. q' is defined as the term, which accounts for the quantum yields of the native and denatured states, the extinction coefficients of each specie and various instrumental parameters. In this study, as employed for the *Trp* aporepressor (Gittelman and Matthews, 1990) and leucine zipper peptide (Wendt *et al.*, 1997), q' was used as a fitting parameter. In unfolding studies, q' has a theoretical limit of -0.5 (Bernasconi, 1976). For these studies q' was typically -0.3. In the refolding studies, the value of q' was near unity. The term $(q'F_i(1-\exp(-t/\tau))$ in the denominator accounts for the contribution of the refolding or unfolding reaction in the transition zone. This term approached zero under conditions that strongly favour the denatured state or the folded state and hence, it approximates to an equation equivalent to a single exponential or to a sum thereof (equation 13).

The deviations of the experimental data from the fitted function (i.e., the residuals) were used to judge the accuracy and quality of the fit.

3.10.3 Urea-dependence of the unfolding/refolding rate

In all unfolding experiments, unmodified/modified native protein/L164A protein (6μM hGST A1-1 in 0M or 3M urea in 20mM sodium phosphate 1mM EDTA and 100mM NaCl buffer, pH 6.5) was diluted six fold (1:5 asymmetric mixing) with 4M - 10M urea, at 25°C. The final urea concentrations were between 3.5M and 8.9M and

the final protein concentration was 1μM. The urea-dependence of the first order rate constants for unfolding (k_u) were fitted to equation 15 which enables the determination of the unfolding rate in the absence of denaturant (an estimate for the rate of dissociation/unfolding *in vivo*) and the change in solvent accessibility of the transition state for unfolding (Tanford, 1970):

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

where k_u is the apparent first order rate constant for unfolding at different concentrations of denaturant, $k_u(H_20)$ is the apparent unfolding rate in the absence of denaturant and m_u is the change in solvent accessibility of the transition state for unfolding.

In the unfolding experiments to assess the influence of ANS on the unfolding pathway, 2μM hGST A1-1 was pre-equilibrated with 0-200μM ANS in sodium phosphate buffer, pH 6.5. The unfolding was initiated by mixing two-fold (1:1 symmetric mixing) with 10M urea at 25°C.

In the refolding kinetic experiments, 6μM hGST A1-1 was unfolded in 6M or 7M urea for at least 1 hour. In the refolding experiments, unfolded protein (6μM hGST A1-1 in 6M or 7M urea in buffer as above) was diluted six fold (1:5 asymmetric mixing) with 0M-3.7M urea. The final urea concentrations were between 1M and 4.25M and the final protein concentration was 1μM. The urea-dependent refolding experiments were performed at three different final conditions: (i) at 25°C using 1μM protein; (ii) at 15°C using 1μM protein and (iii) at 25°C in the presence of 8.3% (w/v) sucrose using 1μM protein.

The urea-dependence of the refolding rate in the presence of 8.3%(w/v) sucrose was performed in an identical manner as described above except that the unfolded protein

was diluted six-fold with urea solutions which contained 10% (w/v) sucrose in the sodium phosphate buffer, pH 6.5.

Kinetic amplitudes for all the unfolding (modified/unmodified/L164A protein) and refolding (unmodified protein) phases were normalised with respect to the fraction unfolded (f_u) measurements obtained in equilibrium studies.

3.10.3.1 Protein-concentration dependence of the unfolding/refolding kinetic studies

For the kinetic unfolding studies, 3μM-30μM of native protein was diluted six fold (1:5 asymmetric mixing) with 6M-10M urea at 25°C. The final protein concentrations were 0.5μM-5μM and the final urea concentrations were between 5.8M and 8.3M.

For the refolding kinetic studies, 3μM-30μM of unfolded protein in 6M urea was diluted six-fold (1:5 asymmetric mixing) with 0M-3.4M urea at 25°C. The final protein concentrations were 0.5μM-5μM and the final urea concentrations were between 1M and 4M.

The change in the apparent unfolding and refolding rate constants with protein concentration were used to predict the order (i.e., first-order (unimolecular) or second-order (bimolecular)) of the reactions. If the rate changes with protein concentration then the reaction is bimolecular as expected for the refolding of a dimeric protein.

3.10.3.2 Kinetic two-state test for unfolding/refolding

Stopped-flow unfolding and refolding experiments were performed at different initial urea concentrations but at the same final urea concentration to assess if the complexities of the unfolding/refolding reactions were the result of a pre-equilibrium

existing in the pre- or post-transition baseline between native or unfolded molecules, respectively.

For the unfolding kinetic experiments, 6μM native protein in 0M, 1.5M or 3M urea was diluted six fold (1:5 asymmetric mixing) with 9.6M, 9.3M or 9M urea, respectively. The final urea concentration was 8M and the final protein concentration was 1μM. For the refolding experiments, 6μM protein in 6M, 7M or 8M urea was diluted six fold (1:5 asymmetric mixing) with either 2.4M, 2.2M or 2M urea, respectively to give a final urea concentration of 3M or with 0.4M, 0.2M or sodium phosphate buffer, respectively to give a final urea concentration of 1.3M.

3.10.4 Temperature dependence of the unfolding/refolding rate

The temperature dependence of unfolding was performed from 10°- 40°C by a six-fold dilution of 6μM native unmodified/L164A protein with 10M urea. The final urea concentration was 8.3M. For modified protein, 6μM native protein was diluted with 9.6M urea to give a final concentration of 8M. The temperature dependence of refolding was performed from 5°- 40°C by a six fold dilution of 6μM protein in 6M or 7M urea to give final urea concentrations of 1M and 3M, respectively. The temperature dependence of the apparent unfolding/refolding rate constants (k) for unmodified/ modified / L164A protein were analysed using the transition state theory (Eyring plot) (Oliveberg *et al.*, 1995; Tan *et al.*, 1996):

$$k = \kappa k_B T/h \exp(\Delta G^{\ddagger}/ -RT) \qquad (16)$$

where k_B is Boltzmann's constant ($1.38 \times 10^{-23} J.K^{-1}$), h is Planck's constant ($6.626 \times 10^{-34} J.s^{-1}$), T is the absolute temperature, R the universal gas constant and $\Delta G^{\ddagger}$ is the activation energy. The transmission coefficient (κ) is assumed to equal 1. There are various factors that limit the use of the transition state theory for protein unfolding/refolding. The values obtained for the activation free energy are highly

dependent on the value of the transmission factor (κ) and the preexponential (k_BT/h) term. Neither of the terms are accessible for protein folding reactions (Kiefhaber, 1995). Thus, the values obtained are apparent activation parameters. The temperature dependence of the apparent unfolding/refolding rate constant (k) reflects the temperature dependence of the free energy of activation $\Delta G^{\ddagger}$ (Oliveberg *et al.*, 1995):

$$\begin{aligned} \Delta G^{\ddagger}(T_o) &= \Delta H^{\ddagger}(T_o) - T\Delta S^{\ddagger}(T_o) \\ &= \Delta H^{\ddagger}(T_o) + \Delta C_p^{\ddagger}(T-T_o) - T[\Delta S^{\ddagger}(T_o) + \Delta C_p^{\ddagger}\ln(T/T_o)] \end{aligned} \quad (17)$$

where T is the absolute temperature, T_o is the standard temperature (298K); $\Delta S^{\ddagger}(T_o)$ is the activation entropy at T_o; $\Delta H^{\ddagger}(T_o)$ is the activation enthalpy at T_o; $\Delta C_p^{\ddagger}$ is the heat capacity difference between the native/denatured states and the transition state.

Substitution of equation (17) into (16) yields:

$$\ln(kh/k_BT) = [\Delta S^{\ddagger}(T_o)/R - \Delta H^{\ddagger}(T_o)/RT - \Delta C_p^{\ddagger}(T-T_o)/RT + \Delta C_p^{\ddagger}/R\ln(T/T_o)] \quad (18)$$

This enables a plot of $\ln(kh/k_BT)$ versus 1/T to be fitted and the thermodynamic parameters of the unfolding and refolding kinetics to be calculated.

3.11 General data analysis and molecular graphic simulations

All least-squares fitting of the data to the appropriate equation was performed using the program Sigma Plot (v 5.0) (Jandel Corporation), unless otherwise stated. The three-dimensional structure of hGST A1-1 was viewed using the molecular visualisation program, Rasmol (v 2.6) (Sayle, 1994). The sequence alignments and the structural superimposition of the alpha helices 6, 7 and 8 for all gene classes (see chapter 5) and the other characteristics of the three-dimensional structure of the glutathione transferases reported were obtained using WPDB (v 2.1) (Shindyalov and

Bourne, 1996). Any interatomic distances reported were calculated by means of the molecular-graphics package, Hyper-Chem (Autodesk LTD).

CHAPTER 4

EQUILIBRIUM AND KINETIC UNFOLDING PROPERTIES OF hGST A1-1

4.1. Fluorescence properties of hGST A1-1

The fluorescence properties of tyrosine and tryptophan (excitation at 280nm) and tryptophan alone (excitation at 295nm) and the extrinsic fluorophore, AEDANS, were used to probe the urea-induced conformational changes of hGST A1-1. The extrinsic fluorophore, AEDANS, was covalently linked to Cys111 (Cys111-AEDANS) (excitation at 295nm or at 340nm).

Human class alpha glutathione transferase A1-1 contains one tryptophan residue (Trp20) per subunit, located at the domain interface in alpha helix one of domain I (figure 8a) (Sinning *et al.*, 1993). The indole side chain of Trp20 protrudes from domain I into domain II and has non-specific hydrophobic interactions with the side chains of Ile158, Glu162 and Tyr165 (all in alpha helix 6) and Val194 and Phe197 (in alpha helix 8). The side chain of Trp20 is inaccessible to solvent (SASA is 4$Å^2$) and can be used as a probe for the dissociation of the two domains (Atkins *et al.*, 1997).

The tryptophan emission spectrum (excitation at 295nm) of native hGST A1-1 has a maximum at 325nm which is red-shifted to 355nm when the protein is denatured (figure 9a). The maximum at 325nm for native hGST A1-1 is consistent with the aromatic nature of the Trp residue and supports crystallographic evidence that Trp20 is buried and at least partially inaccessible to solvent. The region of the Trp20 residue has been implicated to show conformational changes in the presence of substrates or GSH-conjugates (Wang *et al.*, 1993b) and non-substrate hydrophobic ligands (N. Sluis-Cremer, unpublished results) indicating the dynamic nature of this residue. The red- shift to 355nm upon denaturation is similar to that observed for NATA in water (Teal, 1960) and illustrates the complete exposure of the Trp20 side chain to solvent. The red-shift is accompanied by a decrease in fluorescence intensity (figure 9a) which is opposite to that observed for the other cytosolic glutathione transferases (Dirr and

Reinemer, 1991; Erhardt and Dirr, 1995; Kaplan *et al.*, 1997). There are a number of mechanisms, which are known to result in quenching of tryptophan fluorescence in the unfolded state. These include the influence of nearby protonated amines or carbonyl groups, metallic cations, anions or the existence of a variety of conformations each with a different average exposure to the aqueous phase (Gryczynski *et al.*, 1988). An examination of the amino acid sequence and the three-dimensional structure of hGST A1-1 (Sinning *et al.*, 1993), indicate the presence of an arginine residue at position 19. In the folded state, the guanidino group of Arg19 is unable to interact either by collision or via direct hydrogen bonding with the ε-nitrogen of the indole group of Trp20 (interatomic distance 8.47Å) (figure 8b). In the unfolded state, as a result of the greater conformational flexibility of the polypeptide chain, collisional quenching between the two side chains is possible. The phenomenon of fluorescence quenching is described to be the result of the charge effect (Steiner and Kirby, 1969) of the protonated guanidino group of Arg19 with the indole side chain of Trp20. Therefore, the complete unfolding and dissociation of hGST A1-1 results in a decrease in the fluorescence intensity and conversely for complete refolding and association.

In the crystal structure, hGST A1-1 has ten tyrosine residues per subunit. Three are located in domain I (Tyr8, Tyr49, Tyr74), two in the linker region between the two domains (Tyr79, Tyr82) and five in domain II (Tyr95, Tyr132, Tyr147, Tyr165, Tyr166). The excitation of tyrosine and tryptophan (excitation at 280nm) indicates a substantial amount of energy transfer from the tyrosine residues (globally distributed) to the Trp20 residue (figure 9b). In addition, the emission maximum (at 325nm) is identical to that observed for the selective excitation of Trp20. The fluorescence emission maximum at 325nm, as opposed to 305nm as expected for the free tyrosine amino acid, indicates that Trp20 acts as a local reporter of events occurring at the interdomain interface. The unfolding of hGST A1-1 results in a decrease in the fluorescence intensity as the distance between the tyrosine residues and Trp20 becomes too great for efficient energy transfer. The uncoupling of the energy transfer

is apparent as a shoulder in the spectrum at 305nm (for the tyrosine residues) and the Trp emission maximum is red-shifted to 355nm.

The single cysteine residue in hGST A1-1, Cys 111, is located in the turn between the alpha helices 4 and 5 of domain II. Its sulphydryl side chain is exposed to solvent (SASA is 21Å^2) and points into the amphipathic V-shaped cavity between the two subunits (referred to as the L-site because of its ability to bind non-substrate ligands) (figure 8c). The thiol group of Cys111 in subunit 1 is 23Å from the ε-nitrogen of the indole group of Trp20 in the same subunit and 34Å from the ε-nitrogen of the indole group in subunit 2. The location of Cys111 enables it, by means of AEDANS modification, to be employed to monitor changes occurring at/near the cleft between the two subunits.

The tryptophan emission spectrum for native Cys111-AEDANS hGST A1-1 displays an identical emission maximum wavelength (325nm) as that for unmodified hGST A1-1 (figure 9a) but lower fluorescence intensity (figure 10a). The lower fluorescence intensity is the result of the transfer of excitation energy from Trp20 to the AEDANS group. The transfer of energy is possible because of the spectral overlap of the emission spectrum of hGST A1-1 with the AEDANS absorption spectrum. Upon denaturation, the energy transfer from Trp20 to Cys111-AEDANS is uncoupled and the fluorescence intensity increases with a concomitant red-shift in the wavelength to 355nm. The direct excitation of the AEDANS label (excitation at 340nm) indicates an emission maximum at 485nm and 495nm for the native and unfolded protein, respectively (figure 10b). The red-shift and the slight decrease in fluorescence intensity upon denaturation are consistent with a more polar environment for the covalently bound AEDANS fluorophore. The emission maximum for AEDANS in water is at approximately 530nm (Hudson and Weber, 1973).

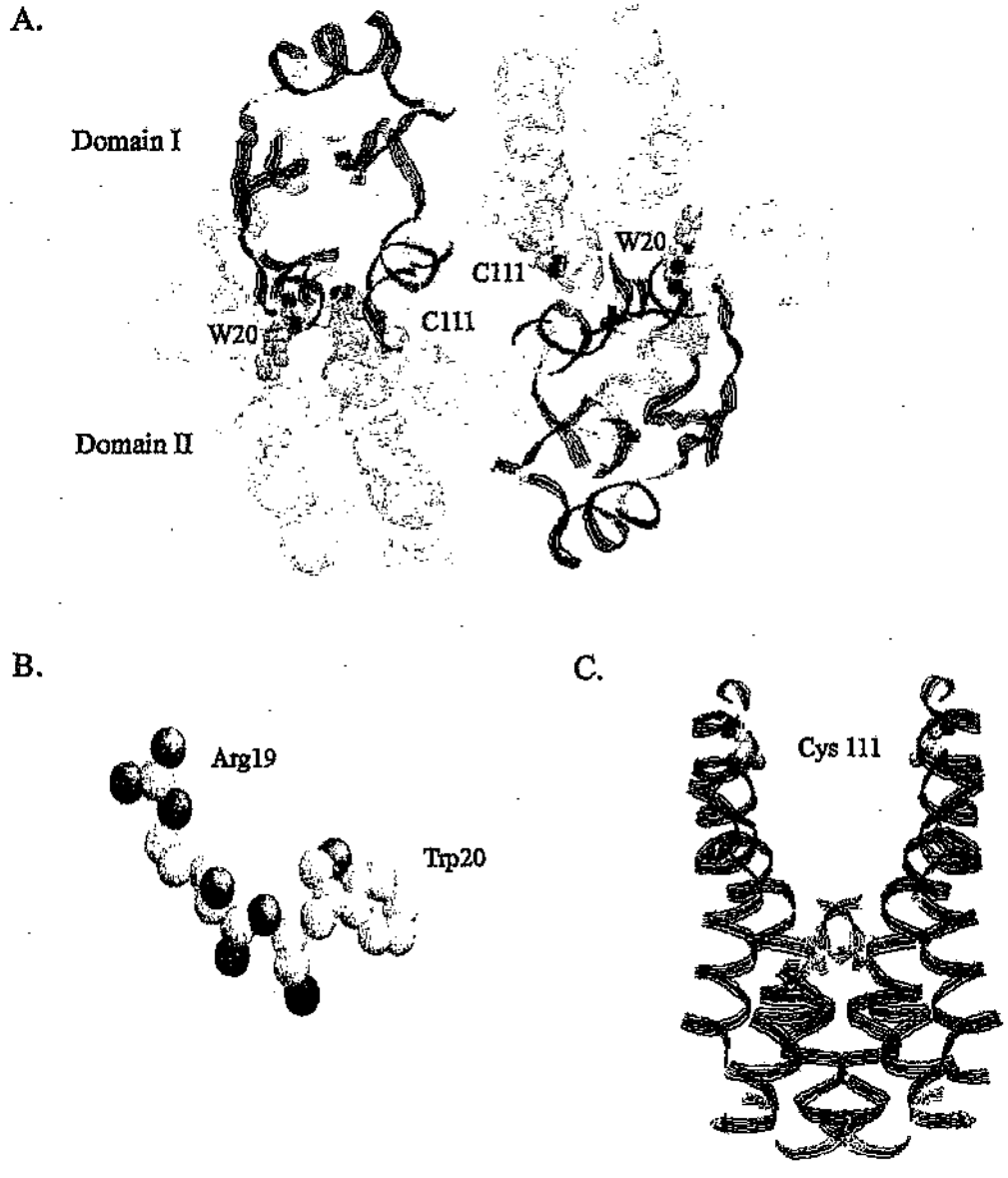


Figure 8. (a) Ribbon representative of the structure of hGST A1-1 (Sinning *et al.*, 1993). The view is parallel to the two-fold axis and Cys111 and Trp20 and the alpha helix 9 over domain I are indicated. (b) The orientation of Arg19 and Trp20 in the folded state and (c) the location of Cys111 at the dimer interface indicating the position of the sulphydryl groups. Figure was generated using Rasmol (Sayle, 1994).

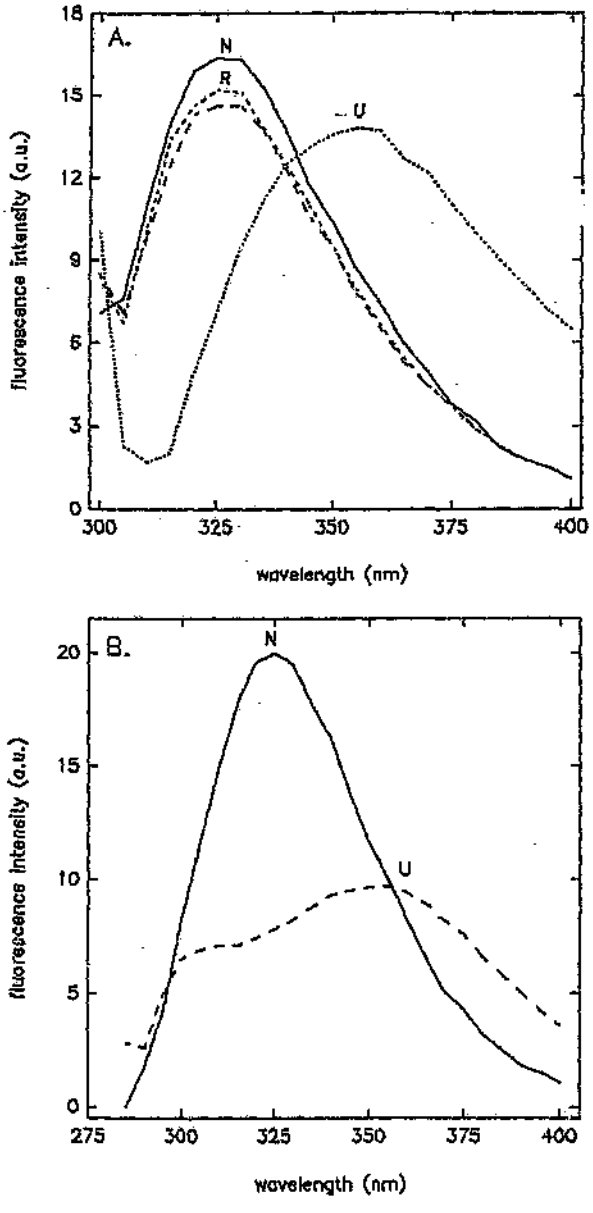


Figure 9. Fluorescence emission spectra for hGST A1-1. (a) Excitation at 295nm. Spectra of 1μM native (N; —) and unfolded protein in 8M urea (U; ——) Reversibility of unfolding (1μM in 0.8M urea) (R; -----) was compared with a control spectrum (---). (b) Excitation at 280nm. Spectra of 1μM native (N; —) and unfolded protein in 8M urea (U;---).

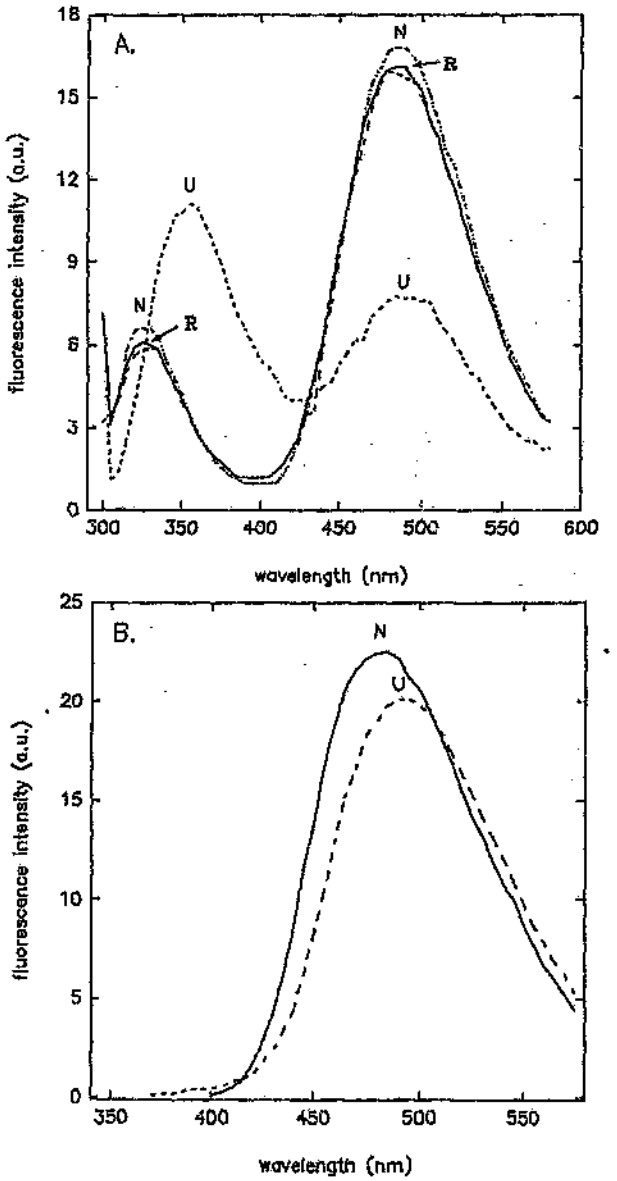


Figure 10. Fluorescence emission spectra of Cys111-AEDANS hGST A1-1. (a) Excitation at 295nm. Spectra of 0.8μM native (N;——) and unfolded protein in 8M urea (U; - - -). Reversibility of unfolding (0.8μM in 0.8M urea) (R; —) compared with control spectrum (-----). (b) Excitation at 340nm. Spectra of 0.8μM native (N; —) and unfolded protein in 8M urea (U; - - -).

4.2 Equilibrium unfolding of hGST A1-1

4.2.1 Reversibility of the unfolding transition

The reversibility of the urea-induced equilibrium unfolding was assessed by a tenfold dilution of denatured hGST A1-1 and monitored using fluorescence spectroscopy (as a structural probe) and enzyme activity (as a functional probe). Refolded unmodified hGST A1-1 has an emission maximum (at 325nm) and intensity identical to that observed for native protein (figure 9a). In addition, the recovery of enzyme activity was 95 to 98%. Similarly, refolded Cys111-AEDANS hGST A1-1 was structurally (100%) and functionally (90%) similar to native protein (figure 10a).

The 100% regain of native protein *in vitro* indicates that the pathway for unfolding/refolding for hGST A1-1, under these conditions, was devoid of competing side reactions such as aggregation. The high degree of reversibility for hGST A1-1 suggests that the folding *in vivo* is a self-assembly process unaided by chaperones or foldases. This is, however, a weak argument since only a very rapid rate of folding will enable a protein to escape chaperones. Similar high refolding recoveries have been documented for the cytosolic glutathione transferases (Dirr and Reinemer, 1991; Erhardt and Dirr, 1995; Kaplan *et al.*, 1997). A high refolding recovery is an essential prerequisite for the detailed characterisation of the equilibrium unfolding transition and for the interpretation of the kinetic mechanism for unfolding and refolding.

4.2.2 Two-state denaturation of hGST A1-1

4.2.2.1 Equilibrium unfolding transitions

The urea-induced equilibrium unfolding for hGST A1-1 was monitored using a variety of structural probes such as steady-state Trp20 fluorescence, tyrosine exposure, anisotropy decays, the binding of the ligand, ANS, and the fluorescence properties of AEDANS covalently linked to Cys111 of hGST A1-1. In addition, the unfolding transition was monitored using enzyme activity as a functional probe.

The equilibrium unfolding reaction for hGST A1-1, induced by urea, shows a single sigmoidal transition indicative of a highly co-operative system involving native dimer and two unfolded monomers (figure 11).

The first line of evidence for support of the two-state transition is the coincidence of the unfolding transition curves obtained using various structural probes (figure 11a). The change in intensity and wavelength of Trp20 and the increase in the extent of tyrosine exposure were used to probe local and global urea-induced structural changes, respectively. In the native state, the percentage solvent accessibility of the tyrosine residues was calculated, from the second-derivative analyses, to be 19%.

The binding of the ligand, 8-anilino-1-naphthalene sulphonate (ANS) was used to monitor structural changes of hGST A1-1. This fluorescent dye binds to hydrophobic regions of proteins and has been used to identify equilibrium intermediates such as molten globules (Fink, 1995). The binding of ANS to hGST A1-1 in the presence of denaturant was used to probe the existence of any species other than the native or unfolded state which would result in a shift in the equilibrium unfolding transition. The binding of ANS to the native state is weak because of the absence of accessible hydrophobic binding sites (Ptitsyn *et al.*, 1995). Similarly, the unfolded state shows weak binding of ANS because the hydrophobic patches are solvated and because the urea competes with the interaction between ANS and the protein (Kumar *et al.*, 1996). The unfolding transitions of hGST A1-1, monitored using Trp fluorescence, in the absence and presence of ANS show identical co-operative sigmoidal transitions indicating the absence of an intermediate that can preferentially bind ANS and hence result in a stabilisation or destabilisation of hGST A1-1.

The anisotropy decays (i.e., the polarisation of Trp fluorescence) were used to measure the structural integrity of hGST A1-1 during the unfolding transition. In general, a tryptophan residue within the folded protein structure is constrained to small amplitude motions, which result in very little depolarisation of the fluorescence

emission with respect to the polarised excitation. The mechanism of depolarisation of the tryptophan residue in the native protein results from the Brownian tumbling of the macromolecule in solution. The disruption of the protein structure upon denaturation results in the greater rotational freedom of the indole side chain of the tryptophan residue around the C_α and C_β bonds and this increases the efficiency of depolarisation (Murphy, 1995). Steady-state anisotropy of a spherical molecule is dependent on temperature, viscosity, the fluorescence lifetime of the fluorophore and on its molecular volume (Lacowicz, 1983). The angular displacement of the fluorophore results in a decrease in the anisotropy values that depends on the rate and extent of rotational diffusion during the lifetime of the excited state. The location of Trp20 at the domain interface enables one to assume that its fluorescence lifetime will not be affected by the possible dissociation of the dimer into a folded monomer. Therefore, under the conditions used in this study, only the molecular volume of the protein affects the anisotropy value. The anisotropy values of 0.082 for native hGST A1-1 decreases to 0.031 upon denaturation and this is consistent with an increase in the rotational freedom of the partially mobile indole side chain of Trp20. The sigmoidal transition with the absence of a plateau at intermediate urea concentrations excludes the existence of a structured intermediate state under equilibrium conditions.

The individual curves obtained for hGST A1-1 and Cys111-AEDANS hGSTA1-1 are coincident indicating that the alkylation of this residue with IAEDANS does not alter the folding pathway of the enzyme (figure 11b). This data in addition to unchanged physicochemical properties, such as Trp fluorescence, specific activity and hydrodynamic volume (SEC-HPLC), indicate that the covalent labelling of Cys111 with the AEDANS group results in little or no gross conformational changes of the protein. An additional probe for tertiary structural changes, occurring at the cleft between the two subunits, was provided using the fluorescence properties of AEDANS-covalently linked to Cys111. The individual unfolding transitions, obtained for the energy transfer from Trp20 to Cys111-AEDANS (at 295nm) and for the direct

excitation of Cys111-AEDANS (at 340nm), are coincident within experimental error (inset of figure 11b) supporting the co-operative unfolding pathway of hGST A1-1.

The coincidence of the equilibrium unfolding transitions for hGST A1-1 obtained from the variety of structural probes described above and illustrated in figure 11a supports the two-state transition between the folded dimer and the two unfolded monomers. It is important to note, however, that the unfolding of a protein can still be described as two-state even though different probes do not give coincident sigmoidal curves (Dill and Shortle, 1991). The reason for this is that the native and unfolded states are not fixed and therefore certain regions of the structure may be more susceptible to external conditions than the others. Therefore, coincident curves give supporting evidence for a two-state (folded dimer/unfolded monomer) transition but non-coincident curves are not definitive evidence against the two-state behaviour (Chan *et al.*, 1995).

The sigmoidal unfolding transition obtained using enzyme activity (functional probe) was non-coincident with that obtained using fluorescence properties (structural probe)(inset of figure 11a). The dimeric structure is important for the catalytic function of glutathione transferases and therefore dissociation of the dimer results in a loss in the function. The unfolding transition for 1μM hGST A1-1 has a midpoint at 4.1M urea for enzyme activity measurements and 4.5M urea for fluorescence measurements. This phenomenon may be the consequence of the unique alpha helix 9 found in the C-terminal region of the GST A1-1 polypeptide chain. This region shows high conformational flexibility in the absence of ligand (Cameron *et al.*, 1995) and is implicated to be an important component of the hydrophobic substrate-binding site (Board and Mannervik, 1991; Cameron *et al.*, 1995). It is not essential for glutathione binding but deletion of this segment does impair catalytic activity (Board and Mannervik, 1991). It is, therefore, feasible that the conformational flexibility of this helix would render it more susceptible to urea than the rest of the protein and the

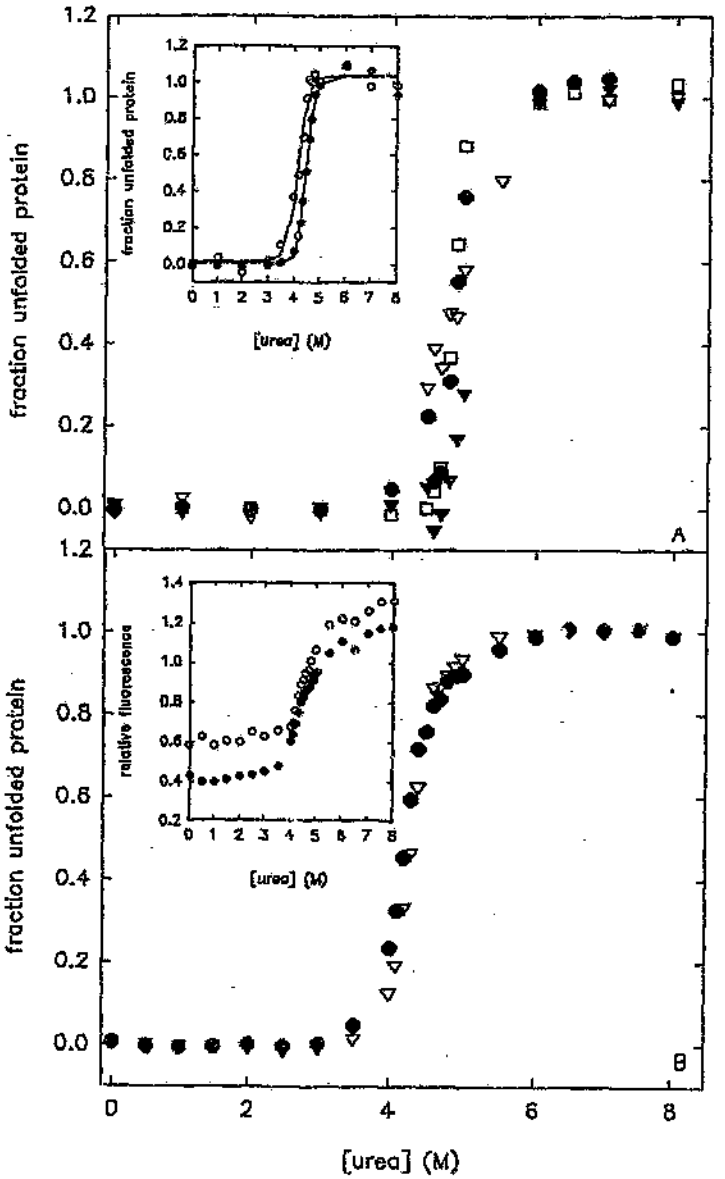


Figure 11. (a) Urea-induced equilibrium unfolding curves for 5μM hGST A1-1 monitored by Trp20 fluorescence (•), anisotropy decays (▽), the binding of ANS (▼) and the extent of Tyr exposure (□). Inset: unfolding transition for enzyme activity (o) and Trp fluorescence (•). (b) Unfolding transition for 0.8μM unmodified hGST A1-1 (▽) and 0.8μM AEDANS-Cys111 hGST A1-1 (•). Trp was excited at 295nm and emission monitored at 325nm and 355nm. Inset: Unfolding transition for Cys111-AEDANS (o; excitation 295nm, emission 485nm, •; excitation 340nm, emission 485nm).

destabilisation of this structural motif would impact on the hydrophobic substrate-binding site and thus the catalytic function.

4.2.2.2 Protein-concentration dependence

The dependence of an unfolding probe on protein concentration is essential for assigning the two-state assumption for unfolding and dissociation to oligomeric proteins (Timm and Neet, 1994). For any dimeric protein, there are two possible descriptions of the equilibrium unfolding reaction which are based on the relative stability of the folded dimer (N_2), the unfolded monomer (U) and the folded monomer (or intermediate)(N).

In the first model, a three-state pathway for unfolding and dissociation is described which involves the dissociation of the native dimer (N_2) into two structured monomers (N) prior to the complete unfolding into two unfolded monomers (U):

$$N_2 \overset{K_1}{\leftrightarrow} 2N \overset{K_2}{\leftrightarrow} 2U$$

where $K_1 = [N]^2/[N_2]$ and $K_2 = [U]/[N]$. In this model, the equilibrium constant K_1, i.e., for the dissociation reaction, is bimolecular and the equilibrium constant K_2, i.e., the unfolding reaction, is unimolecular.

In the second model, a two-state pathway suggestive of a concerted dissociation and unfolding reaction is described and involves only the native dimer (N_2) and the unfolded monomer (U):

$$N_2 \overset{K}{\leftrightarrow} 2U$$

where $K=[U]^2/[N_2]$ represents the equilibrium constant for the concerted unfolding and dissociation reaction and is bimolecular. Therefore, the bimolecular reaction described by K_1 (model 1) and K (model 2) would be protein concentration

dependent. The protein concentration dependence of an unfolding reaction has been used as a diagnostic tool for distinguishing between the two types of models (Bowie and Sauer, 1989).

Figure 12 shows the protein concentration dependence of the midpoint of the equilibrium unfolding transition for hGST A1-1 monitored using tryptophan fluorescence. The stability of hGST A1-1 increases with an increase in protein concentration and is in agreement with a bimolecular reaction governed by the law of mass action.

The satisfactory fit of hGST A1-1 to the two-state assumption enables an estimation of the conformational stability parameters: the free energy change in the absence of denaturant ($\Delta G(H_20)$) and the susceptibility of the free energy for unfolding of hGST A1-1 to urea (the m-value) (table 4). The values for the free energy change in the absence of denaturant, normalised to 1M protein, for these curves are in close agreement and this validates the two-state model. These experimentally determined parameters are in agreement with those predicted by Neet and Timm (1994) and Myers *et al.* (1995). The $\Delta G(H_20)$ calculated by correlating the tertiary/quaternary stabilisation with the size of the dimer ($\Delta G(H_20) = 8.8 + 0.08N$, where N is the number of amino acid residues in the monomer) was 26.4kcal/mol (Neet and Timm, 1994). The large value of $\Delta G(H_20)$ for class alpha and the other glutathione transferases studied thus far (Dirr and Reinemer, 1991; Erhardt and Dirr, 1995; Kaplan *et al.*, 1997) suggests that the tertiary and quaternary interactions are necessary for stabilisation of the native dimer. The theoretical m-value was calculated by relating the amount of surface area buried with the number of amino acid residues (Myers *et al.*, 1995). This correlation ($r = 0.84$ for urea as the denaturant) is described by $m = 374 + 0.11$ (ΔASA) where ΔASA is the change in the solvent-accessible surface area upon unfolding. (For hGST A1-1, the $\Delta ASA = -907 + 93$ (no. amino acid residues for the

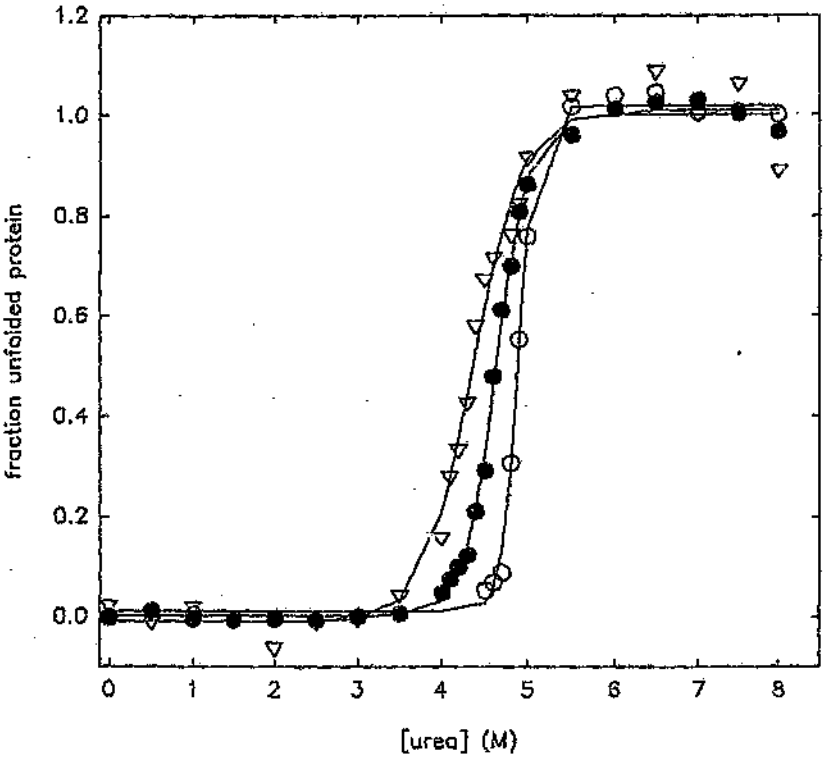


Figure 12. The protein concentration dependence of the unfolding transition for hGST A1-1 at 0.1μM (▽), 1μM (•) and 5μM (○). Trp was selectively excited at 295nm and emission measured at 325nm and 355nm.

Table 4. Conformational stability parameters for individual concentrations of hGST A1-1.

[Protein] (μM)	$\Delta G(H_20)$(kcal/mol)	m (kcal/mol/M)	C_m (M)
0.1	25.98 ± 3.23	4.08 ± 0.70	4.3 ± 0.1
1	27.59 ± 3.80	4.22 ± 0.79	4.5 ±0.1
5	26.68 ± 4.09	4.05 ± 0.61	4.8 ±0.1

dimer) = 40199Å^2). The calculated m-value was 4.79kcal/mol/M of urea. The agreement between the experimental and calculated m-values indicates that the amount of surface area exposed to solvent on unfolding is the major structural determinant of the m-value. The magnitude of the m-value is indicative of the co-operative nature of the interactions maintaining the folded state i.e., the larger the m-value the greater co-operatively between the non-covalent interactions maintaining the folded state.

4.2.2.3 Differential scanning calorimetry (DSC)

Differential scanning calorimetry is the accepted technique for determining the energetics of the protein unfolding/refolding pathway and for evaluation of thermodynamic parameters (Freire, 1995). In addition, it is the most definitive technique for evaluating the absence or presence of an intermediate (Privalov and Potekhin, 1986). The accurate evaluation of thermodynamic parameters obtained from the deconvolution of the DSC profile is limited for the cytosolic glutathione transferases since the glutathione transferase system for thermal unfolding is complicated by irreversibility and aggregation (Kaplan *et al.*, 1997 and this study).

For hGST A1-1, the DSC profile was best fitted to a two-state reaction in order to evaluate the calorimetric enthalpy (ΔH_{cal}) and to compare it with the van't Hoff enthalpy ($\Delta H_{v,H}$), obtained under similar thermally-induced conditions. The validity in calculating the enthalpy values is supported by the fact that experimental conditions were identical for both enthalpic determinations and hence any artifact which results from aggregation and therefore limits the interpretation of the values will contribute to both values and be eliminated when the ratio is calculated. The validity in using UV-melting curves to evaluate an enthalpy value is debatable because the massive increase in absorbance at 280nm is attributed to aggregation of the protein. However, the results of UV-melting do appear to provide evidence for a two-state unfolding mechanism.

The van't Hoff enthalpy calculated from the slope of figure 13b was 358.3kcal/mol and the calorimetrically determined enthalpy that was calculated from the area under the heat capacity profile (figure 13c) was 398.9kcal/mol. The ratio(r) of $\Delta H_{v,H}$ to ΔH_{cal} is used as a measure of the two-state behaviour. For the two-state assumption, r approximates to unity (Marky and Breslauer, 1987). For hGST A1-1, r = 0.90. The correlation supports the two-state approximation i.e., intermediates that differ from the folded or the unfolded protein in enthalpy are not significantly populated at equilibrium.

4.2.2.4 Urea SEC-HPLC

SEC-HPLC, under denaturing conditions is used to characterise any species along the equilibrium unfolding/refolding pathway provided that their rate of interconversion is sufficiently slow enough to be resolved during the time of a chromatographic run (Corbett and Roche, 1984). This technique is the most definitive determination of species along the equilibrium unfolding pathway since it enables the direct observation of the species populated at the midpoint of the transition (Chan *et al.*, 1995).

The elution profile for hGST A1-1, monitored by absorbance at 280nm, in the presence of increasing concentrations of urea is shown in figure 14. At the midpoint of the unfolding transition (i.e., at approximately 4.85M urea) only two species are populated, the native dimer and the unfolded monomer. The midpoint of the transition is consistent with that observed for 5μM protein when the transition was monitored using tryptophan fluorescence (figure 11a). The broad peaks and the increase in retention time at the unfolding transition is the result of solvation of the unfolded protein by denaturant. This has been observed for the class pi GST (Erhardt and Dirr, 1995) and other proteins (Uversky and Ptitsyn, 1994). The folded monomer would have a longer retention time than the native protein and its absence in the chromatographic profile supports the two-state (dimer/unfolded monomer) transition for hGST A1-1.

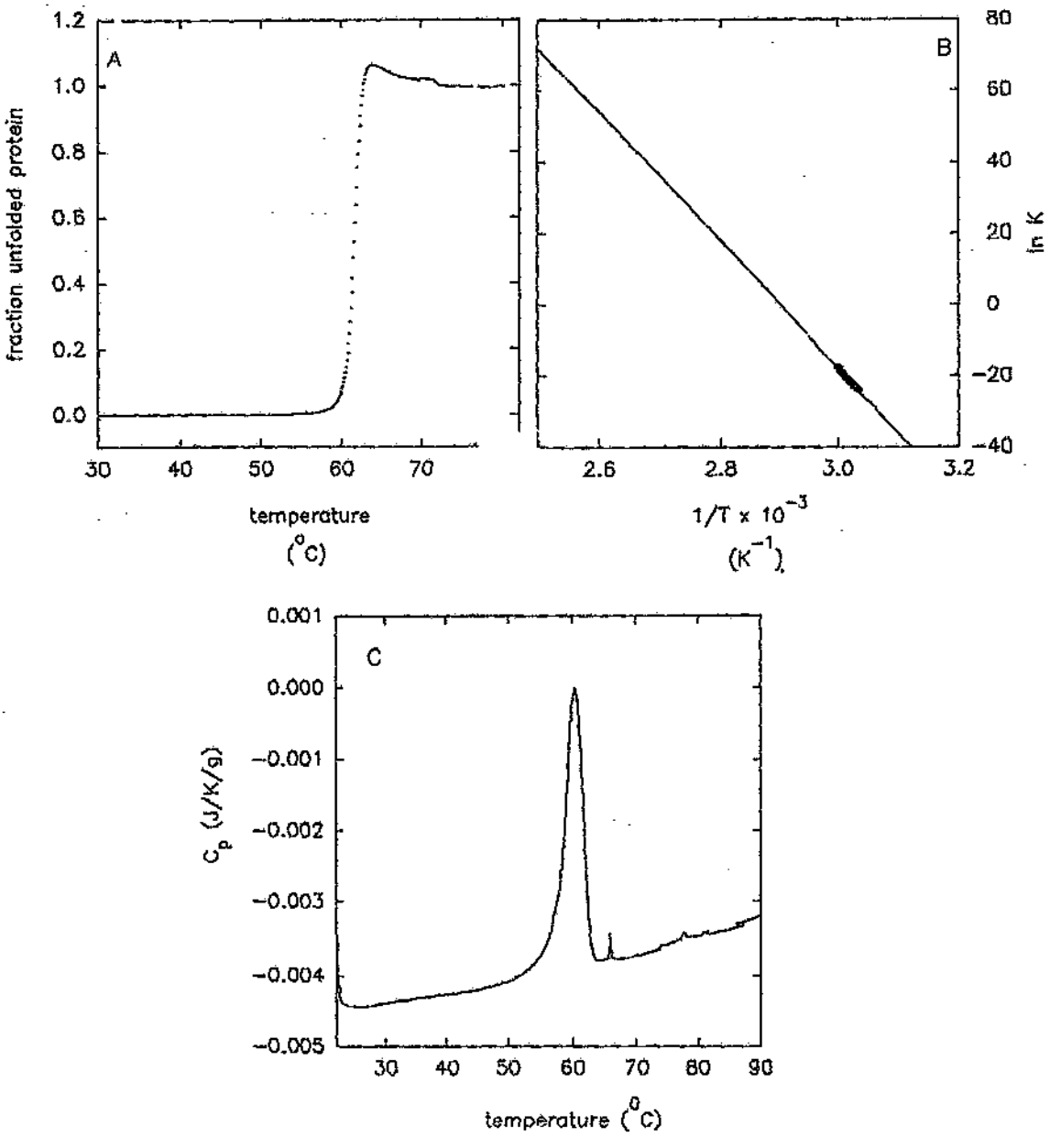


Figure 13. (a) Normalised UV-melting curve for hGST A1-1. Protein was heated at 1°C/min and the absorbance measured at 280nm. (b) van't Hoff plot of UV-melting unfolding transition in (a), the slope provides an estimate for the van't Hoff enthalpy ($\Delta H_{v,H}$)(as described in section 3.9.3). (c) DSC profile for 20μM hGST A1-1.

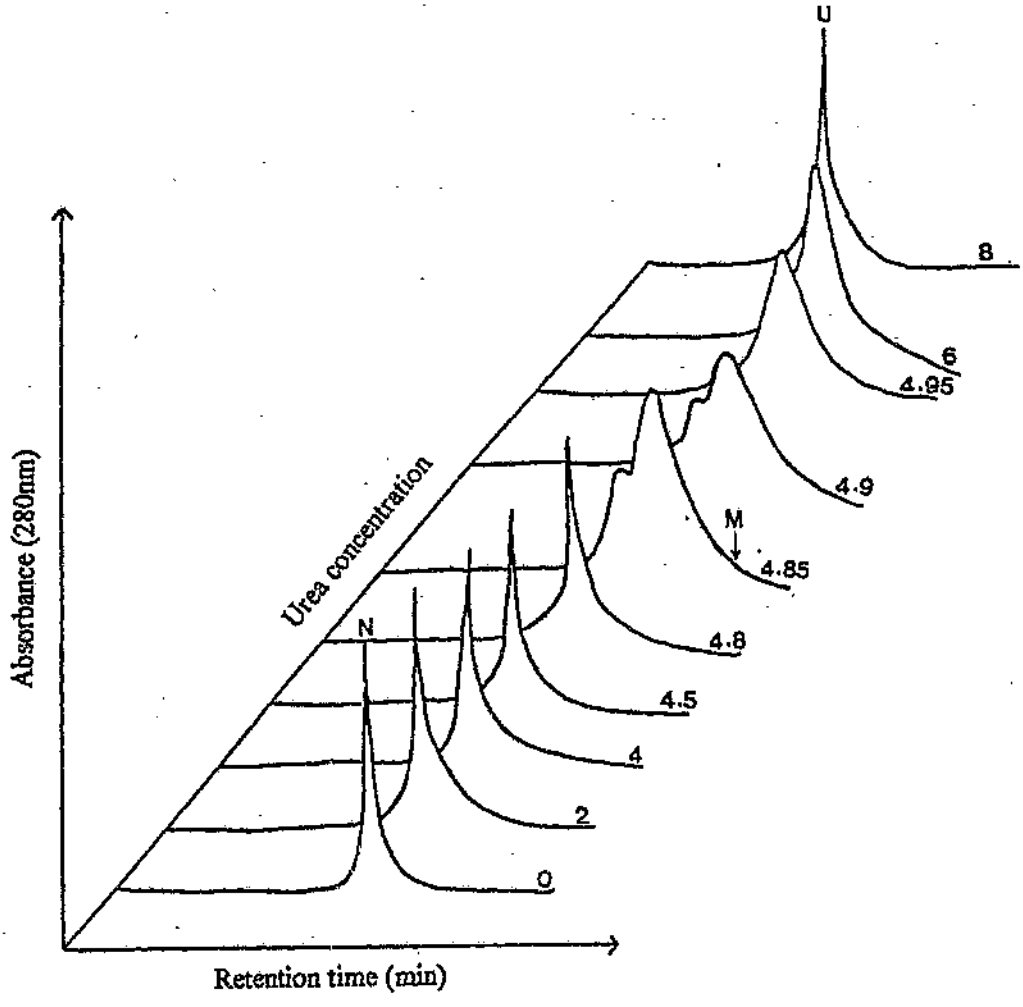


Figure 14. SEC-HPLC elution profile for 5μM hGST A1-1 in the presence of varying concentrations of urea. The profile was obtained by measuring the change in absorbance at 280nm. hGST A1-1 was incubated in the appropriate concentration of urea and applied to a SEC-HPLC-3000 column pre-equilibrated in the same concentration of urea. The position of native dimer (N) and unfolded monomer (U) are indicated. The expected position for a folded monomer (M) is indicated with an arrow.

4.2.2.5 Thermal-gradient gel electrophoresis (TGGE)

TGGE was used to characterise any species that may be trapped along the thermally induced unfolding pathway. The rate of migration of a protein through a gel matrix depends on its molecular dimensions and therefore a compact folded protein will be retarded to a lesser extent than the extended conformation of an unfolded protein. The ability of the gel matrix to trap any species depends on the rate of their interconversion relative to the time taken for the electrophoretic run (Creighton, 1979). The electrophoresis of hGST A1-1 across a temperature gradient from 20°-80°C indicates a single discontinuous transition (figure 15). The transition occurs between 50°C and 60°C and is similar to the transition obtained for the thermal inactivation study (section 5.4). The lack of distinction at the transition midpoint may be the result of either thermally induced aggregation or an intermediate exchange rate between the two species.

4.3 Unfolding kinetics of hGST A1-1

4.3.1 Fluorescence-detected unfolding kinetics

The fluorescence detected unfolding reaction of unmodified hGST A1-1 was followed using an excitation wavelength of 280nm for Tyr and Trp (figure 16a) and an excitation wavelength of 295nm for Trp alone (figure 16b). The unfolding kinetic traces for unmodified protein do not show a simple monophasic change (figure 16a/b and inset). Two kinetic phases were observed for final urea concentrations between 3.5M and 8.9M and over a 10°C to 40°C temperature range. The unfolding phases will be referred to as fast and slow unfolding events. The fluorescence intensity for the fast phase increases rapidly in the millisecond time range following which the fluorescence intensity for the slow phase decreases to its value predicted by equilibrium studies. At the final urea concentrations between 3.5 and 4.5M, the slow unfolding phase was monitored using manual mixing studies since the amplitude was not detectable using the stopped-flow apparatus.

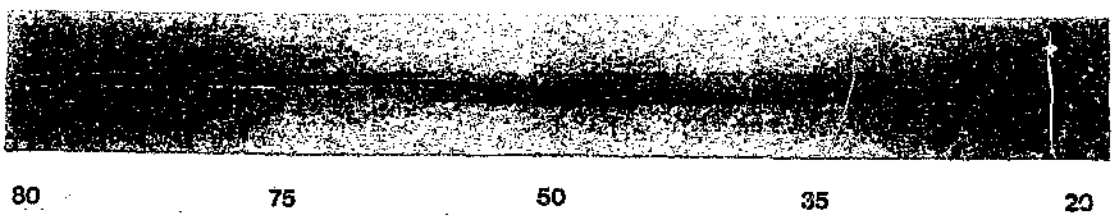


Figure 15. Thermal gradient gel electrophoretogram for hGST A1-1. Native hGST A1-1 was electrophoresed perpendicular to a linear thermal gradient from 20°C-80°C. The transition occurs between 50°C and 60°C.

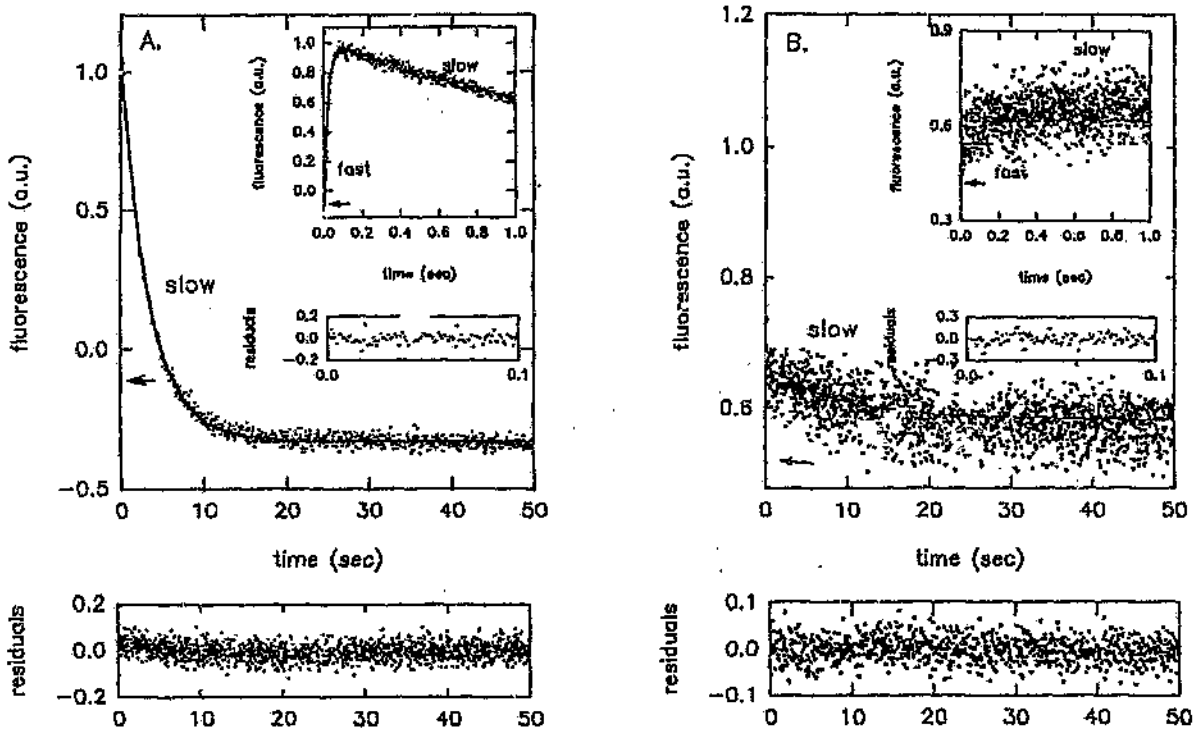


Figure 16. Unfolding kinetics. Final urea concentration of 8.3M, pH 6.5, 25°C. The change in fluorescence upon unfolding 1μM hGST A1-1 was monitored by excitation at (a) 280nm and (b) 295nm and emission monitored above 320nm. The fast and part of the slow phase are indicated in the inset of each panel and the slow phase in panel a or b. The arrows indicate the fluorescence signal for native folded protein unfolding traces were fitted with a single exponential function. The lower panels of each trace show the residuals.

The influence of protein concentration on the apparent unfolding rate constant was used to predict the order of the unfolding reaction. Protein concentration dependent studies were determined for the fast and the slow unfolding phases at three final protein concentrations (0.5μM, 1μM and 5μM) at 5.8M and 8.3M urea (table 5). The tenfold increase in protein concentration had essentially no effect on the unfolding reactions and therefore both reactions are unimolecular (first-order) and best described by a single exponential function. Therefore, the unimolecular reactio reflect unfolding events and not higher order aggregation.

For the final urea concentration of 8.3M, monitored using Tyr/Trp fluorescence, the rate constants for the fast (inset of figure 16a) and the slow phase (figure 16a) are 48.70 (±3.26)s^{-1} and 0.24 (±0.03)s^{-1}, respectively. Similarly at 8.3M urea, for the reaction monitored using Trp fluorescence the rate constants for the fast (inset of 16b) and the slow phase (figure 16b) are 47.60 (±6.09)s^{-1} and 0.25 (±0.06)s^{-1}, respectively. The residuals for the fit of each phase to a single exponential function are indicated in the lower panels of each kinetic trace. The initial and final amplitudes are as predicted by equilibrium experiments indicating that no burst phase occurs within the dead time. The rate constants obtained for the fast and slow phases for the different fluorescence experiments are, therefore, in agreement with each other. This confirms that Trp20 is a local reporter of events occurring at the interdomain interface (section 4.1). The standard deviations of the data obtained using Trp fluorescence as the unfolding probe are larger as a result of the smaller amplitude change and the low signal-to-noise ratio. The unfolding studies of unmodified hGST A1-1 were monitored using an excitation wavelength at 280nm because of the larger amplitude change in fluorescence that occurs upon unfolding.

A number of control experiments established that both phases were real unfolding events and that they were not the results of mixing artifacts. Identical "unfolding" experiments were performed using 2μM N-acetyl-L-tryptopharamide (NATA) and

Table 5. Protein concentration independence of the apparent unfolding rate constants for the fast and slow phases.

Phase	$[Urea]_f$ (M)	$[P]_t$ (μM)	k_{app} (s^{-1})
fast	5.8	0.5	18.74 ± 0.092
		5	18.19 ± 0.510
slow		0.5	0.0192 ± 0.0049
		5	0.0184 ± 0.0069
fast	8.3	0.5	45.15 ± 1.84
		5	51.75 ± 2.52
slow		0.5	0.221 ± 0.0171
		5	0.237 ± 0.0175

20μM N-acetyl-L-tyrosinamide instead of hGST A1-1. "Unfolding" monitored by excitation at 280nm and final urea conditions of 8.3M, yielded a horizontal signal response. In addition, if the fast phase were a mixing artifact or as the result of aggregation it would be observed at any wavelength as a consequence of a change in refractive index. Absorbance at 340nm, used to monitor aggregation in this study, shows a horizontal signal response with no change in amplitude. Finally, all control runs (buffer only, buffer/urea and buffer/protein) showed no signal change. In addition, in order to investigate if residual S-hexylglutathione (inefficiently removed after the purification procedure) was responsible for the biphasic unfolding properties of hGST A1-1, protein was prepared in the absence of ligand (Cameron *et al.*, 1995). The effect of residual ligand on the unfolding kinetics can be excluded since the protein prepared in the absence of ligand displayed identical biphasic behaviour.

The biphasic unfolding kinetics for unmodified hGST A1-1 suggests that more than the two species (i.e., folded dimer/unfolded monomer) predicted by urea-induced equilibrium studies must be present along the unfolding pathway. An explanation for multiple phases in unfolding may be the result of various folded forms in the pre-transition region (Touchette *et al.*, 1986). Since amplitudes depend on the relative initial populations of species present (Hagerman and Baldwin, 1976) it is obvious that any condition which will shift the pre-existing equilibrium will result in a change in amplitude of the kinetic phases. The unfolding amplitudes and rates for hGST A1-1 using the same final urea concentration (8M) but varying the initial urea concentration (up to 3M urea, where the protein is in the folded form) were measured. The change in the initial unfolding conditions had no effect on the amplitudes or the rates on the fast and slow phases (table 6) and this indicates that in the pre-transition baseline, the native molecules are indistinguishable from each other and their relative populations are similar. The absence of different native forms in the equilibrium pre-transition zone confirms that the fast unfolding phase for unmodified protein is an unfolding event.

Table 6. Apparent unfolding rate constants and amplitudes for the fa t and the slow unfolding phases obtained for the final urea concentration of 8M but varying the initial urea concentration.

[Urea] (M)	k_{app} (s^{-1})		Amplitude (A_f-A_i)	
	Fast	slow	fast	slow
0	32.8 ± 2.11	0.174 ± 0.029	-1.334	0.878
1.5	29.33 ± 0.169	0.155 ± 0.037	-1.386	0.9417
3	39.20 ± 3.24	0.178 ± 0.046	-1.28	0.754

Unfolding kinetics of Cys111-AEDANS hGST A1-1 monitored by energy transfer from Trp20 to Cys111-AEDANS (excitation at 295nm) (figure 17a) and the direct excitation of Cys111-AEDANS (excitation at 340nm) (figure 17b) showed a monophasic change. The single phase accounts for all of the amplitude predicted by equilibrium experiments. This single unfolding phase displayed similar urea-dependence properties as those observed for the slow unfolding phase monitored using tyrosine/tryptophan fluorescence. The absence of a detectable lag phase, for the identical time period in which the fast unfolding phase monitored using tryptophan fluorescence was observed, is depicted in the inset of figure 17a and b. For modified hGST A1-1 monitored using energy transfer from Trp20 to AEDANS at 8.3M urea, the rate constant for the single unfolding phase was 0.29 ($\pm$0.04)s^{-1}. Similarly, the rate constant for the single unfolding phase monitored by direct excitation of the AEDANS group covalently linked to Cys111 was 0.27 ($\pm$0.05)s^{-1}.

4.3.2 Urea-dependence of the unfolding rates and amplitudes

The rate of unfolding of unmodified/AEDANS-modified protein increases as conditions favour the unfolded state. For unmodified protein, the urea-dependence of the unfolding data (figure 18a) for both fast and slow phases fit well to a linear regression ($r = 0.998$ and $r = 0.996$, respectively) and similarly for the single phase observed for Cys111-AEDANS hGST A1-1 ($r = 0.991$). The linear dependence of the unfolding rate on denaturant suggests that no changes occur in the rate limiting step and that the unfolded species formed differ from the native state of hGST A1-1 with respect to solvent exposure. The urea-dependence of the rate for the slow unfolding phase and the intermediate refolding phase has a minimum at approximately 4.5M urea (figure 32a). This is consistent with the midpoint of the equilibrium-unfolding curve for 1μM protein (at 4.5M urea) and demonstrates the reversibility of this unfolding/folding system. Refolding studies with hGST A1-1 are consistent with a rate limiting step which is urea and protein concentration dependent (see chapter 6) and the slow unfolding phase for unmodified hGST A1-1 and AEDANS-modified hGST A1-1 is postulated to represent the complete unfolding of the dimer. Therefore,

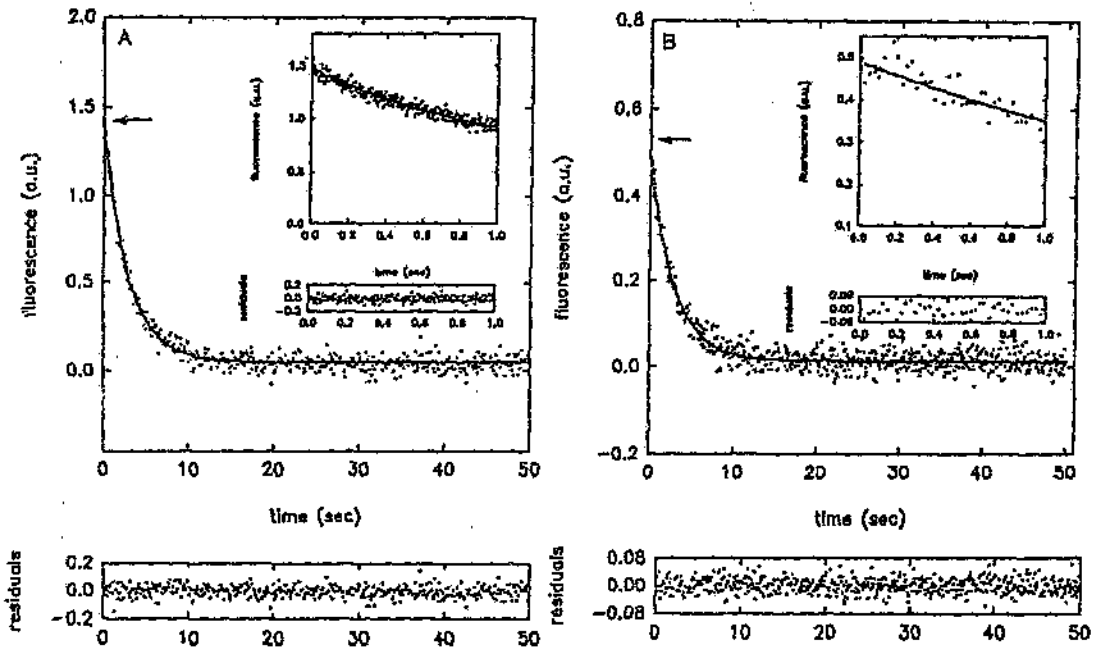


Figure 17. Unf ding kinetics. Final urea concentration of 8.3M, pH 6.5, 25°C. The change in fluorescence upon unfolding 1μM Cys111-AEDANS hGST A1-1 was monitored by excitation at (a) 295nm and emission above 320nm and (b) 350nm and emission above 400nm. The absence of the fast phase is indicated in the inset of each panel and the slow phase in panel a or b. The arrows indicate the fluorescence signal for native folded protein. The unfolding traces were fitted with a single exponential function. The lower panels of each trace show the residuals.

the nature of the rate-limiting step for folding and unfolding involves the transition between two unfolded monomers and one dimer.

The linear dependence of the apparent unfolding rates for the fast and the slow phases for unmodified hGST A1-1 and for the single phase for AEDANS-modified hGST A1-1, even under strongly destabilising conditions, enables the complete characterisation of the unfolding pathway. For unmodified hGST A1-1, the apparent rate constant for unfolding of the fast phase is less urea sensitive than the slow phase (figure 18a). The slopes (m_u) of the fast and slow phases are 119.5cal/mol/M of urea and 270cal/mol/M of urea, respectively. The fast phase, therefore, results in a transition state that is more closely related to the native state of hGST A1-1 in terms of solvent exposure. For modified hGST A1-1, the single unfolding phase has a slope (m_u) of 286.2cal/mol/M of urea. The similar urea-dependence for unmodified and modified hGST A1-1 supports the suggestion that this phase is similar to the slow phase observed for unmodified protein.

The solvent accessibility of the transition states for unfolding were characterised by analysis of the susceptibility of the apparent rate constant to urea. The ratio of m_u/m (where m_u is $RTlogk_u/[urea]$ (figure 18a) divided by the m-value obtained for 1μM hGST A1-1 in equilibrium studies (4.22 (±0.72)kcal/mol per M of urea)) is an indication of the increase in solvent exposure of the transition state relative to the native state, normalised by the increase in solvent exposure between the native and denatured state (Tanford, 1970). The rate limiting transition state for unfolding and refolding can only be characterised for kinetic data that is monoexponential (Doyle *et al.*, 1996). For Cys111-AEDANS hGST A1-1, the monophasic change enables the ratio of m_u/m to be estimated. The ratio of m_u/m is 0.068, which is suggestive of a transition state for the single unfolding phase of modified protein with a solvent accessibility closely related to the native state, i.e., 6.8% of the buried hydrophobic surface in the native state is exposed at the transition state. The slow unfolding phase for unmodified protein has a m_u/m ratio of 0.064, which is comparable with that for

AEDANS-modified protein. The fast unfolding phase for unmodified protein has a m_u/m ratio of 0.028 indicating that this transition state relates more closely to the accessible surface area of the native protein. A ratio that approximates to zero indicates that the solvent accessibility of the native and transition state are similar whilst a ratio near to unity is indicative of a transition state that is as solvent exposed as the denatured state (Doyle *et al.*, 1996). For small proteins, such as the Arc repressor (Waldburger *et al.*, 1995) and the IgG binding domain of protein L (Scalley *et al.*, 1997), m_u/m is typically 0.3. There is, however, considerable variation in the ratio with no significant correlation between the ratio and the proteins structural features, such as compactness of the hydrophobic core (Scalley *et al.*, 1997).

The linearity of the urea dependent data enables an estimation of the rate of unfolding of hGST A1-1 under roughly physiological conditions. For unmodified hGST A1-1, the unfolding rates in the absence of denaturant (k_u (H_2O)) are $0.61s^{-1}$ and $2.92 \times 10^{-5}s^{-1}$ for the fast and slow phases, respectively (as obtained from the zero urea intercepts in figure 18a). These results correspond to time constants of 1.6s and 9.5 hours for the fast and slow phases, respectively. Cys111-AEDANS hGST A1-1 has a similar zero urea intercept ($3.07 \times 10^{-5}s^{-1}$) and time constant for unfolding in the absence of denaturant (9.0 hours) as the slow unfolding phase for unmodified protein. The unfolding and dissociation of proteins *in vivo* varies (Milla and Sauer, 1994). The biological relevance of a slow unfolding rate *in vivo* is related to the stability and physiological function of the protein. The evolutionary pressure placed on the unfolding rate constant varies with the function of the protein (Burton *et al.*, 1996). For example, the fastest folding proteins, ubiquitin, CspB and the Arc repressor, all have regulatory functions. There is, therefore, a selective advantage in reducing the kinetic barrier to unfolding such that the protein may be rapidly degraded at the appropriate time. In general, the slower the unfolding rate the greater the stability of the protein and the less susceptible the protein is to proteolytic degradation.

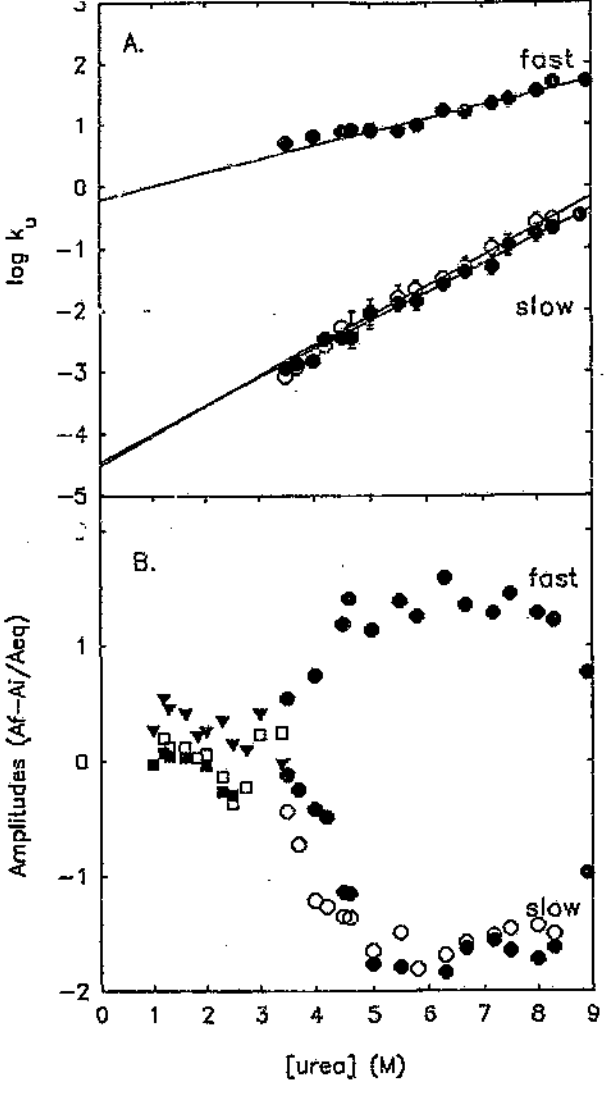


Figure 18. (a) Urea-dependence of the unfolding rate constants for 1μM unmodified hGST A1-1 (•) and 0.8μM Cys111-AEDANS hGST A1-1 (o). The solid lines are the curve fits using equation 15. Unfolding experiments were performed at 25°C, 0.1M NaCl, pH 6.5. Data were fitted to equation 14 for urea concentrations between 4.5M and 5.8M urea and to equation 13 for final urea concentrations greater than 5.8M. The error bars are as indicated or are comparable to the size of the symbol. (b) The urea-dependence of the amplitudes for unfolding (•) and refolding (▼.□.■) of unmodified hGST A1-1 and for the unfolding of Cys111-AEDANS hGST A1-1 (o). The amplitudes were normalised with respect to the fraction-unfolded measurements obtained in equilibrium studies.

The urea-dependence of the amplitudes for the fast and slow unfolding phases for unmodified protein and for the single unfolding phase for modified protein is depicted in figure 18b. The relative amplitudes observed for the unfolding kinetics were about 100% of the total change expected. Both the fast and slow phases occur simultaneously from 3.5M urea. The relative percentages of the two species increased from 3.5M to 5M urea and then became independent of the urea concentration when the final urea concentration exceeded 5M urea. The inverse amplitudes are the result of the different fluorescence properties of each phase. The occurrence of both phases from time zero suggests that the slow phase overlaps the fast phase and it is, therefore, quite likely that the parameters (m_u and $k_u(H_20)$) obtained for the fast unfolding phase are underestimated. The amplitudes for the single unfolding phase for AEDANS-modified protein displayed similar characteristics to those for the slow unfolding of unmodified protein. The amplitudes for the three refolding phases are included to illustrate the consistency of the unfolding/refolding transition for hGST A1-1 (chapter 6).

4.3.3 Temperature dependence of the unfolding rates

The apparent unfolding rate of the fast and slow phases of unmodified hGST A1-1 were characterised from 10°C - 40°C at a final urea concentration of 8.3M. The temperature dependence for the single unfolding phase for AEDANS-modified protein was characterised at a final urea concentration of 8M. For unmodified hGST A1-1, both phases are temperature dependent with the apparent unfolding rate and amplitude of the fast phase increasing to such an extent at 40°C that the fast phase occurred within the dead time.

The rate constant for unfolding was related to the Gibbs free energy ($\Delta G^{\ddagger}$) of activation using the transition state theory (Tan *et al.*, 1996). The temperature dependence of the activation energy enables an estimation of the apparent

thermodynamic properties (i.e., enthalpy, entropy and heat capacity) for the transition state. These properties reflect the flexibility of the polypeptide backbone and side chains as well as the solvation of certain protein moieties in the transition state.

For unmodified hGST A1-1, the plot of ln ($k_u h/k_B T$) versus 1/T (figure 19a) for both phases is almost linear. Both phases fit equally well to equation (18) and to a linear fit (see residuals figure 19b). The best fit for equation (18) to the fast unfolding phase of unmodified protein gives the following thermodynamic parameters: $\Delta C_{pU}^{\ddagger}$ of 253.3cal^{-1}mol^{-1}K^{-1}; $\Delta H_U^{\ddagger}$(298K) of 20.45kcal.mol^{-1}; $\Delta S_U^{\ddagger}$(298K) of 17.45 cal^{-1}mol^{-1}K^{-1}. Fitting the data of the slow unfolding phase of unmodified protein to equation (18) gives: $\Delta C_{pU}^{\ddagger}$ of 475.0cal^{-1}mol^{-1}K^{-1}; $\Delta H_U^{\ddagger}$(298K) of 18.90kcal.mol^{-1}; $\Delta S_U^{\ddagger}$(298K) of 2.28cal^{-1}mol^{-1}K^{-1}. For the single unfolding phase of Cys111-AEDANS hGST A1-1, the temperature dependence of the apparent unfolding rate constant shows distinct curvature when fitted to equation (18), the residuals indicate that the unfolding rate is best fitted to this equation than to a linear function. The curvature occurs at the temperature extremes and may indicate systematic deviation of the data rather than a large heat capacity ($\Delta C_{pU}^{\ddagger}$) change as the protein unfolds. The best fit for equation (18) to the single phase for modified protein gives: $\Delta C_{pU}^{\ddagger}$ of 987.0cal^{-1}mol^{-1}K^{-1}; $\Delta H_U^{\ddagger}$(298K) of 18.46kcal.mol^{-1}; $\Delta S_U^{\ddagger}$(298K) of 0.878cal^{-1}mol^{-1}K^{-1}. The errors from the fit to equation (18) are large since the curvature is marginal in the experimentally accessible temperature range (Tan *et al.*, 1996).

The linearity of the temperature-dependent data, monitored using tyrosine/tryptophan fluorescence, is consistent with a small heat capacity change ($\Delta C_{pU}^{\ddagger} \approx 0$) between the native state and the transition states. The apparent heat capacity value for the fast phase is smaller than that of the slow phase and therefore the transition state for the fast phase is more closely related to the native state in terms of compactness. There are many factors that contribute to the positive heat capacity changes for unfolding such as the increase in configurational freedom of the polypeptide chain due to the

disruption of tertiary interactions. In general, however, the hydration of hydrophobic groups that become exposed upon unfolding contributes the most to the heat capacity change (Privalov and Gill, 1988).

The change in the heat capacity (ΔC_{pN-U}) value for unfolding, as for the m-value, can be estimated by relating the surface area exposed to the size of the protein (Myers *et al.*, 1995). This correlation (r = 0.97) is described by $\Delta C_{pN-U} = -251 + 0.19(\Delta ASA)$ where ΔASA is the change in the solvent-accessible surface area upon unfolding and is 40199Å^2 for hGST A1-1. For hGST A1-1, the heat capacity change for complete unfolding is estimated to be 7.39kcal^{-1}mol^{-1}K^{-1}. To determine the relative heat capacity changes this value is used instead of the value obtained calorimetrically (DSC) since the glutathione transferase system for thermal unfolding is complicated by irreversibility and aggregation (Kaplan *et al.*, 1997 and section 4.2.2.3). The extent of buried surface area for each transition state ($\Delta C_{pU}^{\ddagger}/\Delta C_{pN-U}$) is 3.4% and 6.5% for the fast and slow unfolding phases of unmodified protein, respectively and 13.5% for the single unfolding phase of Cys111-AEDANS hGST A1-1. The compactness of the unfolding transition states is, therefore, in agreement with the urea and temperature data shown above.

The activation energy (calculated using $\Delta H_U^{\ddagger} = E_A - RT$ (Oliveberg *et al.*, 1995) or obtained from a linear Arrhenius plot) is 21.04 (± 1.89) kcal.mol^{-1} and 19.49 (± 1.51) kcal.mol^{-1} for the fast and slow unfolding phases for unmodified protein, respectively. Similarly, for the single unfolding phase for Cys111-AEDANS hGST A1-1 the activation energy is 19.05 (± 2.15) kcal.mol^{-1}. A high activation energy for unfolding has also been found for other proteins, for example, dihydrofolate reductase (Touchette *et al.*, 1986) and T4 phage lysozyme (Chen *et al.*, 1989) and is indicative of the large number of forces required to maintain the native structure (Matthews, 1993). For the slow phase, the time range and the activation energy could be indicative of *cis* X-Pro peptide bond isomerisation. The *cis/trans* isomerisation of a

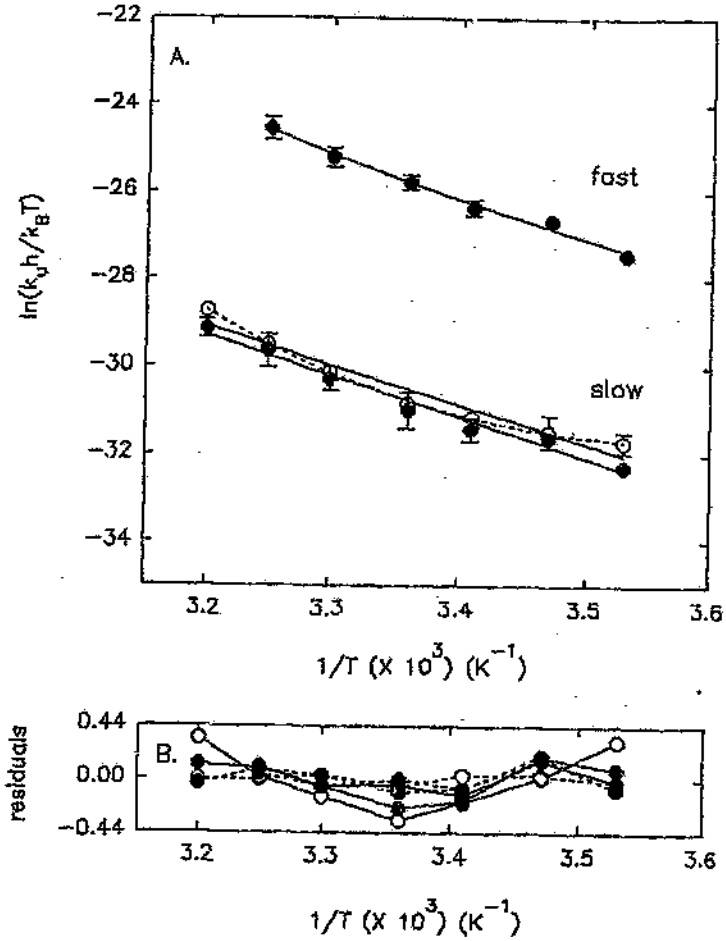


Figure 19. (a) Eyring plots for the temperature dependence of the unfolding rate constant for 1μM unmodified hGST A1-1 (•) and 0.8μM Cys111-AEDANS hGST A1-1 (o). Unfolding experiments were performed at 8.3M urea, 0.1M NaCl, pH 6.5, for unmodified hGST A1-1 and 8M urea, 0.1M NaCl, pH 6.5, for Cys111-AEDANS hGST A1-1. The error bars are as indicated or are comparable to the size of the symbol. The plots for unmodified hGST A1-1 (•) fit equally well to a linear function (—) and equation 18 (---). The plot for Cys111-AEDANS hGST A1-1 (o) is best fitted to equation 18(---). The residuals for fitting to a linear function (solid lines, • unmodified hGST A1-1, o Cys111-AEDANS hGST A1-1) and equation 18 (dashed lines, • unmodified hGST A1-1, o Cys111-AEDANS hGST A1-1) are indicated in panel b.

peptide bond preceding proline is an intrinsically slow process (τ = 10-100sec) with a high activation energy (≈ 20kcal.mol^{-1}) (Nall *et al.*, 1978). However, the urea-dependence of the slow phase excludes the possibility of such an isomerisation.

The large standard deviation of the thermodynamic parameters limits an accurate evaluation of the absolute values for the enthalpic and entropic contributions to unfolding. The enthalpy change is large and positive, consistent with a net gain in enthalpy as the protein unfolds. The net enthalpy change of the transition state for protein unfolding is a balance between a gain in enthalpy due to disruption of van der Waals forces and hydrogen bonding between residues within the protein; a loss in enthalpy due to the solvation of hydrophobic patches by water and a gain in enthalpy when the interactions between the water and the solvated non-polar groups are broken. The solvation of hydrophobic patches and the subsequent ordering of the water around these groups into "icebergs" gives rise to the hydration effect which results in a loss in the enthalpy because of the gain in hydrogen bonding (Kauzmann, 1959). Similarly, the positive entropy gain indicates a balance between the increase in entropy due to disordering of the polypeptide backbone and side chains and an entropy loss due to solvation ("ordering") of water around the non-polar groups.

The apparent activation Gibbs free energy in the absence of denaturant ($\Delta G_U^{\ddagger}(H_20)$) calculated for the fast and slow unfolding phases according to equation 16, are 17.74kcal.mol^{-1} and 23.59kcal.mol^{-1}, respectively. Therefore, the slow unfolding phase represents the transition state for the overall unfolding pathway.

4.4 The non-substrate ligand (ANS) and the unfolding pathway

The dimeric structure of the glutathione transferases results in the formation of a solvent accessible V-shaped cavity between the two subunits (section 2.2.3.3). The cavity (the L-site) has been implicated to be the binding or transport site for the hydrophobic non-substrate ligand, 8-anilino-1-naphthalene sulphonate, ANS (Sluis-

Cremer *et al.*, 1996). ANS binds to the class alpha GST with a relatively high affinity (K_d value ≈ 16μM)(chapter 5).

The urea-induced unfolding curve of hGST A1-1 in the absence and presence of the ligand are superimposable, within experimental error. The coincident data, therefore, indicates that the two-state unfolding pathway for hGST A1-1 is essentially unaltered by the presence of ligand (figure 20a). The conformational stability ($\Delta G(H_20)$) of hGST A1-1 in the absence and in the presence of ANS was estimated to be 27.5 (± 3.79) kcal/mol and 25.55 (± 4.34)kcal/mol, respectively. Similarly, the m-values indicate that the co-operativity of the unfolding transition is unaffected by the presence of ANS. In addition, the DSC profile for hGST A1-1 in the presence of 200μM of ANS indicated a single transition with a similar shape and transition midpoint (≈58°C) as that for uncomplexed hGST A1-1 (figure 20b).

Stopped-flow fluorescence-detected unfolding for hGST A1-1 in the presence of ANS (emission >400nm) indicated that the fast and slow unfolding phases have the same apparent first order rate constant and amplitude as the uncomplexed protein. At a final urea concentration of 6.5M in the absence of ANS, the rates for the fast and slow unfolding phases were 16.24 (±0.57)s^{-1} and 0.028 (±0.002)s^{-1}, respectively. Similarly, in the presence of 100μM ANS, the rate constants for the fast and slow unfolding phases were 17.02 (± 2.36)s^{-1} and 0.025 (± 0.002)s^{-1}, respectively.

This data is consistent with conformational stability studies of the class pi GST in the presence of ANS and the drug, praziquantel (Ohlmeyer; MSc dissertation 1997) and for *S. japonicum* GST in the presence of the drug, praziquantel (Kaplan; PhD thesis, 1997). The results suggest that the interactions involved in sequestering hydrophobic ligands at the L-site and the small conformational changes induced by their binding (Bico *et al.*, 1995; McTigue *et al.*, 1995) does not impact on the equilibrium and

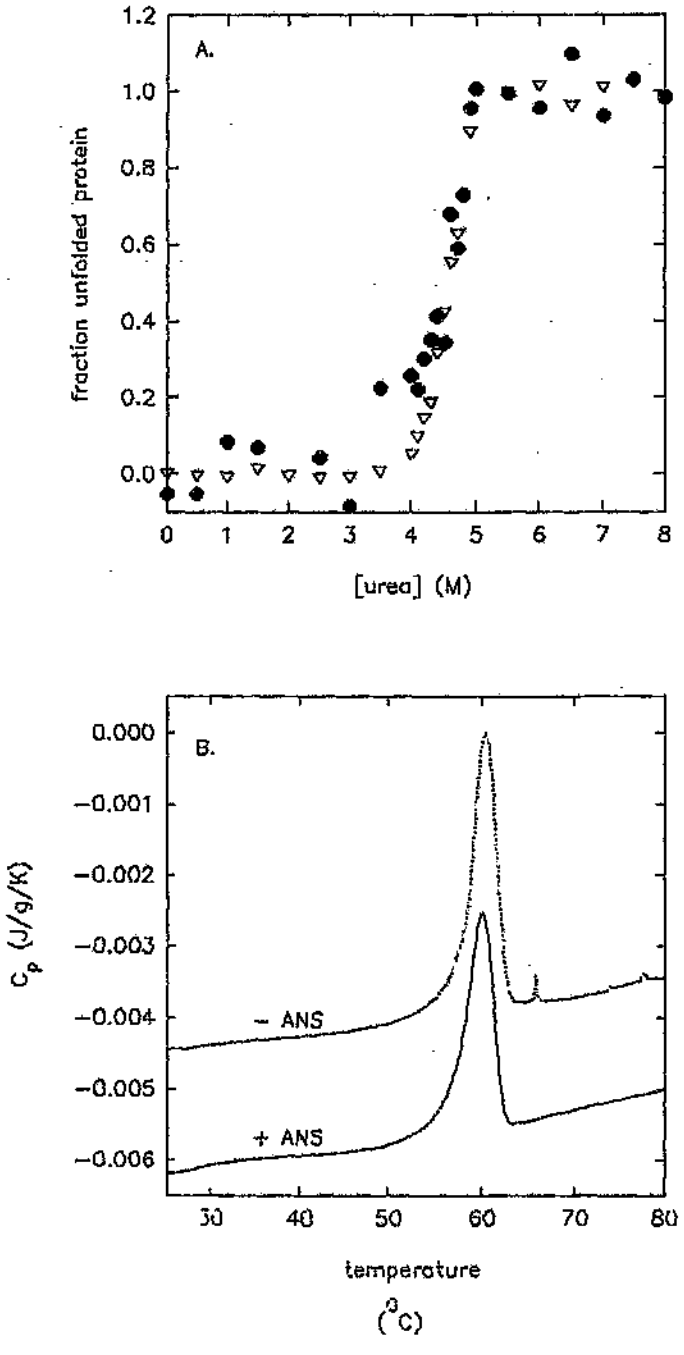


Figure 20. (a) Urea-induced equilibrium unfolding curves for 1μM hGST A1-1 in the absence (▽) and the presence (•) of 200μM ANS. Trp was selectively excited at 295nm and emission measured at 325nm and 355nm. (b) DSC profile for 20μM hGST A1-1 in the absence (...) and presence (—) of 400μM ANS.

kinetic unfolding pathway of glutathione transferases. In addition, the rather large K_d for ANS is consistent with a minimal effect on stability. This is in contrast to conformational stability studies in the presence of G-site ligands (Erhardt and Dirr, 1996; Parsons; MSc dissertation, 1997) where the presence of these ligands influenced the stability of the glutathione transferases.

4.5 Discussion

The equilibrium unfolding of hGST A1-1 is consistent with a two-state transition between native dimer (N_2) and unfolded monomer (U) (i.e., $N_2 \leftrightarrow 2U$). This simple dynamic equilibrium is consistent with an "all-or-none" process for unfolding of a homogeneous population of native protein and is well documented for a large number of proteins (Schmid, 1992) including the glutathione transferases (Dirr and Reinemer, 1991; Erhardt and Dirr, 1995; Kaplan *et al.*, 1997). The single transition, therefore, implies that all molecular conformations of hGST A1-1 can be divided into either folded dimer or unfolded monomers, which are separated by a high-energy barrier. The absence of an observable intermediate indicates that any intermediate(s) are insufficiently populated and too short-lived.

The equilibrium unfolding of several dimeric proteins have been analysed satisfactorily using the two-state denaturation model (Neet and Timm, 1994). The absence of a stable folded monomeric or dimeric intermediate for proteins such as the Arc repressor (Bowie and Sauer, 1989), the *Trp* aporepressor (Gittelmann and Matthews, 1990) and the cytosolic glutathione transferases suggests that each monomer cannot be separated into distinct folding units. All subunits of the glutathione transferases have been described, albeit subjectively by eye, as comprising of two domains. Domain I and II of each subunit were identified by visual inspection of the crystal structure of the porcine class pi glutathione transferase (pGST P1-1) (Reinemer *et al.*, 1991).

An algorithm described by Siddiqui and Barton (1995) was designed to accurately identify domains. The algorithm explicitly locates domains that comprise of one or two continuous segments of the protein chain but may be used to locate domains that include more than two segments. The algorithm identified each subunit of the glutathione transferase as being comprised of only one domain. The reason for this may be attributed to the large number of inter-domain interactions. Therefore, for the cytosolic glutathione transferases the stability of the individual domains may not be significant and the dissociation or the association of the two domains would not be detectable.

For hGST A1-1, as for all cytosolic glutathione transferases studied thus far, the free energy of stabilisation of the native three-dimensional state is high (approximately 25kcal/mol) and the transition between native dimer and unfolded monomer is highly co-operative. This implies that the quaternary structure (i.e., the subunit-subunit interactions) of all cytosolic glutathione transferases is necessary for stabilisation of the tertiary structure of each subunit. All cytosolic glutathione transferases have distinct quaternary structures which are comprised of identical or non-identical subunits with interactions at the subunit interface formed primarily between domain I (alpha helix 2) of one subunit and domain II (alpha helix 3 and 4) of the adjacent subunit. The minimisation of the accessible surface area (~14%) is the driving force for dimer formation (section 2.2.2).

In general, the alpha/mu/pi enzymes display greater stability than their evolutionary forerunners namely the class theta and sigma (unpublished data from our laboratory). This may be explained in part by the differences observed between the subunit interfaces. At the end of the interface two different types of subunit-subunit interactions are apparent (section 2.2.2). For the alpha/mu/pi classes there exists the hydrophobic "lock-and-key" type interface (figure 4) in contrast to the more hydrophilic interface present in the sigma and theta classes (Ji *et al.*, 1995). The "lock-and-key" type interface whilst important for the stabilisation of the native three

dimensional structure is not a criteria for dimer formation, at least for the class theta and sigma GSTs.

The unfolding kinetics of unmodified hGST A1-1 is biphasic with a fast unfolding event and a slower main unfolding event. The increase in intensity for the fast phase indicates t p20 is transiently exposed to solvent and is postulated to represent the partial dissociation of domain I and II of each subunit. Complexities in kinetic unfolding reactions are rare since the transition state for the unfolding pathway is structurally related to the native state. For many monomeric proteins (Kim and Baldwin, 1990) and for the few oligomeric proteins studied (Milla and Sauer, 1994; Liang and Terwilliger, 1991; Gittelman and Matthews, 1990) unfolding kinetics are monophasic, consistent with an "all-or-none" unfolding mechanism for a homogeneous population of protein. The single unfolding reaction implies the existence of a single transition state or a set of energetically similar transition states.

Complex unfolding reactions have been observed for some proteins and originate from either multiple states of the native protein (example, dihydrofolate reductase (DHFR) (Touchette *et al.*, 1986) or from kinetic coupling between peptide bond isomerisation (example, RNase (Houry *et al.*, 1994, Houry and Scheraga, 1996)) and the actual folding reaction (Kiefhaber, 1995). Proteins that unfold via a structural intermediate often display non-linear denaturant dependence of the rate of unfolding (referred to as a "rollover") at high denaturant concentrations. The "rollover" is indicative of a change in the rate limiting step and has been observed for the P22 Arc repressor (Jonsson *et al.*, 1996); Staphylococcal nuclease (Walkenhorst *et al.*, 1997) and *Rhodobacter capsulatus* cytochrome c_2 (Sauder *et al.*, 1996). The absence of a "rollover" for hGST A1-1 suggests that there is no change in the rate limiting step and therefore the fast unfolding event represents a partial dissociation of the two domains rather than the formation of a structured intermediate.

Urea and temperature characterisation of the apparent unfolding rate constants enable the location of the transition states for unfolding and the transition state properties are described by the relative kinetic m-values (m_u/m) and the relative heat capacity values ($\Delta C_{pU}^{\ddagger}/\Delta C_{pN-U}$). The transition state for the fast unfolding event (3%) is more closely related to native protein than the transition state for the main unfolding event (6%). This is consistent with a transition state for each event that has an extensive system of intramolecular interactions (not necessarily the same as those in the native protein) and that has a structure that is compact. The apparent activation Gibbs free energy changes for unfolding in the absence of denaturant, indicate that the slow unfolding phase has the largest apparent Gibbs free energy change (23.59kcal.mol^{-1}) and therefore represents the overall transition state for the unfolding pathway of hGST A1-1.

A sequential unfolding pathway, based on structural and experimental evidence, is proposed to describe the unfolding properties of hGST A1-1:

$$N_2 \rightarrow N_2^* \rightarrow 2U_A$$

where N_2 represents the folded dimeric state, N_2^* represents a native-like dimer or dimeric intermediate and U_A is an unfolded monomeric state of hGST A1-1. The inclusion of an additional isomerisation reaction between different unfolded states of hGST A1-1 is supported by crystallographic evidence (Sinning *et al.*, 1993). The reaction, $2U_A \leftrightarrow 2U_B$, would represent the *cis/trans* isomerisation of X-Pro peptide bonds. *Cis/trans* proline isomerisation was not observed since the conditions used were insensitive to structures that vary only in their bond orientations. There are ten proline residues per subunit of hGST A1-1, one of which is in the *cis* configuration (*cis* Val54-Pro55). The importance of *cis* X-Pro55 in maintaining the structural integrity of the enzyme has been investigated using site-directed mutagenesis (Wang *et al.*, 1993a).

The pathway proposed incorporates the two events observed using Tyr/Trp and only Trp fluorescence as well as the single event monitored by energy transfer to or direct excitation of Cys111, which was covalently labelled with the fluorescent AEDANS group. For the pathway monitored using Tyr/Trp fluorescence, the fast unfolding event can be described by the $N_2 \rightarrow N_2^*$ reaction. This event is proposed to represent the partial dissociation of the two structurally distinct domains, at the domain interface, near the Trp20. The dissociation results in a native-like dimeric state or dimeric intermediate that resembles the folded state in terms of solvent exposure and hydrophobic interactions. The fast unfolding event is followed by a slower unfolding event ($N_2^* \rightarrow 2U_A$) which represents complete dissociation and unfolding of the native-like dimer into two unfolded monomers. This event is also represented by the single phase observed using AEDANS labelled Cys111 hGST A1-1. The two-state (dimer/monomer) equilibrium unfolding transition excludes the presence of a stable folded monomeric state along the unfolding pathway. This slow unfolding event observed using Tyr/Trp fluorescence and the single event observed using AEDANS labelled hGST A1-1 have similar urea and temperature properties suggesting that their transition states occupy similar positions on the reaction profile based on solvent accessibility.

4.6 Conclusion

The unfolding properties of hGST A1-1 are described by a two-step sequential unfolding pathway which incorporates the two-state folded dimer/unfolded monomer equilibrium transition and the fluorescence-detected stopped-flow unfolding kinetics. The pathway involves the formation of a transient dimeric intermediate with partially dissociated domain I and II, at or near the Trp20, of each subunit. The intermediate then unfolds completely into two unfolded monomers. The urea-dependent and thermodynamic activation parameters for the unfolding pathways indicate that the transition states for the unfolding pathway are native-like and compact suggesting that the disruption of the hydrophobic core during unfolding of hGST A1-1 is partly rate-determining.

CHAPTER 5

THE STABILISATION OF THE HYDROPHOBIC CORE OF DOMAIN II OF GLUTATHIONE TRANSFERASES BY A TOPOLOGICALLY CONSERVED RESIDUE IN ALPHA HELIX 6

5.1 Hydrophobic core in domain II of glutathione transferases

The three-dimensional structures for all gene classes of the glutathione transferases show similar archetypal folds (see Dirr *et al.*, 1994b; Wilce and Parker, 1994) (section 2.2). An interesting feature of the three-dimensional structure is the presence of two hydrophobic cores, one located in the smaller N-terminal alpha/beta domain (domain I) and the other located in the larger C-terminal all-alpha-helical domain (domain II) (Sinning *et al.*, 1993).

In domain I, the hydrophobic core is formed by the packing of the alpha helices 1 and 3 against the β-sheet arrangement. The β-sheet is formed by the arrangement of four beta strands in the order beta sheet 2, 1, 3 and 4 with the beta strand 3 aligned anti-parallel to the others. The alpha helix 2 packs against one edge of the β-sheet and the alpha helices 1 and 3 pack against the other edge to form a compact hydrophobic core (Sinning *et al.*, 1993). In domain II, the alpha helices 4, 5, 6 and 7 with their connecting loops pack against each other and the interior forms the hydrophobic core.

In the hydrophobic core of domain II, a hydrophobic amino acid residue (Leu164, alpha; Leu157, pi; Leu163, mu; Leu160, sigma; Leu158, *S. japonicum*; Val163, theta) is topologically conserved throughout the gene classes. A sequence alignment and an overlay of ribbon diagrams of the alpha helices 6 and 7 for all the gene classes is depicted in figure 21a and b, respectively. A comparison of the secondary structural elements of the representative gene classes is shown in table 1. In general, the positions and lengths of the alpha helices 4, 5, 6, 7 and 8 for all classes of GSTs are similar.

The effect of an amino acid replacement in the hydrophobic core of domain II was investigated using hGST A1-1. Leu164 in the class alpha GST has a solvent accessible surface area (SASA) of 2Å^2 (Sinning *et al.*, 1993). The aliphatic residue, Leu164 (in alpha helix 6) is encapsulated in a cavity lined by Phe133 (in alpha helix 5), Leu160, Val161, Val167 (all in alpha helix 6) and Leu181, Leu184 and Lys185 (all in alpha helix 7) (figure 21c). The truncation of the Leu164 side chain by replacement with Ala was designed to minimise gross structural rearrangements of the folded state of the protein. The deletion of the side chain will remove defined van der Waals interactions without introducing new ones and in this way prevent any change in the stereochemistry of the protein. Valine was not chosen as the replacement amino acid because it is the topologically equivalent residue in the theta class glutathione transferase and the substitution to Ile was not considered because of the change in the side chain branching.

5.2 Characterisation of L164A[2] DNA

The pertinent segment of the wild-type nucleotide sequence encoding hGST A1-1 and the corresponding amino acid sequence targeted for mutagenesis are represented in figure 22a.

The transformed *E.coli* X-L1 blue cells were screened for mutant DNA by means of restriction analysis (figure 22b). Mutant DNA restricted with Sac I shows a single linear band of approximately 4000bp in size (lane 3) in contrast to uncut closed circular mutant and wild-type DNA (lane 2 and 6, respectively) which are approximately 3800bp in size. The restriction of the mutant DNA with Sac I, therefore, confirms the successful incorporation of the unique restriction site. In control lanes, mutant (lane 4) and wild-type (lane 8) DNA show identical banding patterns when restricted with Pst I.

[2] L164A refers to the replacement of the wild-type leucine (L) residue at position 164 with alanine (A).

A.

```
        <        Helix 6           >         <  Helix 7  >
       150                        170       179         190
Alpha  GERLSRADIHLVELLYYVEELDSSLISSFPLLKALKTRISNLPTVKKFLQP-------------
Sigma  GHSHTLADLHCYVALEVPLKHTPELLKDCPKIVALRKRVAECPKIAAYLKKRPVRDF-------
Mu     GDKVTYVDFLAYDILDQYHIFEPKCLDAFPNLKDFLARFEGLKKISAYMKSS-----------
Pi     GSQISFADYNLLDLLRIHQVLNPSCLDAFPLLSAYVARLSARPKIKAFLASPEHVNRPINGNGK
Sj26   GDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKRIEAIPQIDKYLKS-------------
Theta  GDSLTVADLALLASVSTFEVAGFDFSKYANVAKWYANAKTVAPGFDENWEGCL-----------
```

B.

C.

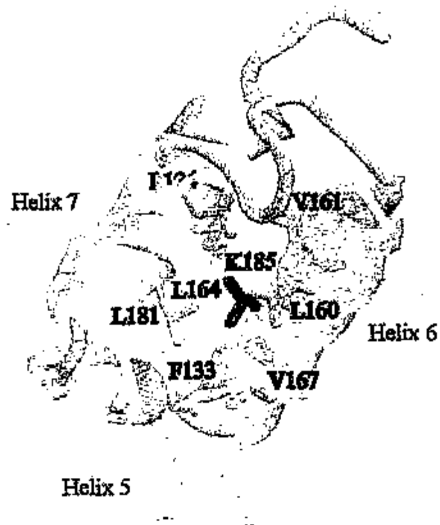


Figure 21. (a) Sequence alignment and (b) overlay of cartoon representatives of the alpha helices 6 and 7 for representative classes of GSTs: class alpha (1guh); class sigma (1gsq); class mu (1gst); class pi (2gsr); *S. japonicum* (1gne) and class theta (*L. cuprina*; 1gstt). The code accession numbers (indicated in parenthesis) were obtained from the Brookhaven Protein Data Bank. The numbering and position of the alpha helices is for the class alpha GST. The spatial localisation of the topologically conserved aliphatic residue is indicated. The alignment and superimposition was performed using WPDB v.2.0. (Shindyalov and Bourne, 1996). (c) Hydrophobic core of GST A1-1 indicating the side chains of Leu164 (in alpha helix 6; bold) Phe133 (helix 5), Leu160, Val161, Val167 (helix 6) and Leu181, Leu184 and Lys185 (helix 7). Figure was generated using Rasmol (Sayle, 1994).

Similarly, mutant and wild-type DNA were both linearised (lanes 5 and 9, respectively) when restricted with Pvu I.

The entire mutant DNA encoding the protein was sequenced and the sequence analysis of the pertinent region is indicated in figure 22c. The Sac I restriction site GAG CTC and the CTC→GCA codon change with the inferred amino acid sequence confirming the mutation are indicated.

5.3 Physicochemical properties of wild-type and L164A hGST A1-1

5.3.1 Expression and purification

The L164A enzyme was expressed in *E. coli* and purified using S-hexylglutathione affinity chromatography in a manner identical to that for the wild-type enzyme. The purified mutant protein has a similar tertiary structure (subunit molecular mass ≈ 27kDa) and quaternary structure (dimeric molecular mass ≈ 55kDa) as the wild-type enzyme. In addition, the single symmetrical peak obtained from SEC-HPLC indicates that the hydrodynamic volume of the mutant protein was not altered.

The intracellular yield for the L164A enzyme was substantially lower than that obtained for the wild-type enzyme. A total of 25mg for the wild type and 15mg for the mutant protein were obtained from 900mls of cell culture. The reason for the lower yield was not investigated since sufficient protein was obtained for the structure-function studies. However, based on the thermal-stability studies (section 5.4) it is possible that protein expression was decreased as a result of temperature sensitivity at 37°C. The decreased yield was not the result of different binding affinities for the affinity matrix (immobilised S-hexylglutathione) because both enzymes showed similar binding properties with S-hexylglutathione (section 5.3.3). The decrease in the yield may be consistent with a greater susceptibility to proteolytic degradation as the native state becomes less stable. Urea-induced equilibrium unfolding studies support a destabilisation of the L164A enzyme (section 5.5).

A.

5' GAG CTG GTG GAA CTT CTC TAC TAC GTC 3'

H L V E L L Y Y V

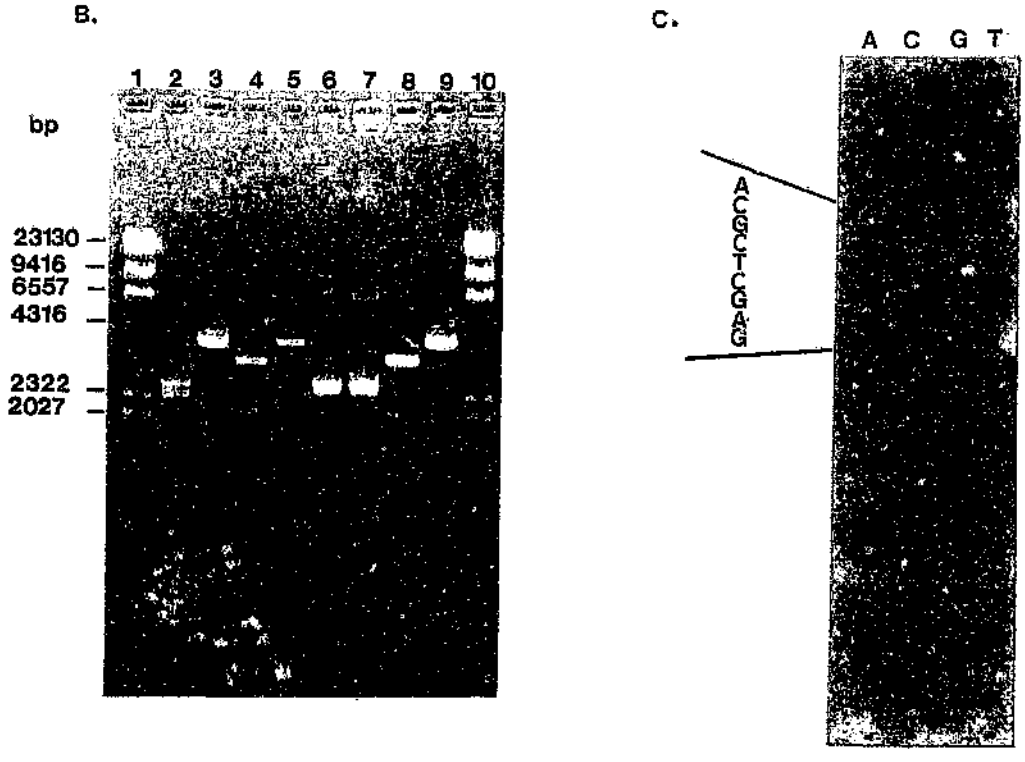


Figure 22. (a) Region of the wild-type nucleotide sequence and corresponding amino acid sequence coding for hGST A1-1 targeted for mutagenesis (red). (b) 1% (v/v) agarose gel of wild-type and L164A DNA. Lane 1 and 10, λDNA restricted with HindIII; lane 2 and 6, undigested mutant and wild-type DNA; lane 3 and 7, mutant and wild-type DNA restricted with SacI; lane 4 and 8, mutant and wild-type DNA restricted with PstI and lane 5 and 9, mutant and wild-type DNA restricted with PvuI. (c) Autoradiogram of pertinent region of L164A DNA: the SacI restriction site (GAG CTC) and the codon change CTC → GCA which results in the Leu-Ala replacement are indicated.

5.3.2 Spectroscopic properties of L164A hGST A1-1

The tryptophan (excitation at 295nm) and the tyrosine/tryptophan (excitation at 280nm) emission spectra for the mutant protein have an emission maximum (at 325nm) and intensity similar to that of native wild-type enzyme (figure 23a and b). Therefore, Trp20 remains the local reporter of events occurring at the interdomain interface and this region is largely unperturbed by the mutation. Second-derivative analyses of the UV-spectra of native wild-type and mutant enzyme are shown in figure 23c. The ratio (r) was calculated to be 1.43 and 1.63 for the mutant and wild type enzyme, respectively. The slight increase ($\approx$ 12%) in the exposure of the tyrosine residues for the mutant enzyme is within experimental error. In general, therefore, the overall tertiary structure and the polarity of the tryptophan and tyrosine environment were unaltered by the amino acid substitution.

5.3.3 Steady-state kinetic properties for wild-type and mutant enzyme

The steady-state kinetic parameters (V_m, K_m and k_{cat}/K_m) for the wild type and mutant enzymes towards the electrophilic substrate (CDNB) and reduced glutathione, are summarised in table 7. The specific activities for the wild type and mutant enzymes are similar and the replacement of Leu164 with Ala does not significantly change the steady-state kinetic parameters towards reduced glutathione and CDNB. The steady-state kinetic parameters (K_m, k_{cat}/K_m) towards reduced glutathione and CDNB vary from those which have been reported for the wild-type enzyme (Stenberg *et al.*, 1991a) as a consequence of different assay conditions.

The kinetic inhibition properties of the wild type and mutant enzyme were compared using the glutathione analogues, glutathione sulphonate and S-hexylglutathione, and the non-substrate ligand, ANS. The inhibition (IC_{50}) values are included in table 7. The replacement of Leu with Ala did not cause significant changes in the inhibition

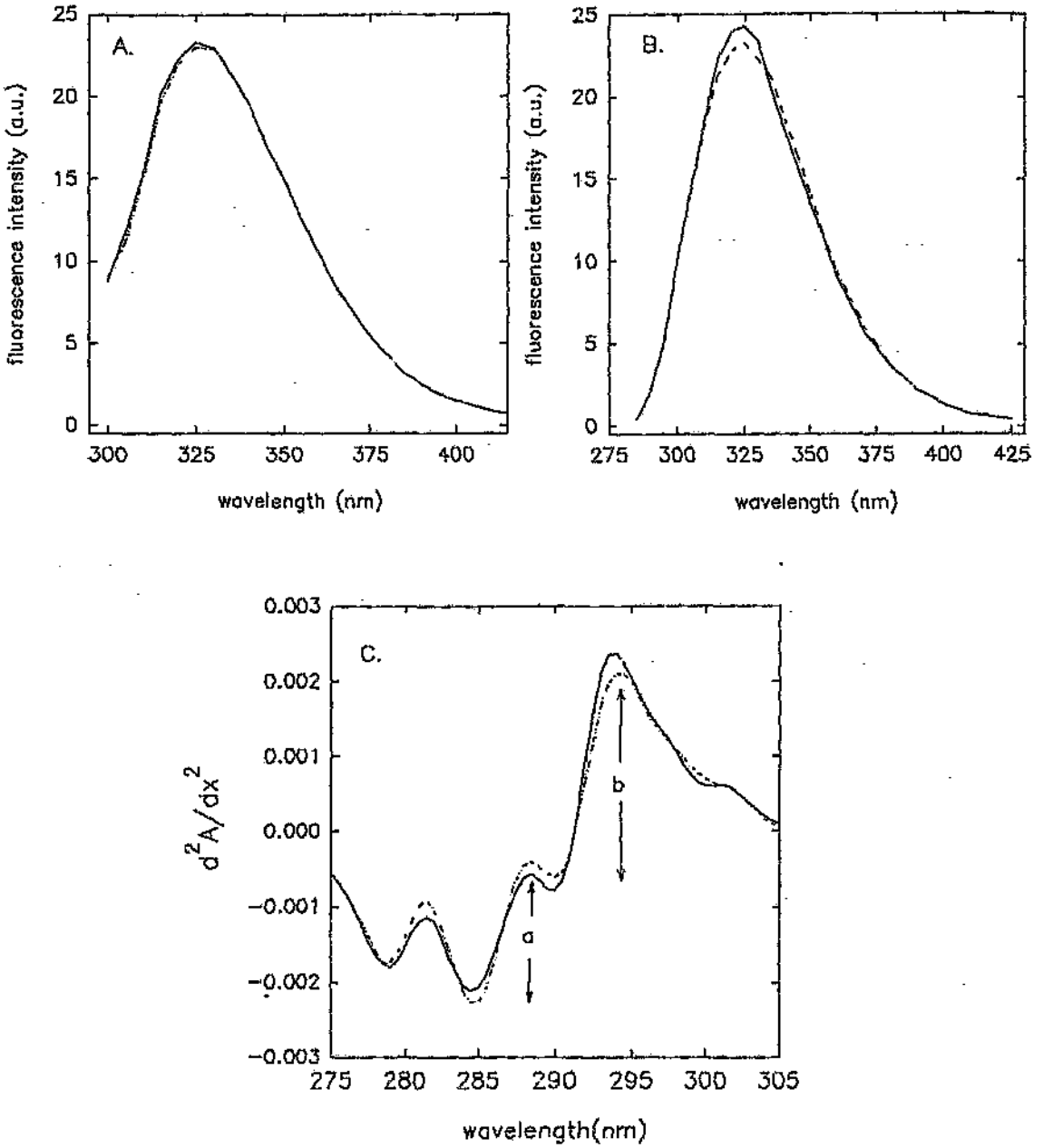


Figure 23. Fluorescence emission spectra. (a) Excitation at 295nm for 5μM wild-type (—) and L164A (– – –) hGST A1-1 and (b) excitation at 280nm for 5μM wild-type (—) and L164A (----) hGST A1-1. (c) Second-derivative UV-spectra for 5μM wild-type (– – –) and L164A (—) hGST A1-1.

characteristics of the enzyme. The CDNB-conjugating activity of the wild type and the mutant enzyme was inhibited competitively by the glutathione analogues, with S-hexylglutathione being the most effective inhibitor. The reason for this is that the hexyl-moiety of the S-hexylglutathione conjugate occupies the H-site and therefore prevents the binding of the electrophilic substrate. The non-substrate ligand, ANS, results in non-competitive inhibition of the CDNB-conjugating activity of the wild type and mutant enzymes. This type of inhibition by ANS has been reported for the class pi (Bico *et al.*, 1995), the class alpha (Sluis-Cremer *et al.*, 1996) and the class mu (Warholm *et al.*, 1983) GSTs. The mode of binding of the drug, praziquantel, to the non-substrate ligand binding site (L-site) of the *S. japonicum* GST, i.e., one molecule per protein dimer, has provided an explanation for the observation of non-competitive inhibition (McTigue *et al.*, 1995). It is, therefore, possible that the presence of the ligand in the L-site blocks one of the active sites at a time and therefore through steric hindrance results in 50% inhibition of the CDNB-conjugating activity. The binding affinities (K_d) of the wild type and mutant enzyme for ANS, assessed by tryptophan quenching, were also unchanged (table 7).

Therefore, the amino acid substitution did not have a significant effect on the protein's overall structure and conformation of the active site and non-substrate ligand-binding site (the L-site).

5.4 Thermal stability

The thermal stability of the wild type and the L164A hGST A1-1 enzymes were assessed using heat-inactivation studies. The thermal inactivation profile for the wild-type enzyme has a steep transition between 50°C-60°C (figure 24). In contrast the mutant enzyme has a broader transition between 40°C-60°C. The temperature at which there was 50% remaining activity (T_m) was 58°C and 52°C for the wild type and mutant enzyme, respectively. The decrease in thermostability may be explained in part by the loss of van der Waals interactions necessary for stabilisation of the

Table 7. Kinetic parameters for wild-type and L164A hGST A1-1. All values are the result of data fitting to the mean of at least three replicates with variation typically 10%.

Parameter	**Wild-type hGST A1-1**	**L164A hGST A1-1**
Specific activity (μmol/min/mg)	54	51
Varied [GSH]: V_m (μmol/min)	0.0123	0.0152
K_m (mM)	0.138	0.104
k_{cat}/K_m ($mM^{-1}s^{-1}$)	218	200.9
Varied [CDNB]: V_m (μmol/min)	0.0113	0.0134
K_m (mM)	0.27	0.24
k_{cat}/K_m ($mM^{-1}s^{-1}$)	100.5	111.7
Binding affinities (μM): S-hexylglutathione (IC_{50})	1.25	1.34
Glutathione sulphonate (IC_{50})	8.9	10
ANS (IC_{50}) (K_d)	100 16	100 16

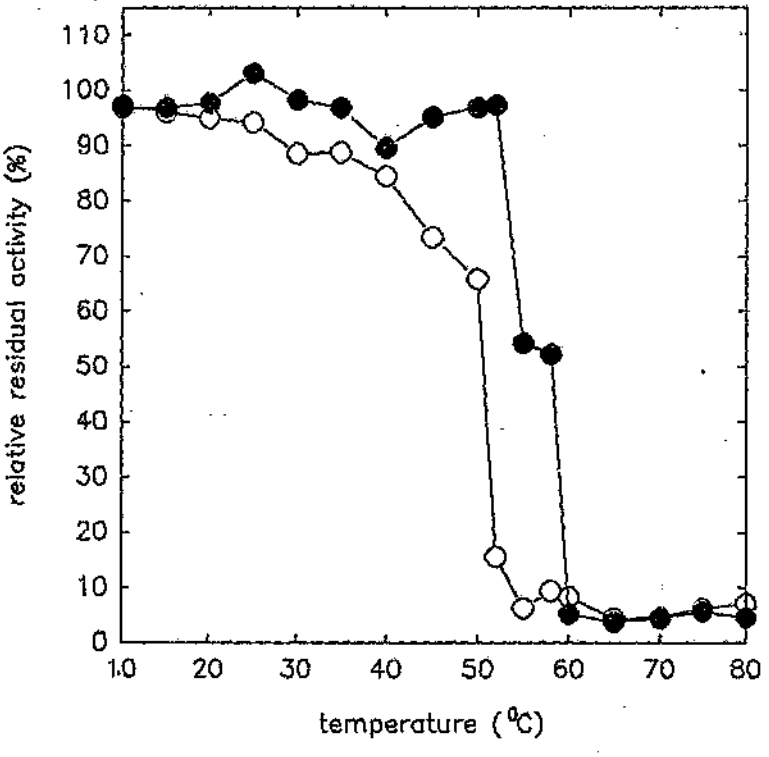


Figure 24. Thermal stability of the wild-type (•) and L164A (o) hGST A1-1 assessed by heat-inactivation assays. 3nM of the protein was assayed after incubation at each temperature for 30minutes.

native state of the protein. In addition, the alpha helix 6 forms non-specific hydrophobic interactions with the alpha helix 1, which is an integral part of the active site. It is, therefore, possible that the decreased thermostability of the L164A hGST A1-1 may be the result of a local conformational change in the alpha helix 6, which has an impact on the active site.

5.5 Equilibrium unfolding of L164A hGST A1-1

5.5.1 Reversibility of the unfolding transition

The reversibility of the urea-induced equilibrium unfolding was assessed by a tenfold dilution of denatured mutant protein and assessed using fluorescence spectroscopy (as a structural probe) and enzyme activity (as a functional probe). The tryptophan emission spectrum (excitation at 295nm) for the native state of the mutant protein has a maximum at 325nm which upon denaturation is red-shifted to 355nm with a concomitant decrease in fluorescence intensity (figure 25). These fluorescence properties are identical to those for the wild-type enzyme (section 4.1). Refolding of the mutant enzyme resulted in the regain of approximately 80% of the fluorescence emission intensity at 325nm. Similarly, 76% of the initial activity of the protein was regained. The reversibility of the unfolding reaction although substantially lower than that for the wild-type enzyme still resulted in a catalytically functional conformation. The lower recovery of refolded protein for the L164A hGST A1-1 may be the consequence of a competing side reaction such as aggregation.

5.5.2 Equilibrium unfolding transition and protein concentration dependence

The urea-induced equilibrium unfolding transition for L164A hGST A1-1 was monitored using steady-state tryptophan fluorescence, anisotropy decays and the binding of ANS (structural probes) and enzyme activity (functional probe). The principles of the techniques, as local and global probes for unfolding, were discussed in section 4.2.2.1.

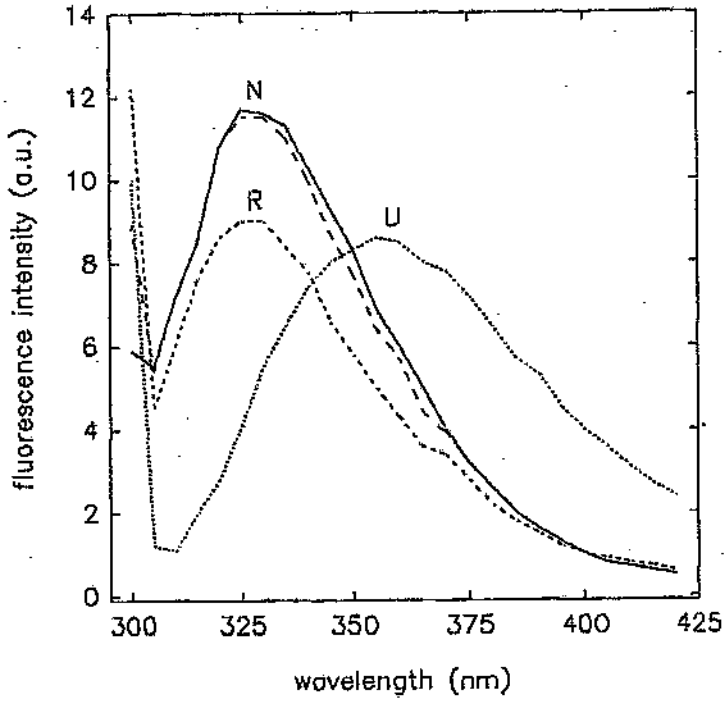


Figure 25. Fluorescence emission spectra for 0.45μM L164A hGST A1-1 in the absence (N; —) and in the presence of 8M urea (U; -). Reversibility of unfolding (0.45μM in 0.8M urea) (R; ---) was compared with a control spectrum (— —).

The unfolding reaction of the mutant enzyme shows a single sigmoidal transition suggestive of a two-state (folded dimer/unfolded monomer) transition. The coincidence of the data obtained for the various structural probes provides support for the two-state unfolding/refolding reaction (figure 26a).

The non-coincidence of the sigmoidal unfolding transition for enzyme activity and Trp fluorescence (inset of figure 26b) is similar to that observed for the wild-type enzyme (section 4.2.2.1). The urea-induced unfolding transition for 1μM L164A mutant hGST A1-1 has a midpoint at 3.0M urea for the enzyme activity measurements and 3.6M urea for the fluorescence measurements.

The protein concentration dependence of the midpoint of the unfolding transition, monitored using tryptophan fluorescence, for L164A hGST A1-1 provides definitive evidence for the two-state reaction (Neet and Timm, 1994) (figure 26b). The conformational stability parameters, $\Delta G(H_20)$ and the m-value, normalised to 1M protein, for the curves are in close agreement (0.1μM protein, $\Delta G(H_20)$ = 17.71 (± 1.02) kcal/mol and m = 2.24 (± 0.167) kcal/mol/M of urea; 3μM protein, $\Delta G(H_20)$ = 18.15 (± 1.33)kcal/mol and m = 2.67 (±0.242)kcal/mol/M of urea) which validates the two-state model. The increased stability with the increase in protein concentration is consistent with a bimolecular reaction i.e., $N_2 \leftrightarrow 2U$, where the equilibrium constant, $K = [U]^2/[N_2]$ is governed by the law of mass action.

The two-state transition for the unfolding and dissociation of the L164A hGST A1-1 suggests that the amino acid replacement does not alter the overall equilibrium unfolding/refolding properties.

5.5.3 Analysis of the equilibrium unfolding transition

A comparison of the unfolding transitions for the wild type and mutant hGST A1-1 indicates a destabilising effect as a consequence of the Leu→Ala164 substitution. The

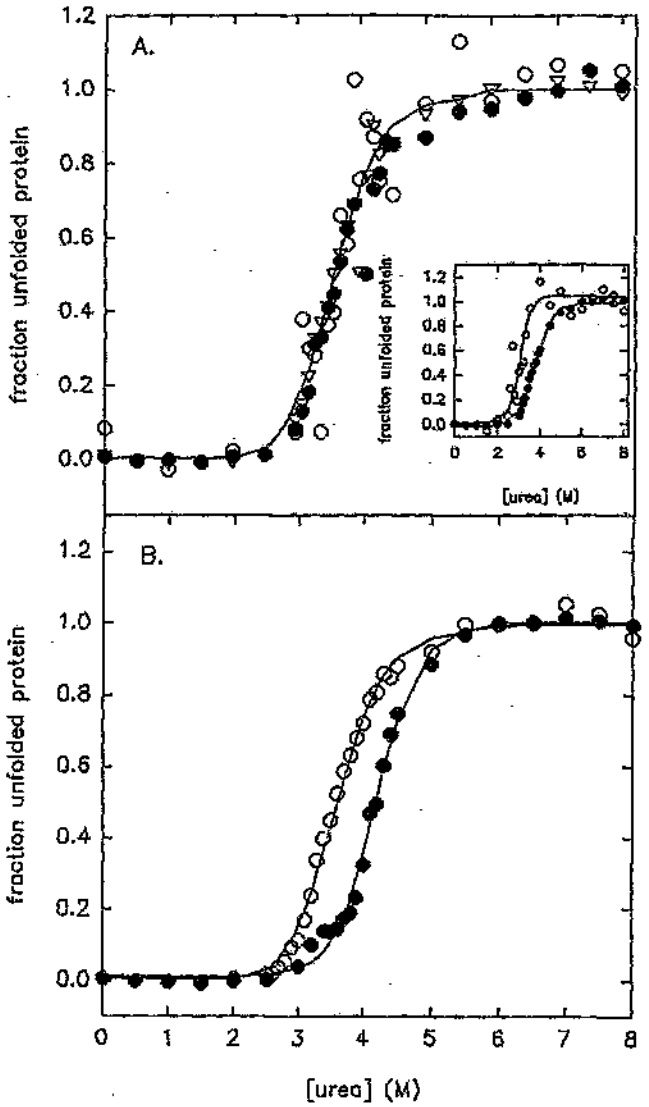


Figure 26. (a) Urea-induced equilibrium unfolding curves for 1μM L164A hGST A1-1 monitored by Trp20 fluorescence (▽), anisotropy decays (o) and the binding of ANS (•). Inset: equilibrium-unfolding transition for enzyme activity (o) and Trp fluorescence (•). (b) The protein concentration dependence of the unfolding transition for L164A hGST A1-1 at 0.1μM (o) and 3μM (•). Trp was selectively excited at 295nm and emission measured at 325nm and 355nm.

concentration of urea that was required to unfolding 50% of the protein (i.e., the C_m) was 4.5M for the wild type and 3.6M for the mutant. The satisfactory fit of the mutant protei to the two-state assumption enables the conformational stability to be estimated and the relative stability's of the wild type and mutant enzyme to be compared. The free energy change in the absence of denaturant ($\Delta G(H_20)$) was estimated using the linear extrapolation method of Pace *et al.* (1989)(chapter 3) $\Delta G(H_20)$ for 1μM of the wild-type enzyme was estimated to be 27.5 (± 3.8) kcal/mol and for 1μM of the mutant enzyme it was estimated to be 16.8 (± 2) kcal/mol.

The decrease in stability as a consequence of the mutation can be quantified. The difference in free energy of unfolding for the mutant and wild type enzyme ($\Delta\Delta G_{N-U}$) can be calculated using three different methods (Kellis *et al.*, 1989; Jackson *et al.*, 1993a). $\Delta\Delta G_{N-U}$ can be calculated from $\Delta\Delta G_{N-U} = (\Delta G(H_20))_{wild\ type} - (\Delta G(H_20))_{mutant}$. This method assumes that the errors resulting from the linear extrapolation method are small. Although, the values of $\Delta G(H_20)$ have been found to be reasonably accurate the precise linearity between the Gibbs free energy of unfolding and denaturant is questionable (Kellis *et al.*, 1989). In addition, the $\Delta G(H_20)$ value may not be sufficiently precise for calculating $\Delta\Delta G_{N-U}$ because the extrapolation from the measurable range (in the transition) is long and the least accurate estimates of K_U and hence ΔG_U are in the pre- (folded) and the post-transition (unfolded) baselines.

The $\Delta\Delta G_{N-U}$ can be measured in the presence of denaturant to prevent the large errors that result from the linear extrapolation method. The presence of denaturant does not affect the interactions of the buried hydrophobic groups in the folded state but may increase the solubility of the hydrophobic groups in the unfolded state. The contribution may be in the range of 0.1-0.2kcal/mol and therefore does not alter the apparent $\Delta\Delta G_{N-U}$ value significantly. For a number of proteins, the value of $\Delta\Delta G_{N-U}$ in the absence and presence of denaturant are in agreement with each other (Kellis *et*

al., 1989; Serrano *et al.*, 1992; Jackson *et al.*, 1993a). The procedure used here takes advantage of the high reproducibility in measuring the urea concentration of the midpoint of the unfolding transition (C_m or $[urea]_{50\%}$). $\Delta\Delta G_{N-U} = 0.5(m_{wildtype} + m_{mutant})[urea]_{50\%}$ where $[urea]_{50\%}$ is the difference between the urea midpoint for the wild type and mutant enzymes. The equilibrium unfolding data for Leu →Ala164 was fitted to this equation because of the change in the m-value as a result of the amino acid substitution.

The truncation of the leucine side chain to alanine results in the deletion of three methylene ($-CH_2-$) groups. The energetic cost of deleting one-to-three methylene groups has been correlated with the free energy loss on creation of a cavity in the hydrophobic core of a protein (Kellis *et al.*, 1988,1989; Shortle *et al.*, 1990). The $\Delta\Delta G_{N-U}$ (4.05M urea) for the Leu→Ala164 mutation in the hydrophobic core of domain II of hGST A1-1 was calculated to be 3065.0cal/mol. Therefore, the mutation destabilised the native state of hGST A1-1 by approximately 0.51kcal/mol per methylene group deleted. The effect of cavity-creating mutations on the stability of a protein has been extensively studied and generally, for a Leu to Ala mutation the stability change was 4.38 (±1.39) kcal/mol (Jackson *et al.*, 1993a and references therein) which is equivalent to 1.0-1.6kcal/mol per $-CH_2-$ group deleted.

In addition to the change in stability, the L164A hGST A1-1 also displays an unfolding transition that is much broader than that of the wild-type enzyme. The m-value for the unfolding transition was 4.22 (± 0.79) kcal/mol/M of urea for the wild-type hGST A1-1 and 2.59 (± 0.32) kcal/mol/M of urea for the L164A hGST A1-1. This value is indicative of the co-operativity of the non-covalent interactions required to maintain the native state and it is proportional to the amount of additional surface area (ΔA) exposed upon denaturation, $\Delta A = A_U - A_N$ where A_U is the surface area exposed of the denatured (U) state and A_N is the surface area exposed the native (N) state (Schellman, 1978).

The change in the m-value can be interpreted in several ways (Lim *et al.*, 1992; Shortle, 1995; 1996). First, it may reflect the difference in the exposed surface area of the native or the denatured state. For SNase, a number of mutations in the hydrophobic core resulted in varying m-value effects. In general, the m-value effect may be described as a result of a significant decrease in the responsiveness of the equilibrium between folded dimer and unfolded monomer to the denaturant. The change in the reversible denaturation reaction is the result of covalent changes in the proteins structure and the basis of the change lies in the energetics of solvation of the denatured state (Shortle, 1995).

The second interpretation is that the change in the m-value may reflect a deviation from the two-state behaviour. If an additional state of intermediate structure and stability was populated the unfolding transition would be broader and the m-value lower. The third interpretation for an altered m-value relates to changes in the interactions of the denaturant with the molecules in the denatured state (Arakawa and Timesheff, 1984). The influence of the action of denaturants is not well understood but evidence does support the view that their action is on the denatured state (Shortle, 1996).

The reason for the decrease in the m-value for the L164A hGST A1-1 is not clear. The decrease in the co-operativity of the unfolding transition may simply be a consequence of a decrease in co-operativity of the non-covalent forces that stabilise the native protein state. Although L164A hGST A1-1 has coincident unfolding curves it is proposed that partially folded species may be populated, to a small extent, in the transition zone. The change in the m-value has been correlated with the change in stability upon substitution of a hydrophobic amino acid residue for the protein, SNase (Shortle *et al.*, 1990).

5.6 Unfolding kinetics of L164A hGST A1-1

5.6.1 Fluorescence-detected unfolding kinetics

The unfolding reaction for L164A hGST A1-1 was monitored using Tyr/Trp fluorescence (excitation wavelength of 280nm) (figure 27a). The kinetic trace for the mutant, when compared with the initial and final equilibrium baseline values indicates that as for the wild-type protein the unfolding reaction cannot be described by a simple monophasic change (figure 27a and inset). However, in contrast to the wild-type protein, the early unfolding phase was not observable i.e., it occurred within the dead time (2ms) of the stopped-flow apparatus. After the burst phase, that by comparison with the native baseline values results in an increase in the fluorescence intensity (the fast phase), the fluorescence intensity decreases to its equilibrium value (the slow phase). To maintain consistency for comparison with the wild-type reaction, the two events are referred to as fast and slow unfolding events. The slow unfolding event is best described by a single exponential function (see residuals figure 27a). Both phases occur throughout the range of denaturant concentrations (3.8M-8.3M urea) and temperatures (10°C-35°C).

Unfolding kinetics of the fast unfolding phase of L164A hGST A1-1 at 10°C using a 5M urea-concentration jump made it possible to observe the fast unfolding event (figure 27b and inset). Under these conditions, the amplitude change observed was about 100% of the total change expected. Therefore, the observation of the fast unfolding phase depends on the conditions of the study. The fast unfolding phase under these conditions was best described by a single exponential function (see residuals, inset of figure 27b). The apparent unfolding rate constant was $109s^{-1}$ which is about 75-fold faster than the apparent unfolding rate constant for the fast unfolding phase for wild-type hGST A1-1 ($k_{app} = 1.49s^{-1}$) under identical unfolding conditions.

The occurrence of the fast phase for mutant protein as a burst phase at higher temperatures due to its increased apparent unfolding rate constant supports the destabilising effect of the amino acid substitution.

5.6.2 Urea-dependence of the unfolding rate and amplitude

The rate of unfolding of the mutant enzyme increased as conditions favoured the unfolded state and the urea-dependence of the data was best fit to a linear function (r = 0.994) (figure 28a). The apparent unfolding rates for the slow phase for the mutant protein were slightly faster than the rates for the wild-type enzyme. Therefore, the mutation has the greatest effect on the fast unfolding event. The denaturant dependence of the slow unfolding phase for the mutant protein (m_u = 259.0cal/mol) is roughly parallel to that of the wild-type protein (m_u = 270.0cal/mol) suggesting that the overall kinetic unfolding pathway is unchanged.

The urea-dependence of the amplitudes for the fast (as a burst phase) and the slow unfolding phase for the mutant protein are indicated in figure 28b. The amplitude change for the burst phase was calculated using the difference in intensity between the native baseline value and the start of the slow unfolding event. Both phases occur from at least 3.8M urea. Their relative percentages increase up to 5M urea and then become independent of the urea concentration when the final concentration is greater than 5M. The overall characteristics of the amplitudes are similar to those for the wild-type enzyme indicating that the mutation has not resulted in a change in the populations of the unfolding species or resulted in the formation of additional unfolded species.

5.6.3 Analysis of the unfolding kinetics

For the slow unfolding phase of the L164A hGST A1-1, the rate of unfolding in the absence of denaturant (k_u (H_20)) is 5.66 x $10^{-5}s^{-1}$ (obtained from the zero urea intercept of figure 28a). This corresponds to a time constant of 4.9 hours for the slow unfolding phase, nearly twofold faster than that for the wild-type enzyme. This is consistent with a decrease in stability and an increase in susceptibility of the mutant to proteolytic degradation. The increase in the apparent rate constant for the mutant suggests that the activation energy for unfolding in the absence of denaturant ($\Delta G^{\ddagger}(H_20)$) is lower. $\Delta G^{\ddagger}(H_20)$ may be calculated using equation 16.

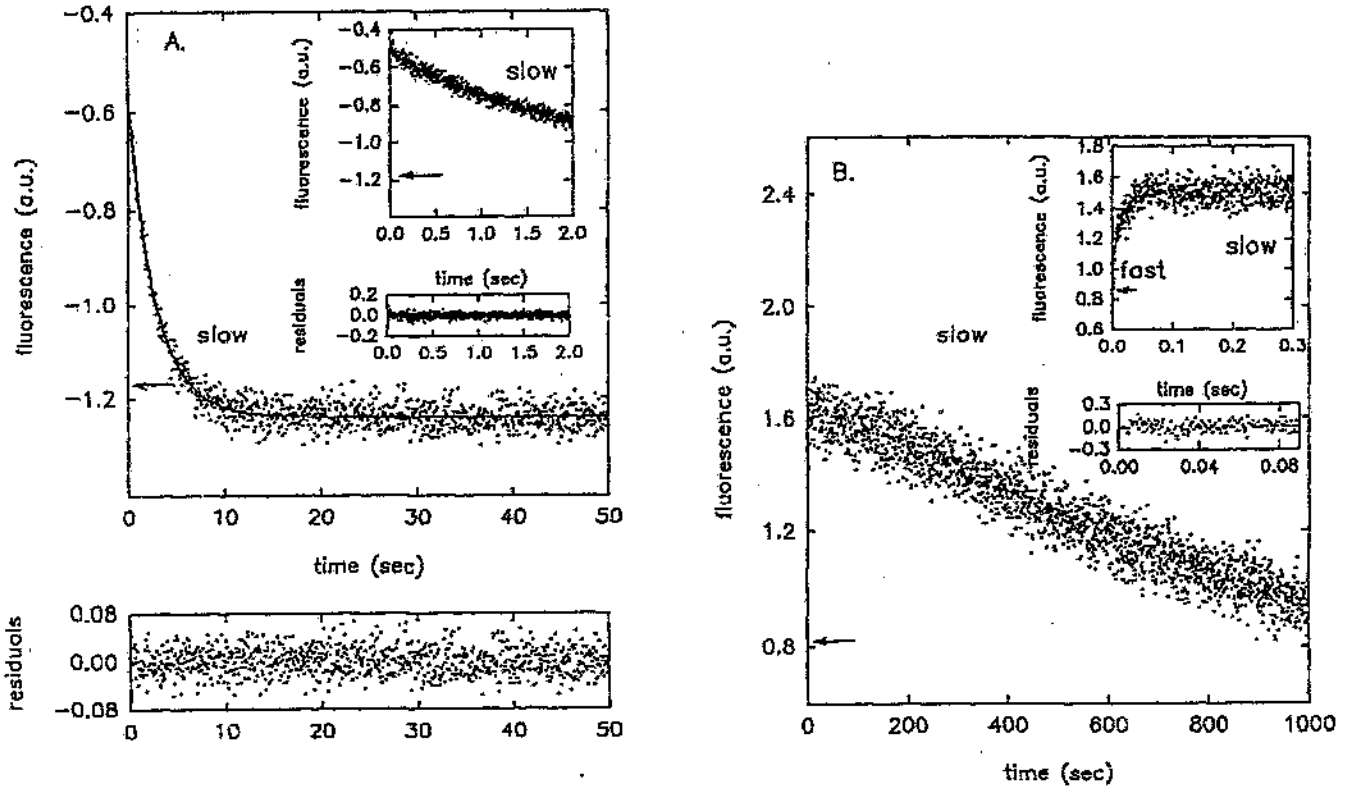


Figure 27. Unfolding kinetics. The change in fluorescence of 1μM L164A hGST A1-1 was monitored at 280nm and emission above 320nm. (a) Final urea concentration of 8.3M, pH 6.5, 25°C. The slow phase and the absence of the fast phase and part of the slow phase (inset) are indicated. (b) Final urea concentration of 5M, pH 6.5, 10°C. The fast phase (inset) and part of the slow phase (inset and main panel) are indicated. The arrows indicate the fluorescence signal for native folded protein. The unfolding traces were fitted with a single exponential function and the lower panels show the residuals.

The change in activation energy for the slow unfolding phase, as a consequence of the mutation, can be calculated by $\Delta\Delta G_U^{\ddagger} = \Delta G_U^{\ddagger} - \Delta G_U'^{\ddagger}$ where $\Delta G_U^{\ddagger}$ and $\Delta G_U'^{\ddagger}$ are the activation energy in the absence of denaturant for the wild type and the mutant enzyme, respectively. For the wild type protein at 25°C in the absence of denaturant, $\Delta G_U^{\ddagger} = 23.59$kcal/mol and for the mutant $\Delta G_U'^{\ddagger} = 23.23$kcal/mol and therefore $\Delta\Delta G_U^{\ddagger}$ is 360.0cal/mol. This value can also be determined using the apparent unfolding rate constant for the wild-type (k_u) and mutant (k_u'), $\Delta\Delta G_U^{\ddagger} = -RT\ln (k_u/k_u') =$ 391.8cal/mol. As mentioned for equilibrium studies the accuracy of the linear extrapolation method is debatable and therefore these values were calculated using the rates in the presence of 4M urea. For the rates of unfolding for the wild-type (k_u) and mutant (k_u') in presence of 4M urea are $2.04 \times 10^{-3} s^{-1}$ and $3.63 \times 10^{-3} s^{-1}$, respectively. Therefore, $\Delta\Delta G_U^{\ddagger}$ (4M urea) = 341.2cal/mol.

The data obtained under the different conditions are within experimental error. The overall change in the activation energy as a consequence of the mutation ($\Delta\Delta G_U^{\ddagger}$) is compared with the change in the Gibbs free energy ($\Delta\Delta G_{N-U}$ from equilibrium studies) to provide structural information on the transition state for the slow unfolding phase (Kellis *et al.*, 1988, 1989; Matouschek *et al.*, 1989; Fersht *et al.*, 1992; Jackson *et al.*, 1993b).

Information on the structure of the transition state for unfolding was calculated using the ratio $\phi_U = \Delta\Delta G_U^{\ddagger}/\Delta\Delta G_{N-U}$ (Kellis *et al.*, 1988). For Leu→Ala164, the value of ϕ_U (4M urea) is 0.112. There are two extreme values of ϕ_U that can be interpreted in a simple manner (Kellis *et al.*, 1988; 1989; Matouschek *et al.*, 1989; Jackson *et al.*, 1993b). A value of 1 means that the region of the mutation is exposed to solvent in the transition state to the same extent as in the unfolded state. In contrast, a value of 0 means that the interaction energies in the transition state are identical to those in the folded state. In general, the fractional values of ϕ_U are harder to interpret. However, for non-disruptive mutations of a side chain without access of water to the site of the

mutation (as commonly found for hydrophobic core residues) ϕ_U can be a measure of the loss of van der Waals interactions (Matouschek *et al.*, 198.· Fersht *et al.*, 1992). This assumes that the difference in solvation energy is small ($\Delta C_{alv} \approx 0$) and in the case of replacement of one hydrophobic residue for another the assumption is valid (Fersht *et al.*, 1992). Most protein engineering studies involving the substitution of a large hydrophobic amino acid residue for a smaller one and for which three-dimensional structures are available have shown that water does not occupy the cavity created (Ericksson *et al.*, 1992; Buckle *et al.*, 1993). The fractional value of ϕ_u for the Leu164 to Ala mutation indicates the partial loss of interactions necessary for stabilisation of the hydrophobic core. The ϕ_U value suggests that 11 -12% of the interaction energy made by the $C^{\delta 1}$, $C^{\delta 2}$, $C^{\gamma 1}$ methyl (ene) groups of Leu164 is lost at the transition state for the slow unfolding phase for hGST A1-1.

Furthermore, the probability of a single water molecule being present in a purely hydrophobic cavity is estimated to be approximately 1:20 000 (Wolfenden and Radzicka, 1994). Recently, a water molecule has been found in the cavity of a mutant protein (Buckle *et al.*, 1996) which raises the possibility that the assumption that all hydrophobic cavities are devoid of water may not be valid in all cases. This, therefore, provides evidence that the alpha helix 6 and therefore the hydrophobic core of hGST A1-1 remain largely intact in the transition state. Typically for other proteins studied, for example SNase A, chymotrypsin inhibitor2, T4 lysozyme and barnase, changing a hydrophobic residue in the core of a protein result in ϕ_u values in the range of 0.3 (± 0.2) (Matouschek *et al.*, 1989 and references therein).

The solvent accessibility of the transition state for the slow unfolding phase for the mutant was characterised using the m_u/m ratio, where m is the value obtained in equilibrium unfolding studies for 1μM L164A hGST A1-1 (2.59 (± 0.32) kcal/mol/M of urea. The slope (m_u) of the slow unfolding phase for L164A hGST A1-1 (figure 28a) is 259.0cal/mol. The ratio is 0.064 and 0.100 for the wild type and

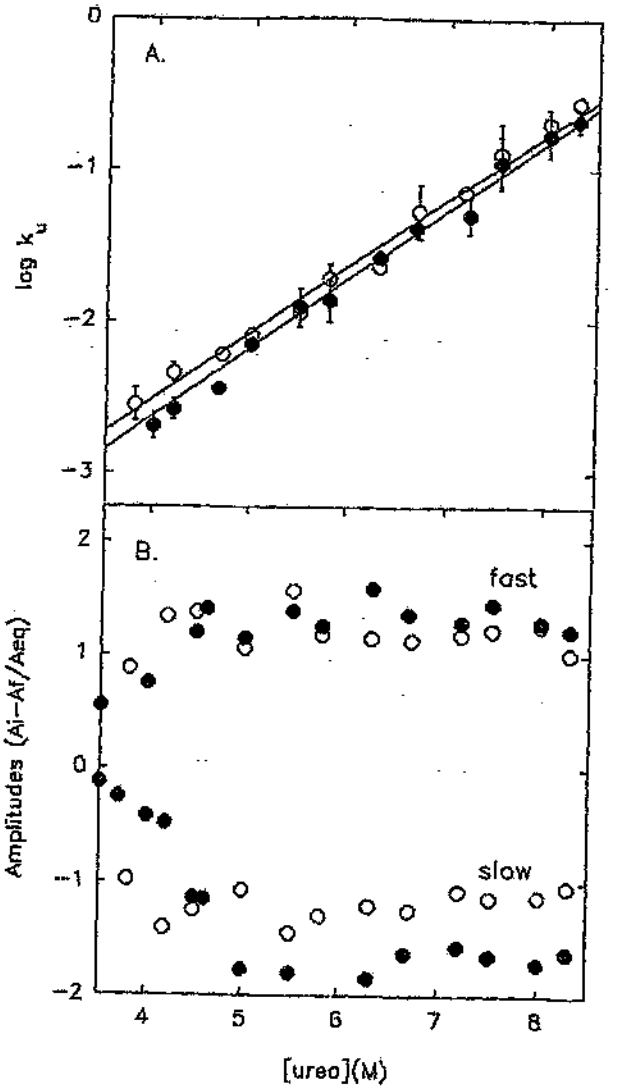


Figure 28. (a) Urea-dependence of the unfolding rate constant for the slow phase for 1μM wild-type (•) and L164A (o) hGST A1-1. The solid lines represent the best fit to a linear function. Unfolding experiments were performed from 3.8M to 8.3M urea at 25°C, 0.1M NaCl, pH 6.5. The data was fitted to equation 14 (3.8M and 5M urea) and to equation 13 (urea concentrations greater than 5M). The error bars are as indicated or are smaller than the size of the symbol. (b) The urea-dependence of the amplitudes for unfolding for wild-type (•) and mutant (o) hGST A1-1. The amplitude of the burst phase was calculated and normalised with respect to the fraction-unfolded measurements obtained in equilibrium studies.

mutant enzymes, respectively. Therefore, 10% of the buried hydrophobic surface area of the native mutant state is exposed at the transition state in contrast to 6.4% for the native state of the wild-type protein. Therefore, the solvent accessibility of the transition state has increased as a consequence of the mutation.

5.6.3 Temperature-dependence of the unfolding rate

The apparent unfolding rate of the slow phase for L164A hGST A1-1 was characterised from 10°C-35°C at a final urea concentration of 8.3M. The unfolding rate constants were related to the Gibbs free energy change in the absence of denaturant according to the transition state theory (Tan *et al.*, 1996). For the mutant enzyme, the plot of ln ($k_u h/k_B T$) versus 1/T (figure 29a) shows distinct curvature and the residuals (figure 29b) indicated the data is best fit to equation 18 than to a linear fit. The reasons for the curvature are unknown but may be the result of data deviation or the consequence of a large heat capacity ($\Delta C_p^{\ddagger}$) as the protein unfolds. The best fit to equation 18 for the slow unfolding phase gives the following apparent thermodynamic activation parameters: $\Delta C_p^{\ddagger}$ of 807.0cal^{-1}mol^{-1}K^{-1}; $\Delta H^{\ddagger}$(298K) of 18.50kcal/mol and $\Delta S^{\ddagger}$(298K) of 0.3367cal^{-1}mol^{-1}K^{-1}.

The change in the heat capacity ($\Delta C_p^{\ddagger}$) (as for the m-value) can be related to the heat capacity change for complete unfolding ($\Delta C_{pN\text{-}U}$ = 7.3kcal^{-1}mol^{-1}K^{-1}) (section 4.3.3; Myers *et al.*, 1995). The extent of buried surface area for the transition state for the slow unfolding phase ($\Delta C_p^{\ddagger}$ /$\Delta C_{pN\text{-}U}$) was calculated to be 6.5% and 11.0% for the wild type and mutant enzyme, respectively. Therefore, the increase in solvent accessibility of the transition state as a consequence of the amino acid substitution is complimented by urea and temperature dependence studies.

5.7 Discussion

The experimental data indicates that the topologically conserved residue (Leu164, alpha; Leu157, pi; Leu163, mu; Leu160, sigma; Leu158, *S. japonicum*; Val163, theta)

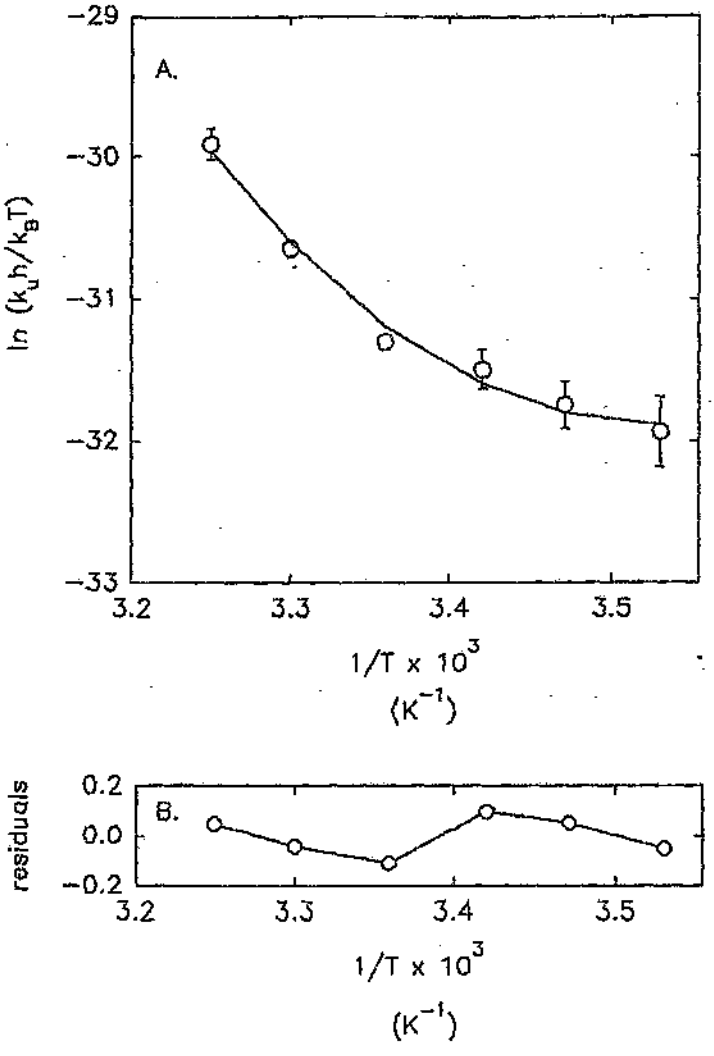


Figure 29. (a) Eyring plots for the temperature dependence of the unfolding rate constant for the slow unfolding phase for 1μM L164A hGST A1-1 (○). The unfolding experiments were performed at 8.3M urea, 0.1M NaCl, pH 6.5. The error bars are as indicated or are smaller than the size of the symbol. (b) The plot is best fit to equation 18.

located in alpha helix 6 of domain II is essential for specifying and stabilising the hydrophobic core of the glutathione transferases.

The functional and gross structural properties of hGST A1-1 were unaltered by the replacement of Leu164 with alanine and the urea-induced equilibrium and kinetic unfolding studies indicate that the overall unfolding pathway was unchanged. Therefore, the sequential unfolding pathway proposed for wild-type hGST A1-1 (section 4.5) that incorporated the equilibrium and kinetic unfolding pathway is still valid. This is a conservative interpretation since it is assumed that the mutant is going through the same unfolding steps as the wild type and that the change in properties is the consequence of the mutation. It is possible that the unfolding pathway could be completely different and that a new path was created as a consequence of the mutation (Waldburger *et al.*, 1996).

The two step sequential unfolding pathway monitored using Tyr/Trp fluorescence is described by:

$$N_2 \rightarrow N_2^* \rightarrow 2U$$

where N_2 is the native dimer, N_2^* is the native-like dimer or dimeric intermediate and U is the unfolded monomer. For the mutant protein, experimental evidence suggests that the amino acid substitution had a greater destabilising effect on the fast unfolding event (which is represented by $N_2 \rightarrow N_2^*$) than on the slow unfolding event. This event was proposed for the wild-type enzyme to be the result of the partial dissociation of domain I and II at the domain interface, near Trp20. Trp20 (figure 8a and b) is located in the alpha helix 1 of domain I but its indole side chain protrudes into domain II and it has non-specific hydrophobic contacts (within 4Å) with alpha helix 6 (section 4.1). One residue with which the indole side chain of Trp20 interacts is Tyr165, located directly alongside Leu164. It is, therefore, feasible that the structure of this region has changed to a greater extent than the rest of the protein

structure. The substantially greater effect of the mutation on the fast unfolding step suggests that the packing is looser in the unfolding intermediate (N_2*) than in N_2. The slower unfolding event described by N_2* $\rightarrow$ 2U, represents the complete dissociation and unfolding of the native-like dimer into two unfolded monomers. The absence of a stable intermediate is supported by the apparent two-state (dimer/unfolded monomer) transition. The urea and temperature dependence studies for this unfolding event describe a small destabilising effect as a consequence of the mutation.

The solvent accessibility of the transition state is described by the relative kinetic m-values (m_u/m) and the relative heat capacity values ($\Delta C_p^{\ddagger}/\Delta C_{pN-U}$). Both values indicate that the mutation has resulted in a shift of the transition state, albeit slightly, towards the unfolded state. The native mutant state has an increased solvent accessibility and is, therefore, not as compact as the native wild-type transition state. The structure of the transition state for the slow unfolding phase was described by the ϕ_U value. For the Leu$\rightarrow$Ala164 substitution, the value of ϕ_U (4M urea) is 0.12 and this indicates that approximately 90% of the non-covalent interaction energies are maintained at the transition state for the slow unfolding phase. Therefore, this implies that the alpha helix 6, domain II and the hydrophobic core of L164A hGST A1-1 are largely intact. The hydrophobic core is weakened but the interactions are still significant enough to maintain the overall structure. Therefore, the disruption of the hydrophobic interactions in domain II is a partly rate-limiting step for the unfolding pathway for hGST A1-1 and by implication for the family of glutathione transferases.

For hGST A1-1, the effect of the Leu $\rightarrow$ Ala substitution was investigated using the unfolding kinetics because the equilibrium and kinetic unfolding properties of the wild-type enzyme are best understood (chapter 4) and the structure of the initial state i.e., the native enzyme, has been well elucidated (Sinning *et al.*, 1993; Cameron *et al.*, 1995). In addition, the refolding pathway for hGST A1-1 displays a number of complexities (chapter 6) and therefore any destabilising effects would be difficult to interpret. However, based on the principle of microscopic reversibility, it is possible

to infer information about the structure of the transition state for the final event of the refolding pathway. From the L164A hGST A1-1 studies, it follows that the last event of the refolding pathway for the glutathione transferases involves the consolidation of interactions in the hydrophobic core of domain II.

5.8 Conclusion

The hydrophobic core mutant of hGST A1-1 indicates that the hydrophobic interactions are important in determining the overall stability of the glutathione transferases. The introduction of a cavity in domain II has a small effect on the physicochemical properties and no overall effect on the folding/unfolding pathway of hGST A1-1. The buried methyl (methylene) groups of Leu164 contribute approximately 3.07kcal/mol to the stability of hGST A1-1. Information on the structure of the transition state for unfolding indicates that the hydrophobic core is weakened and suggests that the breaking of hydrophobic interactions is partly rate-limiting during the unfolding of hGST A1-1 and by implication for the unfolding of the family of glutathione transferases.

CHAPTER 6

REFOLDING KINETIC PROPERTIES OF hGST A1-1

6.1 Refolding kinetics of hGST A1-1

6.1.1 Fluorescence-detected refolding kinetics

The refolding reaction of hGST A1-1 was followed by fluorescence using an excitation wavelength of 280nm for tyrosine and tryptophan. The refolding kinetics were monitored at final urea concentrations between 1M and 4.25M and over a 5°C - 40°C temperature range. The refolding kinetic traces for hGST A1-1 show a number of complexities depending on the final concentration of urea (figure 30). At final urea concentrations between 2.5M and 4.25M the refolding reaction is biphasic whilst at urea concentrations less than 2.5M the reaction is triphasic. The three phases are referred to as fast, intermediate and slow refolding events. Typically, the fluorescence intensity increases for the fast and intermediate events, which occur in the millisecond to second time range, following which the fluorescence intensity decreases for the slow refolding event. At final urea concentrations of 4M and 4.25M, the refolding events were monitored using manual mixing studies since the change in amplitude was not detectable using the stopped-flow apparatus.

The triphasic refolding kinetics indicates that the refolding reaction of hGST A1-1 cannot be described by a simple two-state bimolecular association reaction. The refolding kinetic trace for the final urea concentration of 3M is indicated in figure 30a and the residuals for the biexponential fit in the lower panel. The absence of the slow refolding phase is indicated in the inset of figure 30a. A refolding kinetic trace monitored using Tyr/Trp fluorescence for the final urea concentration of 1M is indicated in figure 30b. The refolding data for the fast and intermediate phases are best described by a biexponential function (see residuals lower panel of figure 30b). The single slow phase is indicated in the inset of figure 30b and is best described by a single exponential function (see residuals).

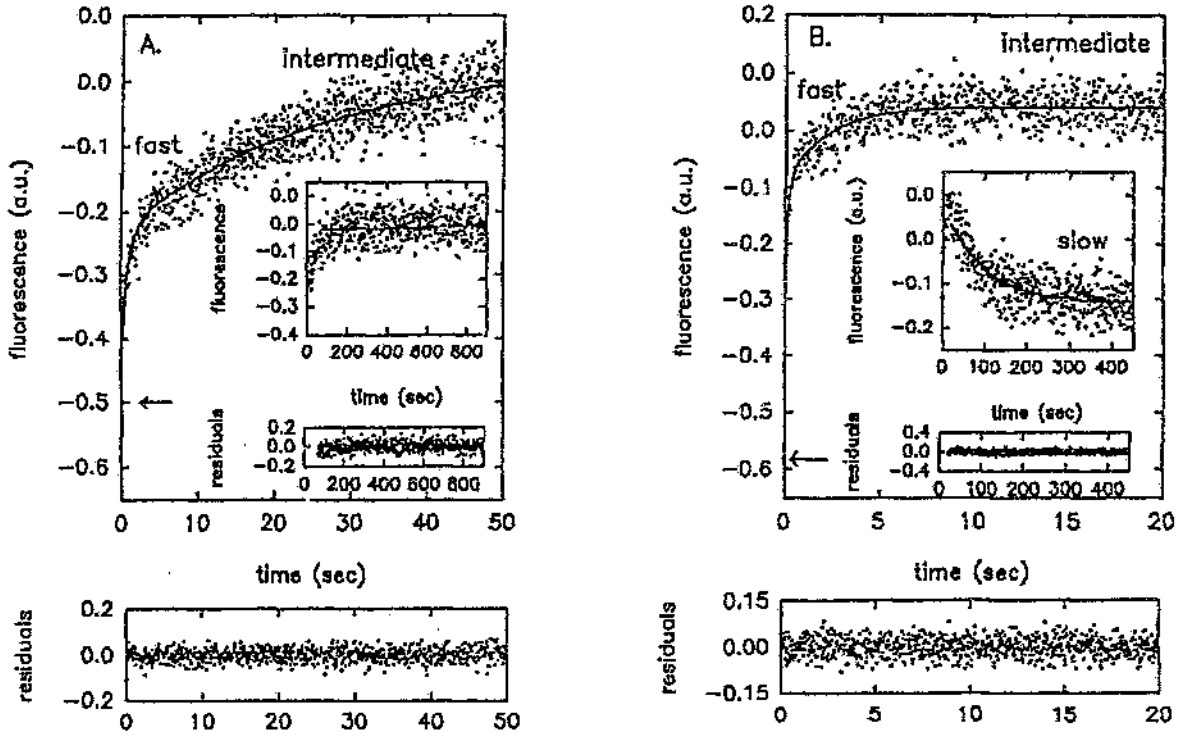


Figure 30. Refolding kinetics. (a) Final urea concentration of 3M, pH 6.5, 25°C. The fast and the intermediate phases are shown. The inset shows the absence of the slow phase and part of the fast and intermediate phase's (b). Final urea concentration of 1M, pH 6.5, 25°C. The fast and intermediate phase and slow phase (inset) are shown. The change in fluorescence was monitored upon refolding of 1μM hGST A1-1 was monitored by excitation at 280nm and emission above 320nm. The arrows indicate the fluorescence signal for unfolded protein. The fast and intermediate phases were fitted to a biexponential function and the slow phase to a single exponential function. The lower panels of each trace show the residuals.

Manual mixing studies over a 60-minute time range support the absence of the slow refolding phase at urea concentrations greater than 2.5M.

A number of control experiments established that the refolding phases were real events and not the result of aggregation, mixing artifacts or photodegradation. Identical "refolding" experiments were performed using 2μM N-acetyl-L-tryptophanamide and 20μM N-acetyl-L-tyrosinamide at a final urea concentration of 1M. "Refolding " monitored by excitation at 280nm yielded a horizontal signal response. In addition, the possibility that the slow refolding phase was the result of possible photodegradation of the tyrosine and tryptophan residues was investigated and excluded since narrowing the slit-width of the monochromator i.e., reducing the intensity of light, did not exclude the observation of this phase. Control experiments to monitor aggregation and baseline (buffer/buffer, buffer/urea and protein/urea) runs were similar to those described in chapter 4 and yielded a horizontal signal response.

Complexities in the refolding of proteins are not unique and indicate a deviation from the "all-or-none" two-state mechanism. To investigate if the multiple phases observed for the refolding of hGST A1-1 are the result of a number of unfolded species in the pre-transition baseline the initial refolding conditions were varied (section 4.3.1). The amplitudes and rates for the three refolding events show a slight but variable dependence on the initial urea concentrations (table 8). This suggests that the various unfolded forms are distinguishable from each other in the post-transition baseline and that changing the unfolding conditions shifts the equilibrium that exists between the various forms. The interpretation of this data is, however, limited because of the complexity of the refolding pathway, which is proposed to consist of parallel and sequential steps with different rates. Therefore, a simple and accurate evaluation of this data is very difficult.

6.1.2 Protein concentration dependence of the refolding rate

The influence of the protein concentration on the apparent refolding rate constants was used to predict the order of the rate-limiting refolding reaction. Protein concentration dependent studies were determined for the three refolding events at final protein concentrations between 0.5μM and 5μM. The final urea concentration was 1M and 3M for the fast and intermediate refolding events and 1M for the slow refolding event. At a final concentration of 1M and 3M urea, both the fast and the intermediate phases proceeded more rapidly at higher protein concentrations. The refolding rate increased linearly from 0.5μM to 3μM of hGST A1-1. This is consistent with a rate-limiting step, which involves a bimolecular (second-order) reaction and is the consequence of an association step involving two monomers (partially folded/folded) to form a dimer.

However, at 1M urea the refolding rate for the fast and intermediate refolding events deviates from linearity at protein concentrations greater than 3μM (figure 31a and b). The deviation of the rate may indicate a change in the rate-limiting step or be the result of inaccurate data collection. At the higher concentrations of protein, it is possible that the refolding rate has increased to such an extent that the part of the amplitude that is missing results in a lower apparent refolding rate. The removal of the first milliseconds of the kinetic trace did not change the calculated rate significantly. The deviation is, therefore, suggested to be the result of the association reaction becoming so rapid that the formation of the dimer is limited by a unimolecular (first-order) reaction which occurs either before (involving the partially folded/folded monomer) or after (involving the dimer) the association event. This phenomena has been observed for the P22 Arc repressor (Milla and Sauer, 1994) and other proteins (Garel, 1992). For hGST A1-1, the deviation does not impact on the interpretation of the results since all refolding studies were performed at a final protein concentration of 1μM for which there is a linear variation of the refolding rate with protein concentration.

Table 8 Amplitudes and apparent refolding rate constants for the fast, intermediate and slow phases obtained for the final (f) urea concentrations of 1.3M and 3M but varying the initial (i) urea concentration.

$[Urea]_i$	$[Urea]_f$	k_{app} (s^{-1})			Amplitudes (A_f-A_i)		
		fast	interme diate	slow	fast	interm ediate	slow
6	1.3	6.03 ± 0.106	0.331 ± 0.020	0.0136 ± 0.0013	0.365	0.186	-0.385
7		7.79 ± 1.615	0.5503 ± 0.165	0.0119 ± 0.0009	0.356	0.113	-0.691
8		8.64 ± 2.86	0.618 ± 0.009	0.0124± 0.001	0.174	0.015	-0.082
6	3	1.881 ± 0.445	0.122 ± 0.010	--------	0.288	0.230	-------
7		2.771 ± 0.248	0.146 ± 0.025	--------	0.188	0.159	-------
8		2.93 ± 0.55	0.148 ± 0.0083	--------	0.141	0.085	-------

Therefore, the rate-limiting step at 1M and 3M urea for the fast and intermediate refolding events is bimolecular.

The increase in protein concentration (0.5μM-3μM) had little effect on the rate of refolding for the slow refolding reaction (inset of figure 31a), therefore, this reaction appears to be unimolecular (first-order) and is suggestive of an isomerisation either between the unfolded monomers or between native-like dimers. The fluorescence properties of this event differ from the expected fluorescence increase for refolding.

An estimation of the rate of refolding (k_f) for each phase was calculated from the slope of the linear portion of the plot of the apparent refolding rate as a function of hGST A1-1 concentration (figure 31a and b). k_f is the second-order rate constant and was calculated to be 8.75×10^5 $M^{-1}s^{-1}$ and 3.75×10^5 $M^{-1}s^{-1}$ for the fast and intermediate refolding phases, respectively, at 25°C, pH 6.5, in the presence of 1M urea.

The rates for folding and dimerization for the dimeric proteins studied vary. For example, the P22 Arc repressor, which is one of the fastest folding proteins, has a rate constant for folding and association ranging from $5 \times 10^6 M^{-1}s^{-1}$ to $2 \times 10^7 M^{-1}s^{-1}$ (Milla and Sauer, 1994). In contrast, the *E.coli Trp* aporepressor (Gittelman and Matthews, 1990) and the GCN4-pI synthetic peptide (from the yeast transcriptional activator, GCN4) (Zitzewitz *et al.*, 1995) have refolding and dimerization rates of 100- to 1000-fold slower than that of the Arc repressor. In general, however, the majority of the second-order rate constants measured for dimeric proteins are in the range of 10^3 to $10^5 M^{-1}s^{-1}$ (Garel, 1992). For hGST A1-1, the rate of the association reaction which results in the dimer is 10^4-fold slower than the diffusion limit of $\sim 10^9 M^{-1}s^{-1}$. However, the refolding rate is still fast and compares with the rate at which protein-protein association reactions occur between stable folded proteins (i.e., 10^5-$10^6 M^{-1}s^{-1}$) (Koren and Hammes, 1976).

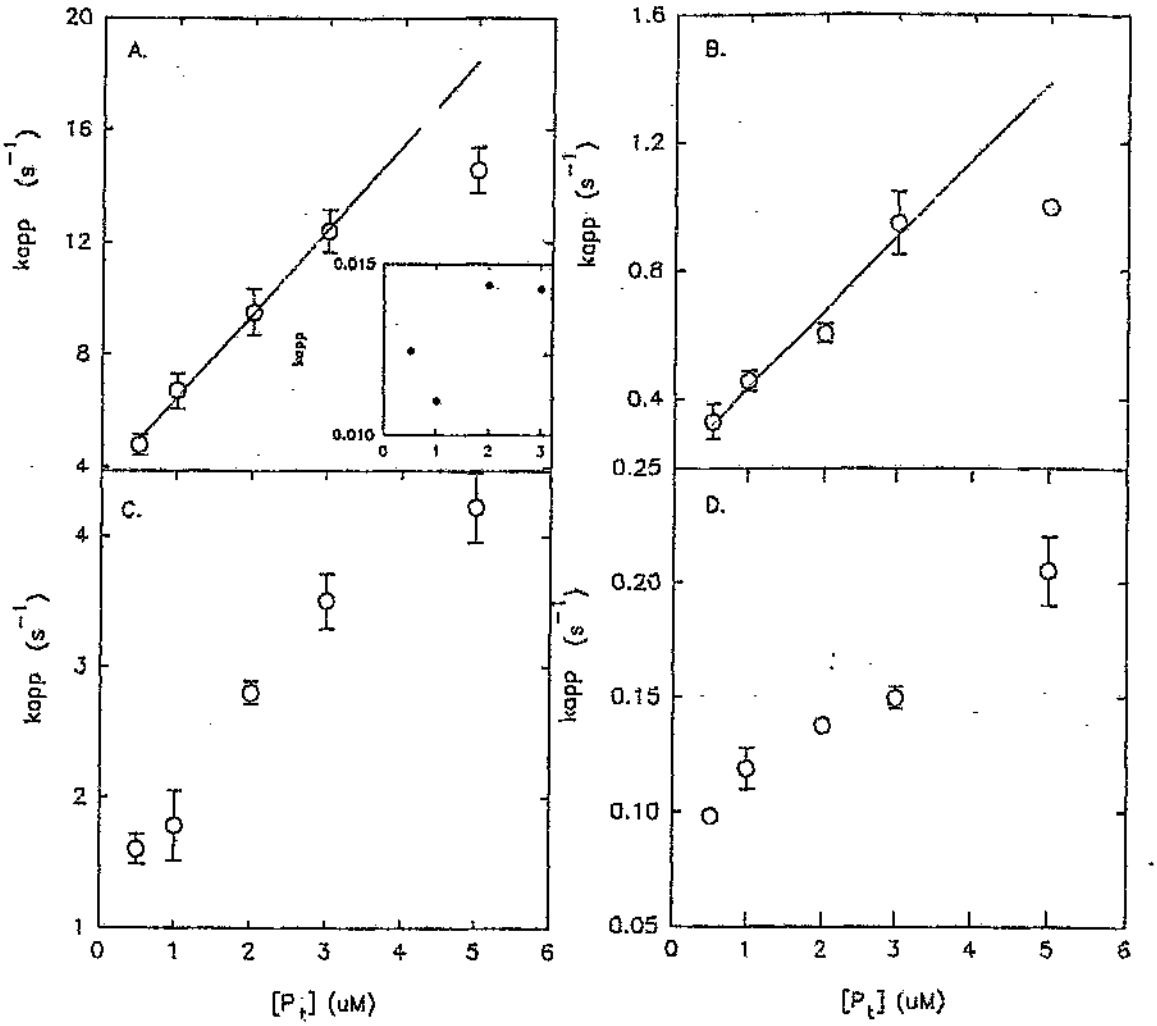


Figure 31. Protein-concentration dependence of the observed rate of refolding for the (a) fast, (b) intermediate and (inset of (a)) slow phase. Final urea concentration of 1M, 25°C. The straight line through 0-3μM hGST A1-1 (panels (a) and (b)) is the best fit according to the equation: $(1/\tau) = k_{app} = 4[P_t]k_f$ where k_f is the actual second-order rate constant. This equation is applicable for homodimeric proteins under experimental conditions which strongly favour the folded state (i.e., $k_u \approx 0$) (Bernasconi, 1976). Protein concentration dependence, at the final urea concentration of 3M at 25°C, for the observed rate of refolding for the (c) fast and (d) intermediate phase.

Koren and Hammes (1976) have proposed and it has since been noted by Zitzewitz *et al.* (1995) that the wide variation in second-order rate constants for the oligomeric systems studied may reflect the extent to which the folding of the monomers and the association reaction are coupled. In general, the larger the fraction of folded monomer or at least partially folded monomer the greater the observed bimolecular rate constant. Based on this one would expect monomer-monomer interactions to form early in the refolding pathway. Therefore, the rate of association approaches the diffusion limit when the monomers are substantially folded despite the complexity of the final structure and the rates are 2-3 orders of magnitude smaller when the monomers are disordered (Chaffotte *et al.*, 1997). For hGST A1-1, the folded monomer is undetected and this is consistent with the proposition in chapter 4 that the subunit-subunit interactions are necessary for stabilisation of the individual subunits. The folding of the individual monomers (subunits) and the association step to form the dimer, therefore, occur concurrently.

6.1.3 Urea-dependence of the refolding rates and amplitudes

The urea-dependence of the logarithm of the refolding rate constant for the three observable kinetic phases is shown in figure 32a. The complex denaturant dependence of the rates is suggestive of multiple refolding pathways for hGST A1-1. The apparent refolding rates for the fast and intermediate phases at urea concentrations greater than 2.5M decrease as the concentration of denaturant increases. The apparent refolding rates for the slow phase are independent of the final denaturant concentration.

The urea-dependence of the amplitudes for the three refolding phases is depicted in figure 32b. The amplitude data indicates that the fast refolding event is the major phase with the largest amplitude change followed by the intermediate phase and then the slow phase. The fast phase results in the greatest population of refolding species (up to 50%). The small amplitude change for the slow refolding phase ($\approx$ 10%)

indicates that this phase is the result of a minor population of molecules and it has little influence on the major refolding phases represented by the fast and intermediate refolding events. At low urea concentrations, up to 35% of the fast refolding event occurs as a burst phase. The denaturant dependence of the apparent rate constants for the fast and intermediate refolding events show distinct deviation from linearity at final urea concentrations less than 2.5M with the deviation being greatest for the fast refolding event. Non-linear refolding behaviour has been observed for a number of monomeric and oligomeric proteins and has been termed a "rollover" (Baldwin, 1996). Typically, the rollover is thought to be the result of accumulation of an intermediate, which is populated at low urea concentrations and results in a change in the rate-limiting step for refolding. The reaction would be unimolecular and suggestive of an isomerisation reaction that involves either the peptide backbone, e.g., *cis/trans* proline isomerisation or the rearrangement of the folded domains/subunits. The event would result in little change in the exposure of the side chains or backbone of the protein to solvent. More recently, the deviation from linearity has been described to result from an off-pathway event, caused by transient aggregation of the denatured protein under native conditions (Silow and Oliveberg, 1997a). A so-called folding intermediate could, therefore, be an artifact of aggregation. The question of whether this is a general problem remains to be determined. The true reason for the rollover in the refolding rate constants for hGST A1-1 was, therefore, investigated using various experimental conditions.

The initial refolding experiments were performed at a final protein concentration of 1μM hGST A1-1 in order to minimise the possibility of aggregation. In addition, control experiments monitoring aggregation suggest that under these conditions aggregation is not occurring. Therefore, it is unlikely that the rollover observed is the result of transient aggregation.

The bimolecular reaction is rate limiting throughout the range of denaturant conditions used in this study (section 6.1.2) and the absence of a change in the rate-

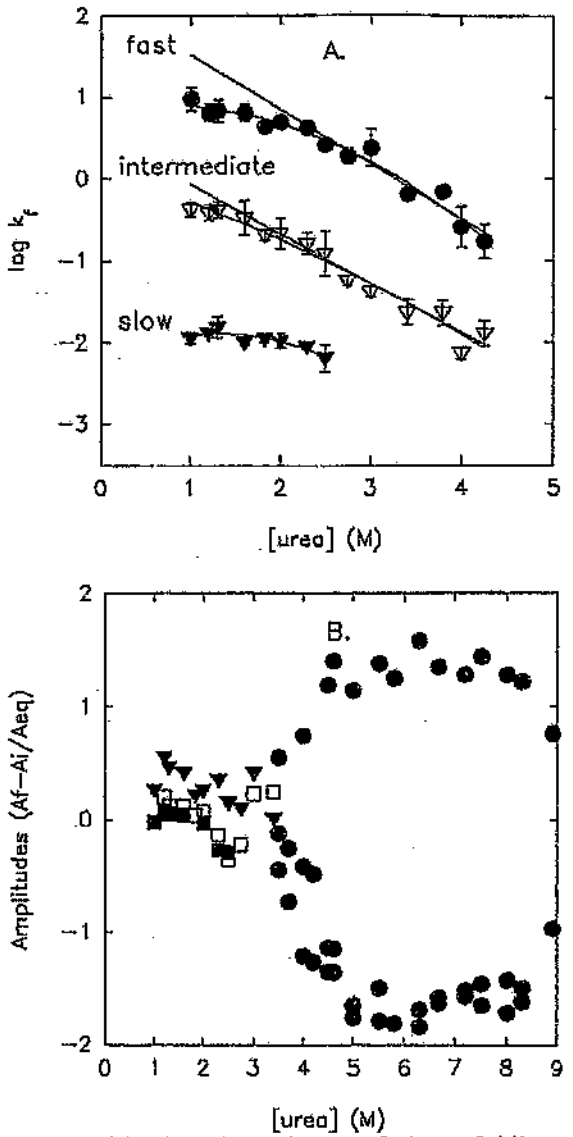


Figure 32. (a) Urea-dependence of the refolding rate constant for the fast (•), intermediate (▽) and slow (▼) phases of 1μM hGST A1-1. The straight-line is the best fit to a linear function and the curved line is the best fit to a polynomial function ($y = -ax^2+bx+c$). Refolding experiments were performed at 25°C, 0.1M NaCl, pH 6.5. Data were fitted to equation 13 for urea concentrations between 1M and 4M and to equation 14 for urea concentrations between 4M and 4.5M. The error bars are as indicated or are comparable to the size of the symbol. (b) The urea-dependence of the amplitudes for the fast (▼), intermediate (□) and slow (■) refolding events and the unfolding events (•) of hGST A1-1. The amplitudes were normalised with respect to the fraction-unfolded measurements obtained in equilibrium studies.

limiting step from bimolecular to unimolecular suggests that the deviation is not the result of accumulation of an intermediate. It is, therefore, possible that the deviation from linearity is the result of an underestimation of the refolding rate constants that occurs because 35% of the fast refolding event occurs as a burst phase.

The urea-dependence of the refolding rate (at concentrations less than 3.8M) was investigated at 15°C (figure 33a) and in the presence of 8.3% sucrose (figure 33b). These conditions enabled all the amplitude change for the fast phase to be collected. The refolding studies performed at 15°C and for those performed in the presence of sucrose, but otherwise identical conditions, displayed similar triphasic properties. For both conditions, the three phases were observed but with no missing amplitude for the fast phase. The fluorescence properties are identical to those described above and the slow refolding phase is again only apparent at final urea concentrations less than 2.5M. In the presence of sucrose, the slow refolding phase was observed but not fitted because the low signal-to-noise ratio prevented the accurate evaluation of the rate constant. The refolding kinetic traces for the fast and intermediate refolding phases under these conditions were best described using a biexponential function.

At 15°C, the apparent refolding rates for the major fast and intermediate refolding phases show a linear dependence on the final urea concentration ($r = 0.978$ and $r = 0.961$, respectively) (figure 33a). The apparent refolding rates decrease as the conditions favour the unfolded state. The data suggests that there are no changes in the rate-limiting step and that both events result in a population of molecules that differ in the extent of exposure of their side chains and/or peptide backbone. The linearity of this data throughout the range of denaturant conditions suggests that the observed non-linearity at 25°C is the result of the missing amplitude for the fast refolding event. Similarly, in the presence of 8.3% sucrose the refolding rates for the fast and intermediate phases show a linear dependence ($r = 0.977$ and $r = 0.895$, respectively) throughout the range of denaturant concentration (figure 33b). The

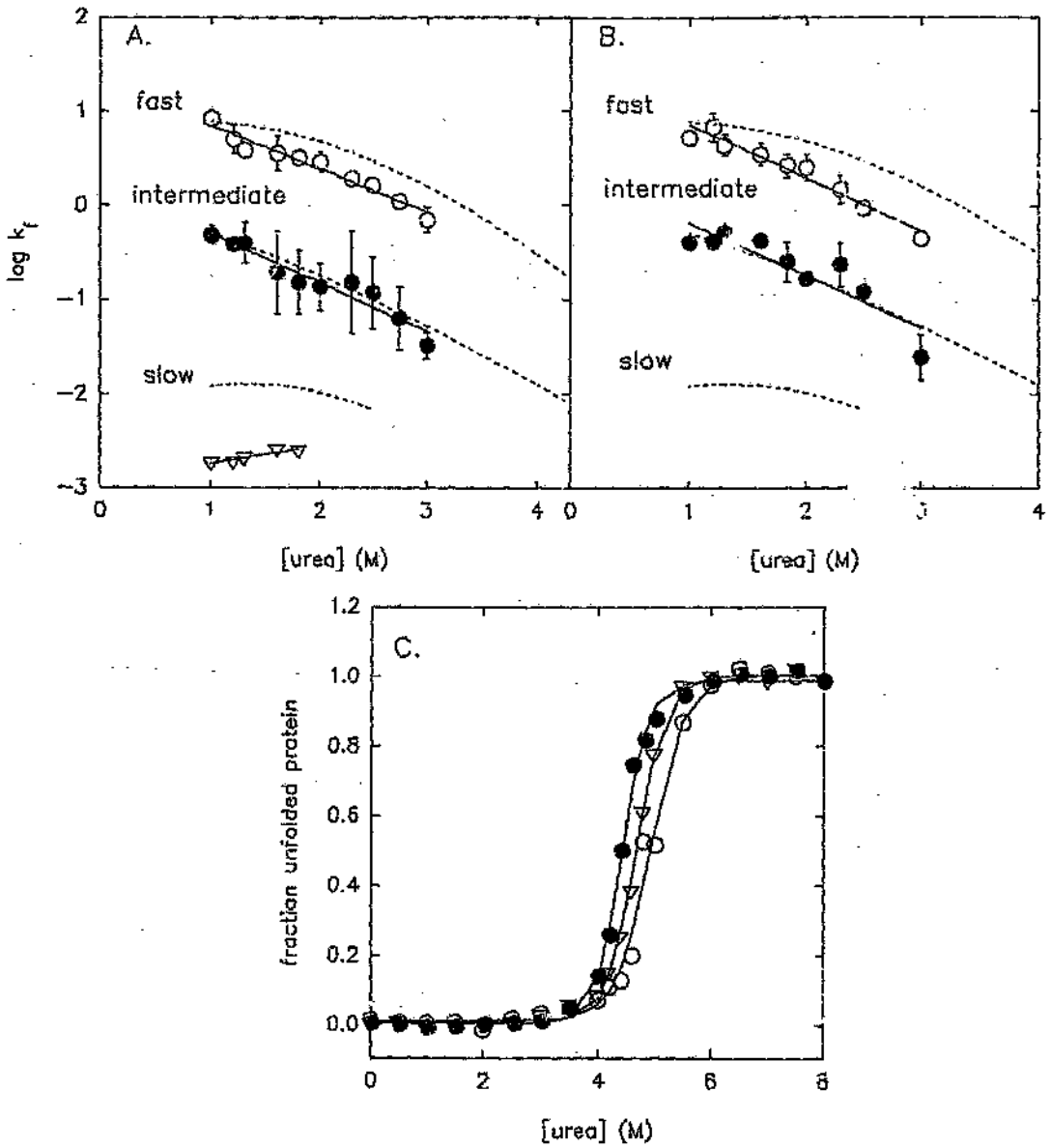


Figure 33. Urea-dependence of the refolding rate constants for the fast, intermediate and slow phases of hGST A1-1. The dashed lines represents the data from figure 32. The straight-line is the best fit to a linear function. (a) Experiments performed for 1μM hGST A1-1 at 15°C, 0.1M NaCl, pH 6.5. (b) Experiments performed for 1μM hGST A1-1 at 25°C, 0.1M NaCl containing 8.3% (v/v) sucrose, pH 6.5. (c) Equilibrium unfolding transitions for 1μM hGST A1-1 in the absence (•) and presence of 10% (▽) and 20% (o) sucrose.

reason for the observed decrease in the rate of refolding in the presence of sucrose is not clear. It may be the result of either an increase in the solvent viscosity which limits the rate of diffusion of the partially folded/folded monomers as they collide to form the native dimer or it may be the result of an increase in the stability as a result of the action of sucrose as a stabilising agent. The conformational stability of hGST A1-1 is altered by the presence of sucrose (figure 33c). The midpoint of the unfolding transition shifts from 4.5M urea (in the absence of sucrose) to 4.7M (in the presence of 10% sucrose), however, there is no evidence of a change in the unfolding/refolding pathway (m-value) or the Gibbs free energy change ($\Delta G(H_20)$). However, the action of sucrose as either a viscogenic agent and/or as a stabilising agent can only be differentiated through a more detailed study (Jacob *et al.*, 1997). It is, therefore, not possible to conclude that the decrease in the refolding rate is the result of diffusion control of the rate-limiting association step. However, from the protein concentration dependent studies this is most likely to be the case.

The urea-dependence of the data obtained for the different refolding conditions are parallel (figure 33a and b) indicating that the overall refolding pathways have not changed and that the decreased rate of refolding is the result of the change in experimental conditions and not the result of a change in the refolding pathway.

The complexities of the refolding data at 25°C and the presence of a multiple refolding pathways limits the accurate evaluation of the transition state solvent exposure (m_f) for each phase relative to the increase in the solvent exposure between the native and unfolded states (the *m*-value) (m_f/m) and the determination of the refolding rate in the absence of denaturant (k_f (H_20)) (Doyle *et al.*, 1996). The occurrence of all phases simultaneously and the overlap of their respective amplitudes will result in a gross underestimation of the refolding parameters.

Information on the transition state(s) for the folding pathway of hGST A1-1 can be inferred from a visual inspection of the sensitivity of the apparent refolding rates of

the fast and intermediate phases to urea (figure 32a). The fast and intermediate refolding events show a similar sensitivity to urea i.e., the parallel refolding rates indicate that the transition states for both phases occupy similar positions along the reaction co-ordinate diagram defined in terms of solvent accessibility. The shallow slopes for the fast and intermediate refolding phases indicate that the transition states for the refolding phases occur early in the refolding pathway, i.e., a small percentage of the native surface area is buried at the transition state.

6.1.4 Temperature-dependence of the refolding rates

The apparent refolding rates of the fast, intermediate and slow phases of hGST A1-1 were characterised from 5°C - 40°C at a final urea concentration of 1M and 3M. All phases are temperature dependent and the amplitude of the fast phase increased significantly such that a substantial percentage occurred in the dead time of the instrument. At 1M urea, at 40°C, only 50% of the fast phase was observable. It is, therefore, quite likely that the underestimation of the apparent refolding rates for the fast phase impacts on the thermodynamic activation parameters (see below) obtained for this phase. At the final urea concentration of 3M and throughout the entire temperature range, all of the amplitude change for the fast phase was detected. The rate constants for refolding for all phases show a complex dependence on the temperature of the reaction and were analysed using the transition state theory (Tan *et al.*, 1996)(section 4.3.3).

At 3M urea, the temperature dependence of the fast and intermediate phases show slight curvature as the temperature increases with the maximum refolding rate between 25°C and 30°C. A complex temperature dependence of the refolding rate has been noted for a large number of proteins. Generally, it is described to be the consequence of the difference in heat capacity between the unfolded state and the native state which results in temperature dependence of the enthalpy change (Oliveberg *et al.*, 1995). The deviation from the simple Arrhenius behaviour may also

be suggestive of a change in the mechanism of the refolding pathway (Oliveberg *et al.*, 1995).

At the final urea concentration of 3M, the plot of ln (k_fh/k_BT) versus 1/T for the fast and intermediate phases (figure 34a) was best fit to equation 18. The best fit for equation (18) to the fast refolding phase gives the following thermodynamic parameters: $\Delta C_{pN}^{\ddagger}$ of -1.36kcal^{-1}mol^{-1}K^{-1}; $\Delta H_N^{\ddagger}$(298K) of 6.26kcal.mol^{-1}; $\Delta S_N^{\ddagger}$(298K) of -41.38cal^{-1}mol^{-1}K^{-1}. Fitting the data of the intermediate refolding phase to equation (18) gives: $\Delta C_{pN}^{\ddagger}$ of -1.33kcal^{-1}mol^{-1}K^{-1}; $\Delta H_N^{\ddagger}$(298K) of 4.73kcal.mol^{-1}; $\Delta S_N^{\ddagger}$(298K) of -41.30cal^{-1}mol^{-1}K^{-1}.

At 1M urea, the curvature of the temperature dependent data is less apparent which is suggestive of a smaller difference in the heat capacity between this transition state and the previous one. For the final concentration of 1M, the plot of ln (k_fh/k_BT) versus 1/T all of the refolding phases (figure 34b) were best fit to equation (18) and a linear fit. The best fit for equation (18) to the fast refolding phase gives the following thermodynamic parameters: $\Delta C_{pN}^{\ddagger}$ of -471.8cal^{-1}mol^{-1}K^{-1}; $\Delta H_N^{\ddagger}$(298K) of 5.81kcal.mol^{-1}; $\Delta S_N^{\ddagger}$(298K) of -35.4cal^{-1}mol^{-1}K^{-1}. Fitting the data of the intermediate refolding phase to equation (18) gives: $\Delta C_{pN}^{\ddagger}$ of -340.4cal^{-1}mol^{-1}K^{-1}; $\Delta H_N^{\ddagger}$(298K) of 6.3kcal.mol^{-1}; $\Delta S_N^{\ddagger}$(298K) of -39.2cal^{-1}mol^{-1}K^{-1}. Fitting the data of the slow refolding phase to equation (18) gives: $\Delta C_{pN}^{\ddagger}$ of -336.6cal^{-1}mol^{-1}K^{-1}; $\Delta H_N^{\ddagger}$(298K) of 13.1kcal.mol^{-1}; $\Delta S_N^{\ddagger}$(298K) of -23.8cal^{-1}mol^{-1}K^{-1}.

The marginal curvature for the refolding phases may be the result of the small number of determinations in the experimentally accessible temperature range (Tan *et al.*, 1996). The slight curvature is, however, suggestive of a difference in heat capacity between the transition state and the unfolded state ($\Delta C_{pN}^{\ddagger} \neq 0$) (Oliveberg *et al.*, 1995) The apparent heat capacity changes at 3M urea for the fast and intermediate phases are greater than at 1M urea and this suggests that multiple transition states

determine the refolding pathway for hGST A1-1 although the intermediates were not significantly populated under these conditions. In general, the negative heat capacity value for refolding is the consequence of a large decrease in exposure of the hydrophobic surface to water as the reaction proceeds from the unfolded state to the transition state for folding. The refolding rate, therefore, decreases as the temperature increases because of disruption of the water molecules surrounding each state.

The change in the heat capacity ($\Delta C_{pN}^{\ddagger}$) for each refolding event can be related to the total heat capacity change for the folding/unfolding pathway ($\Delta C_{pN-U} = 7.39 kcal^{-1}mol^{-1}K^{-1}$)(section 4.3.3) and used to estimate the extent of buried surface area for each transition state. The values calculated are only estimates because of the apparent complexities of the refolding pathway. At 3M urea, the extent of buried surface area ($\Delta C_{pN}^{\ddagger}/\Delta C_{pN-U}$) is 18.0% for both the fast and intermediate refolding phases. At 1M urea, the extent of buried surface area ($\Delta C_{pN}^{\ddagger}/\Delta C_{pN-U}$) is 7.0%, 5% and 4.5% for the fast, intermediate and slow refolding phases, respectively. This data suggests, as proposed by the urea-dependence data in section 6.1.3, that the transition state for the folding of hGST A1-1 is not as compact as the folded state, i.e., the position of the transition state for refolding for each phase occurs early in the reaction co-ordinate diagram. At 3M urea a small percentage of the exposed hydrophobic surfaces are buried, therefore, major structural changes do occur early in the refolding pathway but they are not substantial when compared to the total expected heat capacity change.

The activation energy values at 1M urea (calculated using $\Delta H_N^{\ddagger} = E_A - RT$ (Oliveberg *et al.*, 1995) or obtained from a linear Arrhenius plot) are $8.50 kcal.mol^{-1}$, $9.50 kcal.mol^{-1}$ and $15.08 kcal.mol^{-1}$ for the fast, intermediate and slow refolding phases, respectively. The small values indicate that the barriers for the folding for the fast and intermediate events are low. For the slow refolding phase, the higher activation energy in addition to its slow time constant ($\tau \approx 100 sec$) suggest that this

event may be the result of *cis/trans* isomerisation about the peptide bond (Nall *et al.*, 1978). Because the slow refolding phase only occurs at final urea concentration less than 3M, it is unlikely that the event is the result of *cis/trans* peptide bond isomerisation.

The absolute values for the enthalpic and entropic contributions for refolding cannot be evaluated (section 4.3.3). However, the small positive enthalpy change is consistent with the general characteristics of refolding pathways and is the result of a balance of a number of factors. Similarly, the net entropic changes are small and negative consistent with a balance between an increase in entropy from the loss of water due to the hydrophobic effect and the loss in entropy due to ordering of the polypeptide chain. In general, the major thermodynamic changes for folding involve a decrease (generally large) in the heat capacity, an increase in enthalpy and a decrease in entropy.

6.2 Characterisation of refolded hGST A1-1

The physicochemical properties of native hGST A1-1 in the presence of 1M and 3M urea were compared with hGST A1-1 refolded to a final urea concentration of 1M and 3M. The overall tertiary structure and polarity of the tryptophan residues were compared using fluorescence spectroscopy. The wavelength of maximum emission for refolded hGST A1-1 at 1M and 3M urea are similar to that observed for native hGST A1-1 in 1M and 3M urea, respectively. The intensities are, however, quenched suggesting that either tertiary structural differences exist between these states and native hGST A1-1 or that refolding is only partially complete (figure 35a).

The quaternary structure of refolded hGST A1-1 in 3M urea, assessed using SEC-HPLC, indicates a symmetrical peak with a similar retention time to that for native hGST A1-1 in the presence of 3M urea (figure 35b). The overall hydrodynamic volumes of the native and refolded protein are identical and therefore the subunit-

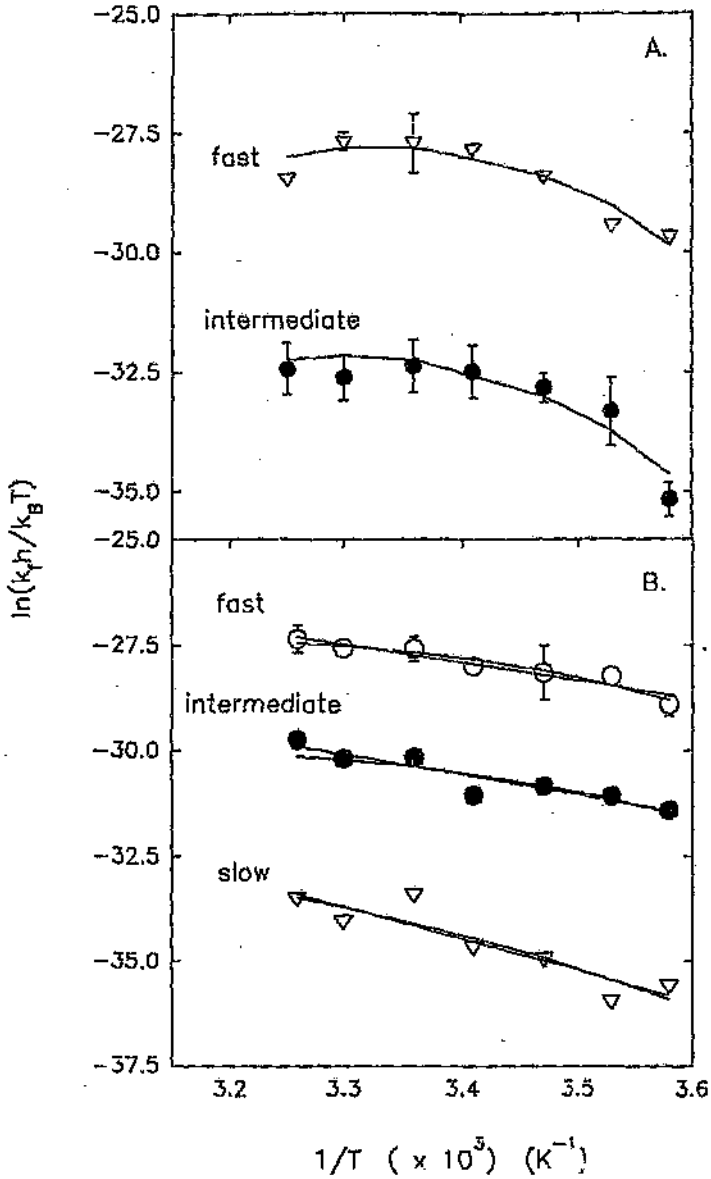


Figure 34. Eyring plots for the temperature dependence of the refolding rate constants for 1μM hGST A1-1. Refolding experiments were performed at (a) 3M and (b) 1M urea, 0.1M NaCl, pH 6.5. The plots for 3M urea are best fit to equation 18 and the plots at 1M urea are fit equally well to a linear function or to equation 18.

subunit interactions that stabilise the dimeric state are intact. In addition, refolded hGST A1-1 has 85% of the CDNB-conjugating activity of native hGST A1-1 in 3M urea and near 100% CDNB-conjugating activity at 1M urea.

The characterisation of the refolded protein, therefore, indicates that at 3M urea the refolded hGST A1-1 is dimeric with CDNB-conjugating activity. The comparison, although crude, indicates that the dimeric structure of hGST A1-1 is formed early in the refolding pathway and that structural rearrangements of the native-like dimer occur between 3M and 1M urea.

6.3 Discussion

The unfolding/folding pathway has been addressed for a number of oligomeric proteins. In general, the investigations of the folding pathway for oligomeric proteins can be divided into two groups based on their structural complexity and size. In the first group, the folding process is very complex with folding of the monomer proceeding to a significant extent before the association reaction. For example, the folding of the dimeric beta subunit of Trp synthase has a sequential unimolecular-bimolecular-unimolecular pathway (Zetina and Goldberg, 1980; Blond and Goldberg, 1985). The dimeric malate dehydrogenase folds through two parallel pathways involving a monomeric and dimeric intermediate (Ruldolph *et al.*, 1986). In the second group, the folding process is less complex with protein folding and association occurring as a concerted process. For example, dimeric P22 Arc repressor (Milla and Sauer, 1994) and the gene V protein of the bacteriophage f1 (Liang and Terwilliger, 1990) where folding is a two-state mechanism under most conditions. The *E.coli Trp* aporepressor dimer is also part of this group although its pathway is characterised by an interchange between unfolded, partially folded and completely folded monomers, followed by two sequential dimerization events (Gittelman and Matthews, 1990; Mann and Matthews, 1993).

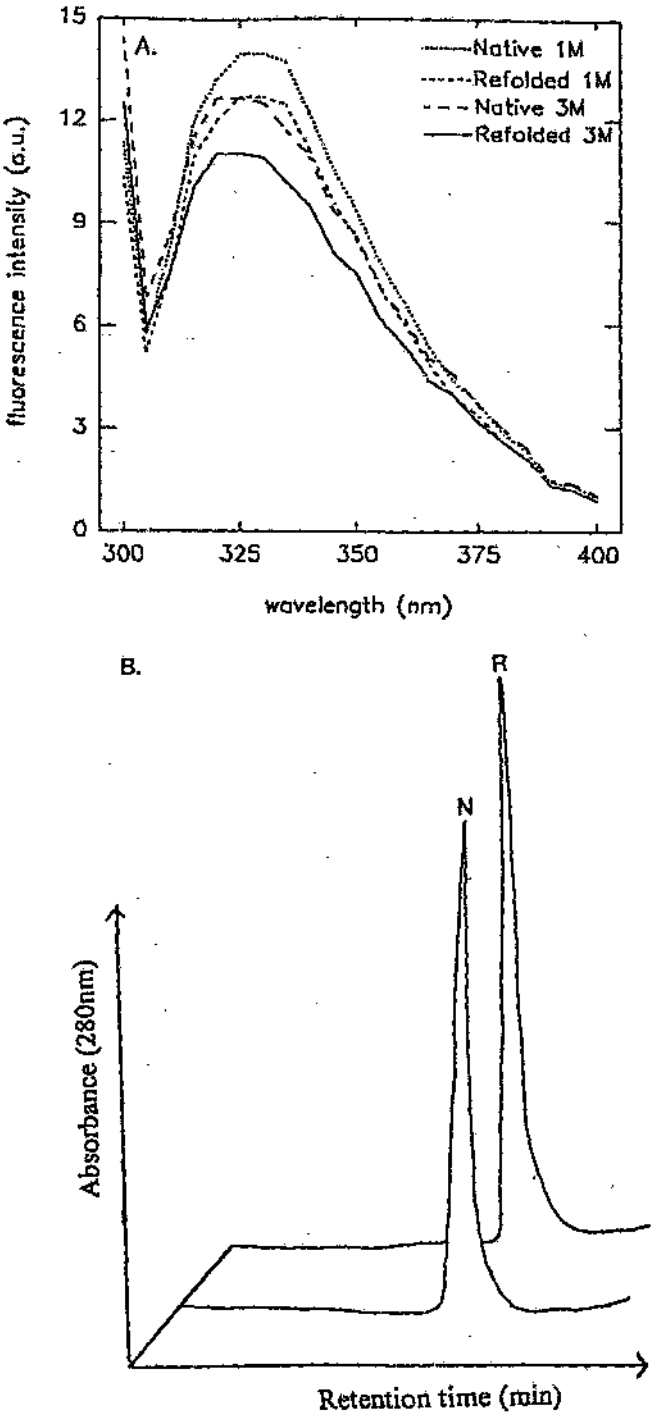


Figure 35. (a) Tryptophan emission spectra of native hGST A1-1 in 1M and 3M and refolded hGST A1-1 to final concentration of 1M and 3M. Trp was selectively excited at 295nm and emission scanned from 300nm-400nm. (b) SEC-HPLC profile of native hGST A1-1 in 3M urea (N) and hGST A1-1 refolded to 3M urea (R).

The folding pathways for oligomeric proteins, therefore, vary in complexity but, in general, appear to be a hierarchical process in which the individual domains first fold very rapidly to an almost native conformation and these folding units assemble to the native state. The co-ordination of these events provides insight into the mechanism of folding/unfolding.

The refolding kinetic studies of hGST A1-1 indicate that the refolding pathway is more complex than the minimal model where the two partially folded/folded monomers combine to form the native dimer in a single phase (i.e., $2U \rightarrow N_2$). The triphasic refolding kinetics, characterised by a fast, intermediate and slow refolding event may be described by a sequential pathway (e.g., $2U \rightarrow I_2 \rightarrow I_2' \rightarrow N_2$) or by multiple parallel pathways (e.g., $2U_1 \rightarrow N_2$; $2U_2 \rightarrow N_2$; etc). For hGST A1-1, the fast and the intermediate refolding events form the major events of the pathway and their similar urea and temperature dependence characteristics suggest that these events are involved in the main conformational changes along the refolding pathway. The parallel behaviour for refolding suggests that the two events are coupled. This, in addition to an undetectable lag phase (using Tyr/Trp fluorescence as the probe), suggests that each event is representative of an individual refolding pathway. The parallel pathway predicts two second-order reactions which is consistent with observed data (figure 31) whilst the sequential pathway predicts only a single pathway. Therefore, a sequential pathway would be incorrect. However, a parallel refolding pathway will only be confirmed when the same pathway is detected using different structural probes.

The urea-independent characteristic and the high activation energy of the slow refolding phase suggest that this phase is an isomerisation reaction resulting from structural heterogeneity of the unfolded state. However, the inverse amplitude of this event suggests that this event may be representative of the burial of Trp20 which is located at the interdomain interface. This event would occur late in the refolding pathway and the change in the fluorescence properties would result from the Trp

residue assuming its normal exposure to solvent that is found in the native conformation. It is possible that the two events may be occurring concurrently and are inseparable under the conditions of this study.

How the refolding pathways described by the fast and intermediate events differ is not clear. In general, different molecules may fold via distinct pathways because of structural heterogeneity of the unfolded states e.g., the result of different isomers of the peptide bond. The other reason is that the multiple pathways are the result of an initial rapid collapse of the protein which results in an intermediate (stable or unstable) which can either continue along the folding pathway or which requires structural reorganisation of the partially folded state (Radford *et al.*, 1992).

The refolding pathway of hGST A1-1 is consistent with a rate-limiting step for the fast and intermediate events, which is bimolecular. The bimolecular reaction, which is an association reaction, results in either a partially folded dimeric intermediate (I_2) or the native dimer (N_2). The dimeric state of hGST A1-1 is formed early in the folding pathway and the process of folding and association is a concerted process, indicating the absence of a monomeric intermediate. The absence of a folded monomer or a partially folded monomer is proposed to be the consequence of the large number of subunit-subunit interactions, which have resulted in the integration of the tertiary structure with the quaternary structure of glutathione transferases.

Urea and temperature dependence of the refolding rate suggests that the transition states for the two events are solvent accessible and closely related to the denatured state. The transition state for both phases involves a dimeric specie and because the transition states are largely solvent accessible in comparison with the native dimer it follows that the species are partially folded dimers or dimeric intermediates which through structural reorganisation form the native dimer. The transition state and hence the partially folded dimer (I_2*) is still compact with a hydrophobic interior but the surface interactions are poorly formed. The intermediates may not be optically

different which would explain why they were not detected under these refolding conditions. The partially folded dimer(s) would represent on-pathway intermediates. This type of pathway is consistent with a reaction profile with a narrow activation barrier with only a limited part of the refolding reaction occurring at the transition state (Silow and Oliveberg, 1997b). The reaction profile would be described by a number of free energy minima (figure 36).

The high solvent accessibility of the transition state is unusual since, in general, the refolding rate is highly sensitive to urea and hence the transition state occurs late in the refolding pathway and is typically solvent inaccessible (Waldburger *et al.*, 1996). Exceptions to this have been observed, for example with the chymotrypsin inhibitor 2 (CI2) (Itzhaki *et al.*, 1995), the acyl CoA binding protein (Kragelund *et al.*, 1995), the MYL mutant of the P22 Arc repressor (Waldburger *et al.*, 1996) and the lambda repressor (Huang and Oas, 1995). In general, there does not appear to be a direct correlation between the extent of folded structure in the transition state and the secondary structure (i.e., all α-helical, mixed α/β or only β-structure) of a protein. In addition, a logical consequence of an early transition state may be as a result of a large number of co-operative interactions that result in the formation of intermediates.

The significant contribution of long-range and middle-range electrostatic forces to the stability of an initial dimeric state without the need for formation of a native-like structure may explain why the transition state for refolding is closer to the unfolded state than to the folded state (Waldburger *et al.*, 1996; Wendt *et al.*, 1997). The probability that long- and middle-range electrostatic interactions are formed early in the refolding pathway is feasible for the glutathione transferases. A statistical method described by Zhu and Karlin (1996) has identified a significant mixed charge cluster at the dimer interface of a glutathione transferase. In general, mixed charged clusters over a long range orient and position the appropriate surfaces between which hydrophobic forces form and the two types of forces stabilise the quaternary structure

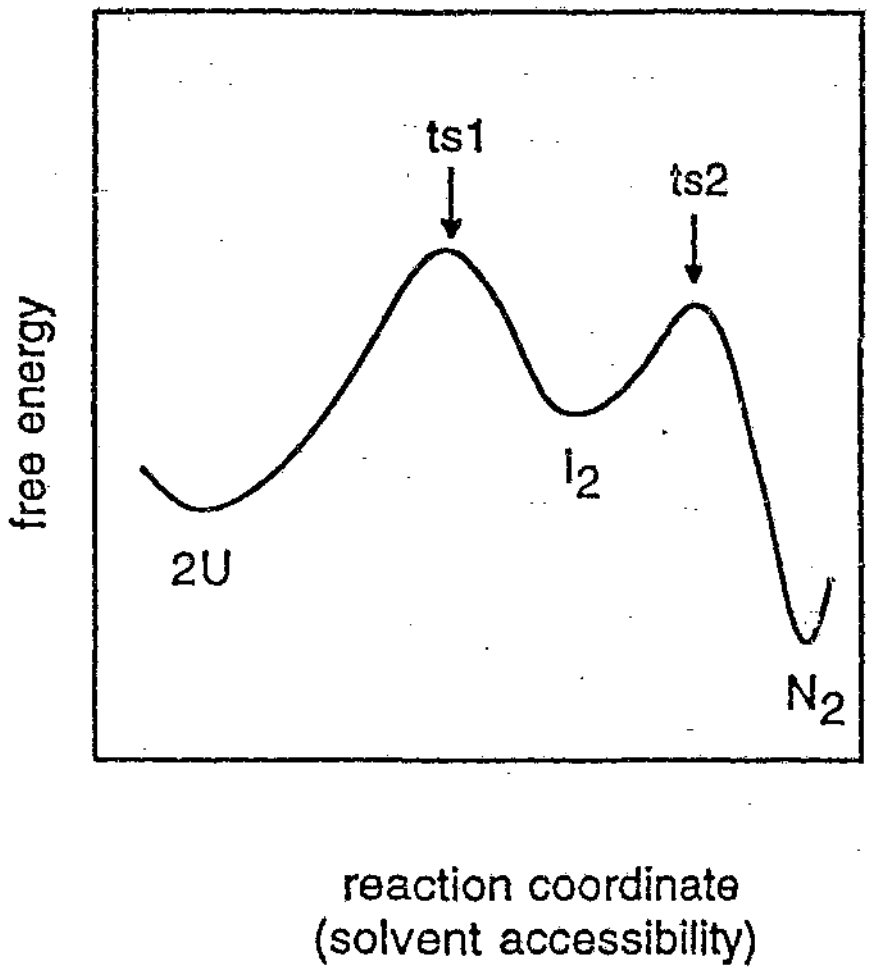


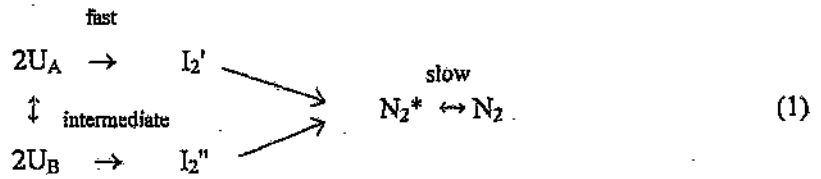
Figure 36. Proposed reaction co-ordinate diagram for the refolding pathway of hGST A1-1. The diagram shows a single route containing denatured monomer (U), a partially folded dimeric intermediate (I_2) and the native dimer (N_2). The narrow activation barrier indicates that a limited part of the reaction occurs at the transition state level.

of the protein. It is, therefore, feasible that the mixed charged cluster contributes significantly to the stability of the initial dimeric state of hGST A1-1.

The view of a poorly formed transition state stabilised by both electrostatic interactions and yet ill-defined hydrophobic interactions fits into the framework of a nucleation-condensation pathway (Fersht, 1995). In this model, which describes the folding properties for some proteins with solvent accessible transition states, a specific nucleus is formed which then directs the rest of the protein to fold around it. This model explains folding pathways for which there are no apparent stable intermediates and is a model which does not require stable elements of structure as a prerequisite for successful folding. For large proteins, it is possible that the folding process is initially a nucleation condensation process that in later stages becomes a hierarchical association of the smaller units.

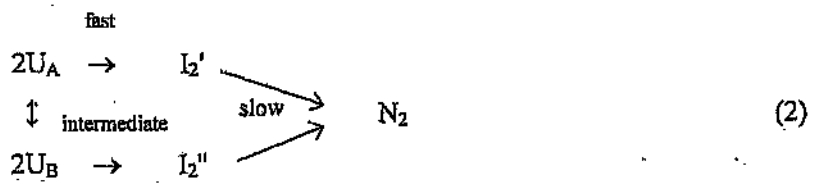
The minimal model for the folding pathway of hGST A1-1 involves a sequential and parallel pathway. The model is based on structural and experimental evidence. All events were not observed in this study but there is experimental evidence to support their inclusion in the pathway.

This model is minimal since it is based on the results obtained using one probe to monitor global tertiary changes of the structure of hGST A1-1. In general, a folding pathway is viewed as an energy landscape with multi-dimensions (Bryngelson and Wolynes, 1987) and a free-energy gradient towards the native structure. The funnel is rough with a large number of local minima, which act as traps during folding. Therefore, the folding pathway for hGST A1-1, described here, represents one route down the folding funnel.



(1)

OR



(2)

where U_A and U_B represent unfolded monomers which vary in their orientation about the peptide bond, I_2' and I_2'' represent partially folded dimeric intermediates, N_2^* represents a folded native-like dimer and N_2 the native dimer. The reactions, I_2' and I_2'' to N_2^* (mechanism (1)) and the $N_2^* \leftrightarrow N_2$ reaction (mechanism (2)) may be optically silent steps.

The pathway incorporates the fast $2U_A \rightarrow I_2'$ and the intermediate $2U_B \rightarrow I_2''$ events where U_A represents the unfolded form(s) in the correct *cis/trans* conformations and conversely U_B represents the unfolded form(s) with the incorrect *cis/trans* conformations. The two pathways result in a native-like dimeric intermediate and converge to form a native-like dimer (N_2^*) which would undergo domain/subunit reorganisation to form the native dimer (N_2). The various native forms (partially folded/folded; I_2' and I_2'') may not be optically distinguishable. The principle of microscopic reversibility suggests that the intermediate refolding event is the reverse of the slow unfolding event and therefore results in a native-like dimeric intermediate with partially dissociated domain I and II of each subunit. The loss in amplitude for

the slow phase at 3M urea suggests that the reaction stops at the dimeric intermediate stage (I_2) i.e., the dimeric intermediate is more stable than the native dimer at 3M urea.

6.5 Conclusion

The fluorescence-detected refolding mechanism for hGST A1-1 is described by a sequential and parallel pathway with evidence for on-pathway intermediates. A limited part of the reaction occurs at the transition state level. The role of long- and middle-range electrostatic forces i.e., a mixed charge cluster at the dimer interface, in stabilising the initial dimeric state is proposed. Studies of the effect of ionic strength on the association and folding of hGST A1-1 are required. The absence of a folded monomer for equilibrium and kinetic unfolding/refolding studies, monitored using tertiary structure probes, indicates that the co-operative interactions formed between the two subunits are necessary for stabilising the tertiary and quaternary structure of the class alpha glutathione transferase. Further investigations to evaluate the proposed model are required using multiple structural probes and amino acid replacements at the dimer interface.

CHAPTER 7
REFERENCES

Abkevich,V.I., Gutin,A.M. and Shakhnovich,E.I. (1994) Specific nucleus as the transition state for protein folding evidence from the lattice model. *Biochemistry* **33**, 10026-10036

Adang,A.E.P., Meyer,D.J., Brussee,J., van der Gen, A., Ketterer,B. and Mulder,G.J. (1989) Interaction of rat glutathione S-transferases 7-7 and 8-8 with γ-glutamyl or glycyl modified glutathione analogues. *Biochem.J.* **264**, 759-764

Adang,A.E.P., Brussee,J., van der Gen, A. and Mulder,G.J. (1990) The glutathione binding site in glutathione S-transferases: Investigations of the cysteinyl, glycyl and γ-glutamyl domains. *Biochem. J.* **269**, 47-54

Anfinsen,C.B. (1973) Principles that govern the folding of protein chains. *Science* **181**,223-239

Arakawa, T. and Timasheff, S.N. (1984) Protein stabilisation and destabilisation by guanidinium salts. *Biochemistry* **23**, 5924-5929

Arcus,V.L., Vuilleumier,S., Freund,S.M.V., Bycroft,M. and Fersht,A.R. (1995) A comparison of the pH, urea and temperature denatured states of barnase by heteronuclear NMR. Implications for the initiation of protein folding. *J. Mol. Biol.* **254**, 305-321

Armstrong,R.N. (1991) Glutathione S-transferases: reaction mechanism, structure and function. *Chem. Res. Toxicol.* **4**, 131-140

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

Atkins,W.M., Wang,R.W., Bird,A.W., Newton,D.J. and Lu,A.Y.H. (1993) The catalytic mechanism of glutathione S-transferase (GST): Spectroscopic determination of the pKa of Tyr9 in Rat A1-1 GST. *J.Biol.Chem.* **268**, 26, 19188-19191

Atkins,W.M., Dietze,E.C. and Ibarra,C. (1997) Pressure-dependent ionization of Tyr9 in glutathione S-transferase A1-1: Contribution of the C-terminal helix to a "soft" active site. *Protein Sci.* **6**, 873-881

Baldwin,R.L. and Esienberg,D. (1987) Protein stability, in *Protein Engineering* (Alan R. Liss ed. Inc) pp.127-148

Baldwin,R.L. (1991) Molten globules: specific or non-specific folding intermediate? *Chemtracts. Biochem. Mol .Biol.* **2**, 379-389

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

Banik,U., Saha [illegible] Mandal,N.C., Bhattacharyya,B. and Roy,S. (1992) Multiphasic denaturation [illegible] repressor by urea and its implications for the repressor structure. *Eur.* [illegible]-21

Barlow,D.J. and [illegible] (1983) Ion pairs in proteins. *J. Mol. Biol.* **168**, 867-885

Barycki,J.J. and Colman,P.F. (1997) Identification of the non-substrate steroid binding site of rat liver glutathione S-transferase, isoenzyme 1-1, by the steroid affinity label, 3β-(Iodoacetoxy)dehydroisoandrosterone. *Arch. Bioch. Biophys.* **345**, 16-31

Benson,A.M., Talalay,P., Keen,J.H. and Jakoby,W.B. (1977) Relationship between the soluble glutathione-dependent Δ^5-3-ketosteroid isomerase and the glutathione S-transferases of the liver. *Proc. Natl. Acad. Sci.USA.* **74**, 158-162

Bernasconi,C.F. (1976) What is a small perturbation? in *Relaxation Kinetics*, pp 76-97 Academic Press, New York

Bernstein,F.C., Koetzle,T.F., Williams,G.J.B., Meyer,E.F.Jr., Brice,M.D., Rodgers,J.R., Kennard,O., Shimanouchi,T. and Tasumi,M. (1977) The protein data bank: a computer-based archival file for macromolecular structures. *J. Mol. Biol.* **112**, 535-542

Bico,P., Erhardt,J., Kaplan,W. and Dirr, H. (1995) Porcine class π glutathione S-transferase: anionic ligand binding and conformational analysis. *Biochim. Biophys. Acta* **1247**, 225-230

Birdsall,B., King,R.W., Wheeler,M.R., Lewis,C.A., Goode,S.R., Dunlap,R.B. and Roberts,G.C.K. (1983) Correction for light absorption in fluorescence studies of protein-ligand interactions. *Anal. Biochem.* **132**, 353-361

Birmes,A., Sättler,A., Maurer,K.H. and Reisner,D. (1990) Analysis of the conformational transitions of proteins by temperature-gradient gel electrophoresis. *Electrophor.* **11**, 795-801

Björnestedt,R., Stenberg,G., Widersten,M., Board,P.G., Sinning,I., Jones,T.A. and Mannervik,B.(1995) Functional significance of arginine 15 in the active site of human class alpha glutathione transferase A1-1. *J. Mol. Biol.* **247**, 765-773

Blond,S. and Goldberg,M.E. (1985) Kinetics and importance of the dimerization step in the folding pathway of the β_2 subunit of *Escherichia coli*. *J. Mol. Biol.* **182**, 587-606

Board,P.G. and Mannervik,B. (1991) The contribution of the C-terminal sequence to the catalytic : ty of GST2, a human alpha class glutathione transferase. *Biochem. J.* **275**, 171-174

Board,P.G., Coggan,M., Wilce,M.C.J. and Parker,M.W. (1995) Evidence for an essential serine residue in the active site of the theta class glutathione transferase. *Biochem. J.* **331**, 247-250

Board,P.G., Baker,R.T., Chelvanayagam,G. and Jermiin,L.S. (1997) Zeta, a novel class of glutathione transferases in a range of species from plants to humans. *Biochem. J.* **328**, 929-935

Bowie,J.U. and Sauer,R.T. (1989) Equilibrium dissociation and unfolding of the Arc repressor dimer. *Biochemistry* **28**, 7139-7143

Boyland,E. and Chasseaud,L.F. (1969) The role of glutathione and glutathione S-transferase in mercapturic acid biosynthesis. *Adv. Enzymol.* **32**, 173-219

Bryngelson,J.D. and Wolynes,P.G. (1987) Spin glasses and statistical mechanics of protein folding. *Proc. Natl. Acad. Sci. USA.* **84**, 7524-7528

Buckle,A.M., Henrick,K. and Fersht,A.R. (1993) Crystal structural analysis of mutations in the hydrophobic cores of barnase. *J. Mol. Biol.* **234**, 847-860

Buckel,A.M., Cramer,P. and Fersht,A.R. (1996) Structural and energetic responses to cavity-creating mutations in hydrophobic cores: observation of a buried water molecule and the hydrophilic nature of such hydrophobic cavities. *Biochemistry* **35**, 4298-4305

Buetler,T.M. and Eaton,D.L. (1992) Glutathione S-transferase: amino acid sequence comparison, classification and phylogenetic relationship. *Environ. Carcino. Ecotox. Rev.* **C10**, 181-203

Burton,R.E., Huang,G.S., Daugherty,M.A., Fullbright,P.W. and Oas,T.G. (1996) Microsecond protein folding through a compact transition state. *J. Mol. Biol.* **263**, 311-322

Cameron,A.D., Sinning,I., L'Hermite,G., Olin,B., Board,P.G., Mannervik,B. and Jones,T.A. (1995) Structural analysis of human alpha-class glutathione transferase A1-1 in the apo-form and in complexes with ethacrynic acid and its glutathione conjugate. *Structure* **3**, 717-727

Chaffotte,A.F., Li,J-H., Georgescu,R.E., Goldeberg,M.E. and Tasayco,M.L. (1997) Recognition between disordered states: kinetics of the self-assembly of thioredoxin fragments. *Biochemistry* **36**, 16040-16048

Chan,H.S., Bromberg,S. and Dill, K.A. (1995) Models of cooperativity in protein folding. *Phil. Trans.R.Soc.Lond.B.* **348**, 61-70

Chen,N.L. and Schellman,J.A. (1989) Low-temperature unfolding of a mutant of phage T4 lysozyme. 1. Equilibrium studies. *Biochemistry* **28**, 689-691

Chen,B-L., Baase,W.A., Nicholson,H. and Schellman,J.A. (1992) Folding kinetics of T4 lysozyme and nine mutants at 12°C. *Biochemistry* **31**, 1464-1476

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

Chothia,C. (1984) Principles that determine the structure of proteins. *Annu. Rev. Biochem.* **53**, 537-572

Collins,K.D. and Washabaugh,M.W. (1985) The Hofmeister effect and the behaviour of water at interfaces. *Q.Rev.Biophys.* **18**, 323-422

Corbett,R.J.T. and Roche,R.S. (1984) Use of high-speed size-exclusion chromatography for the study of protein folding and stability. *Biochemistry* **23**, 1888-1894

Creighton,T.E. (1979) Electrophoretic analysis of the unfolding of proteins by urea. *J. Mol. Biol.* **129**, 235-264

Danielson,U.H. and Mannervik,B. (1985) Kinetic independence of the subunits of cytosolic glutathione transferase from the rat. *Biochem. J.* **231**, 263-267

Dao-pin,S., Anderson,D.E., Baase,W.A., Dahlquist,F.W. and Matthews,B.W. (1991) Structural and thermodynamic consequences of burying a charged residue within the hydrophobic core of T4 lysozyme. *Biochemistry* **30**, 11525-11529

DeJong,J.L., Morgenstern,R., Jornvall,H., DePierre,J.W. and Tu,C-P.D. (1988) Gene expression of rat and human microsomal glutathione S-transferases. *J. Biol. Chem.* **263**, 8430-8436

Dietze,E.C., Wang,R.W., Lu,A.Y.H. and Atkins,W.M. (1996) Ligand effects on the fluorescence of tyrosine 9 in alpha 1-1 Glutathione S-transferase. *Biochemistry* **35**, 6745-6753

Dill,K.A. (1990) Dominant forces in protein folding. *Biochemistry* **29**, 7133-7155

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

Dill,K.A. (1995) Theory for the folding and stability of globular proteins. *Biochemistry* **24**, 1501-1509

Dill,K.A. and Chan, H.S. (1997) From Levinthal to pathways to funnels. *Nat. Struct. Biol.* **4**, 10-19

Dirr,H.W. and Reinemer,P. (1991) Equilibrium unfolding of class π glutathione S-transferase. *Biochem. Biophys. Res. Commun.* **180**, 1, 294-300

Dirr,H.W., Reinemer,P. and Huber,R. (1994a) Refined crystal structure of porcine class pi glutathione S-transferase (pGST P1-1) at 2.1Å resolution. *J. Mol. Biol.* **243**, 72-79

Dirr,H.W., Reinemer,P. and Huber,R. (1994b) X-ray crystal structures of cytosolic glutathione S-transferases. Implications for protein architecture, substrate recognition and catalytic function. *Eur.J.Biochem.* **220**,645-661

Dötsch,V., Wider,G., Siegal,G. and Wüthrich,R.(1995) Interaction of urea with an unfolded protein in the DNA-binding domain of the 434-repressor. *FEBS Lett.* **366**, 6-10

Doyle,D.F., Waldner,J.C., Parkh,S., Alcazar-Roman,L. and Pielak,G.J. (1996) Changing the transition state for protein (un)folding. *Biochemistry* **35**, 7403-7411

Edsall,J.T. (1995) Hsien Wu and the first theory of protein denaturation (1931) *Adv. Prot.Chem.***46**, 1-26

Eklund,H., Ingelman,M.M., Soderberg,B-O., Nordlund,T.U.P., Nikkola,M., Sonnerstam,U. and Joelson,T. (1992) Structure of oxidized bacteriophage T4 glutaredoxin (thioredoxin). Refinement of native and mutant protein. *J. Mol. Biol.* **228**, 596-618

Elöve,G., Chaffotte,A.F., Roder,H. and Goldberg,M.E.(1992) Early steps in cytochrome c folding probed by time-resolved circular dichroism and fluorescence spectroscopy. *Biochemistry* **31**, 6876-6883

Epp,O., Ladenstein,R. and Wendel,A. (1983) The refined structure of the selenoenzyme glutathione peroxidase at 0.2nm resolution. *Eur. J. Biochem.* **133**, 51-69

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

Erhardt,J. and Dirr,H. (1996) Effect of glutathione, glutathione sulphonate and S-hexylglutathione on the conformational stability of class pi glutathione S-transferase. *FEBS Lett.* **391**, 313-316

Ericksson, A.E., Baase, W.A., Zhang, X-J., Heinz, D.W., Blaber, M., Baldwin, E.P. and Matthews, B.W. (1992) Response of a protein structure to cavity-creating mutations and its relation to the hydrophobic effect. *Science* **253**, 178-183

Fersht,A.R., Matouschek,A. and Serrano,L. (1992) The folding of an enzyme. I. Theory of protein stability and pathway of protein folding. *J. Mol. Biol.* **224**, 771-782

Fersht, A.R. (1993) Protein folding and stability: the pathway of folding of barnase. *FEBS Lett.* **325**, 5-16

Fersht,A.R. (1994) Pathway and stability of protein folding. *Biochem.Soc. Trans.* **22**, 267-273

Fersht,A.R. (1995) Optimization of rates of protein folding: the nucleation-condensation mechanism and its implications. *Proc. Natl. Acad. Sci. USA.* **92**, 10869-10873

Fink,A.L. (1995) Molten Globules in *Protein stability and folding theory and practice Methods in Molecular biology. Vol 40.* (Shirley,B.A.ed.) pp.343-360, Humana Press, Totowa, New Jersey

Finkelstein,A.V. and Shakhnovich, E.I. (1989) Theory of co-operative transitions in protein molecules. II. Phase diagram for a protein molecule in solution. *Biopolymers* **28**, 1681-1689

Freire,E. (1995) Differential scanning calorimetry in *Methods in Molecular Biology, Vol. 40: Protein stability and Folding: Theory and Practice* (Shirley,B.A. ed), pp 191-218, Humana Press Inc., Totowa, New Jersey

Garcia-Sáez,I., Párraga,A., Phillips,M.F., Mantle,T.J. and Coll,M. (1994) Molecular structure at 1.8Å of mouse liver class pi glutathione S-transferase complexed with *S*-(*p*-nitrobenzyl)glutathione and other inhibitors. *J. Mol. Biol.* **237**, 298-314

Garel,J-R. (1992) Folding of large proteins: multidomain and multisubunit proteins, In *Protein folding*. (Creighton,T.E., ed), pp.405-454, Freeman and Company, New York.

Gittelman,M.S. and Matthews,C.R. (1990) Folding and stability of *trp* Aporepressor from *Escherichia coli*. *Biochemistry* **29**, 7011-7020

Gryczynski,I., Etfink,M. and Lakowicz,J.R. (1988) Conformation heterogeneity in proteins as an origin of heterogeneous fluorescence decays, illustrated by native and denatured ribonuclease T_1. *Biochim. Biophys. Acta* **954**, 244-252

Habig,W.H. and Jakoby,W.B. (1981) Assays for differentiation of glutathione S-transferases. *Methods Enzymol.* 77, 398-405

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

Hartl,F.U. (1996) Molecular chaperones in cellular protein folding. *Nature* **381**, 571-579

Hayes,J.D. and Wolf,C.R. (1988) Role of glutathione transferase in drug resistance, in *Glutathione Conjugation: Mechanisms and Biological Significance* (Sies,H. and Ketterer,B., eds) pp. 316-356, Academic Press, London

Hoesch,R.M. and Boyer,T.D. (1989) Localization of a portion of the active site of two rat liver glutathione S-transferases using a photoaffinity label. *J.Biol. Chem.* **264**, 17712-17717

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

Houry,W.A., Rothwarf,D.M. and Scheraga,H.A. (1994) A very fast phase in the refolding of disulfide-intact ribonuclease A: implications for the refolding and unfolding pathways. *Biochemistry* **33**, 2516-2530

Houry,W.A. and Scheraga,H.A. (1996) Nature of the unfolded state of ribonuclease A: Effect of cis-trans X-Pro peptide bond isomerisation. *Biochemistry* **35**, 1 1719-11733

Huang,G.S. and Oas,T.G. (1995) Submillisecond folding of monomeric λ repressor. *Proc. Natl. Acad. Sci. USA*. **92**, 6878-6882

Hubbard,S.J., Gross,K.H. and Argos,P. (1994) Intramolecular cavties in globular proteins. *Prot. Eng.* 7, 613-626

Hudson,E.N. and Weber,G. (1973) Synthesis and characterisation of two fluorescent sulphydryl reagents. *Biochemistry* **12**, 4154-4161

Ishikawa,T. (1992) The ATP-dependent glutathione S-conjugate export pump. *Trends Biochem. Sci.* 17, 463-468

Itzhaki,L.S., Otzen,D.E. and Fersht,A.R. (1995) The structure of the transition state for folding of chymotrypsin inhibitor 2 analyzed by protein engineering methods. *J.Mol.Biol.* **254**, 260-288

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

Jackson, S.E., Moracci, M., elMasry, N., Johnson, C.M. and Fersht AR. (1993a) Effect of cavity-creating mutations in the hydrophobic core of chymotrypsin inhibitor 2. *Biochemistry* **32**, 11259-11269

Jackson, S.E., elMasry, N. and Fersht, A.R., (1993b) Structure of the hydrophobic core in the transition state for folding of chymotrypsin inhibitor 2: A critical test of the protein engineering method of analysis. *Biochemistry* **32**, 11270-11278

Jacob,M., Schindler,T., Balbach,J. and Schmid,F.X. (1997) Diffusion control in an elementary protein folding reaction. *Proc. Natl. Acad. Sci. USA.* **94**, 5622-5627

Jaenicke,R. (1984) Protein folding and protein association. *Angew. Chem. Int.Ed.Engl.* **23**, 395-413

Jaenicke,R. and Ruldolph,R. (1986) Refolding and association of oligomeric proteins. *Methods Enzymol.* **131**, 218-250

Jaenicke,R. (1987) Folding and association of proteins. *Prog. Biophys. Mol. Biol.* **49**, 117-237

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

Jaenicke,R. (1995) Folding and association versus misfolding and aggregation of proteins. *Phil.Trans.R.Soc.Lond.B.* **348**, 97-105

Jakoby,W.B. and Ziegler,D.M. (1990) The enzymes of detoxification. *J. Biol. Chem.* **265**, 20715-20718

Ji,X., Zhang,P., Armstrong,R.N. and Gilliland,G.L. (1992) A three-dimensional structure of a glutathione S-transferase from the mu gene class. Structural analysis of the binary complex of isoenzyme 3-3 and glutathione at 2.2Å resolution. *Biochemistry* **31**, 10169-10184

Ji,X., Armstrong,R.N. and Gilliland,G.J. (1993) Snapshots along the reaction co-ordinate of an S_NAr reaction catalysed by glutathione transferase. *Biochemistry* **32**, 12949-12954

Ji,X., von Rosenvinge,E.C., Johnson,W.W., Tomarev,S.I., Piatigorsky,J., Armstrong,R.N. and Gilliland,G.J. (1995) Three-dimensional structure, catalytic properties, and evolution of a sigma class glutathione transferase from squid, a progenitor of the lens *S*-crystallins of cephalopods. *Biochemistry* **34**, 5317-5328

Ji,X., von Rosenvinge,E.C., Johnson,W.W., Armstrong,R.N. and Gilliland,G.J. (1996) Location of a potential transport binding site in a sigma class glutathione transferase by X-ray crystallography. *Proc. Natl. Acad. Sci. USA.* **93**, 8208-8213

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

Kabsch,W. and Sander,C. (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen bonded and geometrical features. *Biopolymers* **22**, 2577-2637

Kaplan,W. (1997) The conformational stability of a detoxification enzyme widely used as a fusion-protein affinity tag. *PhD thesis* University of the Witwatersrand, Johannesburg

Kaplan,W., Hüsler,P., Klump,H., Erhardt,J., Sluis-Cremer,N. and Dirr,H.W. (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 Šali,A. (1995) Theoretical studies of protein folding and unfolding. *Curr. Opin. Struct. Biol.* **5**, 58-73

Karshikoff,A., Reinemer,P., Huber,R. and Ladenstein,R. (1993) Electrostatic evidence for the activation of the glutathione thiol by Tyr7 in the π-class glutathione transferases. *Eur. J. Biochem.* **215**, 663-670

Kauzmann,W. (1959) Some factors in the interpretation of protein denaturation. *Adv. Prot. Chem.* **14**, 1-63

Kellis,J.T.Jr., Nyberg,K., Šali,D. and Fersht,A.R. (1988) Contribution of hydrophobic interactions to protein stability. *Nature* **333**, 784-786

Kellis, J.T.Jr., Nyberg, K. and Fersht, A.R. (1989) Energetics of complementary side-chain packing in a protein hydrophobic core. *Biochemistry* **28**, 4914-4922

Ketterer,B., Tan,K.H., Meyer,D.J. and Coles,B. (1987) Glutathione transferases: a possible role in the detoxification of DNA and lipid hydroperoxides, in Glutathione S-transferases and carcinogenesis (Mantle,T.J., Pickett,C.B. and Hayes,J.D. eds.) p149. Taylor and Francis. London.

Kiefhaber,T. (1995) Kinetic traps in lysozyme folding. *Proc. Natl. Acad. Sci. USA.* **92**, 9029-9033

Kiefhaber,T. (1995) Protein folding kinetics in *Methods in Molecular Biology, Vol. 40: Protein stability and Folding: Theory and Practice* (Shirley,B.A. ed), pp 313-341, Humana Press Inc., Totowa, New Jersey

Kim,P.S. and Baldwin,R.L. (1990) Intermediates in the folding reactions of small proteins. *Annu. Rev. Biochem.* **59**, 631-660

Kolm,R.H., Sroga,G.E. and Mannervik,B. (1992) Participation of the phenolic hydroxyl group of Tyr-8 in the catalytic mechanism of human glutathione transferase P1-1. *Biochem. J.* **285**, 537-540

Kong,K.H., Nishida,M., Inoue,H. and Takahashi,K. (1992) Tyrosine-7 is an essential residue for the catalytic activity of human class pi glutathione S-transferase: chemical modification and site-directed mutagenesis studies. *Biochem. Biophys. Res. Commun.* **182**, 1122-1129

K[illegible]on,R. and Hammes,G.G. (1976) A kinetic study of protein-protein interactions. *Biochemistry* **15**, 1165-1170

Kragelund,P.B., Robinson,C.V., Knudson,J., Dobson,C.M. and Poulson,F.M. (1995) Folding of a four-helix bundle: studies of acyl-coenzyme A binding protein. *Biochemistry* **34**, 7217-7224

Kumar,T.K.S., Jayaraman,G., Lin,W-Y. and Yu,C. (1996) Effect of chaotropic denaturant on the binding of 1-anilino-8-naphthalene sulphonic acid to proteins. *Biochim. Biophys. Acta* **1294**, 103-105

Kuwajima,K. (1977) A folding model of alpha-lactabumin deduced from three-state denaturation mechanism. *J. Mol. Biol.* **114**, 241-258

Lacowicz,J.R. (1983) *Principles of fluorescence spectroscopy.* Plenum Press, New York

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

Lee,B. and Vasmatzis,G. (1997) Stabilisation of protein structures. *Curr. Oppin. Biotechnol.* **8**, 423-428

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

Liang,H. and Terwilliger,T.C. (1991) Reversible denaturation of the Gene V protein of Bacteriophage f1. *Biochemistry* **30**, 2772-2782

Lim,W.A., Farruggio,D.C. and Sauer,R.T. (1992) Structural and energetic consequences of disruptive mutations in a protein core. *Biochemistry* **31**, 4324-4333

Lim,K., Ho,J.X., Keeling,K., Gilliland,G.L., Xinhua,J., Ruker,F. and Carter,D.C. (1994) Three-dimensional structure of *Schistosoma japonicum* glutathione S-transferase fused with a six-amino acid conserved neutralizing epitope of gp41 from HIV. *Protein Sci.* **3**, 2233-2244

Listowski,I. (1993) Glutathione S-transferases: intracellular binding, detoxification and adaptive responses, *In Hepatic transport and bile secretion: physiology and pathophysiology* (Tavoloni,N. and Berk, P.D. eds) pp 397-403. Raven Press, New York.

Liu,S., Zhang,P., Ji,X., Johnson,W.W., Gilliland,G.L. and Armstrong,R.N. (1992) Contribution of tyrosine 6 to the catalytic mechanism of isoenzyme 3-3 of glutathione S-transferase. *J. Biol. Chem.* **267**, 4269-4299

Mach,H., Volkin,D.B., Burke,C.J. and Middaugh.C.R. (1995) Ultraviolet absorption spectroscopy in *Methods in Molecular Biology, Vol. 40: Protein stability and Folding: Theory and Practice* (Shirley,B.A. ed), pp 91-114, Humana Press Inc., Totowa, New Jersey

Manoharan,T.H., Gulick,A.M., Reinemer,P., Dirr, H.W., Huber,R. and Fahl,W.E. (1992) Mutational substitution of residues implicated by crystal structure in binding of the substate glutathione to human glutathione S-transferase π. *J. Mol. Biol.* **226**, 319-322

Mann,C.J. and Matthews,C.R. (1993) Structure and stability of an early folding intermediate of *Escherichia coli trp* Aporepressor by far-UV stopped-flow circular dichroism and 8-anilino-1-naphthalene sulphonate binding. *Biochemistry* **32**, 5282-5290

Mannervik,B., Ali,P., Guthenberg,C., Jensson,H., Tahir,M.K., Warholm,M. and Jornvall,H. (1985) Identification of three classes of cytosolic glutathione transferase common to several mammalian species: correlation between structural data and enzymatic properties. *Proc. Natl. Acad. Sci. USA.* **82**, 7202-7206

Mannervik,B. and Danielson,U.H. (1988) Glutathione Transferases: structure and catalytic activity. *CRC. Crit. Rev. Biochem.Mol.Biol.* **23**, 283-337

Mannervik,B., Awashti,Y.C., Board,P.G., Hayes,J.D., Di Ilio,C., Ketterer,B., Listowski,I., Morgenstern,R., Muramatsu,M., Pearson,W.R., Pickett,C.B., Sato,K., Widersten,M. and Wolf,C.R. (1992) Nomenclature for human glutathione transferases. *Biochem. J.* **282**, 305-306

Marky,L.A. and Breslauer,K.J. (1987) Calculating thermodynamic data for transitions of any molecularity from equilibrium melting curves. *Biopolymers* **26**, 1601-1620

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

Matouschek,A., Kellis,J.J., Serrano,L. and Fersht,A.R. (1989) Mapping the transition state and pathway of protein folding by protein engineering. *Nature* **340**, 122-126

Matsumura,M., Beektel,W.J. and Matthew,S.B.E. (1988) Hydrophobic stabilisation in T4 lysozyme determined directly by multiple substitution of Ile3. *Nature (London)* **334**, 406-410

Matthews,C.R. (1993) Pathways of protein folding. *Annu.Rev.Biochem.* **62**, 653-683

Mei,G., Rosato,N., Silva,N., Rusch,R., Gratton,E., Savini,I. and Finazzi-Agro,A. (1992) Denaturation of human Cu/Zn superoxide dismutase by guanidine hydrochloride: A dynamic fluorescence study. *Biochemistry* **31**, 7224-7230

Meyer,D.J., Xia,C., Coles,B., Chen,J., Reinemer,P., Huber,R. and Ketterer,B. (1993) Unusual reactivity of Tyr-7 of GSH transferase P1-1. *Biochem. J.* **293**, 351-356

McTigue,M., Williams,D.R. and Tainer,J.A. (1995) Crystal structures of a schistosomal drug and vaccine target: glutathione S-transferase from *Schistosoma japonica* and its complex with the leading antischistosomal drug praziquantel. *J. Mol. Biol.* **246**, 21-27

Meyer,D.J., Coles,B., Pemble,S.E., Gilmore,K.S., Fraser,G.M. and Ketterer,B. (1991) Theta, a new class of glutathione transferases purified from rat and man. *Biochem. J.* **274**, 409-414

Milla,M.C. and Sauer,R.T. (1994) P22 Arc repressor: folding kinetics of a single domain dimeric protein. *Biochemistry* **33**, 1125-1133

Murphy,K.P. (1995) Non-covalent forces important to the conformational stability of protein structures in *Protein stability and folding theory and practice Methods in Molecular biology. Vol 40.* (Shirley,B.A.ed.) pp.1-34, Humana Press, Totowa, New Jersey

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 folding. *Protein Sci.* **4**, 2138-2148

Myers,J.K. and Pace,C.N. (1996) Hydrogen bonding stabilises globular proteins. *Biophys. J.* **71**, 2033-2039

Nall,B.T., Garel,J. and Baldwin,R.L. (1978) Test of the extended two-state model for the kinetic intermediates observed in the folding transition of ribonuclease A. *J.Mol.Biol.* **118**, 317-330

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

Neri,D., Billeter,M., Wider,G. and Wuthrich,K. (1992) NMR determination of residual structure in a urea-denatured protein, the 434-Repressor. *Science* **257**, 1559-1563

Oakley,A.J., Rossjohn,J., Lo Bello,M., Caccuri,A.M., Federici,G. and Parker,M.W. (1997) The three-dimensional structure of the human class pi glutathione S-transferase P1-1 in complex with the inhibitor ethacrynic acid and its glutathione conjugate. *Biochemistry* **36**, 576-585

Ohlmeyer,H.R. (1997) The effect of non-substrate ligands on the conformational stability of class pi glutathione S-transferase. *Masters dissertation* University of the Witwatersrand, Johannesburg

Oliveberg,M., Tan,Y-J. and Fersht,A.R. (1995) Negative activation enthalpies in the kinetics of protein folding. *Proc. Natl. Acad. Sci. U.S.A.* **92**, 8926-8929

Orozco,M., Vega,C., Parraga,A., Garcia-Sáez,I., Coll,M., Walsh,S., Mantle,T.J. and Luque,F.J. (1997) On the reaction mechanism of class pi glutathione S-transferase. *Proteins: Struct. Funct. Genet.* **28**, 530-542

Pace,C.N., Grimsley,G.R., Thomson,J.A. and Barnett,B.J. (1988) Conformational stability and activity of ribonuclease T1 with zero, one and two intact disulphide bonds. *J. Biol. Chem.* **263**, 11820-11825

Pace,C.N., Shirley,B.A. and Thomson,J.A. (1989) in *Protein Structure*: a practical approach (Creighton,T.E. ed) 2nd edn, pp 311-330, IRL Press, Oxford University Press, Oxford

Pace,C.N. (1992) Contribution of the hydrophobic effect to globular protein stability. *J. Mol. Biol.* **226**,29-35

Palleros, D.R., Shi,L., Reid,K.L. and Fink, A.L. (1993) Three-state denaturation of DnaK induced by guanidine hydrochloride. Evidence for an expandable intermediate. *Biochemistry* **32**, 4314-4321

Parsons,W. (1997) Conformational stability of glutathione S-transferase in complex with G-site ligands. *Masters dissertation* University of the Witwatersrand, Johannesburg

Parsons,J.F. and Armstrong,R.N. (1996) Proton configuration in the ground state and transition state of a glutathione transferase-catalysed reaction inferred from the properties of tetradeca(3-fluorotyrosyl)glutathione transferase. *J. Am. Chem. Soc.* **118**, 2295-2296

Pemble,S.E. and Taylor,J.B.(1992) An evolutionary perspective on glutathione transferase inferred from class theta glutathione transferase cDNA sequences. *Biochem. J.* **287**, 957-963

Pemble,S.E., Wardle,A.F. and Taylor,J.B. (1996) Glutathione S-transferase (GST) class Kappa: characterisation by the cloning of rat mitochondrial GST and identification of a human homologue. *Biochem.J.* **319**, 754-759

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

Poland,D.C. and Scheraga,H.A. (1970) *Theory of helix-coil transition.* New York. Academic Press

Privalov,P.L. and Potekhin,S.A. (1986) Scanning microcalorimetry in studying temperature-induced changes in proteins. *Methods Enzymol.* **131**, 4-51

Privalov,P.L. and Gill,S.J. (1988) Stability of protein structure and hydrophobic interaction. *Adv. Prot. Chem.* **39**, 191-235

Privalov,P.L. (1996) Intermediate states in the protein folding. *J. Mol. Biol.* **258**, 707-725

Ptitsyn,O.B. (1992) The molten globule state. In *Protein folding*. (Creighton,T.E., ed), pp.243-300, Freeman and Company, New York.

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

Radford,S.E., Dobson,C.M. and Evans,P.A. (1992) The folding of hen lysozyme involves partially structured intermediates and multiple pathways. *Nature* **358**, 302-307

Radford,S.E. and Dobson,C.M. (1995) Insights into protein folding using physical techniques: studies of lysozyme and alpha-lactalbumin. *Phil. Trans. R. Soc. Lond. B.* **348**, 17-25

Ragone,R., Colonna,G., Balestrieri,C., Servillo,L. and Irace,G. (1984) Determination of tyrosine exposure in proteins by second-derivative spectroscopy. *Biochemistry* **23**, 1871-1875

Raghunathan,S., Chandross,R.J., Kretsinger,R.H., Allison,T.J., Pennington,C.J. and Rule,G.S. (1994) Crystal structure of human class mu glutathione transferase GST M2-2: effects of lattice packing on conformational heterogeneity. *J.Mol. Biol.* **238**, 815-832

Ramachandran,G.N. and Sosisekharan,V. (1968) Conformation of polypeptide and proteins. *Adv. Prot. Chem.* **23**, 283-437

Reinemer,P., Dirr,H.W., Ladenstein,R., Schaffer,J., Gallay,O. and Huber,R. (1991) The three-dimensional structure of class π glutathione S-transferase in complex with glutathione sulphonate at 2.3Å resolution. *EMBO J.* **10**, 1997-2005

Reinemer,P., Dirr,H.W., Ladenstein,R., Huber,R., Lo Bello,M., Federici,G. and Parker,M.W. (1992) Three-dimensional structure of class π glutathione S-transferase from human placenta in complex with S-hexylglutathione at 2.8Å resolution. *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 2.2Å resolution: structural characterisation of herbicide-conjugating plant glutathione S-transferase and a novel active site architecture. *J. Mol. Biol.* **255**, 289-309

Richards,F.M. (1977) Areas, volumes, packing and protein structures. *Annu.Rev.Biophys. Bioeng.* **6**, 151-176

Rose,G.D., Geselowitz,A.R., Lesser,G.J., Lee,R.H. and Zehfus,M.H. (1985) Hydrophobicity of amino acid residues in globular proteins. *Science* **229**, 834-838

Royer,C.A. (1995) Fluorescence spectroscopy in *Methods in Molecular Biology, Vol. 40: Protein stability and Folding: Theory and Practice* (Shirley,B.A. ed), pp 65-90, Humana Press Inc., Totowa, New Jersey

Ruddon,R.W., Sherman,S.A. and Bedows,E. (1996) Protein folding in the endoplasmic reticulum: lessons from the human chorionic gonadotropin beta subunit. *Prot. Sci.* **5**, 1443-1452

Ruddon,R.W. and Bedows,E. (1997) Assisted protein folding *J.Biol.Chem.* **272**, 3125-3128

Ruldolph,R., Fuchs,I. and Jaenicke,R. (1986) Reassociation of dimeric cytoplasmic malate dehydrogenase is determined by slow and very slow folding reactions. *Biochemistry* **25**, 1662-1669

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

Sanger,F., Nickeln,S. and Coulson,A.R. (1977) DNA sequencing with chain-termination inhibitors. *Proc. Natl. Acad. Sci. USA* **74**, 5463-5467

Sandberg, W.S. and Terwilliger, T.C. (1991) Energetics of repacking a protein interior. *Proc. Natl. Acad. Sci. USA.* **88**, 1706-1710

Sandermann,H.Jr. (1992) Plant metabolism of xenobiotics. *Trends Biochem. Sci.* **17**, 82-84

Sauder,J.M., MacKenzie,N.E. and Roder,H. (1996) Kinetic mechanism of folding and unfolding of *Rhodobacter capsulatus* cytochrome c_2. *Biochemistry* **35**, 16852-16862

Sayle,R. (1994) *Rasmol 2.5: Molecular Graphics Visualization Tool,* Biomolecular Structures Group, Glaxo Research, Middlesex

Scalley,M.L., Yi,Q., Gu,H., McCormack,A., Yates III,J.R. and Baker,D. (1997) Kinetics of folding of the IgG binding domain of peptostreptoccocal protein L. *Biochemistry* **36**, 3373-3382

Schellman,J.A. (1958) The factors affecting the stability of hydrogen-bonded polypeptide structures in solution. *J.Phys.Chem.* **62**, 1485-1494

Schellman,J.A. (1978) Solvent denaturation. *Biopolymers* **17**, 1305-1322

Schellman,J.A. (1987) The thermodynamic stability of proteins. *Ann. Rev. Biophys.Biophys. Chem.* **16**, 115-137

Schiffer,C.A. and Dötsch,V. (1996) The role of protein-protein interactions in protein unfolding. *Curr. Opin. Biotech.* **7**, 428-432

Schindler,T., Herrler,M., Marahiel,M.A. and Schmid,F.X. (1995) Extremely rapid protein folding in the absence of intermediates. *Nat. Struct. Biol.* **2**, 663-673

Schmid,F.X. (1992) Kinetics of unfolding and refolding of single-domain proteins in *Protein folding* (Creighton,T.ed) pp. 197-241. W.H. Freeman and Company

Seckler,R. and Jaenicke,R. (1992) Protein folding and protein refolding. *FASEB J.* **6**, 2545-2174

Serrano,L., Kellis,J.T.Jr., Cann,P., Matouschek,A. and Fersht,A.R. (1992) The folding of an enzyme. II. Substructure of barnase and the contribution of different interactions to protein stability. *J. Mol. Biol.* **224**, 783-804

Shakhnovich.E.I. and Finkelstein,A.V. (1989) Theory of co-operative transitions in protein molecules. I. Why denaturation of globular proteins is a first-order phase transition. *Biopolymers* **28**, 1667-1680

Shindyalov,I.N. and Bourne,P.E. (1996) The protein data bank through window. The San Diego Supercomputer Centre (SDSC)

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

Shortle, D., Stites, W.E. and Meeker, A.K. (1990) Contributions of the large hydrophobic amino acids to the stability of Staphylococcal nuclease. *Biochemistry* **29**, 8033-8041

Shortle, D. (1995) Staphylococcal nuclease: a showcase of m-value effects. *Adv. Protein Chem.* **46**, 217-246

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

Siddiqui,A.S. and Barton,G.J. (1995) Continuous and discontinuous domains: an algorithm for the automatic generation of reliable protein domain definitions. *Protein. Sci.* **4**, 872-884

Silow,M. and Oliveberg,M. (1997a) Transient aggregates in protein folding are easily mistaken for folding intermediates. *Proc. Natl. Acad. Sci. USA.* **94**, 6084-6086

Silow,M. and Oliveberg,M. (1997b) High-energy channelling in protein folding. *Biochemistry* **36**, 7633-7637

Sinning,I., Kelywegt,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 determination and refinement of human alpha class glutathione S-transferase A1-1, and a comparison with mu and pi class enzymes. *J.Mol.Biol.* **232**, 192-212

Sluis-Cremer,N., Naidoo,N., Kaplan,W., Manoharan,T.H., Fahl,W. and Dirr, H. (1996) Determination of a binding site for a non-substrate ligand in mammalian cytosolic glutathione S-transferases by means of fluorescence resonance energy transfer. *Eur. J. Biochem.* **241**, 484-488

Sosnick,T.R. and Trewhella,J. (1992) Denatured states of ribonuclease A have compact dimensions and residual secondary structure. *Biochemistry* **31**, 8329-8335

Steiner,R.F. and Kirby,E.P. (1969) The interaction of the Ground and Excited states of indole derivatives with Electron Scavengers. *J.Phys.Chem.* **73**, 12, 4130-4135

Stenberg,G., Board,P.G., Carlberg,I. and Mannervik,B. (1991a) Effects of directed mutagenesis on conserved arginine residues in a human class alpha glutathione transferase. *Biochem.J.* **274**, 549-555

Stenberg,G., Board,P.G. and Mannervik,B. (1991b) Mutation of an evolutionary conserved tyrosine residue in the active site of a human class alpha glutathione transferase. *FEBS Lett.* **293**, 153-155

Stenberg,G., Björnestedt,R. and Mannervik,B. (1992) Heterologous expression of recombinant human Glutathione Transferase A1-1 from a hepatoma cell line. *Prot. Exp.Purif.* **3**, 80-84

Tabor,S. and Richardson,C.C. (1987) DNA sequencing analysis with a modified bacteriophage T7 DNA polymerase. *Proc. Natl. Acad. Sci. USA* **84**, 4767-4771

Tan,Y-T., Oliveberg,M. and Fersht,A.R. (1996) Titration properties and thermodynamics of the transition state for folding: comparison of two-state and multi-state folding pathways. *J. Mol. Biol.* **264**, 377-389

Tanford,C. (1968) Protein denaturation, Part A and B. *Adv. Prot. Chem.* **23**, 121-275

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

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

Thornton,J.M., Jones,D.T., Mac Arthur,M.W., Orengo,C.M. and Swindells,M.B. (1995) Protein folds: towards understanding folding from inspection of native structures. *Phil. Trans.R.Soc. Lond. B.* **348**, 71-79

Tirado-Rives,J. and Jorgenson,W.L. (1993) Molecular dynamic simulations of the unfolding of apomyoglobin in water. *Biochemistry* **32**, 41275-4184

Tomarev,S., Zinovieva,R.D., Guo,K. and Piatigorsky,J. (1993) Squid glutathione S-transferase. Relationship with other glutathione S-transferases and S-crystallins of cephalopods. *J. Biol. Chem.* **268**, 4534-4542

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

Touchette,N.A., Perry,K.M. and Matthews,C.R. (1986) Folding of dihydrofolate reductase from *Escherichia coli*. *Biochemistry* **25**, 5445-5452

Tsuchida,S., Izumi,T., Shimizu,T., Ishikawa,T., Hatayama,I., Satoh,K. and Sato,K.(1987) Purification of a new acid glutathione S-transferase, GST-$Y_{n1}Y_{n1}$, with a high leukotriene-C4 synthase activity from rat brain. *Eur. J. Biochem.* **170**,159-164

Tsuchida,S. and Sato,K. (1992) Glutathione transferases and cancer. *CRC. Crit. Rev. Biochem. Mol. Biol.* **27**, 337-384

Utiyama,H. and Baldwin,B.L. (1986) Kinetic mechanism of protein folding. *Methods Enzymol.* **131**, 51-68

Uversky,V.N. and Ptitsyn,O.B. (1994) "Partly-folded" state, a new equilibrium state of protein molecules: four-state guandinium chloride-induced unfolding of β-lactamase at low temperature. *Biochemistry* **33**, 2782-2791

Vince,R., Daluge,S. and Wadd,W.B. (1971) Studies on the inhibition by glyoxalase by *S*-substituted glutathione. *J. Med. Chem.* **14**, 402-404

Vogt,G. and Argos,P. (1997) Protein thermal stability: hydrogen bonds or internal packing? *Folding Design* **2**, S40-S46

Waldburger,C.D., Jonsson,T. and Sauer,R.T. (1995) Are buried salt bridges important for protein stability and conformational stability. *Nat. Struct. Biol.* **2**, 122-128

Waldburger,C.D., Jonsson,T. and Sauer, R.T. (1996) Barriers to protein folding: Formation of buried polar interactions is a slow step in acquisition of structure. *Proc. Natl. Acad. Sci. USA.* **93**, 2629-2634

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

Wang,R.W., Newton,D.J., Pickett,C.B. and Lu,A.Y.H. (1991) Site directed mutagenesis of glutathione S-transferase YaYa: nonessential role of histidine in catalysis. *Arch. Bioch. Biophys.* **268**, 2, 574-578

Wang,R.W., Newton,D.J., Huskey,S.W., McKeever,B.M., Pickett,C.B. and Lu,A.Y.H. (1992a) Site directed mutagenesis of glutathione S-transferase YaYa: Important roles of tyrosine 9 and aspartic acid 101 in catalysis. *J. Biol. Chem.* **267**, 19866-19871

Wang,R.W., Newton,D.J., Pickett,C.B. and Lu,A.Y.H. (1992b) Site directed mutagenesis of glutathione S-transferase YaYa: functional studies of histidine, cysteine and tryptophan mutants. *Arch. Bioch. Biophys.* **297**, 1, 86-91

Wang,R.W., Newton,D.J., Johnson,A.R., Pickett,C.B. and Lu,A.Y.H. (1993a) Site directed mutagenesis of glutathione S-transferase YaYa: mapping the glutathione binding site. *J.Biol.Chem.* **268**, 23981-23985

Wang,R.W., Bird,A.W., Newton,D.J., Lu,A.Y.H. and Atkins,W.M. (1993b) Fluorescence characterisation of Trp21 in rat glutathione S-transferase 1-1: Microconformational changes induced by S-hexylglutathione. *Protein Sci.* **2**, 2085-2094

Warholm,M., Guthenberg,C. and Mannervik,B. (1983) Molecular and catalytic properties of glutathione transferase μ from human liver: An enzyme efficiently conjugating epoxides. *Biochemistry* **22**, 3610-3617

Weiner,M.P., Costa,G.L., Schoettlin,W., Cline,J., Mathur,E. and Bauer,J.C. (1994) Site -directed mutagenesis of double stranded DNA by the polymerase chain reaction. *Gene* **151**, 119-123

Wendt,H., Leder,L., Härmä,H., Jelesarov,I., Baici,A. and Bosshard,H.R. (1997) Very rapid, ionic strength-dependent association and folding of a heterodimeric leucine zipper. *Biochemistry* **36**, 204-213

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

Widersten,M., Björnestedt,R. and Mannervik,B. (1994) Contribution of amino acid residue 208 in the hydrophobic binding site to the catalytic mechanism of human glutathione transferase A1-1. *Biochemistry* **33**, 11717-11723

Widersten,M., Björnestedt,R. and Mannervik,B. (1996) Involvement of the carboxyl groups of glutathione in the catalytic mechanism of human glutathione transferase A1-1. *Biochemistry* **35**, 7731-7742

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

Wilce,M.C.J., Board,P., Feil,S.C. and Parker,M.W. (1995) Crystal structure of a theta-class glutathione transferase. *EMBO J.* **14**, 2133-2143

Wolfenden,R. and Radzicka,A. (1994) On the probability of finding a water molecule in a non-polar cavity. *Science* **265**, 936-937

Wolynes,P.G., Onuchic,J.N. and Thirumalai,D. (1995) Navigating the folding routes. *Science* **267**, 1619-1620

Yu,Y., Monera,O.D., Hodges,R.S. and Privalov,P.L. (1996) Ion-pairs significantly stabilize the coiled-coil in the absence of electrolyte. *J. Mol. Biol.* **255**, 367-372

Zetina, C.R. and Goldberg, M.E. (1980) Reversible unfolding of the b2 subunit of *Escherichia coli* tryptophan synthetase and its proteolytic fragments. *J. Mol. Biol.* **137**, 401-414

Zhu,Z-Y. and Karlin,S. (1996) Clusters of charged residues in protein three-dimensional structures. *Proc. Natl. Acad. Sci. USA.* **93**, 8350-8355

Zimm,B.H. and Bragg,J.K. (1959) Theory of the phase transition between helix and random coil in polypeptide chain. *J.Chem.Phys.* **31**, 526-535

Zitzewitz,J.A., Bilsel,O., Luo, J., Jones,F E. and Matthews,C.R. (1995) Probing the folding mechanism of a leucine zipper peptide by stopped-flow circular dichroism. *Biochemistry* **34**, 12812-12819