UNIVERSITY OF THE WITWATERSRAND
JOHANNESBURG

Implications of N-capping motifs for folding and design of human glutathione transferase A1-1

Tessa Little

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

Johannesburg, October 2005

Declaration

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

Signed this __ day of _____ 20__

_____

Tessa Little.

Abstract

It is well documented that N-capping motifs are stabilising local motifs for α-helices. N-capping motifs have been identified within hGST A1-1 at the N-terminal ends of α-helix 9 and α-helix 6. The conservational role of these two motifs in protein stability, folding and function was investigated. α-Helix 9 is a unique structural feature to class Alpha GSTs that is important for its catalytic functioning. This amphipathic helix is highly dynamic, where upon ligand binding at the active-site, the delocalised C-terminal region becomes immobilised to form a structured helix forming a "lid" over the active-site. The specific role of the Asp N-cap motif toward the stability and dynamics of α-helix 9 was determined by substituting the Asp-209 for a Gly. ANS binding and urea-induced activity studies showed that by removing the N-cap motif of α-helix 9 in hGST A1-1, the α-helix 9 is destabilised rendering a less hydrophobic binding site compared to that in the wild-type. The helical content of the peptide, corresponding to α-helix 9 in the C-terminal region of hGST A1-1 (208 -222), decreased significantly upon the removal of the N-cap motif. The explanation for the conservation of the Asp N-cap residue can be found in its stabilising role of the C-terminal region of class Alpha GSTs. This stabilising role was however less apparent in context of the protein compared to that in the peptide. Majority of the atomic contacts owing to the stability of α-helix 9 appear to be governed by non-local tertiary interactions rather than local interactions, such as the N-cap motif. These tertiary interactions are likely to include short and long range contacts between residues on the surface of the protein that are already known to contribute towards the stability of the C-terminal region. In this study, the ligand displacement-studies and the molecular docking results strongly suggest that 8-aniline-1-napthalene

sulfonate binds at the H-site in hGST A1-1.

The N-capping motif of α-helix 6 identified in class Alpha GSTs is located within the core of domain 2. This motif is a common feature found amongst almost all GST-like proteins and is thought to be the folding nucleation site (Stenberg *et al.* J. Biol. Chem. 275 (2000), 10421-10428). The N-cap (Ser-154) and N3 (Asp-157) residues were each substituted with an Ala in hGST A1-1 to investigate the role of this motif in the folding of hGST A1-1. Both substitutions resulted in thermal sensitive mutants compared to that of the wild-type. The N3 substitution (D157A) was however too disruptive, where the yields of this mutant were insufficient for any further studies to be carried out. For the N-cap mutant (S154A), the unfolding kinetic studies revealed a significantly destabilised core in domain 2 compared to that of the wild-type. The kinetic folding studies monitored by fluorescence spectroscopy, revealed that the N-cap motif contributes to the efficient folding and dimerisation of the subunits, and to a far lesser extent towards the final tight packing and reorganisation of tertiary interactions in hGST A1-1. Since no changes in the burst-phase of S154A was evident compared to that of the wild-type, it seems unlikely that this motif is a folding nucleation site in hGST A1-1. These results do not exclude the possibility that this motif contributes to the rapid formation secondary structure during the burst-phase of folding. Due to the highly conserved region surrounding α-helix 6 , the role of this motif contributing to the stability of hGST A1-1 could be a general feature for GSTs and GST-like proteins. In this study, further insight into the mechanism of folding for hGST A1-1 was gained. The hydrophobic core packing surrounding α-helix 6 occurred as a late folding event, that is during the final packing and reorganisation of tertiary interactions of the protein. The N-cap motif is an important structural feature for the fast folding of domain 2. This N-cap motif is a unique structural feature important for the efficient folding of the monomers, which is exclusive to its role in stabilising α-helix 6 in hGST A1-1.

Acknowledgements

I would like to thank my supervisor, Prof. Heini Dirr, for the opportunities and challenges he provided. I also wish to thank my colleagues and friends, from the Protein Structure-Function Research Programme and the School of Molecular Cell Biology, for their support and helpful discussions over the years, where insight and knowledge was nurtured.

I would like to extend my appreciation to Dr Louise Wallace for her continued guidance during the course of my work in using the stopped-flow instrument and the data analyses involved.

I would like to thank Dr Yasien Sayed for extending his knowledge and understanding of isothermal titration calorimetry, which proved valuable.

I would like to thank Fourie Joubert from the Bioinformatics and Computational Biology Unit in the Department of Biochemistry, University of Pretoria, for the use of Cirius[2], and to Tjaart de Beer for his time in assisting me with the molecular docking.

I acknowledge the University of the Witwatersrand and the National Research Foundation of South Africa for their financial support.

I dedicate this work to the Universe!! and of course...ALL those in my heart who I am so fortunate to be sharing with in this delicious experience called Life!

Contents

List of Figures

List of Tables

Abbreviations

AGADIR:	helix/coil transition algorithm
ANS:	8-anilino-1-napthalene sulfonate
CD	circular dichroism
CDNB	1-chloro-2,4-dinitrobenzene
ΔC_p	change in heat capacity
D209G	replacement of Asp at position 209 with a Gly in hGST A1-1
EDTA	ethylenediaminetetraacetic acid
GSH	glutathione
G-site	glutathione binding site
GST	glutathione transferase
Grx2	glutaredoxin-2
$\Delta G(H_2O)$	the change in Gibbs free energy of unfolding in the absence of denaturant
ΔH	change in enthalpy
hGST A1-1	human class Alpha glutathione transferase with two type one subunits
hGST P1-1	human class Pi glutathione transferase with two type one subunits
H-site	hydrophobic binding site
IPTG	isopropyl-thio-galactoside
ITC	isothermal titration calorimetry
$k_{\mathrm{u}}(H_20)$	apparent unfolding(u)/refolding(f) rate in the absence of denaturant
k_{u}	apparent unfolding(u)/refolding(f) rate constant
K_{d}	dissociation constant
K_{m}	Michaelis constant
$k_{\mathrm{cat}}/K_{\mathrm{m}}$	catalytic efficiency
m-value	co-operativity of unfolding transition
N_2	native folded dimer
N_2*	dimeric native-like intermediate

ns	nanoseconds
NMR	nuclear magnetic resonance
μs	microseconds
PDB	protein data bank
PCR	polymerase chain reaction
ΔS	change in entropy
S154A	replacement of Ser at position 154 with an Ala in hGST A1-1
D157A	replacement of Asp at position 157 with an Ala in hGST A1-1
SDS-PAGE	sodium dodecylsulfate polyacrylamide gel electrophoresis
TFE	2,2,2-trifluoroethanol
U	unfolded monomer
Ure2	yeast prion protein from *Saccharomyces cerevisiae*
UV	ultraviolet

The IUPAC-IUBMB one and three letter codes for amino acids are used. The enzyme: glutathione transferase (EC 2.5.1.18)
The nomencelature N", N', N-cap, N1, N2 and N3 is used to assign amino acids at the N-terminal end of α-helices according to Richardson and Richardson (1988), where N1 is the first helical residue.

Chapter 1

Introduction

1.1 Linear sequence to three-dimensional shapes of proteins

Over the past four decades much research has been focused on deciphering the mechanism by which proteins acquire their precise folded three-dimensional structures. To solve this complex problem, two central questions have emerged; what are the forces that determine and hence stabilise the native protein structure and, how does a polypeptide fold to its native state? A thorough understanding of the principles that govern folding of proteins *in vitro* is a prerequisite for elucidating protein folding in the cell, which has far-reaching biological and medical implications. It is known that the unique native conformation of the protein determines its function. By extension, the function of a protein depends entirely on the ability of the protein to fold rapidly and reliably to its native structure.

The structure of a protein can be described in four different hierarchical levels: primary, secondary, tertiary and quaternary. The secondary structure refers to the regular, re-occurring spatial arrangements of adjacent amino acid residues

of a polypeptide chain (primary structure), such as: α-helices, β-sheets, β-turns, β-barrel etc. The tertiary structure, that is the native protein, is defined by the three-dimensional arrangement of secondary or super-secondary structures. The quaternary structure defines the three-dimensional arrangement of prosthetic groups and solvent with respect to the native protein. It is the spatial arrangement of the atoms in the polypeptide chain, defined by the dihedral bond angles of the peptide bonds, that governs the precise and reproducible patterning of hydrophobic and hydrophilic components, positive and negative charges, hydrogen-bond donors and acceptors, that ultimately defines the unique structure. The atomic interactions within proteins can be local, between residues close in sequence, or non-local, being medium-range and long-range interactions.

To date the prediction of any particular secondary structure within an amino acid sequence of a protein has an accuracy close to 80% (de Brevern *et al.*, 2002). The field has yet to produce consistent and reliable *ab initio* structure prediction protocols (Bonneau and Baker, 2001). As the understanding in the formation of the unique three-dimensional structure of proteins is increasing, so new goals in protein design are also emerging. One such area of development in protein structure design, has been the design of specific amino acid sequences that are not necessarily similar to known proteins, nonetheless will fold into a compact unique three-dimensional structure when synthesised *in vitro* or *in vivo*. This approach will provide a means for the major advancement in drug design. To date α-helical, β-sheet, and simple α/β proteins have been synthesised. Site-directed mutagenesis is also a common approach to increase the understandings of the forces that contribute to the conformational stability of globular proteins in atomic detail (Matthews, 1993). A greater understanding of these forces, and key features in relation to protein folding, is providing for the ability to design new proteins (Betz *et al.*, 1995).

1.2 α-Helices and the N-capping motif

By identifying common key features that define a homologous fold or simple secondary structural elements, such as in α-helices and β-sheets, is one area under investigation attempting to ascribe most common features that can define secondary structural elements/or conserved folds.

1.2.1 α-Helices and their putative motifs

Several scales have been derived in defining the intrinsic helical tendencies for the different amino acids determined from thermodynamic analysis of proteins (Horovitz *et al.*, 1992; Blaber *et al.*, 1993), or peptides (Lyu *et al.*, 1990; O'Niel and DeGrado, 1990; Horovitz *et al.*, 1992; Chakrabartty *et al.*, 1994; Munoz and Serrano, 1995b,a; Rohl *et al.*, 1996). Although there are some discriminating differences between one scale to another, in general, there is a reasonable correlation. The best correlation of these scales is best described in the algorithm, AGADIRs-2 (Lacroix *et al.*, 1998). AGADIR has been derived from analysing multiple peptides with different amino acid compositions. The different helical propensities of the 20 amino acids in a central helical position have been explained by taking into account the changes in solvation, enthalpy, and side-chain entropy between the denatured and the helical states (Luque *et al.*, 1996). This program is useful in predicting the α-helix propensities of a given amino acid sequence. Another point to take into account with regards to the amino acid propensity, is that every amino acid appears to have different helical propensities in a position-dependent manner.

Whilst understanding that there are different sequence-dependent factors that contribute to the formation and stability of α-helices, it was found that through statistical analysis of the frequency of occurrence of amino acids in different secondary structures, certain side chains have distinct conformational preferences (Chou and Fasman, 1978). These preferences were termed amino acid propensities, which lead to the discovery of local putative motifs that contribute largely to secondary structure stability. One such statistical analysis

has shown that there are specific amino acid preferences at or near the N-cap and C-cap positions (Richardson and Richardson, 1988; Presta and Rose, 1988), see Figure 1.1. Three motifs at the N-terminus have been described: the N-capping box, the hydrophobic staple motif (Munoz and Serrano, 1995a; Munoz *et al.*, 1995) and Pro-box motif (Viguera and Serrano, 1999). The nomenclature of the N-capping region is denoted N'-N'-N-N1-N2-N3, where N or N-cap is the residue with non-helical ϕ and ψ angles, immediately preceding the first amino acid (N1) of the α-helix (Richardson and Richardson, 1988). At the C-terminus, the Schellman motif or C-capping box (Viguera and Serrano, 1995) and a C-terminal proline box (Kim and Kang, 1999) have been identified. The nomenclature of the C-capping region is denoted C3-C2-C1-C-cap-C'-C". The C-cap is the residue immediately proceeding the last residue (C") of the α-helix. The rational behind amino acids having different propensities in a position-dependent manner is explained, in part, where the non-helical N-cap side-chain can fulfil a hydrogen bond with the amino group of the first three helical residues. These residues do not have main-chain hydrogen bond partners as apposed to those within the α-helix. Asp, Ser, Thr and Asn are all good N-cap residues where the side chain groups form hydrogen bonds with one or two main-chain amino groups of residues at position N1, N2, N3 and/or N4. The N' and N-cap residues have shown to adopt dihedral angles in the β-region of the Ramachandran plot (Munoz and Serrano, 1995b) and not α-helical conformation, and thus is part of the loop preceding the α-helix. The proline box at the C-terminal end of helices is defined by a proline at C, creating the last turn of the alpha-helix to be slightly unwound (3_10 helix) (Kim and Kang, 1999). The proline box at the N-terminal end of helices has a Pro residing at the N-cap position, and is often accompanied with an Ile or Leu at position N', Val at position N3 and a hydrophobic residue at position N4 (Viguera and Serrano, 1995). These N-capping motifs contribute significantly to helix stability, whereas the C-capping motifs contribute significantly less (Viguera and Serrano, 1995). This emphasises the importance of the N-capping region as a stabilising feature for helices, which in turn contributes to the overall helical propensity of a given peptide sequence.

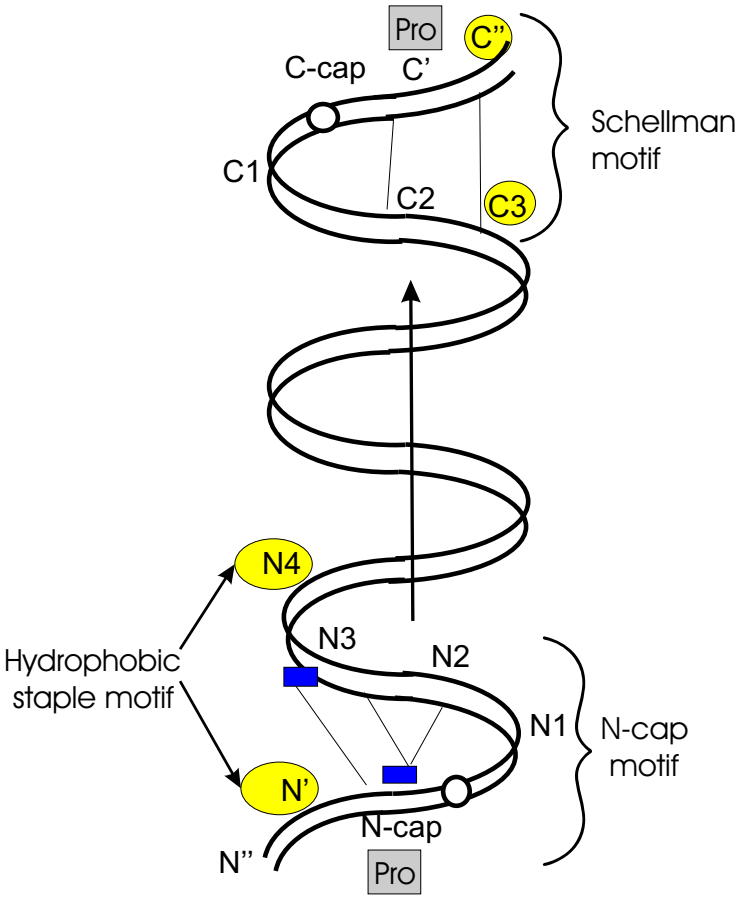

Figure 1.1: Schematic diagram showing putative N- and C-terminal capping motifs of α-helices. The nomenclature for the capping residues are shown. N1 and C1 are the first and last helical residues, respectively. The lines represent hydrogen bonds, hydrophobic residues are highlighted in yellow. An example of an N-capping motif is shown where hydrogen bonds between the side-chain (blue boxes) of N3 and N-cap residue form with the main-chain NH or CO group of residues within the first helical turn. The hydrophobic staple motif and C-terminal Schellman motif are defined by van der Waals contacts formed between hydrophobic residues. The position of the proline residue (grey) forming a Pro-box motif, either at the C- or N-terminal end of helices, is shown. The δ - and δ + symbols indicate the partial negative and positive charge of the helix dipole, respectively. The arrow indicates the direction of the N-Cα peptide bonds.

1.2.2 Stabilising role of N-capping motifs toward α-helices

Whilst understanding why small segments of secondary structures are relatively more stable than the random coil state, it was found that the choice of the amino acids within the capping motif have a major effect on the helical content of peptides (Chakrabartty *et al.*, 1994; Cochran and Doig, 2001; Cochran *et al.*, 2001; Forood *et al.*, 1993; Lyu *et al.*, 1993; Zhou *et al.*, 1994; O'Niel and DeGrado, 1990; Petukhov *et al.*, 1999) and towards the global stability of proteins (elMasry and Fersht, 1994; Chen and Fersht, 1994; Serrano *et al.*, 1992; Bell *et al.*, 1992; Wilcock *et al.*, 1998; Zhukovsky *et al.*, 1994; Parker and Hefford, 1998), and see review by Scholtz and Baldwin (1992). The free energy preference for all 20 amino acids from the N-cap to the N3 positions have recently been shown to correlate well to the statistical preferences of amino acids at their corresponding positions of α-helices in proteins (Cochran and Doig, 2001; Iqbalsyah and Doig, 2004; Cochran *et al.*, 2001; Rohl *et al.*, 1996).

The stabilising effect of an N-capping box is well established, which consists of a hydrogen bond between the N-cap side-chain and the N3 main-chain amide group, accompanied by a reciprocating hydrogen bond between the side-chain of the N3 residues and the main-chain carbonyl group of the N-cap residue (Richardson and Richardson, 1988; Presta and Rose, 1988). Both the N-capping box and staple motif (hydrophobic residues flanking the N-cap motif) together have an additive effect on α-helix stability (Munoz and Serrano, 1995a) and also contribute to the helical rigidity. In addition to the importance of the hydrogen bonding patterns at the capping box, other factors also play a role in the stabilisation of helical conformations of short peptides. Such factors include interactions among side chains (Jimenez *et al.*, 1994; Padmanabhan and Baldwin, 1994), electrostatic interactions among polar and charged groups at the helical ends with the helix macrodipole (Lockhart and Kim, 1992, 1993). The helix-dipole effects are likely to be present at all sites, however this can be overwhelmed by strong hydrogen bonds, such as the N-cap motif (Cochran *et al.*, 2001).

Even though the N-capping box is reported to be a recurrent hydrogen bond

motif in proteins (Presta and Rose, 1988; Harper and Rose, 1993), it occurs only in 5% of α-helices (Wan and Milner-White, 1999). Recently it has been shown that a third of all α-helices have either a ST-motif (Ser or Thr at the N-cap position) or Asx-motif (either Asp or Asn at the N-cap) (Wan and Milner-White, 1999). These motifs are similar to the N-capping box, in that the side-chain of the N-cap residue hydrogen bonds with the main-chain amino group of N2 or N3, however they lack the reciprocal N3 or N4 side-chain hydrogen bond with the main-chain N-cap CO group. In fact, it is proposed that the N-capping box could be regarded as a subset of the ST- or Asx-motif, where an Asx-motif found at the N-terminal end of an α-helix is termed an Asx N-cap motif. Moreover, it is now known that motifs such as the Asx and ST-motif, not only contribute to the stability of α-helices, but also to β-bulge and β-hairpins. Interestingly, the ST-motif is more prolific amongst proteins than the Asx-motif (Wan and Milner-White, 1999). Work by Wan and Milner-White (1999) show that for the N-capping box, is it more common to find a Ser or Thr at the N-cap position than an Asp (Wan and Milner-White, 1999). In this study, the hydrogen bonding pattern is defined according to Wan and Milner-White (1999).

A reasonable estimation of the enthalpy change for the helix-coil transition in water is $\sim$5.4 kJ.mol^{-1} per residue (Scholtz *et al.*, 1991). This supports the explanation that the α-helix backbone conformation is stable in water. It is well know that hydrogen bonds stabilise globular proteins, and in the case of helices, the enthalpy contribution for each intramolecular hydrogen bond is approximately between 4.2 and 9.2 kJ.mol^{-1} (Myers *et al.*, 1995). In saying this the hydrogen bond may have differing contributions in a peptide compared to that of a full length protein, since peptides lack tertiary interactions unlike proteins, and therefore this can complicate the interpretation of the results. An example of this is the work done by Peterson *et al.* (1999) where they attempted to introduce an N-capping box in a small phosphocarrier protein. Instead they introduced a new tertiary hydrogen bond between an α-helix and β-sheet, which resulted in more stable protein.

A comparative study of the N-capping residues in proteins and helices, is a good approach to further ones understanding of the local (short range) and

non-local (long range) contacts governing the stability and folding of proteins. Helices in peptides tend to fray at their ends, whereas those in proteins are constrained by the rest of the structure and they are thus better defined structurally. A comparative study of a dimeric helical peptide, where its ends were designed not to fray (O'Niel and DeGrado, 1990), suggest that like others (Horovitz *et al.*, 1992), the α-helix forming tendencies determined from studies on model peptides may not, to a large extent, be relevant to proteins. Nonetheless, promising statistical mechanic approaches to the prediction of helical propensities of short peptides were reported (AGADIR) to have excellent accuracy in the prediction of the helical content of peptides, which include a very wide range of lengths and sequences (Doig *et al.*, 1997; Lacroix *et al.*, 1998).

1.3 Implications of stabilised secondary structure in protein folding

1.3.1 Mechanism of protein folding

The protein folding pathway is commonly modelled as an energy landscape that resembles a funnel starting with a diverse multitude of unfolded conformations at the rim, going "down hill" with the native-state at the bottom consisting of a tight ensemble (Chan and Dill, 1998; Dill and Chan, 1997). Along the interior of the funnel represents an infinite number of possible pathways, where each pathway has its own distinct energetic topography. Figure 1.2 shows three theoretical folding funnels illustrating the differences of the interior of a folding funnel as a result of different folding mechanisms of proteins. The width of the funnel represents conformational entropy, with maximum entropy in the random coil state and minimum entropy in the final state, being the native protein. The height of the funnel represents the free energy change. The simplest kinetic behaviour for protein folding is the two-state kinetics *via* a single exponential process, where the interior of the funnel tends to be smooth, Figure 1.2 (Dill and Chan, 1997).

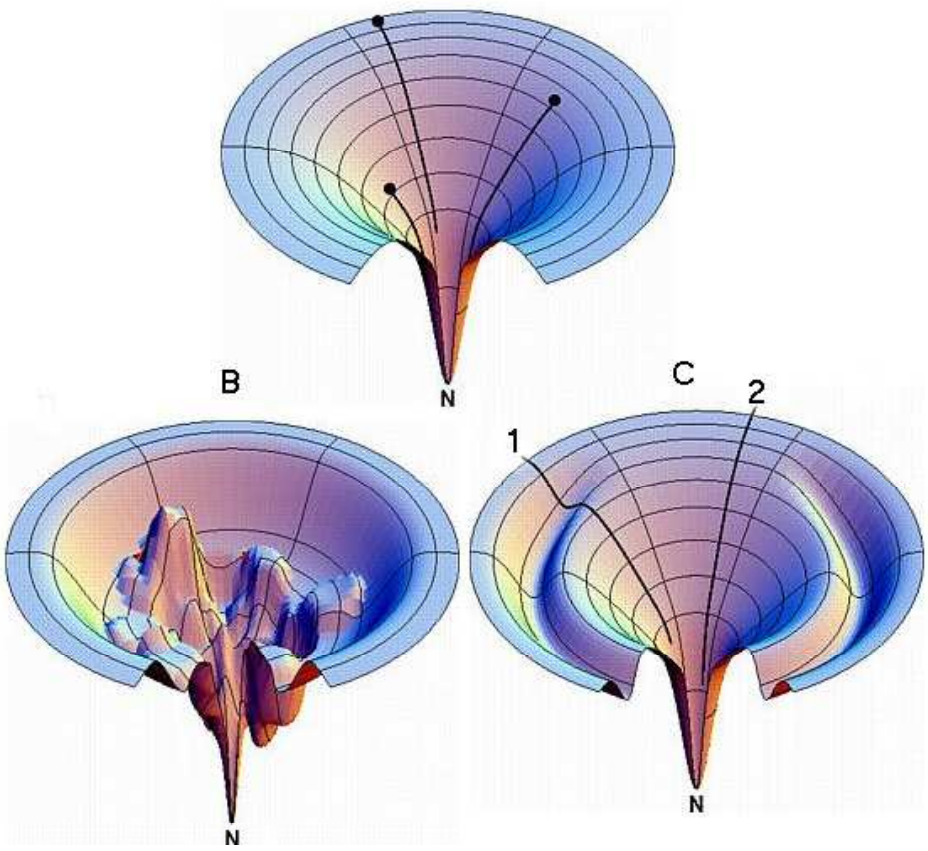

Figure 1.2: Different folding scenarios shown as the energy surface of a folding funnel. The vertical axis is the internal free energy. Each conformation is represented as a point on the landscape. The two horizontal axes represent the many chain degrees of freedom. Scenario A shows an idealised funnel landscape of two-state (un)folding reaction. B shows a rugged energy landscape with kinetic traps, energy barriers, and some narrow throughway paths to native (N) protein. This folding can be multi-state. C is a "moat" landscape to illustrate how a protein could have a fast-folding throughway process (2), in parallel with a slow-folding process (1) involving a kinetic trap. Figures are reproduced with permission from Dill and Chan (1997).

The protein folds directly from the denatured state to the native folded state without any detectible intermediates. The transition state of a folding reaction is between the native and denatured state. In this scheme the transition state is not considered as a single obligatory step through which all molecules pass, but rather it is an ensemble of many different chain conformations (Onuchic *et al.*, 1996). Many small proteins have been observed to fold *via* two-state pathway such as chymotrypsin inhibitor 2, SH3 domain of α-spectrin (Viguera *et al.*, 1994), IgG binding domain of streptococcal protein G (Alexander *et al.*, 1992) and cold-shock protein B, see review by Jackson (1998). However, other small proteins, and in fact most proteins, fold *via* more complex multistate kinetics with transient intermediates, such as apomyoglobin (Kuwajima, 1977), barnase (Matouschek *et al.*, 1990; Wong *et al.*, 2000), CheY (Lopez-Hernandez *et al.*, 1997) and cytochrome c (Sosnick *et al.*, 1997; Sauder and Roder, 1998), and for larger proteins dihydrofolate reductase (Jones *et al.*, 1994; Jennings *et al.*, 1993), heterodimer luciferase (Clark *et al.*, 1997; Fedorov and Baldwin, 1997) and dimeric Trp repressor (Gloss and Matthews, 1998; Gloss *et al.*, 2001), to name a few. Each step in the reaction will have exponential kinetics describing the time-dependent interconversion between the stable states adding to the roughness of the interior of the funnel (Onuchic *et al.*, 1995; Wolynes *et al.*, 1996). Here a protein molecule may follow the steepest (fastest folding) path, or a more round-about route passing through several local minima (intermediates) and maxima (transition states). The transition state can be described within the interior of the folding funnel as the "spurs" or "hills" and the intermediates as being the "valleys". Thus, there are potentially a plethora of routes to the native state. A common characteristic feature of the transition state, observed from both mutational and computational studies, is that it tends to be closer to the native structure than unfolded. It appears that some native contacts are more probable than others due to the polymeric nature of the peptide chain and the topology of the native state (Daggett and Fersht, 2003; Uversky and Fink, 2002). The main discrepancy between different folding theories in the "new view" is in relating the conformational breadth of the transition state ensemble of proteins that fold *via* two-state mechanism with existing proteins that fold *via* a multistate mechanism (Dill and Chan, 1997; Onuchic *et al.*, 1996; Abkevich *et al.*, 1994). The folding of a protein following

a two-state folding mechanism occurs without detectable and obligatory intermediates, whereas a multistate folding mechanism involves intermediates. As techniques for detecting structure during folding have improved with respect to both spacial and temporal resolution, as well as monitoring folding on faster time scales of nanoseconds to microseconds, so a unifying view is emerging providing answers to these discrepancies (Fersht, 1997). In the unified view, the delicate balance of both secondary and tertiary structural elements along the folding pathway of proteins is highlighted, which appears to be an inherent predisposition in determining the mechanism of folding.

1.3.2 Implications of local and non-local interactions in protein folding

To date, the current studies that are involved in characterising the structural elements in the folding reactions have highlighted the importance of both secondary and tertiary structural elements. From this two models have emerged describing the mechanism of folding. This being the nucleation-condensation (Fersht, 1997, 2000) and diffusion-collision (Baldwin, 1999) model. The former model is based on the concept that a rapid collapse in the polypeptide chain occurs upon which further stable secondary structure exist either concurrently or after the condensation of a collapsed chain (Fersht, 1997, 2000) forming an extended nucleus (Itzhaki *et al.*, 1995a). The diffusion-collision model is based on the early formation of native-like secondary structure first, followed by the diffusional docking of these structured elements, and hence further organisation of tertiary structure (Baldwin and Rose, 1999). The diffusion collision model predicts that the rate of folding is dependent on the probability of a successful collision between two macrodomains or secondary structural units. A third model is the hydrophobic collapse. Here the expulsion of water molecules through hydrophobic interactions is the driving force for the collapsing of the polypeptide chain (Baldwin, 1989; Kauzmann, 1959). In cases where the native-like hydrophobic interactions that stabilise, otherwise, weak secondary structural elements, during the hydrophobic collapse, in fact fits more to the nucleation-condensation model (Itzhaki *et al.*, 1995a; Fersht,

1995, 1997).

It appears as if the nucleation-condensation and diffusion-collision models may in fact fit the two extremes of what appears to be manifestations of the same mechanism of folding, where in the case of inherently more stable and better formed secondary structural elements, the diffusion model best describes the folding mechanism rather than the nucleation-condensation model (Bilsel and Matthews, 2000; Daggett and Fersht, 2003; Fersht, 2000). The persistence of the secondary structure elements in relation to stabilising native-like and non-native tertiary interactions along the folding pathway still remains to be seen. It is by investigating the early folding events of the transition state and intermediate conformations at atomic resolution, that one can probe for such interrelationships of secondary and tertiary interactions and their relative importance in defining the rate and mechanism of folding. As this work focuses on investigating the likely existence of a folding nucleus defined by an N-capping motif, this discussion will focus on describing the two folding models in light of other experimental results in order to illustrate the current views in describing the folding mechanism of proteins.

How important local interactions are in driving the formation of secondary structure elements, is still an open question. An interesting set of results illustrated by Bilsel and Matthews (2000) show that for cold shock proteins (a β-sheet protein) the native-like transition state is attained *via* a higher-order mechanism (Perl *et al.*, 1998; Schindler and Schmid, 1996), i.e. involving nonlocal interactions predicted by the nucleation-condensation model. On the contrary an all-helical monomer λ repressor is described to fold according to the diffusion-collision model where the preformed helical structures dock/associate (Burton *et al.*, 1998). These results also suggest that the stabilisation or destabilisation of these local structural elements can shift the folding transition state (Burton *et al.*, 1996). It is apparent that there is a bias toward stability of local structural units, where in the former case, it is improbable that the β-strand or complete β-sheet formation occurs in isolation, whereas this is highly probable in the latter case for the all-helical protein. Work by Viguera *et al.* (1996b) demonstrated that by increasing the helical propensities, the rate of folding increased. In some instances the stabilisation of secondary structure elements

decreases the folding rate, especially when the intermediate is a kinetic trap (Lopez-Hernandez *et al.*, 1997). It has been proposed that the stabilisation of intermediate structure can either enhance or retard folding depending on its nature and content of native-like interactions (Chiti *et al.*, 1999). It can be suggested that the nature of the transition state and hence the mechanism of folding is defined by the propensity of secondary structure formation (Viguera *et al.*, 1996a; Lopez-Hernandez *et al.*, 1997). However, work by Munoz *et al.* (1996) have shown that even though secondary structure is important, it does appear more likely that tertiary interactions play a larger role in defining the native structure. Contrary to this, work by Daggett and Fersht (2003) have shown that during the early phases of folding of chymotrypsin inhibitor 2, the formation of secondary and tertiary interactions occur concurrently supporting the notion of the equivocal role of both local and non-local contacts for the rapid folding of a protein (Daggett and Fersht, 2003). It appears that these factors are in some way contained inherently within the amino acid sequence of a protein. It would seem likely that persistent secondary structural elements are likely to occur during the earlier phases in folding forming stable structural units independent of tertiary interactions, such as α-helices and β-hairpins. However, in general, majority of peptide fragments that do tend to populate as native-like conformations, do not occur to a large extent, but more often occur as random coil conformations (Munoz and Serrano, 1996; Gay *et al.*, 1995).

1.4 Proposed role of stabilised α-helices in folding nucleation

1.4.1 Evidence of folding nucleation sites

The selective pressures arising within protein folding that is under some evolutionary control would be evident as conserved residues or key structural motifs. Assuming a folding nucleus promotes fast folding, it is hypothesised that different (even non-homologous) sequences that fold into the same structure may

have similar folding nuclei and can be identified by conserved residues (Mirny *et al.*, 1998; Ptitsyn, 1998). A folding nucleus for the protein CheY has been identified (Lopez-Hernandez and Serrano, 1996), where the key residues are conserved within the superfold family to which CheY has a common topological fold (Mirny and Shakhnovich, 2001; Tseng and Liang, 2004). This conservation of a folding nucleation site implies a mechanism whereby certain amino acids (folding nucleus) form majority of their contacts in the transition state (Mirny *et al.*, 1998) and are key determinants for the fast folding of proteins (Mirny and Shakhnovich, 1999; Li *et al.*, 2000; Mirny and Shakhnovich, 2001). Even though it is suggested that some residues are conserved for kinetic reasons, the limited kinetic data available prevent any verification of this relationship between sequence conservation across superfamilies and their participation in the folding nucleus for many proteins.

The folding of small proteins in theoretical simulations *via* the two-state mechanism are shown to have a single folding nucleus, whereas for larger multidomain proteins, multiple nucleation sites are suggested to exist (Dinner and Karplus, 1998). Kinetic refolding studies for some proteins, such as barnase (Chu and Bai, 2002), GCN4 (Moran *et al.*, 1999) and equine lysozyme (Morozova-Roche *et al.*, 1999) show that no strictly definable region exists within the polypeptide chain which can been identified as an initiating point in folding. Contrary to these results, a single nucleation site has been identified for CheY (Lopez-Hernandez and Serrano, 1996), chymotrypsin inhibitor 2 (Itzhaki *et al.*, 1995b; Ptitsyn and Ting, 1999), and immunoglobulin domain protein (Fowler and Clarke, 2001) to name a few. It has been demonstrated that a defined folding nucleus is more often not evolutionary conserved (Plaxco *et al.*, 1999; Larson *et al.*, 2002; Tseng and Liang, 2004). These contradicting results may indicate that the process involved in the initiation of folding appears to be multifactorial, and is inherently defined in the sequence (Morozova-Roche *et al.*, 1999). One set of results that appear to suggest this, is a computational study of circular permutations in a lattice protein. It was demonstrated that by disrupting the original folding nucleus of a wild-type protein by incision, a new folding nucleus existed accompanied by a novel, but slower, folding mechanism to the original folding pathway (Li and Shakhnovich,

2001).

1.4.2 α-Helices involved in early folding events

The "speed limit" to protein folding is determined by how fast a protein can collapse (Hagen *et al.*, 1996). Studies on the folding of peptides using both computational simulations and from experimental work has thus far shown that the essential elements in the initial folding include the hydrophobic collapse, formation of hydrogen bonds and side-chain packing. Experimental results have shown that the folding rates of λ repressor (Burton *et al.*, 1996), Arc repressor (Robinson and Sauer, 1996) and cytochrome c (Chan *et al.*, 1997; Sosnick *et al.*, 1997) are under some rate-limiting step during the initial collapse of the polypeptide chain. If this collapse is rate limiting, the next question is whether this is attributed to interactions that are specific and deliberate, or non-specific?

Answering this question requires knowing the intrinsic time scales and driving forces for all these processes. There are currently two views to date. In the first view the chain collapses to form an unstructured globule, from which native secondary structure and topology will form (Sadqi *et al.*, 2003). The second view is that the secondary structure form concurrently to chain collapse. Whether folding is initiated by the formation of secondary structure or hydrophobic collapse is still open to debate (Baldwin, 2002). Some α-helices form in $\sim$600 nanoseconds measured by laser-induced temperature jump experiments (Werner *et al.*, 2002; Williams *et al.*, 1996), whereas β-hairpin formation occurs in microseconds (Munoz *et al.*, 1998, 1997) and loops fold in 1 milliseconds (Eaton *et al.*, 1997). However, current research has shown that the hydrophobic collapse and intermolecular contacts occur at much faster time scales than the formation of stable β-hairpins, and slightly faster than the formation of helices Sadqi *et al.* (2003). It is possible that very stable helices may compete with chain collapse. The extent of interactions that occur during the collapse will most likely depend predominantly on the protein's sequence, and the solvent environment. This is in accordance with the unified view in

protein folding explaining the delicate balance of both local and non-local interactions, which appears to be an inherent predisposition that determines the precise mechanism of folding (Daggett and Fersht, 2003).

1.4.3 Proposed role of N-capping motifs in α-helix nucleation

The level of organisation needed to initiate favourable energetic contacts, requires the protein to first pay an entropic price before the downhill tendency of the energy landscape can be manifested (Higo *et al.*, 2001). It has been noted that unfolded polypeptides chains do have persistent structural regions such as partially folded α-helices and turn structures or, less frequently, β-structures in the form of β-hairpins (Dobson, 1992; Shortle, 1996). It is thought that these structures may represent initiating points in structure formation, and hence play an important organising role during the earliest stages along the protein folding pathway (Thompson *et al.*, 1997; Williams *et al.*, 1996). This has been shown for β-hairpin formation, where the loss in conformational entropy is compensated for by strong stabilising interactions occurring early along the reaction coordinate thus lowering the free energy barrier (Blanco *et al.*, 1998; Munoz *et al.*, 1998; Maness *et al.*, 2003). In saying this, it is possible that a protein sequence may provide local or non-local interactions during the chain collapse which can counteract the energy input of these motifs, invalidating this possibility. The folding mechanism of β-hairpin is initiated by forming a loop whereby side-chain and main-chain hydrogen bonds form, and hence maintain the β-hairpin structure (Munoz *et al.*, 1998). Helices tend to fold differently, in that they require an initial local helical turn to initiate helical growth along it axis (Williams *et al.*, 1996; Thompson *et al.*, 1997), see Eaton *et al.* (2000) for most recent review on fast folding events. Therefore, α-helices tend to represent a mechanism of folding involving more local interactions, whereas β-hairpins and loops undergo a different mechanism of folding, involving local and non-local interactions. Current protein folding models attribute local interactions as playing a more critical role in defining the protein-folding pathway (Dill *et al.*, 1993).

The free energy barrier to helix formation is purely entropic (Shimada *et al.*, 2001), and the rate appears to be dominated by a rotational diffusive search for a nucleation event (Hummer *et al.*, 2001). The formation of helices incorporates two steps, that is helix nucleation followed by a series of helix propagation steps. Once the nucleus is formed, the formation of the rest of the helix is rapid. Ser is a good helix nucleation sites for helix growth (Shimada *et al.*, 2001), owing to its absence of side-chain entropy, thus lowering the free energy cost of constraining this residue in a helical conformation. From a sequence and structural based study, there appears to be a pronounced directionality (N to C terminus of the polypeptide chain) of helix growth defined by the polypeptide sequence (Pal *et al.*, 2003). Helix simulation show that helix nucleation at the N-terminal end propagates helix growth towards the C-terminus (Shimada *et al.*, 2001). The Phe-Thr/Ser-X-X-Glu/Asp-Leu fingerprint sequence, consisting of the N-cap residue (Thr/Ser) and N' staple residue (Phe), have also been shown to determine the growth direction of the polypeptide chain forming helices (Munoz and Serrano, 1995b).

Thus, the N' residues of a staple motif together with the Ser N-cap, could represent the nucleation site and direction for α-helix growth. In the nucleation-condensation model it is proposed that in the unfolded chain, putative secondary structural motifs, such as N- and C-capping motifs may define the starting point in folding (Fersht, 2000). N-terminal motifs of helices, such as ST and Asx motifs (see section 1.2.1) are more common than the C-terminal motifs (Pal *et al.*, 2003). Thus, these putative motifs may represent common folding nucleation sites and/or stabilising features for the formation of secondary structural elements, and hence the folding of proteins. The evolutionary conserved N-capping motifs, identifiable from structure and sequence based alignments of homologous proteins, may represent a folding nucleation site of a particular topological fold and/or an obligatory structural element during the early stages of folding (Tseng and Liang, 2004).

Chapter 2

N-capping motifs in class Alpha glutathione transferases

GSTs are a good model for investigating the evolutionary role of N-cap motifs in context of protein function, structure stability and folding. hGST A1-1 has two N-capping motifs per subunit; (1) at the N-terminal end of the catalytically important α-helix 9 for class Alpha GSTs (see section 2.2.1), and (2) at the N-terminal end of α-helix 6 which is positioned within the hydrophobic core of domain 2 (Aceto *et al.*, 1997). From the literature it is well known that N-cap motifs serve to promote helix stability. Additionally, the evolutionary conservation of these N-cap motifs does suggest their importance in playing a role as a stabilising structural feature promoting helical stability and in initiating helix formation.

2.1 Nomenclature, function and structure of cytosolic glutathione transferase

Glutathione transferases (GSTs, EC 2.5.1.18) are a superfamily of complex and multifunctional enzymes found in all aerobic organisms (Hayes and Pulford,

1995; Reinemer *et al.*, 1996; Ketterer, 2001). This superfamily of mainly cytosolic enzymes is further subdivided into an ever-increasing number of species-independent classes based on a variety of criteria such as; immunological properties, archetypal fold, intracellular location and biological activity. To date there exist at least 15 gene classes: Alpha, Beta, Delta, Kappa, Lambda, Mu, Omega, Phi, Pi, Sigma, Theta, Tau, Zeta and metalloglutathione transferases (see review by Sheehan *et al.* (2001) and references therein). Members of each enzyme class have been isolated and characterised from mammalian tissues, plants, insects, fungi, bacteria and helminths. Although GST isozymes of each class exhibit relatively broad substrate selectivity, most have unique catalytic attributes that are important in defining their physiological significance (Sheehan *et al.*, 2001; Armstrong, 1991, 1994; Hayes and Pulford, 1995; Armstrong, 1997).

GSTs play an important role in cellular defense against endogenous or exogenous toxic chemicals. This is achieved by their role as phase II detoxification enzymes in the mercapturate pathway. This pathway involves rendering toxic compounds more water soluble allowing for their eventual elimination from the cell (Mannervik, 1985; Hayes and Pulford, 1995). This is achieved by catalysing the conjugation of reduced glutathione (GSH) *via* its sulphur atom to a variety of toxins containing an electrophilic functional group (Armstrong, 1991, 1997). Products of oxidative metabolism such as organic hydroperoxides, epoxides, quinones, and activated alkenes are possible "natural" substrates for the glutathione transferases (Mannervik and Danielson, 1988). GSTs also function to bind a wide range of endogenous and exogenous ligands non-catalytically. This ligand binding facilitates the transport and storage of a variety of hydrophobic non-substrate compounds such as steroids (Barycki and Colman, 1997), drugs (Listowsky *et al.*, 1988) and hormone metabolites (Danger *et al.*, 1992; Ketley *et al.*, 1976). GSTs have also been implicated in a variety of resistance phenomena involving cancer chemotherapy agents (Tew, 1994; McLellan and Wolf, 1999), insecticides (Tang and Tu, 1994; Ranson *et al.*, 1997), herbicides (Edwards *et al.*, 2000; Dixon *et al.*, 1998) and microbial antibiotics (Arca *et al.*, 1997; Bernat *et al.*, 1997).

The cytosolic GSTs exist either as homo- or heterodimers (Armstrong, 1997).

Within a particular class of these dimeric enzymes, subfamilies may sometimes be identified representing unique subunit types that often share a high sequence identity (Hayes and Pulford, 1995; Rowe *et al.*, 1998). The topological fold and structural scaffold is well preserved along the evolutionary development of GSTs (Wilce *et al.*, 1996; Dirr *et al.*, 1994; Sheehan *et al.*, 2001). A representation of the structural fold of class Alpha GSTs (hGST A1-1) is shown in Figure 2.1. A nomenclature system has been proposed such that hGST A1-1 refers to human GST from class Alpha composed of two type 1 subunits (Mannervik *et al.*, 1992). This nomenclature will be used throughout this work. To date crystal structures resolved for all the representative classes of GSTs indicate that cytosolic GSTs are dimeric consisting of two chains or subunits. Each subunit consists of two distinct domains, namely; domain 1 (N-terminal domain) and domain 2 (C-terminal domain).

Domain 1 encompasses approximately the first 80 residues, comprising of four β-sheet and three α-helices that is homologous to the thioredoxin fold (Martin, 1995). This fold can be subdivided into two motifs, being the N-terminal $\beta\alpha\beta$ motif and a C-terminal $\beta\beta\alpha$ motif connected by a loop of residues that incorporates a third α-helix. Domain 1 of GSTs comprise of an active-site, which for the more ancient GSTs (Omega/Theta) is a redox active-site consisting of either one or two cysteine residues (much like thioredoxin and glutaredoxin-2) or a Ser (Theta). For the more evolved GSTs (Alpha/Mu/Pi/Sigma) the equivalent active-site residue is a highly conserved Tyr, which is the active center facilitating GSH conjugation to toxic chemicals (Dirr *et al.*, 1994; Armstrong, 1997; Mannervik, 1985). A characteristic structural hallmark to the thioredoxin-like fold is a conserved Pro in the *cis* configuration located in β-turn immediately before β-strand 3, of which is adjacent to the active-site, shown in Figure 2.2. In GSTs, domain 1 is highly conserved and constitutes most of the GSH binding site.

Domain 2 of GSTs (residues 80 to 214 residues, on average) is connected to domain 1 by a short linker sequence. Domain 2 is specific to GSTs and has an all-helical topology referred to as the GST-like fold, where in most cases consists of 5 helices (Dirr *et al.*, 1994; Sheehan *et al.*, 2001). For the more

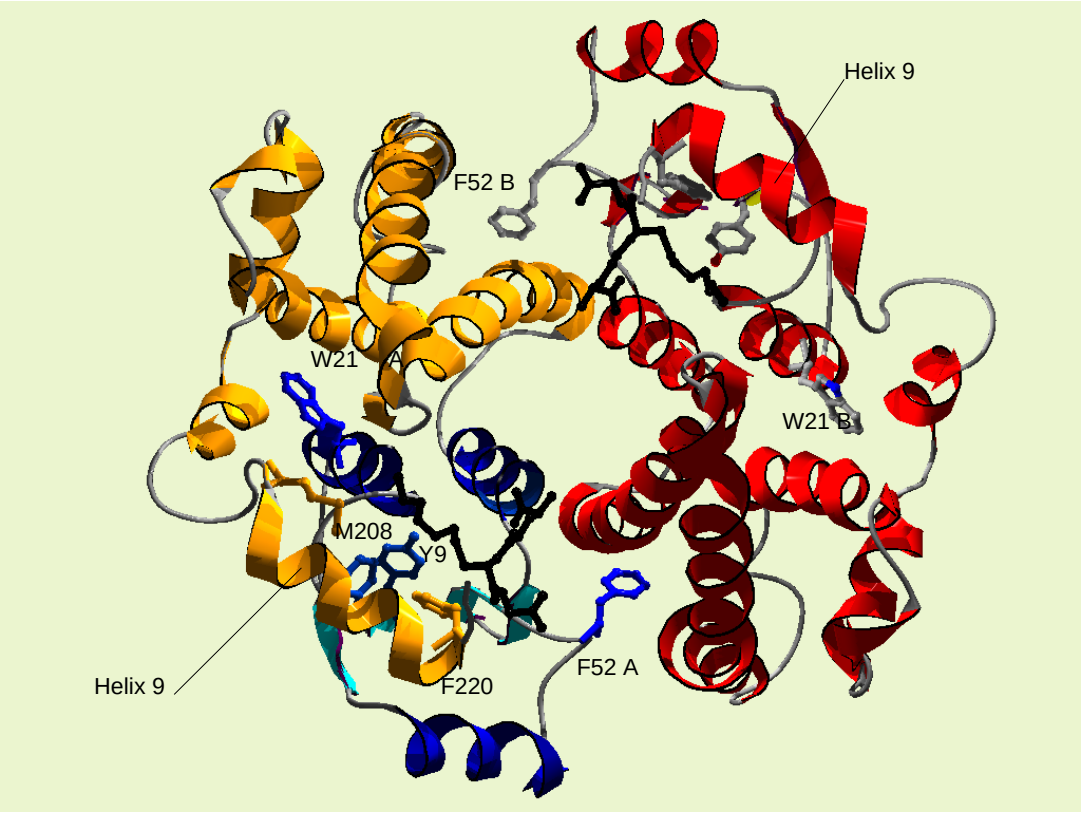

Figure 2.1: Ribbon representation of the structure of homodimeric class Alpha GST, hGST A1-1 (Le Trong *et al.*, 2002). Viewing down the two fold axis, subunit A is represented in red, and subunit B in blue and orange. Each subunit has a thioredoxin-like domain (domain 1, blue) at the N-terminal region, connected to a GST-like domain (domain 2, orange). The active-site ligand (in black, S-hexyl GSH) complexed to hGST A1-1 is located in the glutathione (G-site) and hydrophobic (H-site) binding pockets. The catalytically important residues; Tyr-9, Phe-10, Met-208 and Phe-220 and structurally important residues; Phe-52 and Trp-21 are represented in ball-and-stick. The C-terminal region (α-helix 9) of class Alpha GSTs forms a "lid" over the active-site. The figure was generated using Swiss-PdbViewer (Guex and Peitsch, 1997).

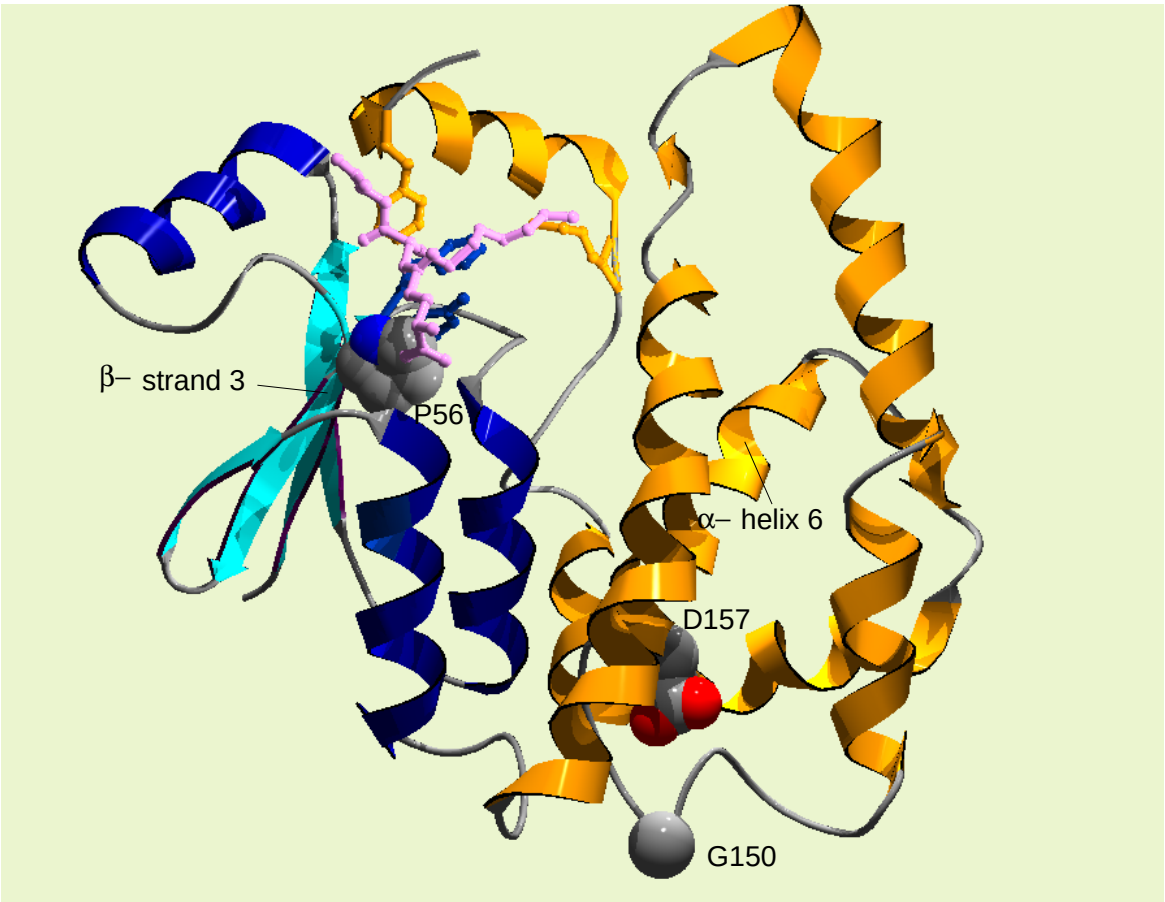

Figure 2.2: Ribbon representation of one subunit of class Alpha GST (hGST A1-1) (Le Trong *et al.*, 2002). The N-terminal domain 1 (blue) has a thioredoxin-like fold, which is connected to the C-terminal domain 2 (orange) which has a GST-like fold. Residues and S-hexyl GSH (pink) complexed with hGST A1-1 are represented in ball-and-stick in order to highlight the active-site. Conserved residues amongst the supergene family of GSTs, that are non-functional, are represented in van der Waals contacts, namely; Pro-56 in domain 1, and Asp-157 and Gly-150 in domain 2. The figure was generated using Swiss-PdbViewer (Guex and Peitsch, 1997).

modern GSTs (Alpha/Mu/Pi/Sigma) this domain makes up the second active-site, a hydrophobic binding pocket (H-site) that lies adjacent to the G-site of domain 1 (Wilce *et al.*, 1996). This GST-like fold is more divergent in terms of its amino acid sequence and topological fold compared to the thioredoxin-like domain (Sinning *et al.*, 1993; Dirr *et al.*, 1994).

The dimeric structure of GSTs forms primarily through interactions between domain 1 of one subunit and domain 2 from the second subunit. This results in a V-shaped cleft at the intersubunit interface, and is exposed to the bulk solvent. For more ancient proteins, such as class Omega and Theta GSTs, the two interacting subunits form a surface interface that is flat (Board *et al.*, 2000; Reinemer *et al.*, 1996), whereas for more evolved GSTs a more round surface interface exists. A key distinguishing feature at the subunit interface of the evolved GSTs separate from the more ancestral GSTs, is the hydrophobic "lock-and-key" interaction (Dirr *et al.*, 1994; Sheehan *et al.*, 2001). An aromatic residue (Phe-52 in hGST A1-1) acts as a "key" extending from the loop preceding β-3 which fits into a hydrophobic "lock" provided by α-helix 4 and 5 of the other subunit (Sayed *et al.*, 2000), shown in Figure 2.1.

2.2 N-cap motif of α-helix 9 in class Alpha GSTs

2.2.1 Identification of an N-capping motif of α-helix 9

A structural feature unique to class Alpha GSTs is α-helix 9 located at the C-terminal end of the protein. The α-helix forms an integral part of both the G- and H-site and is a defining structural feature for the catalytic functioning of this class of enzymes (Board and Mannervik, 1991; Sinning *et al.*, 1993; Cameron *et al.*, 1995). A sequence alignment of the C-terminal end of class Alpha GSTs, represented in Table 2.1, shows the conservation of the N-cap (Asp-209) residue of α-helix 9. This N-cap residue forms a class 3 Asp N-cap motif according to Wan and Milner-White (1999), shown in Figure 2.3.

Table 2.1: Sequence alignment of the C-terminal end of class Alpha GSTs

Gi number	Type	Source	PDB	Sequence
				N' N N1 N2 N3
121730	A1	human	1K3Y	M **D** E K S L E E A R K I F R F
66611	A1	rat	1EV9	M **D** A K Q I E E A R K I F K F
11514503	A1	mouse	1F3B	M **D** A K Q I Q E A R K A F K I Q
1170081	A1	rabbit		M **D** E K N L E K A K K I F K I P
2495106	A1	bos taurus		T **D** E K K I E E A R K V F K F
30749486	A2	mouse	1ML6	M **D** A K Q I E E A R K V F K F
121733	A2	human		M **D** E K S L E E S R K I F R F
6166190	A3	human		A **D** A K A L E E A R K I F R F
7767212	A4	human	1GUM	P **D** E I Y V R T V Y N I F R P
3114387	A4	mouse	1GUK	P **D** G P Y V E V V R I V L K F

The sequence alignment of the C-terminal region begins two residues preceding the first α-helical residue (N1) of α-helix 9. The underlined residues correspond to those located in α-helix 9 resolved from crystal structures for GST A1-1. Type refers to the subunit type of class Alpha GST. The conserved N-cap residue, Asp, is highlighted in bold.

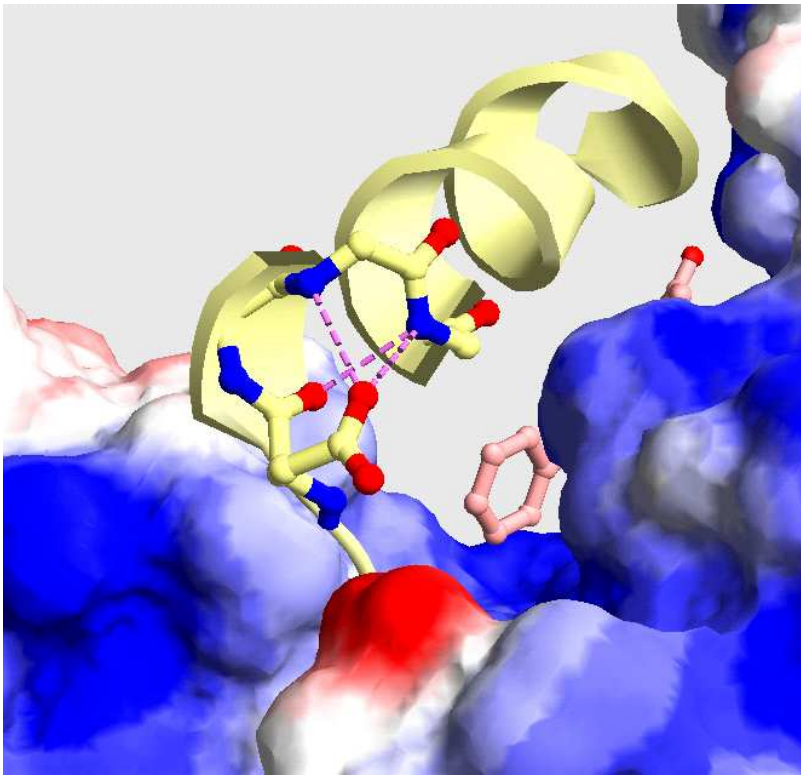

Figure 2.3: Graphical representation of hGST A1-1 (pdb code:1KY3) showing Asp-209 residue forming an Asp N-cap motif of α-helix 9 (shown in ribbon) and the surface of the protein is represented in its electrostatic surface potential. Here the Asp N-cap side-chain CO group forms a bifurcated hydrogen bond with the main-chain NH groups of the i+2 (i being Asp-209) and i+3 residues. A second bifurcated hydrogen bond pattern occurs between the Asp N-cap main-chain CO group and the main-chain NH group of i+3 and i+4 residues. The latter hydrogen bonding pattern is typical for back-bone α-helical hydrogen bond. In the presence of the ligand S-benzyl GSH (pink), the α-helix 9 is situated over the active-site where this ligand occupies both the H- and G-site of hGST A1-1. The figure was generated using Swiss-PdbViewer (Guex and Peitsch, 1997).

The side-chain CO group of residue Asp-209 of hGST A1-1 (pdb code: 1KY3) forms two hydrogen bonds with the main-chain NH groups of N2 and N3. This N-capping motif, however does not form an N-capping box, since there is no reciprocal N3 (Ser-212) side-chain to main-chain N-cap hydrogen bond (Figure 2.3). Upon inspection of all resolved structures of the class Alpha GSTs (see section 3.2 for listed PDB codes), the Asp N-cap main-chain CO group forms the first helical hydrogen bond of α-helix 9. In addition, the side-chain carboxylate group of Asp, always fulfils at least one hydrogen bond with the main-chain NH group of N3, thus forming an Asp N-cap motif at the N-terminal end of α-helix 9 of class Alpha GSTs. In some instances amongst class Alpha GSTs the Asp side-chain forms a second side-chain to main-chain hydrogen bond with either N1, N2 or N4 residue, forming a bifurcated hydrogen bond pattern. This is seen in rat GSTA2-2 (Gu *et al.*, 2003) and GST A1-1 from human (Sinning *et al.*, 1993), mouse (Gu *et al.*, 2000), and rat (Adman *et al.*, 2001).

2.2.2 Role of α-helix 9 towards the catalytic mechanism of hGST A1-1

Considerable amount of research has contributed in determining the precise catalytic mechanism of class Alpha GSTs. Unique to this class of enzymes, is the C-terminal region which includes α-helix 9 forming part of the active-site (Board and Mannervik, 1991). For GSTs each active-site per subunit, of this dimeric protein, functions independently of the other (Mannervik and Danielson, 1988). In saying this, it is only in the native dimeric form, where the subunits are sufficiently stabilised maintaining their native conformation, that GSTs are fully functional. The G-site of hGST A1-1 shares a common feature to the other classes of GSTs (Mu/Pi/Sigma), in that it is highly specific for the binding of the tripeptide, glutathione (γ-Glu-Cys-Gly). Most of the polar atoms in glutathione (GSH) are involved in hydrogen bonds or salt links shown in Figure 2.4. The cysteinyl sulfur atom of GSH forms close contacts with the hydroxyl group of Tyr-9 (numbering according to hGST A1-1). The Tyr-9 residue at the G-site contributes to the catalytic mechanism of GSH

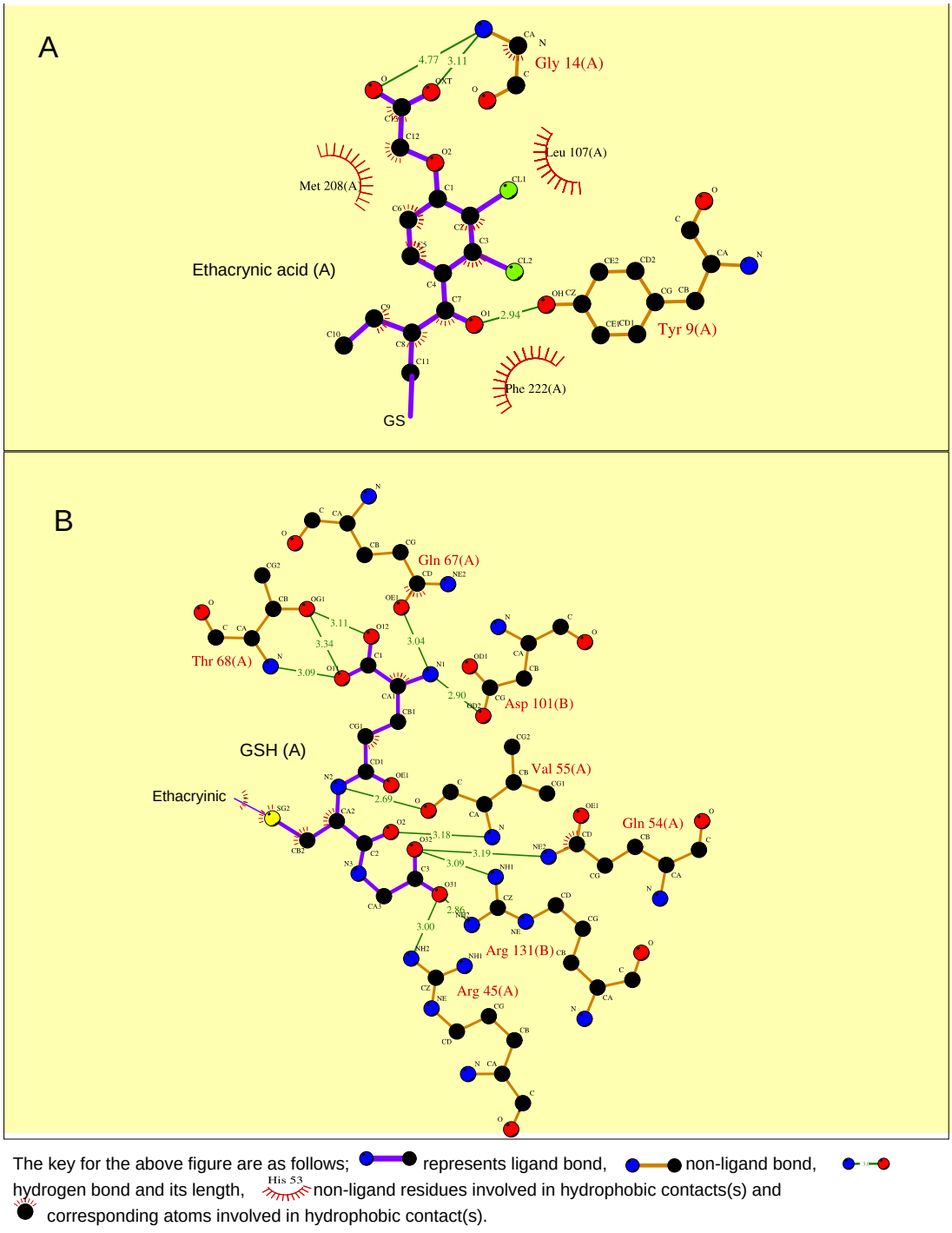

Figure 2.4: Schematic representation of amino acid residues of hGST A1-1 forming van der Waals contacts and hydrogen bonds with ethacrynic-GSH conjugate. For clarity, the ethacrynic portion bound to the hydrophobic binding pocket is shown in panel A. The GSH portion, bound to the G-site, is shown in panel B. The key for the figure is shown below panel B. The figure was generated using Ligplot (Wallace *et al.*, 1995)

conjugation reactions by modulating the pK_a and the nucleophilic reactivity of the enzyme bound GSH. This is achieved by stabilising the thiolate anion through hydrogen bonding as Tyr-OH $\cdots^-$SG (Stenberg *et al.*, 1991; Kolm *et al.*, 1992). By stabilising the thiolate ion, the Tyr-9 is directing the sulfur atom into position suitable for nucleophilic attack on the second substrate situated at the H-site. In the substrate-free form of class Alpha GSTs the catalytic tyrosine is "activated" by several nearby residues that lower the pK_a of the phenolic hydroxyl group from 10.3 (the pK_a value of free tyrosine in solution) to 8.2 (Bjornestedt *et al.*, 1995). This lowering of the pK_a of Tyr-9 is an important feature for its catalytic role in promoting GSH thiolate formation. In the crystal structures GSH always binds in an extended conformation and positioned such that the sulfur atom extends into the electrophilic substrate binding site. Residues at the active-site that contribute to the binding of the second hydrophobic substrate involve hydrophobic residues, Phe-10, Arg-15, Met-208, Ala-216, Ala-219, Phe-220 and Phe-222. Figure 2.4 shows the residues at the H-site that form van der Waals contacts with ethacrynic acid. Of the above mentioned residues forming part of the H-site, the latter five are located at the C-terminal region of the protein, including α-helix 9. It is known that α-helix 9 provides for the specificity between the subclasses of class Alpha GSTs (Gu *et al.*, 2000; Bruns *et al.*, 1999). Met-208 is situated at the "bottom" of the electrophile binding pocket where its side-chain can interact with the second substrate and possibly contributes to the specificity of binding to this site (Widersten *et al.*, 1994). Moreover, this residue is suggested to primarily affect the aromatic substitution reaction by means of interacting with the delocalised negative charge of the substituted ring structure in the transition state (Widersten *et al.*, 1994).

X-ray crystal structures of hGST A1-1, in the presence of GSH product conjugates, demonstrate that the C-terminal region is a structured α-helix positioned over the active-site of the protein. This conformation brings the Phe-220 in close contact with the catalytically important Tyr-9 (Sinning *et al.*, 1993; Cameron *et al.*, 1995), and is involved in "finely tuning" the ionisation of the Tyr-9 (Lian, 1998; Allardyce *et al.*, 1999). Interestingly, the docking the α-helix 9 over the active-site in hGST A1-1, the Phe-220 and Phe-10 (adjacent

to Tyr-9) compete for the same location in the active-site, and hence cannot be accommodated simultaneously. It is the side-chain of the Phe-10 in apo form, that changes its orientation so as to create room for Phe-220 and thereby making it possible for the proper accommodation of the α-helix 9 to dock tightly over the active-site. This altered orientation of the phenolic ring of Phe-10 provides a molecular switch which plays a critical role in the direct coupling in the ionisation of Tyr-9 (Ibarra *et al.*, 2001), active-site solvation, and conformational dynamics of α-helix 9 (Ibarra *et al.*, 2001). This molecular switching is absent for other isozymes of class Alpha GSTs, such as rGSTA1-1 (Adman *et al.*, 2001) and mGSTA1-1 (Gu *et al.*, 2000), since the Phe-10 (α-helix 9 residue) occupies a different location, avoiding steric clashes with the active-site Phe-220. Phe-10 with Phe-220 directly affect the lowering of the Tyr-9 pK_{a} in the apo conformation as well as in the enzyme's complexed form (Ibarra *et al.*, 2001). Importantly, in the ligand free form, the Phe-220 of the C-terminal helix does contribute to the environment of Tyr-9 (Ibarra *et al.*, 2001). An important feature to the H-site and catalysis for hGST A1-1 is the desolvation of the active-site relative to bulk solvent as a result of the docking of α-helix 9 over the active-site. This increased hydrophobicity of the H-site is important in directing and facilitating the degree of substrate specificity and turn-over rate of product formation (Allardyce *et al.*, 1999; Dirr and Wallace, 1999; Bruns *et al.*, 1999; Gu *et al.*, 2003).

2.2.3 Non-substrate ligand binding to class Alpha GST

An interest in the non-substrate ligand binding of GST A1-1 lies in the ability of these compounds to elicit noncompetitive inhibition. The reports that certain hydrophobic compounds (particularly endogenous anions such as prophyrins and bile acids) act as non-competitive inhibitors. This lead to the assumption that they bind at a discrete site which is dubbed the ligandin site, i.e. the non-substrate ligand site. The anitihelmenthitic drug praziquantel binds within the solvent-accessible inter-subunit cleft region of GSTs from *Schistosoma japonicum* judged from crystal structures (McTigue *et al.*, 1995).

A variety of organic anions, including dyes and acidic drugs, have been characterised as inhibitors of different GST isoforms. Two such examples are cibacron blue and bromosulfophthalein, both of which are non-substrate ligands and non-competitive inhibitors of GST P1-1 (Mannervik and Danielson, 1988). In crystallographic studies, however, these ligands are clearly localised to the hydrophobic substrate binding site; and the observation of noncompetitive inhibition appears to be due to the large size of this site which can simultaneously be occupied by substrate and inhibitor (Oakley *et al.*, 1999).

Another anionic dye, that has been demonstrated to bind to GSTs is ANS. The usefulness of ANS as a fluorescent probe of non-polar binding-sites in studies with enzymes and proteins has been demonstrated by a number of investigators (Stryer, 1965; Semisotnov *et al.*, 1991). It has been proposed by Sluis-Cremer *et al.* (1996) from their fluorescence resonance energy transfer studies that ANS binds to this amphipathic cleft of hGST A1-1 with a stoichiometry of two ANS molecules per dimeric protein. It is generally assumed that neither the negative sulfonate charge on the ANS, nor charges on the protein, participate significantly in ANS-protein interaction. Titration calorimetry studies of ANS binding to hGST A1-1 show that the C-terminal α-helix 9 contributes towards an increasing affinity of binding, and that the Phe-222 residue alone contributes a significant hydrophobic component of the binding of ANS. Whether ANS binds at the amphipathic cleft or at/near the H-site for GST A1-1 remains to be elucidated (Mosebi *et al.*, 2003; Sayed *et al.*, 2002). It has been suggested by Le Trong *et al.* (2002) that extended ligand binding surfaces for non-substrate ligands may exist for class Alpha GSTs, adjacent to the G-site. Interestingly, no crystal structures of GST A1-1 have emerged in the presence of non-substrate ligands despite the fact that this isoform is most closely associated with ligandin function.

2.2.4 Dynamics and conformational transition of α-helix 9

The structural order of the C-terminal α-helix 9 is highly dependent on the occupancy of the two binding pockets at the active-site. The C-terminal region in hGST A1-1 is a well-ordered helical structure located over the active-site when the protein is complexed with S-hexylglutathione (Le Trong *et al.*, 2002), S-benzylglutathione (Sinning *et al.*, 1993), ethacrynic acid or ethacrynic acid-GSH conjugate (Cameron *et al.*, 1995; Lian, 1998; Zhan and Rule, 2004). The C-terminal helix has been observed, in an altered position, with glutathione sulfonate bound to rGSTA1-1. The stabilisation of this region appears to occur due to the crystal packing (Adman *et al.*, 2001). Very recently, from NMR studies, it has been reported that the C-terminal region in the presence of GSH is a well ordered helix, however the α-helix 9 is more flexible at its ends compared to that bound with ethacrynic acid-GSH conjugate (Zhan and Rule, 2004). From crystal structures, in contrast to the ligand complexed forms of hGST A1-1, the C-terminal region is "invisible" in the apo form (Cameron *et al.*, 1995). Even though the information of the structure of this disordered C-terminal region is limited, NMR studies by Zhan and Rule (2004) report that the C-terminal region resembles an α-helical conformation. Their results exclude the possibility that the C-terminal region in hGST A1-1 persists as a random coil conformer (Zhan and Rule, 2004), but rather samples a heterogeneous ensemble of dynamic helices, whereby on average no single position can be described (Cameron *et al.*, 1995; Adman *et al.*, 2001; Zhan and Rule, 2004). The structural changes of the C-terminal region, upon ligand binding, can be best described as a transition between a disordered to a well-ordered α-helix situated over the active-site of the protein, forming a "lid" (Zhan and Rule, 2004). Urea-dependent activity studies of hGST A1-1 have shown that when GSH is bound to hGST A1-1, there is an increased stability of α-helix 9 compared to that of the apo form (Dirr and Wallace, 1999; Mosebi *et al.*, 2003).

The NMR data demonstrate that the greatest heterogeneity of the "non-uniformly" dynamic C-terminal region is at the Phe-220 position (Lian, 1998;

Zhan and Rule, 2004). To date results have shown that the stabilised tyrosinate form of Tyr-9, that is hydrogen bonded with the free thiol of GSH, appears to contribute in part toward stabilising/sequestering a more closed helical conformation. More detailed studies have revealed that the coupled Phe-10/Tyr-9 interaction functions as a whole in directing the sequestering of a well-ordered α-helix 9 over the active-site (Ibarra *et al.*, 2001; Atkins *et al.*, 1997). Another residue, Phe-220 also appears to have some contributing role towards α-helix 9 stability (Atkins *et al.*, 1997), however the effects appear to be less pronounced as compared that of the Phe-10 residue. To date it appears that the ionisation states of Tyr-9 and GSH and the conformation of α-helix 9 are all intimately linked. A more detailed model illustrating the order and the direct role of these specific interactions remain to be elucidated. Additionally, and very importantly the flexibility of the C-terminal region is not observed for all subtypes of class Alpha GSTs. For A4-4 the crystallographic data show that α-helix 9, in the absence of ligands, is a well ordered α-helix positioned over the active-site of the protein (Bruns *et al.*, 1999). This raises the question, what factors contribute to the extent of a well-ordered C-terminal helix positioned over the active-site of GST A1-1 given that, for GST A4-4 the α-helix 9 is not disordered?

2.3 N-capping motif of α-helix 6 in class Alpha GSTs

2.3.1 Identification of the N-capping motif of α-helix 6

A sequence alignment of all GSTs within the GST supergene family reveal the existence of a Ser/Thr N-cap and Asp N3 (Allocati *et al.*, 1999). This motif is also common to chloride intracellular channel 1 (Clic1) (Harrop *et al.*, 2001), and Ure2 yeast prion protein from *Saccharomyces cerevisiae* (Bousset *et al.*, 2001). These two proteins have recently been identified as GST homologs, and are functionally unrelated to mammalian GSTs. The Ser N-cap and a conserved Gly in the loop preceding α-helix 6, shown in Figure 2.2, is located

in the hydrophobic core of domain 2 of GSTs (Aceto *et al.*, 1997; Cocco *et al.*, 2001; Rossjohn *et al.*, 2000). The N-capping motif of α-helix 6 in class Pi and Alpha GSTs forms an extensive hydrogen bond network, which has been shown to contribute to the rate of folding and towards the thermal stability of these proteins (Cocco *et al.*, 2001; Dragani *et al.*, 1997). Flanking the N-capping motif in α-helix 6 is a hydrophobic staple motif identified in class Alpha and Pi GSTs (Cocco *et al.*, 2001; Stenberg *et al.*, 2000).

The N-capping motif of α-helix 6, forms a reciprocal main-chain side-chain hydrogen bond between the N-cap (Ser-154) and N3 (Asp-157) residues (Dragani *et al.*, 1997; Cocco *et al.*, 2001; Aurora and Rose, 1998), as shown in Figure 2.5. Here, the side-chain CO group of Asp-157 hydrogen bonds with the main-chain NH group of Ser-154. The first helical hydrogen bond forms between the Asp main-cain carbonyl group and the NH group of Ser-154. The side-chain hydroxyl group of Ser-154 reciprocates a hydrogen bond with the main-chain NH group of Leu-156. The hydrogen bond pattern between N-cap and N3 forming the N-capping motif is defined as a class 5 Ser-motif, according to the classification by Wan and Milner-White (1999). An interesting feature of the N-capping motif of α-helix 6 includes an additional hydrogen bond between the side-chain of the N-cap and N3 residues formed *via* a water molecule. This water molecule also mediates hydrogen bonds with the NH group of Leu-148 and Val-149, and the main-chain carbonyl group of Val-149 and Lys-152. The side-chain carboxylate group of Asp-157 forms an additional hydrogen bond with the main-chain amino group of Leu-148. This extensive hydrogen bonding network, described above, has also been identified for class Pi GSTs (Dragani *et al.*, 1997; Cocco *et al.*, 2001). A second helix stabilising motif, the staple motif, in GST A1-1 is present consisting of two topologically conserved hydrophobic residues, being Leu-153 (N') and Ile-158 (N4). The side chains of these two residues, bracketing the N-capping motif, form hydrophobic interactions which contribute to the stability of α-helix 6 of class Alpha and Pi GSTs (Cocco *et al.*, 2001; Stenberg *et al.*, 2000).

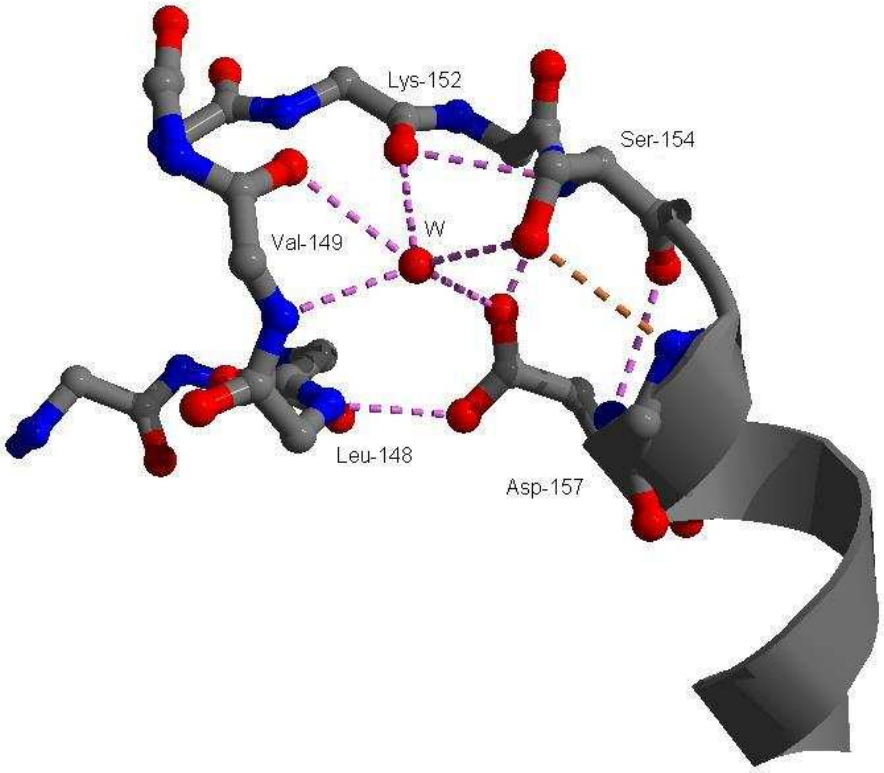

Figure 2.5: Graphical representation of hGST A1-1 (pdb code:1KY3) showing the N-capping interactions at the N-terminal end of α-helix 6, which is characterised by a network of 8 hydrogen bonds coordinated by the N-cap (Ser-154) and N3 (Asp-157) residues. The hydrogen bonds are indicated as dashed lines in purple. The side-chain OH group of Ser-154 forms a bifurcated hydrogen bond with a water molecule (W) and the main-chain NH group of Asp-157. One of the side-chain CO groups Asp-157 forms a bifurcated hydrogen with a water molecule (W) and the main-chain NH group of Ser-154. The second carboxyl moiety of Asp forms a hydrogen bond with the main-chain NH group of Leu-148.

2.3.2 N-capping motif of α-helix 6 as a folding nucleation site in GSTs

Since the N-cap motif of α-helix 6 is highly conserved amongst GSTs and GST-like proteins, this does imply the possibility that this conserved region is a folding nucleation site (Mirny and Shakhnovich, 2001), and section 1.4.1. The kinetics of folding of GSTs monitored by catalysis is limited in that the conformational detail and hence the mechanism of folding for these mutants cannot be resolved. In other words, the faster folding events i.e., the fast and intermediate folding phases monitored by fluorescence spectroscopy (Wallace and Dirr, 1999) remain unresolved. It has been proposed that N-capping motif of α-helix 6 is a folding nucleation site for class Alpha and Pi GSTs (Aceto *et al.*, 1997). This has been proposed on the basis that the N-capping motif contributes to the helical content of the peptide corresponding to the α-helix 6 of GST P1-1, and that this α-helix folds autonomously in more apolar solvents. However, the role of the N-capping motif as a folding nucleation site (Dragani *et al.*, 1997) and/or as an important structural feature involved in the early folding events of GSTs has not been validated. A more detailed study, such as by stopped-flow fluorescence spectroscopy, is therefore required to ascertain the significance of this motif in the mechanism of folding for GSTs.

2.3.3 Conformational stability and folding of GST A1-1

GSTs fold *in vitro* without the aid of other proteins or enzymes such as chaperones, etc. Over the past decade unfolding mechanisms for the different classes of GSTs (Alpha, Mu, Pi, Sigma and Sj26) have been elucidated and have recently been reviewed by Dirr (2001). Work by Wallace *et al.* (1998b) have reported that hGST A1-1 unfold *via* a cooperative two-state mechanism, indicating that only the folded dimer and unfolded monomer are present in significant concentrations at equilibrium. However, the kinetic folding reactions of hGST A1-1 reveal a more complex multistate folding pathway with the formation of intermediate species (Wallace *et al.*, 1998b; Wallace and Dirr, 1999).

The unfolding mechanism of hGST A1-1 proposed by Wallace *et al.* (1998a), involves a sequential 3-state pathway, described below,

$$N_2 \rightarrow N_2^* \rightarrow 2U \tag{2.1}$$

where $N_2 \rightarrow N_2{}^*$ represents the rapid and partial loss of the compact packing of the protein, also resulting in the dissociation of the flexible C-terminal helix from the surface of the protein. This is followed by the global unfolding of the dimeric protein to its unfolded monomeric form, $N_2{}^* \rightarrow 2U$. To date, the kinetics of unfolding of variant forms of hGST A1-1 have been carried out demonstrating the stabilising role of Pro-56 (Nathaniel *et al.*, 2003), Leu-164 (Wallace *et al.*, 1998a) and Trp-21 (Wallace *et al.*, 2000), where the transition state of unfolding remained unaltered. The folding mechanism of these variant forms have not been resolved.

The folding and assembly of hGST A1-1 proceeds *via* a more complex multi-state mechanism compared to the unfolding reaction, and is described below in scheme 2.2 proposed by Wallace and Dirr (1999):

$$\begin{array}{lllll} 2U_c \rightarrow & 2M_c \rightarrow & N_2{}^* \rightarrow & N_2{}' \rightarrow & N_2 \\ \updownarrow & \uparrow & & & \\ 2U_t \rightarrow & 2M_t & & & \end{array} \tag{2.2}$$

The unfolded monomers (U) fold rapidly (within the 2 ms) to an ensemble of intermediates (M) that have a loosely packed secondary structure and high degree of solvent exposed residues. This burst-phase intermediate species ($2M_c$ and $2M_t$) proceeds *via* two parallel folding channels leading to the formation of an inactive intermediate dimer ($N_2{}^*$). In scheme 2.2, the ensemble of U and M species with a *cis* or *trans* configuration of the X-Pro-56 peptide bond, are represented as subscript c and subscript t, respectively. The distinction between the two channels is as a result of the kinetic partitioning of a slower rate limiting step, involving the *trans* to *cis* X-Pro-56 isomerisation ($2M_t \rightarrow 2M_c$), from the faster folding reaction where the X-Pro-56 pre-exists in its native *cis*

(M_c) configuration and thus not requiring the isomerisation reaction. After dimerisation the intermediate N_2^* follows a slow unimolecular folding event leading to the final docking of the monomers and the formation of a localised C-terminal region (N_2'). N_2 represents, an additional conformational change involving the α-helix 9 where in the presence of GSH or GSH conjugates, the α-helix 9 is sequestered over the active-site forming a "lid".

2.3.4 Folding mechanism of other proteins homologous to GSTs

Recently, the folding mechanism of Ure2 (Galani *et al.*, 2002; Thual *et al.*, 2001), class Pi GSTs and glutaredoxin-2 (Grx2) (Wallace *et al.*, in preparation) have been reported. The monomeric Grx2 and dimeric Ure2 are homologous to GSTs (Galani *et al.*, 2002; Thual *et al.*, 2001; Xia *et al.*, 2001). Of the above mentioned proteins, all have an N- and C-terminal domain consisting of the thioredoxin fold and GST-like fold, respectively. Briefly, the distinctive features shared in the folding mechanisms of GSTs includes a fast burst-phase involving the formation of well structured monomer species, followed by the tertiary packing of the monomeric species *via* one (Ure2) or two (Alpha/Pi/Grx/Trx) parallel channels leading to native protein. This tertiary packing of the well-ordered monomeric species is coupled with dimerisation (Ure2/Pi/Alpha), and the *trans* to *cis* isomerisation of the X-Pro (position 56 for class Alpha GSTs). Whereas for class Pi GSTs and Grx2, the X-Pro isomerisation occurs during the slow folding phase (Wallace *et al.*, in preparation), i.e. the final folding phase. The slow phase includes the reorganisation and tight packing of the protein involving predominantly tertiary interactions. It has been proposed that the predominant folding mechanism for the GST homolog corresponds more to the thioredoxin folding mechanism (Wallace *et al.*, in preparation), on the basis that the conserved folding of GST homologs are the same as for thioredoxin. This suggests that domain 2 may be less important in the folding mechanism of GSTs. However, it has been reported that domain 2 in GST P1-1 has an intrinsically higher stability than the thioredoxin-like domain, and that this region is able to refold autonomously before subunit

reassociation (Aceto *et al.*, 1995; Dragani *et al.*, 1998).

The Pro isomerisation step appears not to be an essential event for the correct folding of hGST A1-1 but possibly as a rate-limiting step (Nathaniel *et al.*, 2003). Separate studies have investigated the folding of thioredoxin, where it has been shown that the two subdomains within thioredoxin, representing the thioredoxin-like domain of GSTs, fold and associate as a coupled event. Here a *trans/cis* X-Pro (equivalent to X-Pro-56 of hGST A1-1) isomerisation occurs in a compact intermediate with secondary and tertiary structure (Chaffotte *et al.*, 1997). Further, a separate study shows that the preceding β-strand 3 of Pro-56 and the β-strand 1 are involved in the association of the two subdomains of thioredoxin, where the anti-parallel β-sheet 3 is suggested to be preformed and is an early folding event (Tasayco *et al.*, 2000). The requirement of the X-Pro-56 bond in the *cis* configuration for the rapid formation of dimers of class Alpha GSTs has been proposed based on the fact that the driving force of dimerisation appears to be reliant upon the high specificity and complementarity of the two monomers (Wallace and Dirr, 1999; Stevens *et al.*, 1998). Thus, the possible role of the slow *trans/cis* X-Pro-56 isomerisation towards the mechanism of folding of the hGST A1-1, can be explained in that the X-Pro-56 bond in its native *cis* configuration forces the proceeding loop towards the dimer interface that forms the "lock-and-key" motif (Wallace and Dirr, 1999; Dirr *et al.*, 1994). The hydrophobic "lock-and-key" motif (Phe-52 of hGST A1-1) has been shown to compromise the stability of the protein, without the formation of monomeric species (Sayed *et al.*, 2000). Therefore, the lock-and-key" motif appears not to be essential for dimerisation since more stabilising interactions of a charged cluster persist. No refolding studies have investigated whether this motif impacts upon the refolding rates of the protein. The role of the charge cluster in class Mu GSTs contributing to the stability and folding of the dimeric protein has been reported to be important (Luo *et al.*, 2002).

Even though the folding mechanism of proteins homologous to the GST-supergene family have been proposed, much remains to be answered in terms of the detailed mechanism of folding for this family of proteins. Questions such as: (1) which residues or structural features are essential for dimerisation

of GSTs, and to what extent are the monomers folded prior to dimerisation? (2) Do tertiary interactions play an important role in the early stages of folding aiding the partial or complete formation of secondary structures? (3) Do domains 1 and 2 of the monomeric subunit fold independently prior to association where one domain predominates the folding of the subunits, or are the folding of these two domains concurrent events.

2.4 Objectives

This work aims to address the question of the importance of the conserved N-cap motif toward the stability of α-helix 9 and α-helix 6 in class Alpha GST A1-1. In order to investigate the role of the N-cap motif, the N-cap Asp-209 is substituted with a Gly. Any changes in the structure and stability α-helix 9 can be monitored by catalysis and ligand binding studies, and compared to that of the wild-type. Since the folding and assembly model of hGST A1-1 has been proposed (Wallace and Dirr, 1999), this protein is a good model for investigating the role of the N-capping motif of α-helix 6 and the importance α-helix 6 in the folding mechanism of this protein. In this study single site-directed mutagenesis is employed in order to disrupt the N-capping motif, where the N-cap (Ser-154) and N3 (Asp-157) are substituted with an Ala. To investigate the role of the N-cap motif as a folding nucleation site for GSTs, refolding kinetic studies are to be carried out of these N-cap mutants and compared to that of the wild-type. Further, by investigating the role of Leu-164 with respect to its role in the folding of GSTs will provide a more detailed picture into the mechanism of folding with respect to the hydrophobic core packing surrounding a stabilised α-helix 6 along the folding pathway of hGST A1-1. Thus, these results together with the folding studies of the N-cap mutants will give further insight into the role of the N-cap motif and the implications of a stabilised α-helix 6 in the mechanism of folding in hGST A1-1.

Chapter 3

Experimental procedures

3.1 Materials

The High Pure Plasmid Isolation kit, *Dnase*I, *Xba*I *Sma*I were obtained from Roche (Basel, Switzerland). Polypeptides, with sequences corresponding to the C-terminus of the hGST A1-1 polypeptide (208 to 222), were synthesised by Alpha Diagnostic Int. (San Antonio, USA), where their molecular masses were confirmed by MALDI-TOF mass spectrometry, and were determined to be 100% pure. QuikChange and ExSite site-directed mutagenesis kits were from Stratagene (La Jolly, California, USA). Ultrapure urea was purchased from Merck laboratory supplies (Darmstadt, Germany). Glutathione was from ICE Biomedical Inc. (Aurora, Ohio, USA). TFE (99 + % grade), 8-aniline-1-naphthalene sulfonic acid, ethacrynic acid, glutathione sulfonate, *p*-bromobenzyl GSH and 1-chloro-2,4-dinitrobenzene were purchased from Sigma-Aldrin (St. Louis, MO, USA). All other reagents were of analytical grade.

3.2 Molecular graphics and sequence alignment

All sequence alignments were carried out using BLASTP 2.2.6 (Altschul *et al.*, 1997). The calculated root mean square deviation of superimposed structures, the calculated possible hydrogen bonding and inter-atomic distances were carried out using Swiss PDB Viewer (ver. 3.7) (Guex and Peitsch, 1997), unless otherwise stated. Crystal structures were retrieved from the Brookhaven Protein Data Bank (www.rcsb.org) *via* anonymous file transfer. The following PDB files were used for the crystal structures of human GST A4 (1GUM and 1GUL; (Bruns *et al.*, 1999)), murine GST A4 (1GUK; (Krengel *et al.*, 1998)), murine GST A1 (1F3A; (Gu *et al.*, 2000)), rat GST A1 (1EV4; (Adman *et al.*, 2001)), murine GST A2 (1ML6; (Gu *et al.*, 2003)), human GST A1 (1K3Y and 1K3O; (Le Trong *et al.*, 2002)). Representations of the various crystal structures were generated by Molscript (ver. 2.0) (Kraulis, 1991), or Swiss PdbViewer (Guex and Peitsch, 1997). Unless stated otherwise, all least-squares data fitting were performed using the program SigmaPlot (ver. 5.0) (Jandel Corporation).

3.3 α-Helix 9 peptide studies

3.3.1 Peptide design and predicted helical content

Two peptides were designed to investigate the role of Asp-209 (N-cap residue) towards the helical content of the C-terminal region of hGST A1-1. The wild-type peptide (WT-pep: Acetyl-YGM**D**EKSLEEARKIFRF), consisted of the sequence of the C-terminal region, residues 207 to 222, of hGST A1-1. The mutant peptide (DG-pep: Acetyl-YGM**G**EKSLEEARKIFRF) was derived from the WT-pep, where the N-cap residue (Asp in bold), was replaced by a Gly. The Tyr followed by a Gly was included at the N-terminal end of the peptide sequence in order to facilitate the concentration determination of the peptides by absorbance spectroscopy at 275 nm. The Gly was included as a helix breaker (Rohl *et al.*, 1996) ensuring Tyr does not form part of the helical conformation

of the peptide. The predicted helical content of the peptides were calculated using the AGADIR1s-2 helix-coil transition algorithm (Lacroix *et al.*, 1998), available at http://www.emblheidelberg.de/Services/index.html#5. The conditions were set at pH 6.5, temperature 293 K and an ionic strength of 0 M.

3.3.2 Helical content by far-UV circular dichroism

Peptide stock solutions were made in water. The concentrations of WT-pep and DG-pepwere determined by UV absorbance of the Tyr, using an extinction coefficient of 1450 $M^{-1}cm^{-1}$ at 275 nm (Brandts and Kaplan, 1973). Far-UV CD measurements were at 20°C in a 1 mm path length cuvette using a Jasco model 810 CD spectropolarimeter. Averaged CD signals, corrected for solvent baseline, were converted to mean residue ellipticity,

$$[\Theta] = \frac{100\ X\ \Theta}{Cnl} \tag{3.1}$$

where C is the peptide concentration in millimolar, and Θ is measured ellipticity in millidegree, n is the number of residues (17 in this case), and l is the path length (cm). The percentage helical content was calculated from the mean residue ellipticity at 222 nm, using the equation,

$$\%Helix = \frac{[\Theta] - [\Theta]_{coil}}{[\Theta]_{helix} - [\Theta]_{coil}} \tag{3.2}$$

where $[\Theta]_{helix}$ and $[\Theta]_{coil}$ represent the mean residue ellipticity of a helix [-42500(1 - (3/n)] and random coil (+640), respectively (Rohl *et al.*, 1996). The units of mean residue ellipticity are $deg.cm^2dmol^{-1}$.

The relationship of ellipticity at 222 nm and increasing peptide concentrations (0-100 μM) in 5 mM phosphate buffer, pH 6.5 was assessed. The peptide conformation was monitored by far UV CD with increased concentrations of a

helix propagating co-solvent, 2,2,2-trifluoroethanol (TFE), ranging from 0-63% (v/v).

3.4 Site-directed mutagenesis of hGST A1-1

The plasmid (pKHA1), which contains the encoded gene for human glutathione transferase A1-1 (hGST A1-1) from Hepatoma HepG2 cells (Stenberg *et al.*, 1992), was a gift from B. Mannervik (Department of Biochemistry, University of Uppsala, Sweden). The plasmid pKHA1 was used for the design and expression of variant forms of hGST A1-1. Plasmids were purified using the High Pure Plasmid Isolation Kit. The concentration of extracted DNA was approximated and checked for its correct size by 0.8% agarose gel electrophoresis (AGE) using molecular weight markers. Single restriction digestions were used to characterise pKHA1, where the restriction enzyme *EcoR*I and *Pvu*I linearise the plasmid. Commercial oligonucleotide site-directed mutagenesis was carried out to construct all mutant plasmids encoding for hGST A1-1. The ExSite mutagenesis kit was used to construct D157A, and the QuikChange kit for all other mutant plasmids. A set of primers were required for each kit that were designed according to the guidelines provided by the kits. The designed primers were then synthesised by Integrated DNA Technologies Inc, USA. The following primers were used for creating the S154A pKHA1, were:
5-GAATGTCAG*CCCG**G**G**C*CAGCTTGTTGC and
5-GCAACAAGCTG**G*C**C*CGGG*CTGACATTC.

Primers designed for the construction of D209G pKHA1 were;
5-GCCTCCATG**G**G**T**GAGAAATC*T*C*TAGA*AGAAGC and
5-GCTTCT*TCTA*G*A*GATTTCTC**A**C**C**CATGGGAGGC,

and for D257A pKHA1,
T*C**CCGGG*CT**G**C**A**ATTCATCTGGTG, and
P*CAGTTCATTGCCAACGAG,
where the latter primer was phosphorylated at the 5' end (P*).

The underlined nucleotides are the mutated nucleotides. The nucleotides in bold represent the new codons creating the desired single site directed mutant hGST A1-1. In the case of S154A, the mutation was at position 154; coding for an Ala, where it was originally a Ser. For D209G, the mutation was at position 209; coding for a Gly in place of an Asp. For D157A the mutation was at position 157; coding for an Ala, where it was originally a Asp. The nucleotides in italics represent an introduced unique restriction site, which is a translationally silent mutant sequence.

The mutagenesis reactions consisted of 30 ng of full-length parental pKHA1 plasmid DNA, being the DNA sequence template. The thermal cycler: PCR Express, Hybraid was used for the thermal cycling reactions. The PCR cycling parameters are summarised in Table 3.4. After the PCR reaction, in the case of the QuickChange kit, an additional step was included, including an hour incubation of 10 Units of methylated DNA restriction enzyme (*Dpn*I) at 20°C. Treated PCR product was transformed into super-competent *Escherichia coli* cells (XL1-Blue) by heat shock, which were then grown over-night on agar plates made with 2 x YT-culture media and 100 μg/ml ampicillin. The successfully transformed cells were screened for mutant pKHA1 by restriction analysis and agarose gel electrophoreses. A unique restriction site (*Sma*I) was used to screen for S154A and D157A pKHA1, and for D209G pKHA1 (*Xba*I). The isolated mutant plasmids were then sequenced using the ABI Prism sequencer to ensure that the respective mutations were correct and that no newly introduced mutations existed.

Table 3.1: Summary of the PCR reaction cycles

Desired Mutant	PCR reactions	Melting step	Annealing step	Elongation step
D157A	x 1 cycle	240 sec, 94°C	120 sec, 54°C	120 sec, 72°C
	x 8 cycles	60 sec, 94°C	120 sec, 94°C	120 sec, 72°C
D209G	x 1 cycle	30 sec, 95°C	60 sec, 70°C	
	x 16 cycles	30 sec, 95°C	480 sec, 68°C	
S154A	x 1 cycle	30 sec, 95°C	60 sec, 70°C	
	x 16 cycles	30 sec, 95°C	60 sec, 70°C	

The number of reaction cycles are indicated, with the corresponding temperature and length of time for each step that was performed. One cycle is defined as involving all three steps beginning with the melting step and completing with the elongations step.

3.5 Expression and purification of hGST A1-1

The mutant and parental plasmids were transformed into BL21 (DE3) pLysS *Escherichia coli* cells, which were utilised for the over-expression of hGST A1-1 by induction with IPTG. The preferred cell growth medium was 2 x YT broth for efficient growth. Prior to over-expression of the hGST A1-1 gene, all cell cultures were grown at 280 r.p.m to an optical density to approximately 0.3 units at 600 nm, after which hGST A1-1 was over-expressed with 0.5 mM IPTG. The optimisation of over-expression of soluble mutant protein was investigated by varying the time of induction and growth temperatures, in this case 37 and 20°C were used. Cells were harvested and resuspended in 10 mM phosphate buffer, pH 7.4 and lysed by freeze/thawing with 5 mg of DNA digestive enzyme (*Dnase*I) per initial litre of cell culture. The degree of solubility for each over-expressed mutant protein was assessed, where the soluble and insoluble protein were partitioned by centrifugation and compared by visual inspection by SDS-PAGE (Laemmli, 1970). The expected molecular mass of the subunit of hGST A1-1 was 27 kDa.

Wild-type and mutant forms of hGST A1-1 were purified from the supernatant of lysed cells on a CM-Sepharose column, using a 0 to 0.3 M NaCl salt gradient in 10 mM phosphate buffer, pH 7.4 ,as described by Wallace and Dirr (1999). The protein was stored at 4°C in 20 mM phosphate buffer, pH 6.5, with 100 mM NaCl, 1 mM EDTA, 0.02% sodium azide. The purity of the protein samples were judged by SDS-PAGE and size-exclusion high performance chromatography (SEC-HPLC). The oligomeric molecular mass of dimeric hGST A1-1 was determined by SEC-HPLC, and expected to be about 57 kDa (Wallace and Dirr, 1999). The concentration of the dimeric protein was determined using the molar extinction coefficient of 38 200 $M^{-1}cm^{-1}$ at 280 nm (Perkins, 1986).

3.6 Spectroscopic methods

3.6.1 Far-UV CD spectroscopy

Far-UV circular dichroism (CD) measurements were carried out using a Jasco model 810 CD spectropolarimeter, with a 2 mm path-length cuvette. Far-UV CD spectra (from 200 nm to 250 nm) of protein samples were obtained from an average of 8-12 scans ranging. The final concentration of protein for each scan was 2.5 μM in 5 mM sodium phosphate buffer, pH 6.5 at 20°C. Single ellipticity values were taken at 222 nm and were reported as molar ellipticity ([Θ]) using the equation:

$$[\Theta] = \frac{100\ X\ \Theta}{Cl} \tag{3.3}$$

where [Θ] is the ellipticity in degrees after subtraction of the solvent baseline, C is the molar concentration of protein and l is the path length in cm.

3.6.2 Fluorescence spectroscopy

Fluorescence measurements were performed at 20°C using a Hitachi model 850 fluorescence spectrofluorimeter. Trp fluorescence of hGST A1-1 was monitored by selectively exciting a single tryptophan residue (Trp-21) per subunit dimeric protein using an excitation wavelength of 295 nm and emission wavelengths of 330 nm. Excitation and emission slit-widths of 5 nm was chosen for all fluorescence studies, unless otherwise mentioned.

The anionic dye, 8-aniline-1-napthalene sulfonate (ANS) has previously been reported to bind to hGST A1-1 (Sluis-Cremer *et al.*, 1996). ANS fluorescence was monitored using an excitation wavelength of 390 nm and 480 nm emission. The fluorescence properties of ANS bound to 1 μM hGST A1-1 were monitored in the presence of 100 μM of ANS. The contribution of ANS fluorescence free in solution, uncomplexed with the protein, was subtracted from the ANS-protein

fluorescence signal.

3.7 Active-site tyrosinate formation

The ionisation of the Tyr-9 in the active-site of hGST A1-1 was determined by UV absorption difference spectroscopy measured at 20°C on a Jasco V-550, UV/VIS double-beam spectrophotometer. The final protein concentration ranged from 20 to 50 μM in 0.1 M buffer consisting of sodium acetate, sodium phosphate, glycine. The pH range used was from 5.5 to 9.0. A difference spectra was obtained at each pH, with the use of a reference sample consisting of an equivalent protein concentration at pH 5.5. Spectral peaks at 300 nm, corresponding to tyrosinate of Tyr-9, were recorded and the concentration of tyrosinate ion formation was calculated using $\Delta\ \varepsilon$ of 2350 $M^{-1}cm^{-1}$ (Atkins *et al.*, 1997; Bjornestedt *et al.*, 1995). The pK_a of the Tyr-9 was obtained by fitting the following equation using Origin 5, Microcal Inc software,

$$\Delta A = \Delta A_{max}/[1 + 10^{(pK_a - pH)}] \tag{3.4}$$

where ΔA is the concentration of tyrosinate ion formation per protein subunit (Atkins *et al.*, 1997; Bjornestedt *et al.*, 1995).

3.8 Steady state enzyme kinetics

Steady state enzyme assays were performed by determining the rate of conjugation of CDNB to GSH catalysed by glutathione transferase (Habig and Jakoby, 1981). The rate of the enzyme-catalysed reaction was monitored on a Jasco V-550, UV/VIS double-beam spectrophotometer. The progression of the reaction was monitored at 340 nm, at 20°C for product formation (1-(S-glutathionyl)-2,-4-dintotrobenzene), which has an extinction coefficient of 9600 $M^{-1}cm^{-1}$. All rates were corrected for the spontaneous reaction of GSH and

CDNB. Activity measurements were performed in triplicate and the standard error between assays was less than 10%. A standard assay in order to determine the specific activity of hGST A1-1 consisted of 1 mM CDNB and 1 mM GSH final concentration in 0.1 M phosphate buffer, pH 6.5; to which 3 nM final concentration of enzyme was added. For k_{cat}/K_m^{GSH} determination; 0.05-0.15 mM GSH was used with 1.5 mM CDNB, and for k_{cat}/K_m^{CDNB}; 0.05-0.15 mM CDNB was used in the presence of 5 mM GSH. Parameters used for the determination of K_m^{GSH} was 0.05-6.5 mM GSH with 1.5 mM CDNB, and for K_m^{CDNB}; 0.15-1.5 mM CDNB with 5 mM GSH was used. Saturating curves were fitted to a hyperbolic function of the Michaelis-Menten equation, by non-linear regression analysis using Origin Microcal Inc software.

3.9 Binding studies

3.9.1 Dissociation constants determined for GSH and ANS to hGST A1-1

The binding affinity of GSH and ANS to hGST A1-1 was determined by tryptophan fluorescence quenching and ANS fluorescence enhancement, respectively (Bico *et al.*, 1995; Sluis-Cremer *et al.*, 1996). Fluorescence measurements were obtained on a Hitachi model-850 fluorescence spectrophotometer, using a 5 nm excitation/emission slit-widths. The K_d of ligand binding to hGST A1-1 was determined by titrating aliquots of ligand into the protein sample. For GSH binding, the quenching of the intrinsic Trp-21 fluorescence at 325 nm (excited at 295 nm) is utilised as a probe which reports conformational changes occurring upon ligand binding (Dirr and Wallace, 1999). For ANS binding, the fluorescence enhancement of ANS excited at 390 nm (emission wavelength of 480 nm for hGST A1-1) was monitored. The fluorescence signal of each sample (F_{obs}) was average over a period of 20 s and corrected for the inner filter effect according to (Birdsall *et al.*, 1983):

$$F_{corr} = \mathrm{F}_{obs} \text{ x } 10^{(Aex+Aem)/2} \tag{3.5}$$

where Aex and Aem denote the absorbance of the ligand at the excitation wavelength and emission wavelength respectively. The sum of Aex and Aem were kept below 0.2 to ensure the inner filter effect is kept to a minimum. The quenching data was analysed according to

$$1/F = K_d/(\mathrm{F}_{max}[\mathrm{L}]) + 1/\mathrm{F}_{max} \tag{3.6}$$

where F is F_{corr}, F_{max} the maximum ligand induced change in fluorescence and [L] is the concentration of the ligand after correcting for the dilution effect of titration (the volume was $\leq$ 10% of initial sample volume). The linear regression of $1/\Delta$F versus 1/[L] yields an intercept on the abscissa -$1/K_{\mathrm{d}}$. The same data was fitted according to:

$$F = F_{max} \text{ - } \mathrm{K}_d(\mathrm{F}/[\mathrm{L}]) \tag{3.7}$$

The obtained dissociation constants (K_{d}) from the fitted data using both equations were compared to ensure for accuracy of the fits.

3.9.2 ANS binding measured by isothermal titration calorimetry

Isothermal titration calorimetry (ITC) was used to determine the thermodynamic parameters, binding affinity and stoichiometry of binding, by the direct measurement of the released or absorbed heat of the interactions. The titration experiments were performed on a VP-ITC calorimeter from MicroCal (Northampton, MA, U.S.A) as described previously (Lopez-Hernandez *et al.*,

1997; Brokx *et al.*, 2001). Protein was dialysed, in 3500 Da molecular cut-off dialysis tube, exhaustively against 20 mM phosphate buffer, pH 6.5. The stock solution of 4.1 mM ANS was prepared in a sample of the final dialysed buffer. Titrations were performed by injecting 5 μl of ANS stock into the ITC sample cell containing 0.04 mM D209G hGST A1-1. Heats of dilution were determined by titrating ligand into buffer and the total observed heats of binding were corrected for heats of dilution prior to data analysis. Concentrations of protein and ANS were determined spectrophotometrically, and any effects due to light scattering were corrected for according to Winder and Gent (1971). Raw data were integrated and processed using ORIGIN 5 analysis software (MicroCal, U.S.A). The independent variables for non-linear least-squares fitting of titration curves with ORIGIN 5 are K_{a}, ΔH, and N. Once K_{a} and ΔH were obtained, the free energy and entropy of binding were calculated using the following equations:

$$\Delta G = \Delta H - \mathrm{T}\Delta S \tag{3.8}$$

where T is the absolute temperature.

3.9.3 Molecular docking

Molecular docking programs, Cerius^2 molecular modelling system (Accelrys Inc, San Diego, U.S.A) and LIGIN software (Sobolev *et al.*, 1996), were used to predict the ANS binding site in hGST A1-1 (PDB code: 1K3Y). Molecular modelling with Cerius^2 was performed under the supervision of Fourie Joubert at the Department of Biochemistry, University of Pretoria. The molecular docking program, Cerius^2, treated ANS as a flexible ligand and the receptor as a rigid body. However, steric overlap were allowed between the ligand and residues of the protein surface without penalty during the docking procedure, thus providing a simple means of treating the receptor as a flexible molecule. The program excludes ligands complexed with the protein in the crystal structure during the docking procedure. Various ligand conformations and binding

locations were searched by Cerius^2. The results were ranked according to the conformational energy minima, with the lowest being the most probable binding complex. The output of the results was given in PDB format. A dynamic molecular simulation followed by an energy minimisation was then carried out, using GROMOS (van Gunsteren and Berendsen, 1990) in the software package WHAT IF, ver 2.0 (Vriend, 1990). The output of the results was in PDB format, including the calculated energy minima of the structure.

Details of the docking algorithm implemented in LIGIN is described by Sobolev *et al.* (1996). Surface complementarity between ligand and receptor was the guiding principle for predicting the ANS binding site in hGST A1-1 using LIGIN. In summary, each of these starting positions is subjected to a complementarity optimisation procedure. The docked positions obtained are further refined by optimising the lengths of hydrogen bonds formed with the ligand. The entire protein was searched for binding sites by dividing it into non-overlapping cubes. Then a number of random ligand positions and orientations were generated within each cube so that the number of starting points corresponded to a density of 5 per Å^3 to ensure that global conformational minima was achieved. ANS was docked to hGST A1-1 in the presence of GSH (PDB code: 1GSE excluding ethacrynic acid), glutathione sulphonate (PDB code: 1EV4) and in apo form (PDB code: 1K3Y). The output of the docking results give surface complementarity between ligand and protein, predicted hydrogen bonds and close contacts. The change in accessible surface area of the ligand is also given. These results are also reported in the format of a PDB file, and in a table format listing van der Waals contacts and predicted hydrogen bonds stabilising the ligand-protein complex (Sobolev *et al.*, 1996).

3.10 Thermal denaturation of S154A and wild-type hGST A1-1

Thermal transitions were monitored by far-UV CD at 222 nm described in section 3.6.1. The temperature was controlled with a PTC-348WI peltier-type

cell holder attached to a Jasco model 810 CD spectropolarimeter. Thermal denaturation transitions were carried out by increasing the temperature of the samples from 10 to 80°C, where different heating rates were employed, i.e. 0.5, 0.1 and 2.0°C min^{-1}. The final concentration of protein was 1 μM in 20 mM sodium phosphate buffer, pH 6.5, with 1 mM EDTA, 100 mM NaCl, 0.02% sodium azide. Renaturation profiles were recorded from 80 to 10°C after the denaturation transition had been completed; cooling rates were also controlled through the peltier accessory, at the same unfolding rates.

3.11 Urea-induced unfolding studies

Samples of 1 μM hGST A1-1 in storage buffer, i.e. 20 mM sodium phosphate buffer, pH 6.5, with 1 mM EDTA, 100 mM NaCl, 0.02% sodium azide were incubated in urea (0-8 M). All studies were at 20°C, unless otherwise mentioned. For equilibrium unfolding studies, protein samples were incubated for at least 30 min to ensure equilibrium was achieved prior to analyses using various probes. Urea stock solutions were made fresh since urea solutions form cyanate, both of which can react with the amino groups of proteins (Stark, 1965). The concentration of urea stock solutions were checked by refractometry (Warren and Gordon, 1966).

Establishing the reversibility of the unfolding reaction is essential as it indicates whether the reaction can achieve equilibrium. Protein samples of 6 or 10 μM final concentration, were unfolded in 6 M urea for a minimum of 16 min and up to 6 hours at 20°C. Unfolded protein samples were refolded upon dilution into buffer without urea, giving a final concentration of 1 μM protein in 1 M urea. The protein was allowed to refold for a minimum of 30 min before monitored by the various probes. The degree of refolded native protein was compared to equivalent final concentration of protein.

Changes in the structural conformation of hGST A1-1 were monitored using various spectroscopic probes as well as enzyme activity. The activity of the enzyme was monitored using the general CDNB assay described in section

3.8. The progress curves were linear for the duration of the assay, where the reactivation of unfolded hGST A1-1 was kept less to than 10%. The secondary structural content of hGST A1-1 was monitored by far-UV circular dichroism at 222 nm described in section 3.6.1. The tertiary conformation of hGST A1-1 was monitored by the Trp-21 fluorescence emission (see section 3.6.2) at 330 nm for folded protein and 355 nm for unfolded protein. The Trp-21 emission maximum (λ_{max}) was obtained from fluorescence spectra ranging from 320 to 420 nm. The formation of protein aggregates was monitored by light scattering from fluorescence excitation and emission wavelengths of 340 nm, using slit-widths of 2.5 nm.

3.12 Analysis of equilibrium unfolding data

The equilibrium unfolding parameters were determined by fitting the fraction of unfolded protein using the equation 3.9, to a two-state unfolding transition, according to equation 3.10 previously described for hGST A1-1 (Wallace *et al.*, 1998b). The y_{obs} in equation 3.9 is any measured property of the conformation of hGST A1-1, be it the Trp fluorescence emission (where the fluorescence signal is expressed as a ratio of fluorescence emission at 355 nm and 325 nm), ellipticity, specific activity or ANS fluorescence emission signal. The experimentally observed signal (y_{obs}) was converted to the fraction of the population of protein in the unfolded form f_U, using the relationship:

$$f_U = (y_F - y_{obs})/(y_F - y_U) \tag{3.9}$$

y_F and y_U denote the signal of the probe for the native and unfolded states, respectively. The values for the observed folded (y_F) and unfolded (y_U) protein were obtained by fitting the pre- and post-transition baselines observed in the equilibrium unfolding transition curve to a linear equation, where y_F and y_U are the y-axis ordinates. For the two-state model of dimeric proteins (Neet and Timm, 1994; Bowie and Sauer, 1989), consisting of a monophasic transition

from the native dimer (N_2) to two unfolded monomers (U), the equilibrium constant (K_{eq}) is defined as,

$$K_{eg} = [U]^2/[N_2] = 2[P_t](f_U{}^2)/(1 - f_U) \tag{3.10}$$

where P_t is the monomer concentration of protein. The linear relationship of the Gibbs free energy change calculated from ΔG = -$RT\ln K_{eq}$, from unfolded to folded protein and the urea concentration is as follows (Greene and Pace, 1974);

$$\Delta G = \Delta G(H_2O) \text{ - } m[urea] \tag{3.11}$$

where m is expressed as the dependence of the protein susceptible to the denaturing effects of urea and can be represented as the change in solvent exposure of the protein from the native to unfolded conformation. $\Delta G(H_2O)$ is the extrapolated Gibbs free energy change of the native protein and it unfolded conformation in the absence of denaturant and can be used as an estimation of the conformational stability of the protein.

3.13 Unfolding and refolding kinetics studies

3.13.1 Stopped-flow kinetics

Stopped-flow kinetic studies were performed in order to monitor the refolding and unfolding pathway of hGST A1-1, in the presence of the denaturant (urea). The folding events were monitored using fluorescence detection methods using an Applied Photophysics SX-18MV stopped-flow instrument (UK). Stopped-flow methods allow for the detection of intermediate kinetic events, being within the millisecond time range to as slow as 1000 seconds. The

dead-time of the instrument, being the unobserved reaction occurring within the mixing time of the experiment, was determined according to the methods described for the SX-18MV instrument. In this case the reaction between 6-dichlorophenolindophenol (DCIP) and L-ascorbic acid was used, where a dead-time of approximately 2 ms was consistently determined.

The kinetic experiments designed to monitor the (un)folding reactions were based on the urea-induced equilibrium unfolding transition for hGST A1-1. The folding events were monitored using fluorescence detection methods, where the fluorescence emission of the Tyr and Trp residues of hGST A1-1 were monitored (excitation at 280 nm). The emission intensity was recorded using a 320 nm cut-off filter, and a path-length of 10 mm for the sample. The optimum excitation band-width chosen for all the fluorescence experiments was 2.32 nm (0.5 mm for both entry and exit bandwidth). The chosen slit width resulted in little or no observed Trp or Tyr photodecomposition. The data were recorded at various time intervals depending on which phases of the folding/unfolding events were being studied. All solutions were made up in 20 mM phosphate buffer, pH 6.5, with 100 mM NaCl, 1 mM EDTA, 0.02% sodium azide.

For all refolding reactions, protein samples were unfolded in 6 M urea between 16 to 30 min. For (un)refolding stopped-flow kinetics of hGST A1-1, 6 μM of native or unfolded protein samples were rapidly mixed with varying concentrations of urea (8-5 M) in the sample cell. Asymmetric mixing was used, i.e. two different sized syringes were used resulting in a mixing ratio of 1:5. Upon mixing of the protein samples were diluted 6-fold in varying concentrations of urea to initiate the unfolding or folding reaction of the protein. Prior to the refolding reactions, the protein samples were incubated in 6 M urea for a minimum of 16 min where equilibrium of unfolded protein is achieved. The dependence of apparent rate constant of unfolding with increasing final protein concentration was investigated. The protein in 6 M urea were diluted 6-fold to a final protein concentration range of 0.5 to 5 μM into non-denaturing buffer. For both unfolding and refolding kinetic experiments, Rayleigh scatter was monitored for the introduction of any mixing artifacts, protein aggregation or crystallisation of urea during the experiments. This was done by having repeated the

same experiments as for the actual experiments, where the Rayleigh scatter was monitored using an excitation wavelength at 340 nm, and the emission fluorescence was recorded using a 320 nm cut-off filter.

3.13.2 Manual mixing kinetics

The manual mixing experiments for the kinetics of (un)refolding of hGST A1-1 was recorded using a Perkin Elmer, Luminescence spectrometer, model LS50B. Upon the addition of protein to buffer solution, the samples were manually mixed with a magnetic stirrer. The time zero for the manual mixing correlated to the point upon addition of the protein sample to the buffer. The fluorescence emission of Tyr and Trp residues of hGST A1-1 were used to monitor the refolding events (280 nm excitation). The excitation and emission bandwidths were set at 2.5 nm, which was optimised for best signal-to-noise ratio without introducing photodecomposition. The folding events were recorded over the required period of time at 330 nm, being 1000 sec to 2500 sec.

Kinetics of unfolding of 1 μM S154A in 3 M was carried out at 20 and 10°C. The unfolding reactions were monitored using various probes mentioned in section 3.11. For all stopped-flow and manual mixing experiments, the temperature of the reactions were regulated within 0.1°C of the required temperature using a thermostated water bath connected to the instruments.

3.13.3 Kinetic data analysis

All kinetic data obtained from the stopped-flow instrument were 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 8). For each point represented in the chevron (where the observed rate of the reaction phase is plotted against final urea concentration) entailed an average of at least two

separate experiments, where for each experiment at least 3 sample runs were performed and averaged. The standard deviation represented in the chevron was derived from the separate experiments.

The observed kinetic rates (k_{obs}) for each kinetic phase represented in the chevron plot were derived by the best fit according to the sum of the exponentials, as follows,

$$A_t = \sum [A_i + e^{-k_i t}] + A_o \tag{3.12}$$

where t is the time in seconds, A_t is the total fluorescence amplitude change of the reaction signal, A_o is the final end-point value of the reaction and A_i is the amplitude change for the kinetic phase i with the observed rate, k_i. The apparent rate constant for a reaction phase can be expressed as a time constant (τ) which is the inverse of the k_{obs} i.e. k_i. The accuracy and quality of the derived fitted function for the experimental data were judged according to the residuals of the fit. The folding phases for hGST A1-1 have been previously reported for both the refolding and unfolding pathway of hGST A1-1 (Wallace *et al.*, 1998b,a). The derived folding phases and the proposed folding model for hGST A1-1 have been discussed in section 2.3.3.

The dependence of urea for each kinetic phase for hGST A1-1 is represented according to the equation:

$$\log k_{\text{obs}} = \log k_{\text{obs}}(\text{H}_2\text{O}) + \text{m}[urea] \tag{3.13}$$

where k_{obs} is the apparent rate constant for kinetic phase at different urea concentrations, $k_{\text{obs}}(\text{H}_2\text{O})$ is the apparent rate in the absence of denaturant and m is the change in solvent accessibility of the transition state for that particular phase.

Chapter 4

Significance of Asp N-capping motif in α-helix 9 stability

4.1 α-Helix 9 peptide studies

It is well documented that an N-cap motif is a stabilising feature for α-helices, as discussed in section 1.2.2. For this work, the role of the N-cap of α-helix 9 was investigated by replacing Asp residue with a Gly and thus removing the stabilising N-cap hydrogen bonds formed by the Asp-209 side-chain. The amino acid substitution was carried out both in context of the protein, hGST A1-1, and in a peptide corresponding to the sequence of the C-terminal region (residues 208-222). Since the disruption in the stability of α-helix 9 is desired, a Gly substitution is preferred rather than an Ala, because (1) Gly based peptides have a lower tendency to form helices compared to Ala based peptides (Creamer and Rose, 1994; Chakrabartty *et al.*, 1994), and (2) Gly is a strong helix breaker (Rohl *et al.*, 1996; Aurora *et al.*, 1994).

4.1.1 Predicted α-helix 9 content from the helix-coil transition algorithm, AGADIR

The helix-coil transition algorithm AGADIR (Lacroix *et al.*, 1998) is a useful means to predict the helical content of a particular amino acid sequence. For this study AGADIR was used to derive the amino acid propensities of residues corresponding to the C-terminal helix (α-helix 9) of hGST A1-1 (208 - 222). The amino acid propensity for Asp-209 (25%) and Ser-212 (7%) are relatively high in comparison to the amino acid propensities of the other residues in α-helix 9, which ranges between 1-5%, Table 4.1. These results highlight the importance of the N-cap Asp-209 and Ser-212 at N3 position of α-helix 9, which together form an N-capping motif. Moreover, the greater propensity of Asp suggests a predominant role of this residue toward the stability of the helix compared to Ser-212. This is likely to be the case, since the side-chain hydroxyl group of Ser-212 does not form any stabilising hydrogen bonding patterns, whereas Asp-209 residue forms a class 3 Asp N-cap motif, as discussed in section 2.2.1. In addition, the results from AGADIR highlight the presence of a staple motif, involving Met-208 (N') and Leu-212 (N4) which flank the N-cap motif. The sequence alignment, shown in Table 4.1, shows that the staple motif is a conserved structural feature for class Alpha GSTs. The predicted overall helical content for wild-type α-helix 9 is 20% (Table 4.1). Upon substituting the Asp with a Gly, the predicted helical content drops to 8%. With this, the results from AGADIR predict a 2.4 loss in helical content upon the removal of the Asp N-cap motif.

4.1.2 Far-UV CD spectroscopic studies on α-helix 9 peptides

The helical content of WT-pep and DG-pep peptides were investigated by far-UV CD. The design of the peptides is described in section 3.3.1. The amino acid sequence of WT-pep corresponds to the C-terminal region (208-222) of the protein hGST A1-1, whereas the sequence of DG-pep, that is derived

Table 4.1: Amino acid propensity scale of the C-terminal region of hGST A1-1 as predicted by AGADIR (Lacroix *et al.*, 1998)

	Wild-type				Asp-209 to Gly substitution			
Residue	aa	Helix	N-cap	H-staple	aa	Helix	N-cap	H-staple
208	M	0.6	0.18	27.17	M	0.7	0.07	4.11
209	**D**	0.8	27.17	0.00	**G**	0.7	4.11	0.00
210	E	27.9	0.83	0.00	E	4.8	0.69	0.00
211	K	28.7	7.75	0.00	K	5.5	0.64	0.00
212	S	29.3	10.29	0.00	S	6.1	10.29	0.00
213	L	37.0	0.73	1.11	L	16.4	0.97	1.48
214	E	37.4	1.11	0.00	E	17.3	1.48	0.00
215	E	37.3	0.33	0.00	E	18.5	0.43	0.00
216	A	36.5	0.02	0.00	A	18.6	0.03	0.00
217	R	31.3	0.00	0.00	R	16.4	0.00	0.00
218	K	21.6	0.00	0.00	K	11.7	0.00	0.00
219	I	21.2	0.00	0.00	I	11.4	0.00	0.00
220	F	7.3	0.00	0.00	F	4.0	0.00	0.00
222	R	2.2	0.00	0.00	R	1.1	0.00	0.00
223	F	0.0	0.00	0.00	F	0.0	0.00	0.00
Helical content for sequence: 19.97%					Helical content for sequence: 8.34%			

The output of the data is given as percent amino acid propensity (helix) for each amino acid (aa), predicting the contribution of each amino acid towards the helical content of the sequence corresponding to the C-terminal residues 208 - 220 (residue) in hGST A1-1. The predicted amino acid propensities and helical content for the sequence of the C-terminal region, where the Asp-209 was substituted with a Gly, is also shown. The conditions set were, pH 6.5, temperature at 293 K and an ionic strength of 0 M, i.e. water as the solvent. The predicted helical content for the sequence is given as percent helical content. The amino acid Asp-209 and Ser-212 were predicted to have high N-cap propensities (N-cap). Met-210 and Leu-213 are predicted to form a hydrophobic staple motif (H-staple).

from WT-pep, includes a replacement of an Asp with a Gly, corresponding to position 209. Figure 4.1 shows far-UV CD spectra of both wild-type (WT-pep) and mutant (DG-pep) peptides in water. The spectra of these two peptides had an ellipticity minima at 222 nm and 201 nm and ellipticity maxima at 218 nm and 194 nm. The 222 nm minima and 194 nm maxima is indicative of a helical conformation, and the 201 nm minima and a small trough at 225 nm is characteristic of a random-coil. The findings of a linear dependence with increased helicity with increasing peptide concentrations, i.e. the fractional helical content is independent of peptide concentration, is indicative that the peptides were monomeric. Thus, it is evident that both WT-pep and DG-pep are a mixture of helical and random coil conformations.

Results in Figure 4.1 show that the far-UV spectrum of WT-pep shared the same wavelength maximum and minimum as compared to DG-pep, showing only a shift in the degree of ellipticity. The difference in the two spectra can be interpreted as a result of the changes in the ratio between helical and random coil conformations. The helical content for WT-pep and DG-pep can be derived from the intensity of the far-UV CD signal at 222 nm (section 3.3.2), with the above assumption that the peptides undergo a two-state helix-coil transition. The lower residual ellipticity of DG-pep at 222 nm compared to WT-pep indicate a lower helical content for this peptide. The calculated fraction of helical content for WT-pep and DG-pep in water (Figure 4.1) was 10% and 4%, respectively. This is a 2.5-fold decrease in helical content of DG-pep compared to WT-pep. The experimental value of the helical content of WT-pep in water was 2-fold less than what was predicted by AGADIR (see section 4.1.1). Nevertheless, the contribution of the N-cap residue in α-helix 9 peptide was correctly predicted by AGADIR.

Changes in the intensity of the far-UV CD signal at 222 nm of helical peptides is also dependent upon the length of helical formation, where increasing length of helices along a peptide chain enhances the CD signal (Gans *et al.*, 1991). Therefore, not only can the intensity in signal be interpreted in terms of helical content but also the extent to which helices tend to fray at their N- or C-terminal ends. It is well documented that capping motifs increase helical propensity by means of initialising and/or stabilising the helical propagating

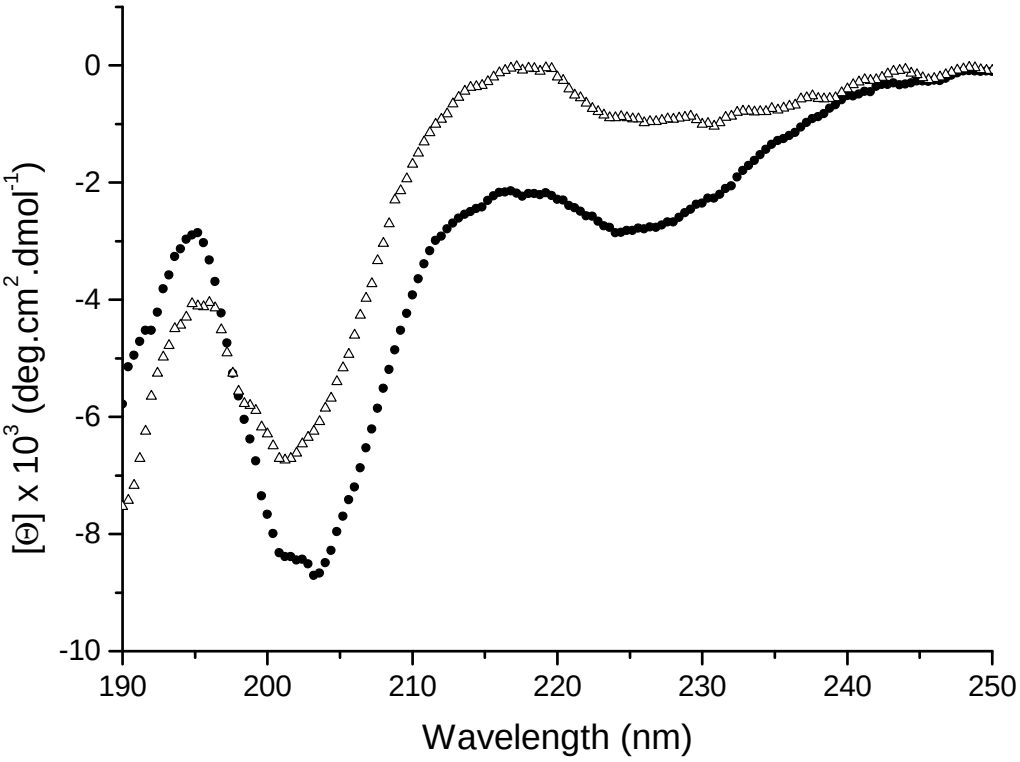

Figure 4.1: Far-UV CD spectra of DG-pep (△) and WT-pep (•). The samples consisted of 70 μM of peptide in water.

tendencies at their ends (Richardson and Richardson, 1988; Doig *et al.*, 1997; Dasgupta and Bell, 1993). Thus, the reduced CD signal of DG-pep could also be interpreted in the lines of increased fraying at the N-terminal end, which in turn decreases the helical length of the mutant peptide compared to that of WT-pep. For either scenario, from the CD spectra of WT-pep and DG-pep in water, DG-pep has a lower helical propensity compared to its wild-type sequence.

TFE, as a co-solvent, is known to enhance the helical content of peptides with intrinsic helical properties (Rohl *et al.*, 1996; Kumaran and Roy, 1999; Myers *et al.*, 1998). Figure 4.2 shows far-UV CD spectra of WT-pep and DG-pep with increasing concentrations of TFE. The TFE dependence of the CD spectra for each peptide exhibits an isodichroic point at 202-204 (Figure 4.2). This isodichrioc point is indicative of a two-state helix-coil transition (Brown and Klee, 1971). The inset in Figure 4.2 illustrates the calculated helical content for each peptide. A plateauing effect is evident for both WT-pep and DG-pep with 40% to 70% (v/v) TFE, where the maximum helical content for both peptides is approximately 35%. It is a well known feature that a variety of peptides reach their maximum helical content in 40% (v/v) TFE (Kumaran and Roy, 1999; Rohl *et al.*, 1996). In this case, the convergence of helical content of DG-pep and WT-pep, in approximately 30% TFE, can be explained in terms of the non-linear increase in helical propensities among amino acids in water up to 40% TFE (Rohl *et al.*, 1996; Myers *et al.*, 1998). More specifically, the amino acid propensity of Asp at the N-cap position in water decreases with 40% TFE, whereas Gly remains as a strong helix breaker both in water and TFE (Rohl *et al.*, 1996). Thus, the equivalent helical content of WT-pep and DG-pep in 35% TFE, can be explained in terms of this convergence of the amino acid propensity of Asp and Gly at the N-cap position (Rohl *et al.*, 1996).

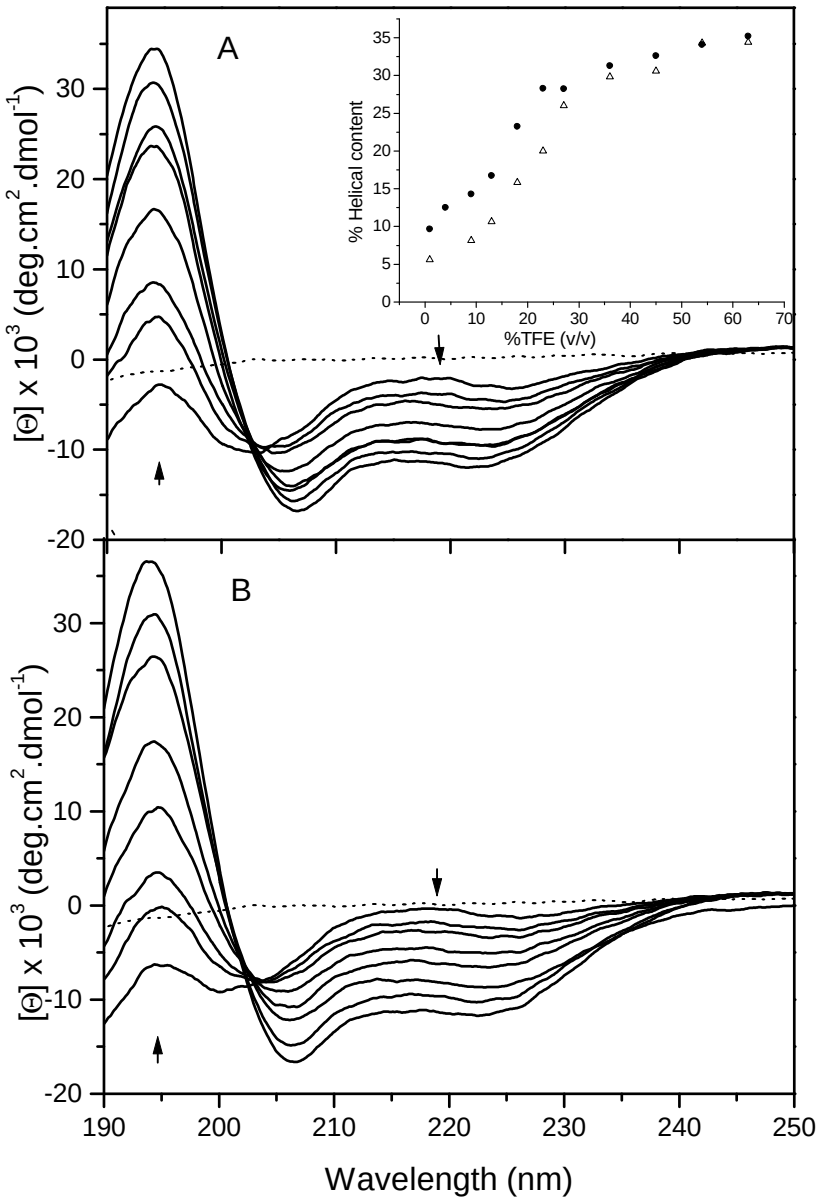

Figure 4.2: TFE concentration-dependence of the far-UV CD spectra of WT-pep (A) and DG-pep (B). The spectra are represented in the order of increasing concentrations of TFE depicted by the arrow; 0, 10, 13, 18, 23, 27, 45, and 63% (v/v). The dotted line is the buffer spectrum. The inset shows the dependence of calculated percent helical content of DG-pep (△) and WT-pep (•) with increasing concentrations of TFE. Peptide concentration is 50 μM in 5 mM sodium phosphate buffer, pH 6.5.

4.2 Construction and purification of D209G

The substitution of the residue Asp-209 with a Gly in hGST A1-1 by site-directed mutagenesis (section 3.4) proved successful. DNA sequencing confirmed the desired mutation (Figure 4.3). As judged by SDS-PAGE (Figure 4.5) and size-exclusion HPLC (Figure 4.4) the purified D209G hGST A1-1 had the expected molecular masses of 26 kDa (for subunit) and 55 kDa (for the dimeric quaternary conformer).

4.3 Comparison of native D209G and wild-type

Figure 4.6 shows far-UV CD spectra of D209G and wild-type hGST A1-1. The spectra for both proteins have two ellipticity minima at 222 nm and 210 nm, characteristic of a predominantly helical protein. The CD spectra of D209G and wild-type are coincident, indicating that no changes occurred in the overall secondary structure of the mutant. In fact, the deletion of the C-terminal region results in far-UV spectra which is identical compared to that of the wild-type. Far-UV CD is an insensitive probe in detecting the α-helix 9 (Dirr and Wallace, 1999), where it constitutes only a small percentage of the ellipticity signal at 222 nm in a predominantly helical protein.

Trp-21 is an intrinsic fluorescent probe, used to report global conformational changes in hGST A1-1 (Dirr and Wallace, 1999; Nieslanik *et al.*, 2001). hGST A1-1 has only one Trp per subunit (Trp-21) that is located at the interface between domain 1 and 2 of each subunit. Figure 4.7 shows the fluorescence emission spectra of Trp-21 in D209G and wild-type hGST A1-1. The fluorescence intensity of D209G is about 91% that of wild-type. Changes in Trp-21 fluorescence can be related to conformational differences of α-helix 9, where the docking of the α-helix over the active-site results in a quenching effect on Trp-21 (Dirr and Wallace, 1999). Upon removing α-helix 9, an increase in Trp-21 fluorescence intensity accompanied by a red-shift in emission maximum

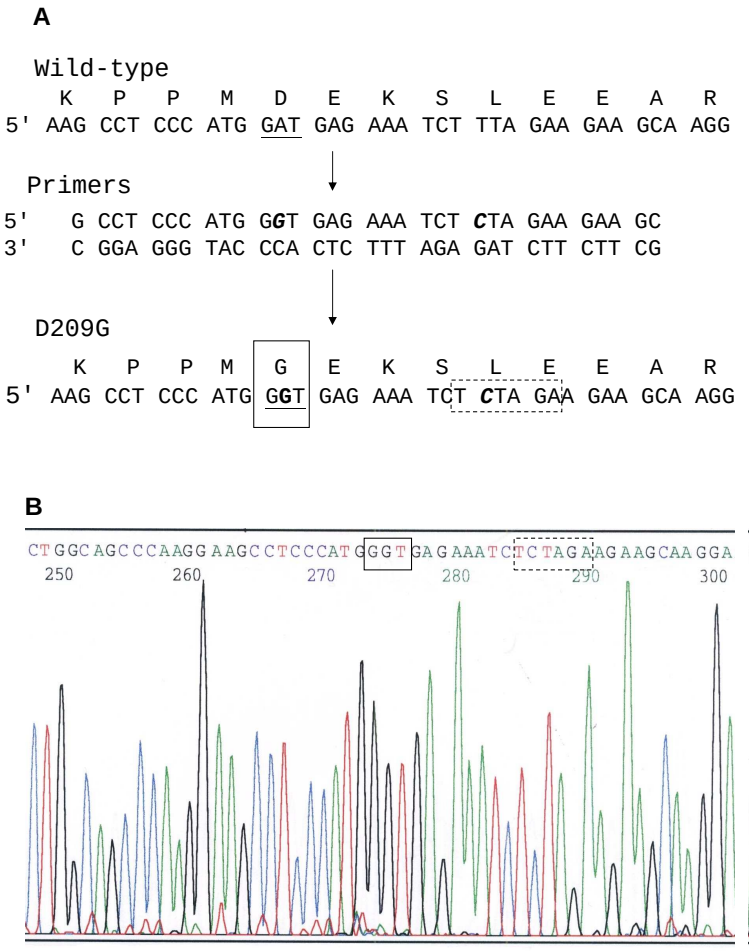

Figure 4.3: (A) A region of wild-type nucleotide sequence indicating the GAT codon for Asp-209 in hGST A1-1 (underlined) that was selected for mutagenesis. These changes in codons were designed into the primers (changed nucleotides are highlighted in bold and italics). The codon change for D209G mutation is GAT to GGT, where GGT codes for an Gly. The codon TTA in the wild-type sequence corresponds to the codon to CTA in the primers. This was a translationally silent mutation, incorporating a unique restriction site into the mutated nucleotide sequence. (B) A portion of the nucleotide sequence obtained from DNA sequencing reaction of the D209G mutant plasmid DNA. The nucleotide sequence demonstrates that an Ala was encoded for at position 209 and a unique restriction site, *Xba*I, was incorporated (highlighted in squares).

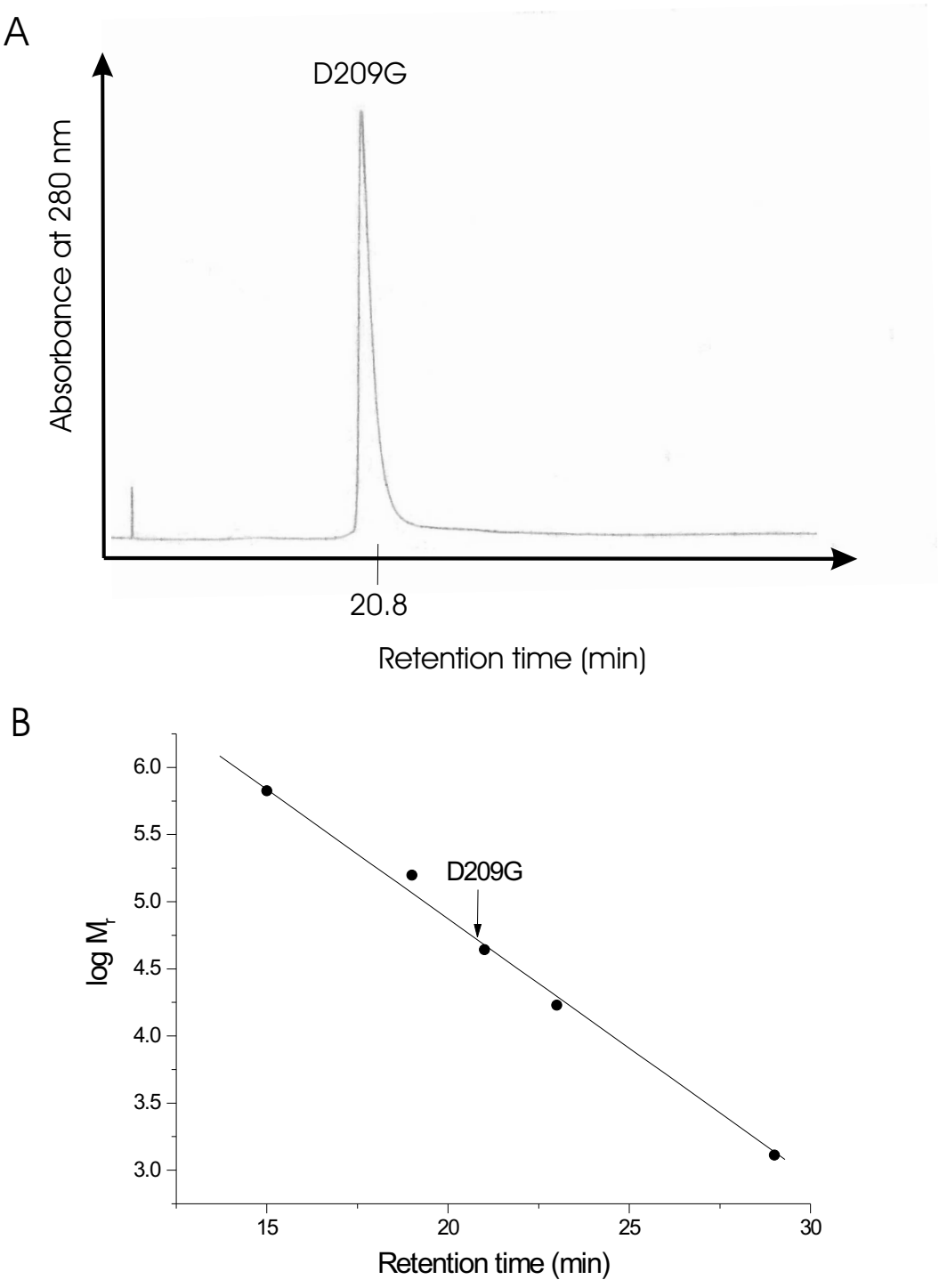

Figure 4.4: (A) SEC-HPLC elution profile of D209G hGST A1-1. (B) The calibration curve indicated the dimeric molecular mass of 55 kDa for D209G.

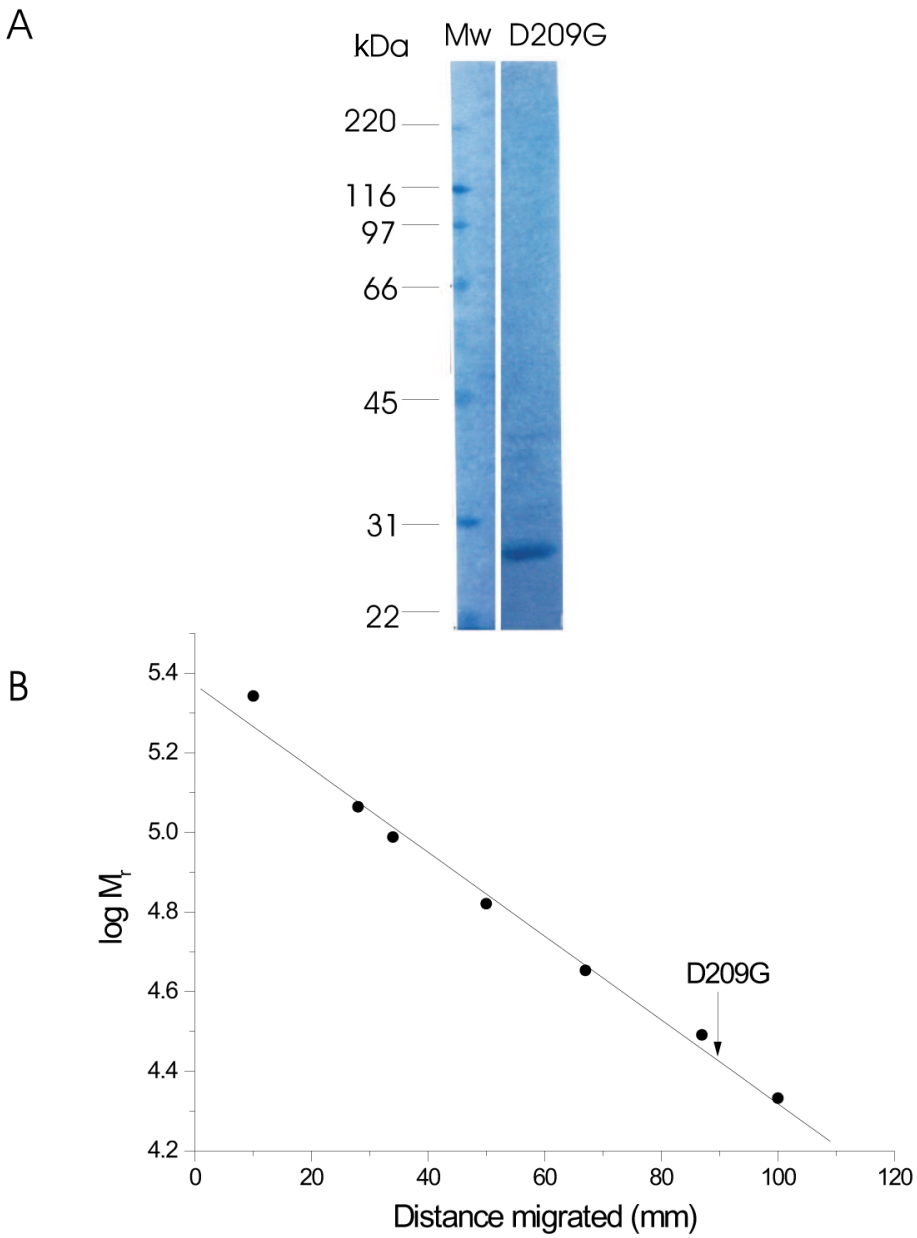

Figure 4.5: (A) SDS-15% PAGE of purified D209G. MW indicates the lane with molecular weight markers, and D209G consisting of purified D209G. (B) Calibration curve constructed using the marker proteins indicated the monomeric mass of D209G to be 26 kDa.

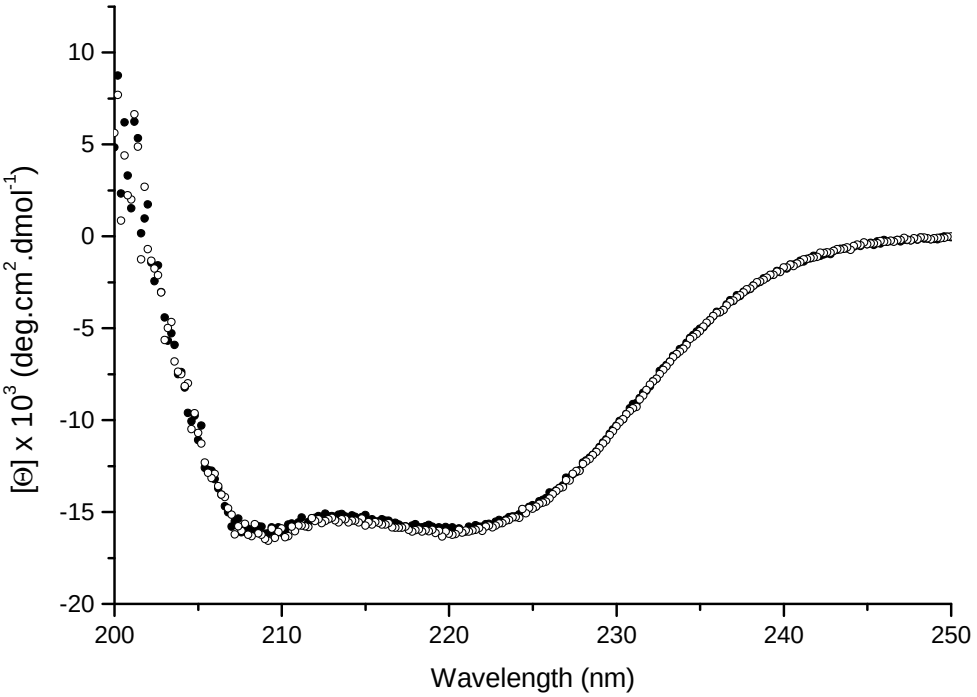

Figure 4.6: Far-UV CD spectra of D209G (△) and wild-type (•) hGST A1-1. Samples measured consisted of 2.5 μM protein in 5 mM phosphate buffer, pH 6.5.

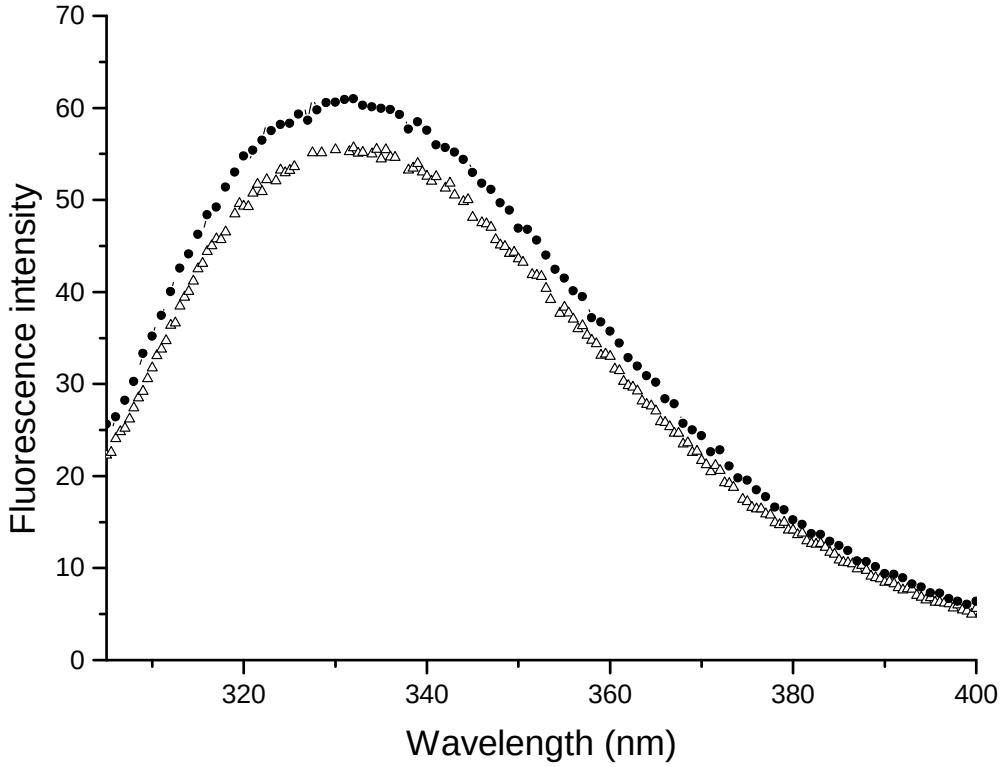

Figure 4.7: Intrinsic fluorescence emission spectra of D209G (△) and wild-type (•) hGST A1-1. Trp-21 of hGST A1-1 was selectively excited at 295 nm. The fluorescence emission maximum for both forms is 330 ± 2 nm. Samples consisted of 1 μM of protein in 20 mM phosphate buffer, pH 6.5, with 100 mM NaCl, 1 mM EDTA, 0.02% sodium azide.

has been reported (Dirr and Wallace, 1999). The insignificant lower Trp-21 fluorescence of D209G compared to its wild-type form, suggests only a minor conformational difference in the tertiary packing of the proteins. Therefore, in the absence of the N-cap motif of α-helix 9, the conformation of the C-terminal region appears not to be significantly different compared to that of the wild-type.

Unlike far-UV and Trp-21 fluorescence, the enzyme's activity is a sensitive probe in monitoring any significant perturbations in the conformation and dynamics of α-helix 9 (Dirr and Wallace, 1999). This is possible as the α-helix 9 forms part of the active-site of hGST A1-1 (Board and Mannervik, 1991; Sinning *et al.*, 1993). The specific activity, with CDNB as substrate, of wild-type was 47 $\pm$ 6 mmol.min^{-1}.mg^{-1} and that of the mutant form was 54 $\pm$ 8 mmol.min^{-1}.mg^{-1}. The similar specific activities of both proteins indicates that the active-site conformation of D209G and wild-type were comparable. From these results it can be concluded that there is no significant differences in the overall tertiary packing of the protein and the active-site conformation (including the α-helix 9).

4.4 Conformational stability of D209G and wild-type

Equilibrium unfolding studies with hGST A1-1 were carried out in the absence and presence of 10 mM of GSH. Figure 4.8 shows the urea-induced unfolding curves of D209G and wild-type hGST A1-1 monitored by far-UV CD, Trp-21 fluorescence and enzyme activity. The unfolding transitions for both proteins displayed a single sigmoidal transition between native and unfolded conformers. A comparison of the unfolding transitions for wild-type and D209G, monitored by ellipticity at 222 nm and fluorescence spectroscopy, showed that both unfolding transitions are coincident. The unfolding thermodynamic parameters of hGST A1-1 were obtained from the unfolding transition curves that were monitored by ellipticity at 222 nm and fluorescence.

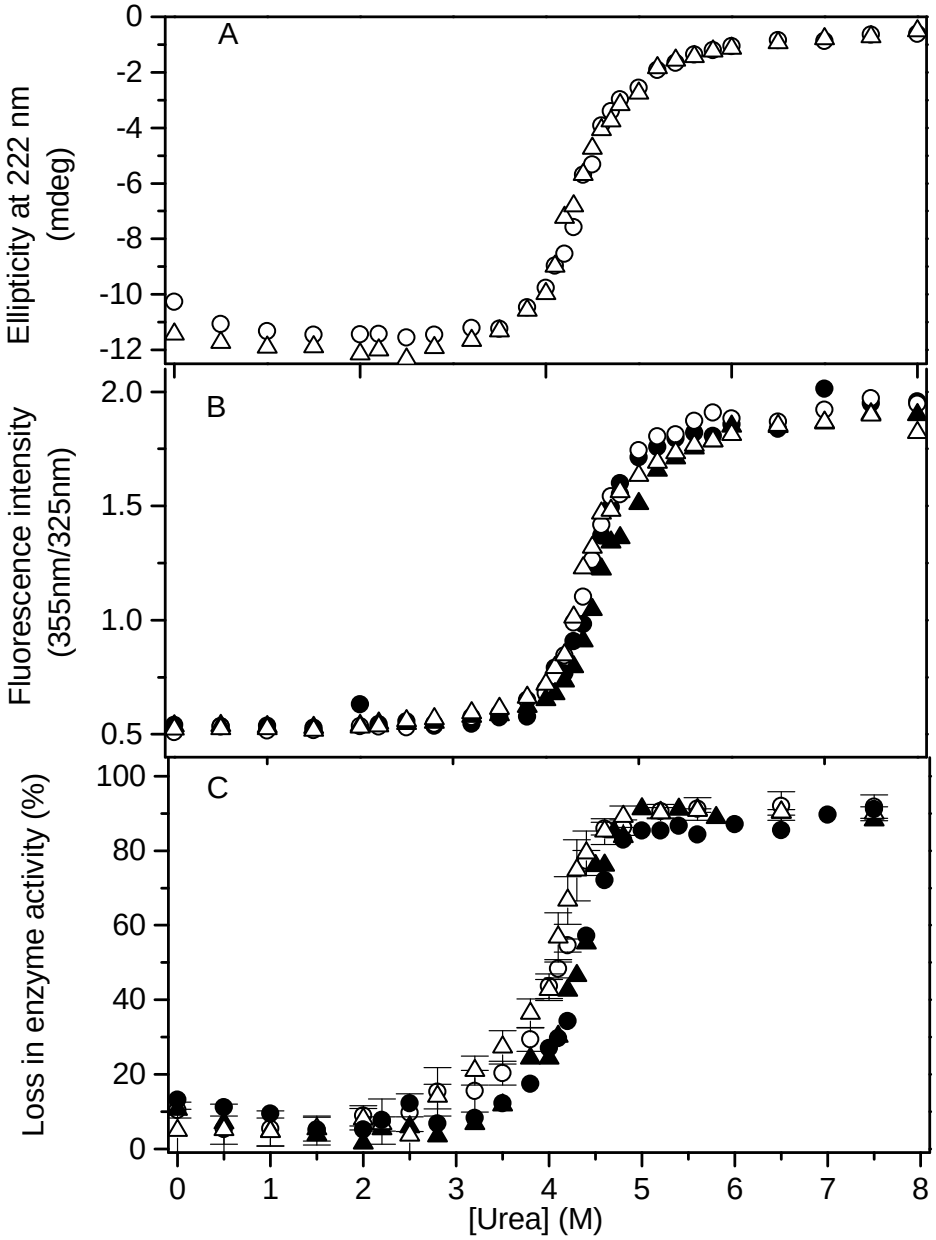

Figure 4.8: Urea-induced unfolding curves wild-type (circles) and DG209 mutant (triangles) hGST A1-1, monitored by: (A) circular dichroism at 222 nm, (B) Trp-21 fluorescence and (C) loss in enzyme activity. CDNB and GSH were the substrates used to measure the activity of hGST A1-1. Samples consisted of 1 μM protein in the absence (open symbols) and presence (closed symbols) of 10 mM GSH, in 20 mM sodium phosphate buffer, pH 6.5, 1 mM EDTA, 0.1 M NaCl, 0.02% sodium azide.

The free energy ($\Delta G(H_2O)$) for the unfolding of wild-type was 112 ± 4 kJ.mol^{-1} and 117 ± 4 kJ.mol^{-1} for D209G. The m-values, assuming a two-state model, was 18 ± 0.4 kJ.mol^{-1}.M^{-1} urea for wild-type and 19.2 ± 0.4 kJ.mol^{-1}.M^{-1} urea for D209G. The mid-point of the unfolding transition for wild-type and D209G was 4.5 ± 0.1 M. The thermodynamic parameters of hGST A1-1, from the fluorescent equilibrium unfolding studies in the presence and absence of GSH, were identical for both wild-type and D209G. The similar thermodynamic stability of D209G to that of the wild-type hGST A1-1 demonstrate that the mutation at position 209 in the C-terminal region of D209G did not affect the stability of the protein. This was to be expected since the C-terminal region does not contribute to the overall stability of the protein (Dirr and Wallace, 1999). Results in Figure 4.8 C demonstrated that, unlike for far-UV and Trp-21 fluorescence studies, the mid-point of the unfolding transition observed by enzyme activity was 4.2 M urea in the absence of 10 mM GSH. Moreover, this mid-point shifted to 4.5 M urea in the presence of GSH (Figure 4.8 C). Previous studies, have shown that a significant loss in activity of wild-type hGST A1-1, in the pre-transition region of the unfolding curve in the absence of ligand, is directly related to the destabilisation of α-helix 9 (Dirr and Wallace, 1999; Wallace *et al.*, 1998b; Mosebi *et al.*, 2003). Important work by Zhan and Rule (2004) used NMR to characterise the structure and dynamics of the C-terminal region of hGST A1-1. In their work the C-terminal region was shown to be more structured with GSH bound to the G-site compared to that of the apo form. Thus GSH bound to this protein contributes to an increased stability of the helical conformation at the C-terminal region. In comparing the catalytic activity, as derived from unfolding transition curves of D209G and wild-type in apo form (Figure 4.8 C), the activity of the mutant (D209G) is slightly less between 3 to 4 M urea in comparison to that of the wild-type. However, in the presence of GSH, the unfolding transitions of both proteins were the same. One would expect a loss in enzyme activity at low concentrations of urea in the case of a destabilised α-helix 9 (Mosebi *et al.*, 2003). Thus, in the absence of GSH, the lower activity of D209G compared to that of the wild-type suggests that the stability of α-helix 9 in D209G is marginally lower in comparison to that of the wild-type. Thus, the Asp N-cap residue appears to play some favourable role towards the helical stability of

α-helix 9 in hGST A1-1.

4.5 GSH binding and steady state kinetics of D209G and wild-type

The fluorescent binding assays and steady state kinetics permitted the characterisation of the structural fidelity of the active-site in hGST A1-1. This includes elucidating the structural conformation of α-helix 9. The binding affinity of GSH to hGST A1-1 was measured by fluorescence quenching methods (the dissociation constant, K_{d}). The K_{d} of GSH to D209G was 0.18 ± 0.01 mM, and 0.16 ± 0.01 mM to wild-type. Since the binding affinity of GSH is independent of the α-helix 9, the equivalent K_{d} value for both proteins was expected (Dirr and Wallace, 1999; Allardyce *et al.*, 1999; Nieslanik *et al.*, 1999b). The steady-state kinetic parameters for wild-type and D209G are summarised in Table 4.2, and are in agreement with work published by other groups (Dirr and Wallace, 1999; Allardyce *et al.*, 1999; Board and Mannervik, 1991; Gustafsson *et al.*, 1999; Widersten *et al.*, 1994). D209G has approximately 1.2 and 1.4-fold lower $k_{\mathrm{cat}}/K_{\mathrm{m}}$ values relative to those obtained for the wild-type for CDNB and GSH, respectively. The $K_{\mathrm{m}}^{\mathrm{CDNB}}$ and $K_{\mathrm{m}}^{\mathrm{GSH}}$ values were 1.2 and 1.4-fold higher, respectively, for D209G compared to those obtained for the wild-type (Table 4.2). The marginal differences in the steady-state kinetics of the N-cap mutant (D209G) suggests that only modest changes have occurred at the active-site of D209G compared to that of the wild-type (Table 4.2).

The slight enhanced K_{m} values for CDNB and GSH of D209G and reduced $k_{\mathrm{cat}}/K_{\mathrm{m}}$ values, suggests that the active-site environment of D209G is less favourable for the optimum binding of substrates when compared to that of the wild-type enzyme. Since the binding affinity of GSH was the same for wild-type and D209G, as mentioned previously, it seems likely that the binding site of GSH for D209G remains intact, i.e. it is the same as for the wild-type. Since the N-cap residue does not form part of the active-site of hGST A1-1,

Table 4.2: Steady-state kinetic parameters of D209G and wild-type hGST A1-1

		D209G	Wild-type
K_{m} (mM)	GSH	0.29 ± 0.02	0.21 ± 0.01
	CDNB	0.35 ± 0.06	0.31 ± 0.08
$k_{\mathrm{cat}}/K_{\mathrm{m}}$ ($\mathrm{mM}^{-1}.\mathrm{s}^{-1}$)	GSH	115 ± 17	161 ± 4
	CDNB	73 ± 2	86 ± 8

the replacement of the Asp-209 with Gly, which contributes to a slightly less favourable active-site environment, is most likely attributed to conformational changes occurring at the C-terminal region. Work by Nieslanik *et al.* (1999a) have reported that the rate-limiting step of the CDNB conjugating reaction is not product release controlled as for ethacrynic acid conjugation reaction which involves the dynamics of α-helix 9. However, Allardyce *et al.* (1999) have demonstrated that the docking of α-helix 9 over the active-site plays an essential role in facilitating the correct and productive binding mode of CDNB, at the H-site, for its conjugation to GSH. Previous studies have shown that by significantly increasing the dynamics of α-helix 9, where residues at position Phe-220 and Met-208 were substituted for Ala, the catalytic efficiency of the CDNB conjugating reaction was significantly decreased (Gustafsson *et al.*, 1999; Widersten *et al.*, 1994; Ibarra *et al.*, 2001). Hence, the relationship of a reduced $K_{\mathrm{m}}^{\mathrm{CDNB}}$ and enhanced $k_{\mathrm{cat}}/K_{\mathrm{m}}$ of GST A1-1, is evidence for an increase in α-helix 9 mobility and/or a disruption at the H-site. Thus the slight reduction in the catalytic efficiency observed for D209G is most likely due to a more solvent exposed H-site attributed to marginal conformational differences at the C-terminal region of D209G compared to that of the wild-type.

4.5.1 Active-site tyrosinate formation

The spectrophotometric determined pK_{a} value of the catalytic Tyr-9 for wild-type and D209G were determined from the data shown in Figure 4.9. The Tyr-9 of hGST A1-1 has an unusually low pK_{a} that distinguishes it spectroscopically from the other tyrosine residues in the protein (Bjornestedt *et al.*, 1995). The pK_{a} of the active-site Tyr-9 determined for D209G and wild-type hGST A1-1 was 8.2 $\pm$ 0.2 (Figure 4.9). These values are in accordance with those reported in the literature (Bjornestedt *et al.*, 1995; Gustafsson *et al.*, 1999). A lower Tyr-9 pK_{a} in GST A1-1 has been observed upon replacing Met-208 or Phe-220 with Ala, which was shown to be in relation the dislocation of the C-terminal region of GST A1-1, leading to an increased solvent exposed active-site (Gustafsson *et al.*, 1999; Atkins *et al.*, 1997). Contrary to this, in the presence of a more destabilised C-terminal region, monitored by

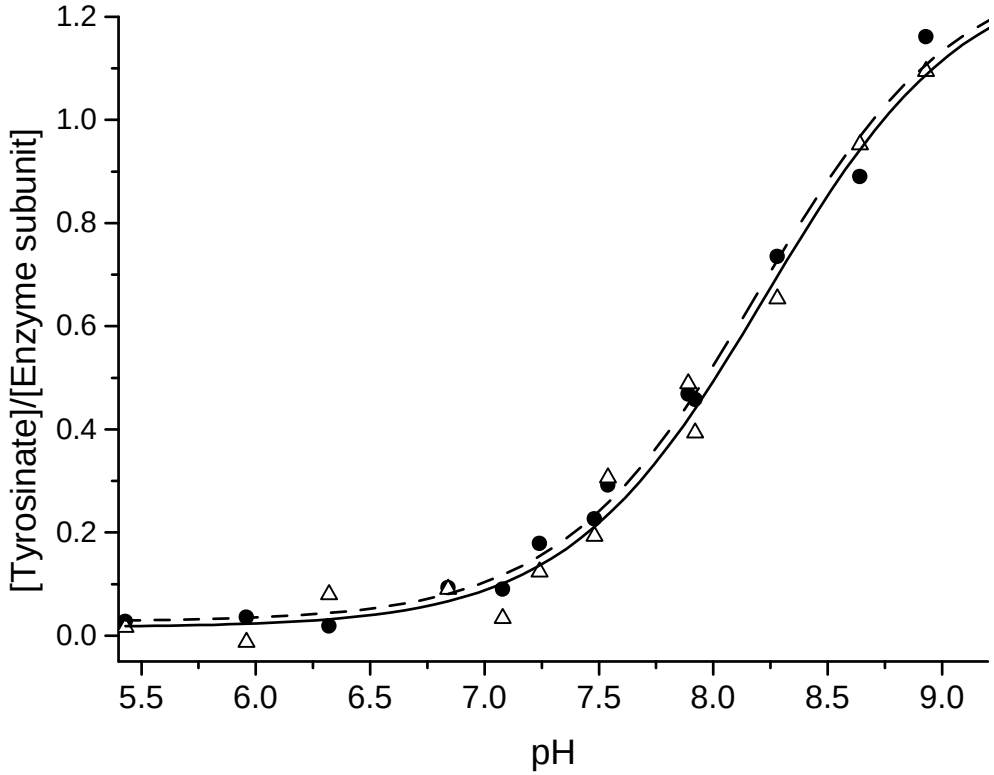

Figure 4.9: Titration of the tyrosinate formation in the active-site of wild-type (•) and D209G (△) hGST A1-1. The number of tyrosinates per subunit is plotted as a function of pH. The pK_{a} values were obtained by fitting $\Delta A = \Delta A_{max}/[1 + 10^{(pK_a - pH)}]$.

the activity of the enzyme, the pK_a of this active-site tyrosine was maintained (Mosebi *et al.*, 2003). Thus, the equivalent Tyr-9 pK_a of D209G and wild-type suggests that the dynamics of the C-terminal region in D209G, in the absence of active-site ligands, is the same compared to that of the wild-type.

4.6 ANS binding to hGST A1-1

4.6.1 ANS fluorescence properties

The fluorescence properties of the anionic dye, ANS, bound to hGST A1-1, as well as its binding affinity has been used to report the conformation of the active-site and the C-terminal region of hGST A1-1 (Dirr and Wallace, 1999; Allardyce *et al.*, 1999; Mosebi *et al.*, 2003). The binding affinity, i.e. the dissociation constants (K_d) of ANS to hGST A1-1 was determined by fluorescent enhancement methods. The K_d^{ANS} value for D209G (29 $\pm$ 7 μM) was similar to that for wild-type (18 $\pm$ 2 μM). This is approximately in accordance with the K_d obtained from isothermal titration calorimetry (section 4.6.2). The similar ANS binding affinities suggests that the binding site of ANS of D209G and wild-type are most likely the same. In the presence of 10 mM GSH, the K_d^{ANS} value of D209G (25 $\pm$ 5 μM) and wild-type (15 $\pm$ 2 μM) were similar. As no significant changes in the binding affinities of ANS occurred upon the addition of GSH, these results show that the ANS binding affinity is independent of G-site occupation. From fluorescence energy transfer studies, it has been reported that GSH does not displace ANS that is bound to hGST A1-1 (Sluis-Cremer *et al.*, 1996).

Figure 4.10 shows the fluorescence spectra of ANS bound to hGST A1-1 in the absence and presence of GSH. The ANS fluorescence emission maximum is 475 nm for wild-type and D209G in apo form. The fluorescence intensity of ANS bound to D209G is 1.6-fold lower compared to that bound to the wild-type. In the presence of GSH, the fluorescence intensity of ANS decreased with a 5 nm blue shift in the emission maximum of the dye. On the other hand the

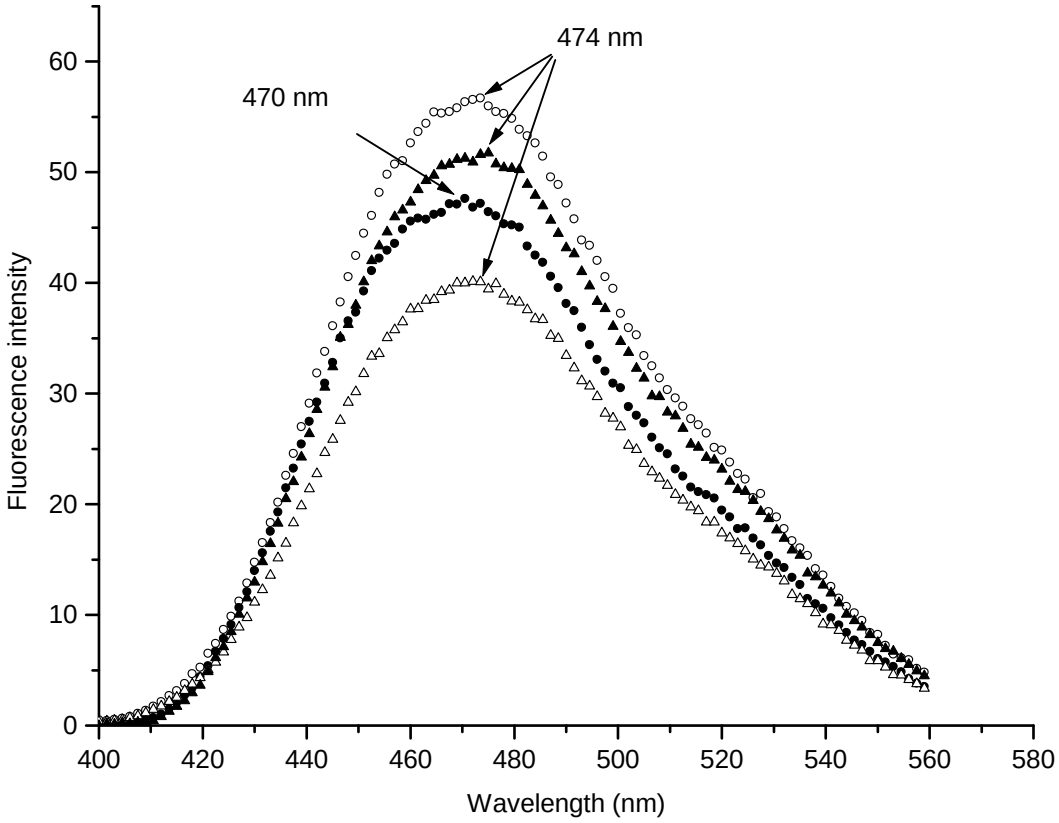

Figure 4.10: Fluorescence emission spectra of ANS complexed with hGST A1-1. Solutions consist of 100 μM ANS and 1 μM wild-type (circles) and D209G (triangles) hGST A1-1 in the absence (open symbols) and presence of 10 mM GSH (closed symbols).

binding of GSH to D209G increased the fluorescence intensity of ANS without a shift in emission maximum. Therefore, differences observed for the fluorescent properties of the dye bound to D209G and wild-type can be explained in terms the differences in the microenvironment at/or surrounding the ANS binding site. Previous studies have shown that changes in the fluorescent properties of ANS can be related to differences in conformation of the C-terminal region of hGST A1-1 and the degree of hydrophobicity at the H-site (Dirr and Wallace, 1999).

The lower fluorescence intensity of ANS bound to D209G compared to that bound to wild-type can be explained by the increased quenching effect of water (Kirk *et al.*, 1996). A more delocalised α-helix 9 over the surface of the protein can give reason for a more solvent exposed ANS binding site. In Figure 4.10 the blue shift in ANS wavelength emission (wild-type) and the increased fluorescence intensity (D209G), with GSH bound at the G-site of hGST A1-1, is indicative of increasing hydrophobicity at the ANS binding site. As mentioned previously (section 4.4), GSH bound to hGST A1-1 increases the stability of the C-terminal region of hGST A1-1 forming a well-ordered helical structure situated over the active-site (Zhan and Rule, 2004). There still remains some degree of flexibility at the ends of the helical structure in comparing to that when ethacrynic acid is bound to the protein (Zhan and Rule, 2004). Even though the structure of the C-terminal region in the apo form of hGST A1-1 cannot be resolved either by crystallography (Cameron *et al.*, 1995) nor NMR (Lian, 1998; Zhan and Rule, 2004), results from this work support the notion that this region persists as an ensemble of helical structures moving on a type of hinge, rather than as a random coil conformer. The increased stability of a well-ordered and structured α-helix 9, in the presence of bound GSH, demonstrates increased stability of this region compared to that of a disordered C-terminal helix (Dirr and Wallace, 1999; Mosebi *et al.*, 2003). An increase in hydrophobicity at the ANS binding site of hGST A1-1, in the presence of GSH, is likely as a result of the induced conformational change of the C-terminal region forming a well-ordered helical conformation positioned over the active-site. This sequestering of the α-helix 9 is likely to result in the expulsion of solvent molecules rendering a more hydrophobic active-site. The

absence of a band shift in the fluorescence of ANS bound to D209G, in the presence of GSH, indicates that the ANS binding site of D209G is less hydrophobic compared to that of the wild-type. A more solvent exposed ANS binding site of D209G, both in the presence and absence of GSH, could be explained in terms of a more disordered and/or delocalised α-helical conformation. Here the degree of expulsion of water molecules is less in comparison to a tightly packed α-helix 9 over the active site, as in the case of the wild-type. A second possibility, is that the α-helix 9 of D209G is ordered but is in a novel location where the active sight is partially exposed to solvent. The former seems more likely, since the stability of the C-terminal region of D290G was marginally less stable in comparison to that of the wild-type.

4.6.2 Energetics of ANS binding: an isothermal titration calorimetry study

To obtain a more detailed picture of the molecular environment of a more solvent exposed ANS binding site of D209G compared to wild-type; the thermodynamics of ANS binding to D209G was investigated by isothermal titration calorimetry. The binding energetics of ANS to wild-type hGST A1-1 has previously been reported by Sayed *et al.* (2002). Figure 4.11 A, shows a typical calorimetric titration curve of ANS binding to D209G, where the binding of the organic anion ANS to D209G is exothermic over the temperature range studied (i.e., 5 - 25°C).

All the data fit well to a model describing one binding site per monomer (Figure 4.11 B) which is consistent with that reported for wild-type (Sayed *et al.*, 2002). The association reaction of ANS with D209G is characterised by a linear dependence in the observed change in enthalpy (ΔH) and entropy ($\mathrm{T}\Delta S$) of the reaction with increasing temperature (Figure 4.12). An enthalpy-entropy compensation is also indicated by a reasonably invariant ΔG with temperature. The linear dependence of the change in enthalpy with temperature yields a heat capacity change of +0.3 $\mathrm{kJ.mol^{-1}.K^{-1}}$ for ANS binding to D209G (Figure 4.12).

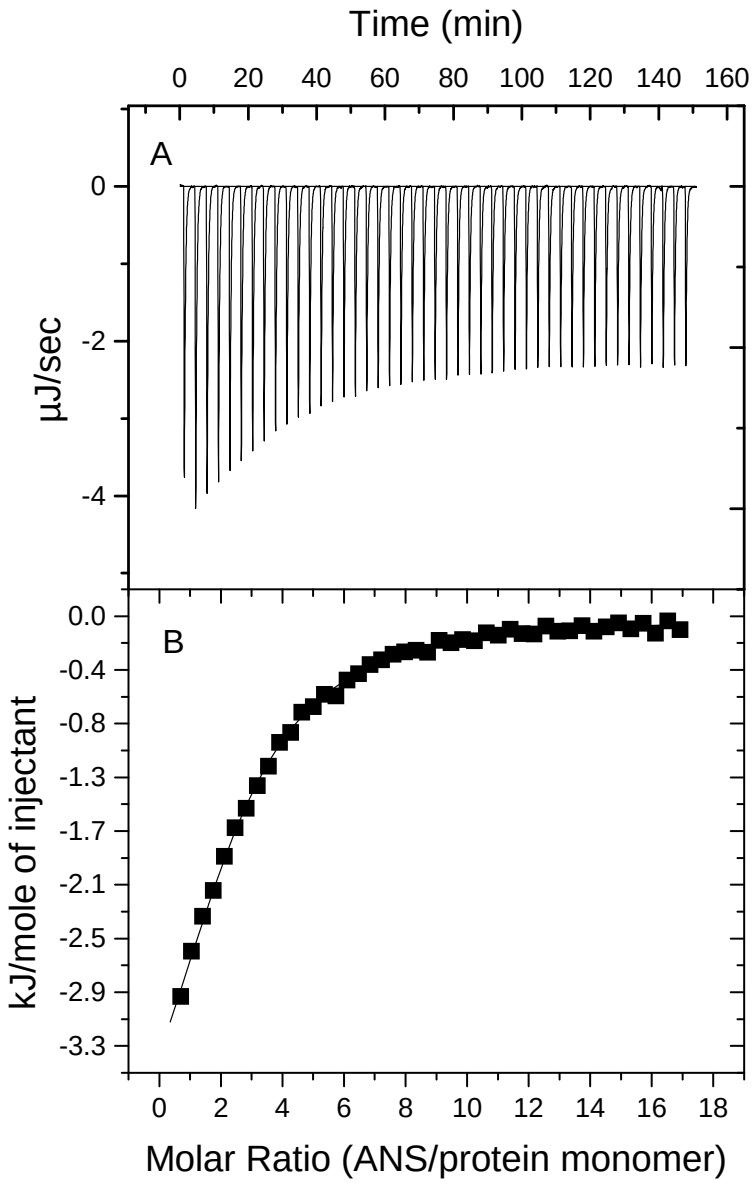

Figure 4.11: Calorimetric titration profile of the binding of ANS to D209G hGST A1-1. This experiment was performed at 25°C. Panel A shows the exothermic heat effects associated with the addition of ligand to the protein sample. Panel B shows the binding isotherm corresponding to results in panel A where the heats of dilution have been corrected for. The line represents the best fit to a single binding site per monomer hGST A1-1.

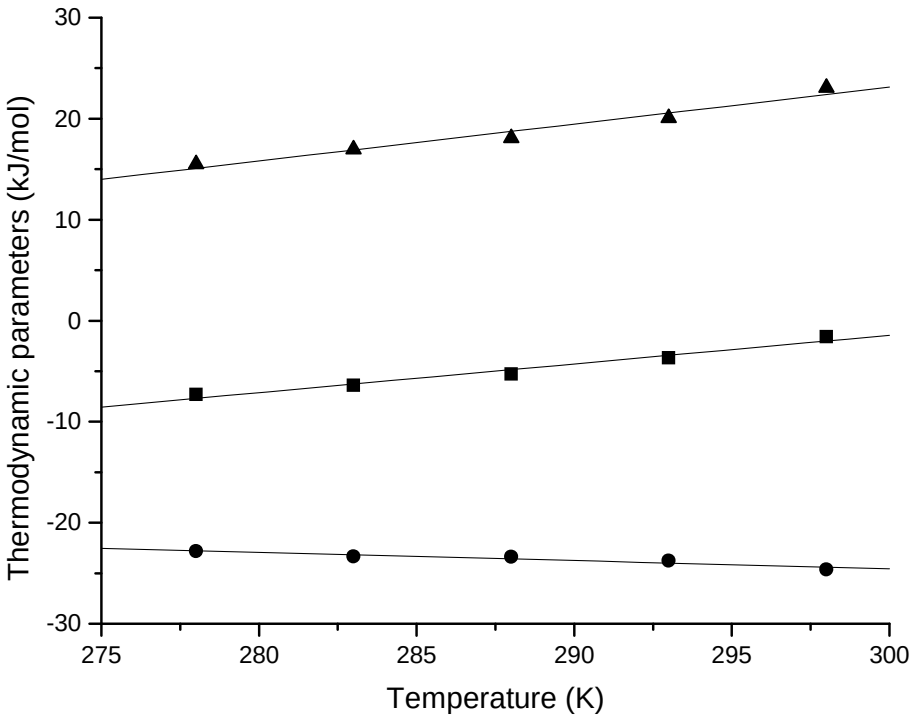

Figure 4.12: Thermodynamic parameters for the binding of ANS to D209G GST A1-1 from the isothermal titration calorimetry data, where the ΔG (•), ΔH (■) and TΔS (▲) values at varying temperatures were fitted to a linear regression.

The major contributor towards ANS binding would be the entropic component which increased with increasing temperature. The binding affinities were in micromolar range (47 - 57 μM) and appeared not to be dependent upon the temperatures in this study, Table 4.3. The K_{d} value for ANS to D209G at 20°C is 57 μM which corresponds well with that of wild-type (65 μM) (Sayed *et al.*, 2002). Thus, the binding affinity of ANS to D209G and wild-type were not significantly different. seeing that the binding affinities of ANS to wild-type and D209G are comparable, it seems unlikely the binding of ANS to D209G is less specific compared to that of the wild-type.

The change in heat capacity (ΔC_p) upon ANS binding was +0.3 kJ.mol^{-1}.K^{-1} for D209G, whereas for wild-type the ΔC_p was reported to be -0.84 kJ.mol^{-1}.K^{-1} (Sayed *et al.*, 2002). The negative ΔC_p obtained for the ANS binding to wild-type hGST A1-1 is comparable to those values obtained for ANS binding to other proteins such as I-FABP (-1.18 kJ.mol^{-1}.K^{-1}) and A-LABP (-0.922 kJ.mol^{-1}.K^{-1}) (Kirk *et al.*, 1996). An increase in heat capacity change (+0.53 kJ.mol^{-1}.K^{-1}) has also been observed for the binding of ANS to the Phe-222 deletion mutant of hGST A1-1 (Sayed *et al.*, 2002). This residue, positioned at the end of the C-terminal region of hGST A1-1, contributes a hydrophobic interaction with substrate bound at the H-site (such as the ethacrynic acid; Cameron *et al.*, 1995) and, hence the hydrophobicity of the H-site (Sayed *et al.*, 2002).

A negative ΔC_p suggests a reduction in solvent accessible non-polar surfaces, whereas a positive ΔC_p suggests the burial of polar groups (Loladze *et al.*, 2001). The negative ΔC_p exhibited by ANS complexed with the wild-type is an attribution to a decrease in the exposure of the hydrophobic surface area (Sayed *et al.*, 2002). Since, the transferal of non-polar compounds to water results in a positive heat capacity change, an increased or positive ΔC_p upon protein-ligand interactions could suggest the exposure of non-polar surface area to solvent. The positive ΔC_p for ANS binding to D209G, suggests that there is a significant hydrophobic component that remains exposed to bulk solvent. This may result in the amount of buried polar surface area in binding to be greater than that of the nonpolar surface. Results from the ANS fluorescent binding studies has provided evidence, and is in agreement, that the binding

Table 4.3: Energetics of the interaction between D209G and ANS at different temperature conditions. All these values are based on a single experiment and the standard deviation were obtained by fitting the titration data using the ITC programme (ORIGIN 5.0)

Temperature (K)	K_{d} (μM)	N‡	ΔG (kJ.mol^{-1})	ΔS (J.mol^{-1}K^{-1})	ΔH (kJ.mol^{-1})
278	51 ± 1.4	2.1 ± 0.05	-22.8	56	-7.3 ± 0.7
283	49 ± 1.7	2.2 ± 0.03	-23.3	60	-6.4 ± 1.2
288	57 ± 3.5	2.4 ± 0.1	-23.4	62	-5.3 ± 0.9
293	57 ± 4.1	2.4 ± 0.2	-23.8	68	-3.7 ± 1.5
298	47 ± 2.8	2.2 ± 0.02	-24.6	77	-1.5 ± 1.1

‡ N is the stoichiometry of binding of ANS per dimer hGST A1-1.

of ANS to the wild-type is accompanied by a more hydrophobic/less solvent exposed binding site compared to that of D209G, as discussed in section 4.6.1.

The free energy change (ΔG) for ANS binding to D209G (-25 kJ.mol^{-1}) and wild-type (-24 kJ.mol^{-1}) were similar, at 25°C. At the same temperature the favourable change in entropy (ΔS) seen for the binding of ANS to D209G (77 J.mol^{-1}.K^{-1}) was unlike wild-type (-23 J.mol^{-1}.K^{-1}), where the change in ΔS was unfavourable. The change in enthalpy (ΔH) of ANS binding to both wild-type and D209G were favourable, however for wild-type (-30 kJ.mol^{-1}) this enthalpic term contributed predominantly to the free energy binding of ANS. Whereas for D209G the enthalpic contribution (-1.57 kJ.mol^{-1}) was minimal (Figure 4.12).

The changes in entropic (ΔS) and enthalpic (ΔH) contributions towards the favourable free energy change are the sum of contributions from hydration effects, polar interactions (hydrogen bonding and ion pairs) and non-polar interactions (van der Waals contacts). ΔS is attributed to the changes in the degree of hydration of the enzyme upon binding, and to the loss of translational and rotational degrees of freedom upon formation of the complex, and any additional changes in rotational and vibrational entropy as a result of a loss in conformational flexibility of the enzyme and/or the ligand. The ΔH reflects the formation or removal of hydrogen bonds, van der Waals forces, and electrostatic interactions between the protein, ligand, and solvent molecules. Thus, the enhanced enthalpic contribution to the binding of ANS to wild-type can be attributed to the formation of hydrogen bonds and/or hydrophobic contacts. The change in free energy of solvent can be explained in terms of the two-state entropy change of water. Water molecules form well-ordered structures (clathrate-like structures) surrounding hydrophobic surfaces and are more ordered than bulk solvent (Franks, 1975). Additionally, water molecules are also structured in instances of forming hydrogen bonds with well-organised polar regions found on the surface of proteins. Hence, in the case of the binding of two hydrophobic interfaces, the dehydration/desolvation of such interfaces would contribute to an increase in free entropy of solvent. In the case of ANS binding to hGST A1-1, the burial of hydrophobic groups would include the surface interface of ANS that is bound to the surface of the protein, and the

interface between α-helix 9 and the H-site. Additionally, upon ligand binding one would expect losses in translational, rotational, and vibrational degrees of freedom in the protein, ligand and the newly electrostricted water molecules hydrating the solvent exposed surface area, all of which would contribute unfavourably to the free entropy change.

In the case of a disordered/delocalised α-helix 9 positioned over the active-site the mechanism of binding is entropically favourable with an unfavourable enthalpic contributions. In contrast, the slow isomerisation step is enthalpically favourable, where the α-helix is well-ordered and docked over the active-site. Importantly, in the absence of α-helix 9, the mechanism of binding of ethacrynic acid is enthalpically driven with unfavourable entropic contributions. The partitioning of these entropic and enthalpic components illustrates that the binding energy can be best explained in terms of the conformational differences of α-helix 9. Work by Nieslanik *et al.* (2001) proposed that the binding of ethacrynic acid-GSH conjugate to GST A1-1 proceeds *via* a two-step mechanism, where α-helix 9 undergoes an intermediate transition conformation that is unlike the delocalised and the well-ordered/docked conformation. This intermediate conformation is realised during the initial docking step of the binding mechanism, prior to the isomerisation which involves the closure of α-helix 9 over the active-site (Nieslanik *et al.*, 1999b). Work by Sayed *et al.* (2000) have reported that increased enthalpic contributions can be explained in terms of increased hydrophobic binding, as a result of increased closure of the active-site from bulk solvent owed to a well-ordered and tightly packed α-helix 9 over the active-site of the protein. Thus, the entropically favourable binding of ANS to D209G is best explained by a more disordered α-helix 9 conformation, much like the intermediate conformation described by Nieslanik *et al.* (2001). Whereas, the enthalpically favourable binding of ANS to wild-type is a result of increased hydrophobic interactions due to a more well-ordered helical conformation. Thus an entropically favourable and enthalpically less favourable binding of ANS to D209G, compared to the wild-type, is best explained in that the ANS binding site of D209G is more exposed to solvent compared to that of the wild-type (Velazquez-Campoy *et al.*, 2000). The degree of solvent exposure of the active-site is expected to be greater in the case

of a more disordered or flexible α-helix 9. These results are in agreement with the those obtained form the ANS fluorescence studies, see section 4.6.1.

Since the α-helix 9 is essential for the turn-over and productive binding of CDNB for its conjugation to GSH (Gustafsson *et al.*, 1999), changes in the conformation of α-helix 9 is expected to directly impact upon the catalytic functioning of this enzyme. Thus a partially delocalised/increased disorder of α-helix 9 of D209G, provides as a good explanation for the marginal decrease in the steady-state turn over of the conjugation of CDNB to GSH (section 4.5) when compared to that of the wild-type.

4.6.3 Predicted ANS binding site to hGST A1-1

Figure 4.13 shows the relationship of ANS fluorescence intensity with the addition of increasing amounts of active-site ligands to hGST A1-1 in the presence of ANS complexed to the protein. The observed change in ANS fluorescence intensity when bound to wild-type and D209G, with the addition of GSH, has been explained in relation to the closure of the α-helix 9 over the surface of the protein, rendering a more hydrophobic active-site, see section 4.6.1, Figure 4.10. Ethacrynic acid and GSH bound to hGST A1-1 are known to induce a tightly packed α-helix 9 forming a "lid" over the active-site (Cameron *et al.*, 1995; Zhan and Rule, 2004). With the addition of increasing amounts of glutathione sulfonate, ethacrynic acid and *p*-bromobenzyl GSH, the fluorescence of ANS that is bound to hGST A1-1, decreases substantially compared to that with saturating amounts of GSH. This substantial loss in the fluorescence signal of ANS does suggest for the competitive displacement of ANS bound to hGST A1-1 by active-site ligands. This is likely, since the binding site of ANS has previously been proposed to be near the H-site of hGST A1-1 (Sayed *et al.*, 2002). The observation that GSH does not displace ANS bound to hGST A1-1, is consistent with a previous study investigating the ANS binding to GSTs by fluorescence-resonance energy transfer (Sluis-Cremer *et al.*, 1996). Recently, displacement studies monitoring the decrease in fluorescence of ANS bound to

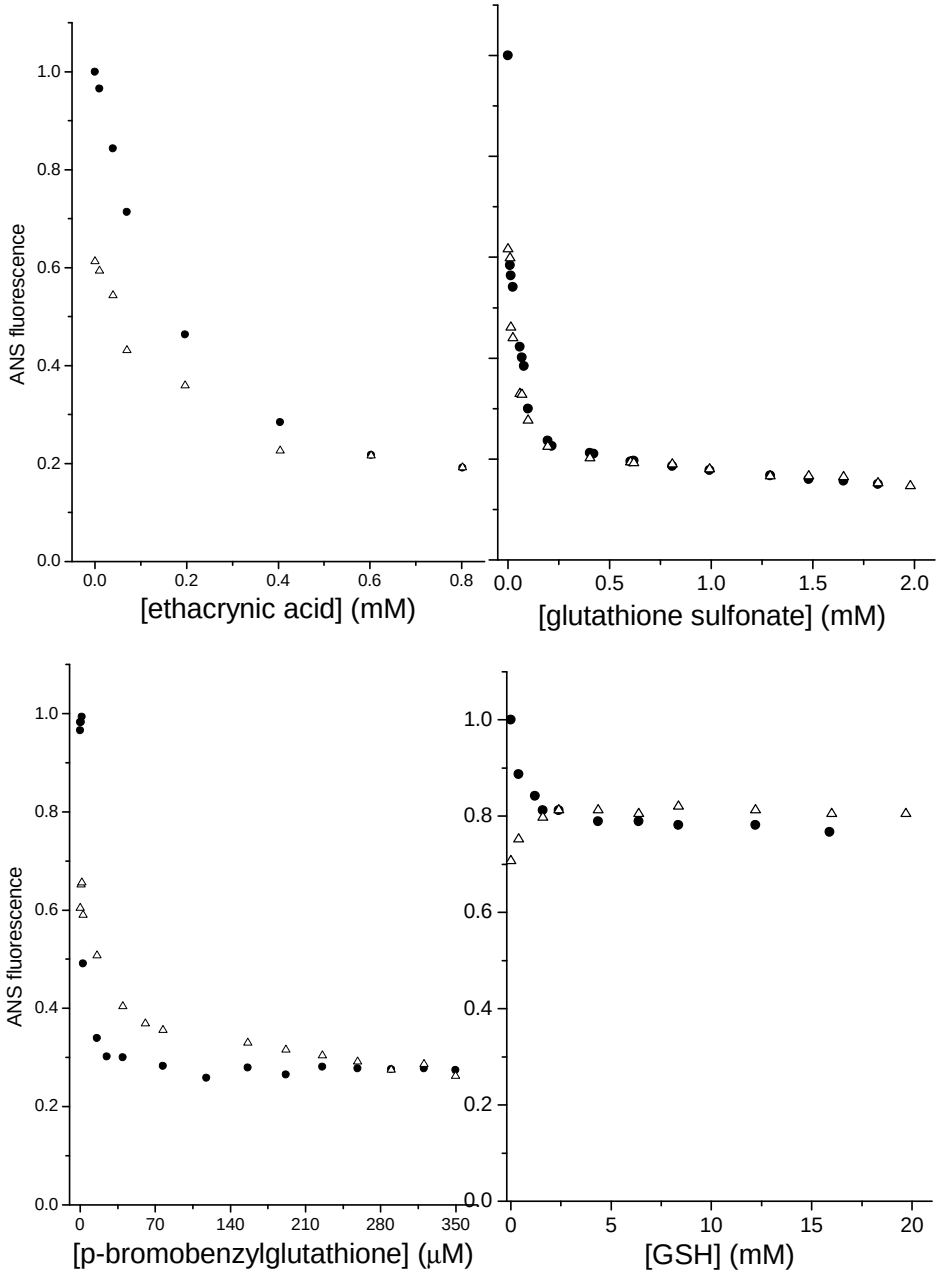

Figure 4.13: Fluorescence emission of ANS when complexed with D209G (open triangles) and wild-type hGST A1-1 (closed circles). Samples consisted of 100 μM ANS and 1 μM protein. Active-site ligands were titrated in μM amounts into samples containing ANS bound to the protein. The fluorescence intensity is normalised, expressed as a fraction of the initial fluorescence signal in the absence of ligand.

the monomeric form of GST P1-1, lacking an interface cleft, was observed (Abdalla *et al.*, 2002). Ligands such as bromosulfophthalein, CDNB and ethacrynic acid displaced ANS, suggesting that the ANS binding site may not be at the interface cleft proposed by Sluis-Cremer *et al.* (1996). The K_d values reported for ethacrynic acid (7 μM, Nieslanik and Atkins, 2000), glutathione sulfonate (4 μM, Sayed, 2001), and *p*-bromobenzyl GSH (14 μM, Kolobe *et al.*, 2004) and GSH (170 μM, section 4.5) are all within the order of ten. The K_d of ANS (65 μM, Sayed *et al.*, 2002) is also within the same magnitude as the above mentioned ligands, and hence has a similar binding to the other active-site ligands used in this study. The fact that a concentration of 10 mM of GSH is sufficient to saturate the G-site of hGST A1-1, gives reasonable evidence that the displacement of ANS by active-site ligands that share similar binding affinities to GSH, is due to steric clashes and/or electrostatic repulsion.

Molecular docking of ANS to hGST A1-1 was performed in order to predict the ANS binding site. The software package Cerius^2 and LIGIN were used (see section 3.9.3 for a detailed description of the docking procedure). The advantage in using Cerius^2 was that the ligand was treated as a flexible molecule, and hence all ligand configurations were searched during the docking procedure. Contrary to this, the program LIGIN treated both ligand and protein as rigid bodies. Nonetheless, an advantage in using LIGIN for this study was the possibility of docking ANS to a ligand-protein complex. The docking results obtained from Cerius^2 (in red) and LIGIN (in CPK colour) are shown in Figure 4.14. Both programs predicted the ANS binding site to be at the H-site of hGST A1-1. No steric clashes were observed. The predicted loss of the accessible surface area of ANS to bulk solvent upon binding to the H-site was 80%. Amino acid residues forming van der Waals contacts with ANS, are shown in Figure 4.14, which include; Val-111, Leu-107, Met-208, Leu-213, and Ala-216, and close contacts with Leu-108, Phe-220 and Phe-222. These residues are commonly reported in forming contacts with other electrophilic substrates bound at the hydrophobic pocket (H-site) of GSTs (see section 2.2.2). Other contacts include Phe-10, Tyr-9 and Arg-15. These residues constitute as part of the G-site. The hydroxyl group of Tyr-9 and the main-chain CO group of Val-55 are predicted to form hydrogen bonds with the sulfonate group of ANS.

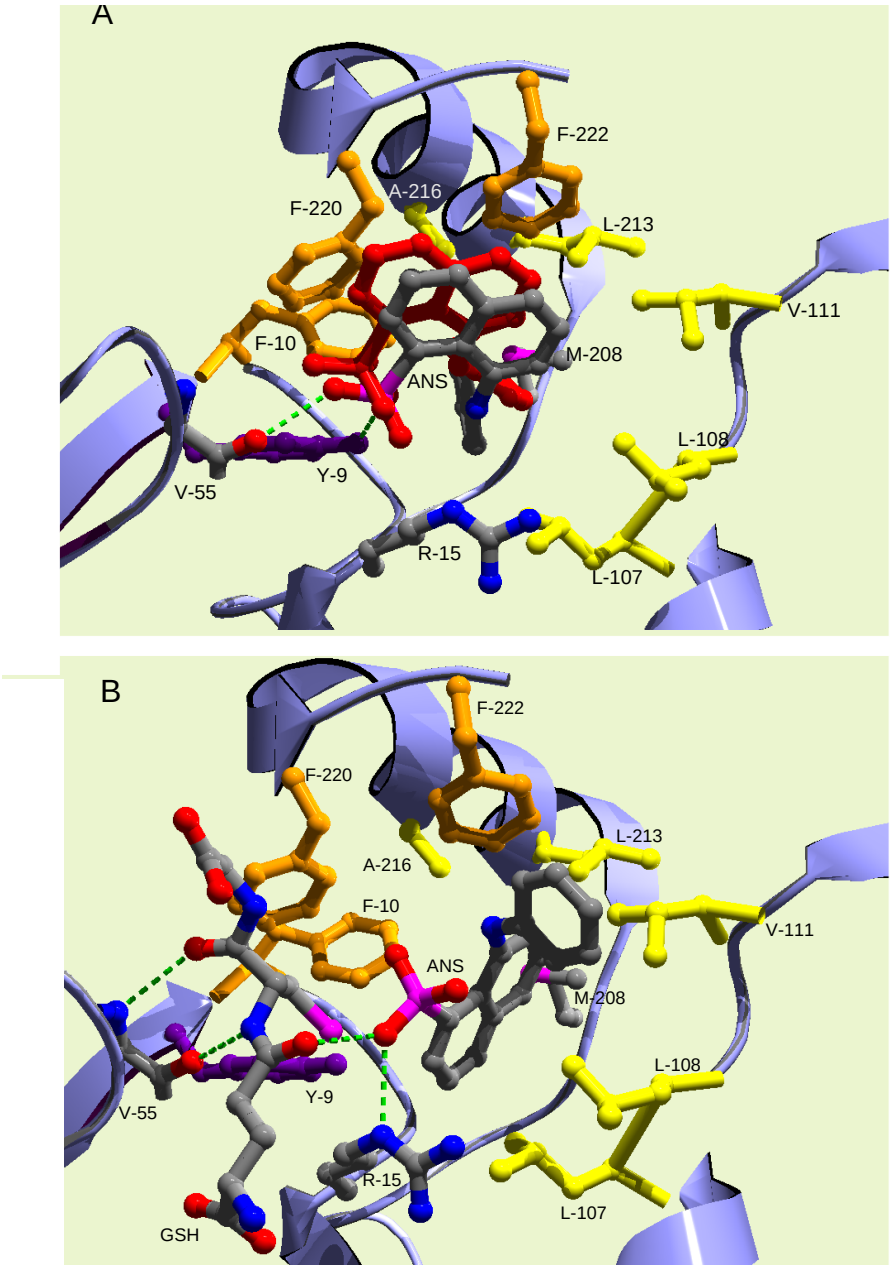

Figure 4.14: Molecular docking results showing the predicted binding site of ANS to hGST A1-1 in apo form (A), and in the presence of GSH pre-complexed with the protein (B). The predicted ANS binding to hGST A1-1 obtained from the molecular docking program Cerius [2] is shown in red, and from LIGIN is shown in CPK colour. Residues that form van der Waals contacts with ANS are shown. Hydrophobic residues are represented in yellow, Phe in orange and Tyr in purple. Hydrogen bonds are depicted as green dotted lines. The figures were generated using Swiss-PdbViewer (Guex and Peitsch, 1997).

The sulfonate of ANS, averaged from both docking results, is positioned approximately 3.5 Å from the guanidino group of the positively charged Arg-15. This close proximity of Arg-15 to the active-site, would appear to satisfy the negative sulfonate group of ANS. Other evidence indicating that ANS binds at the H-site, are the similar the atomic contacts formed between the protein and ANS when compared those formed by other known electrophilic substrates, such as ethacrynic acid (see section 2.2.2). Moreover, the binding of ANS at the H-site of hGST A1-1 occurs with sufficient geometric complementarity. These results are supported experimentally in that the K_d of ANS is similar to that of ethacrynic acid, as mentioned previously in this section. Moreover, the burial of accessible surface area of ANS upon binding does suggests for a stabilised ANS-hGST A1-1 complexation. Residue Phe-220 and Phe-222 in the C-terminal region, form the majority of hydrophobic contacts with the ANS molecule, as judged by the greatest degree of surface area contacts with ANS compared to other hydrophobic residues. In fact, the contribution of Phe-222 to the binding specificity of ANS to hGST A1-1 has been previously reported (Dirr and Wallace, 1999; Sayed *et al.*, 2002).

The question, however, is whether ANS binds to hGST A1-1 in the presence of GSH and/or glutathione sulfonate. The program Cerius^2 could not include a pre-complex of ligand bound to the protein during a docking procedure. Thus, the program LIGIN was used to predict the ANS binding site with GSH or glutathione sulfonate pre-complexed to hGST A1-1 (see section 3.9.3 for method). The predicted atomic contacts and shape complementarity of ANS docked to hGST A1-1 by LIGIN and Cerius^2 were comparable, shown in Figure 4.14. In the presence of GSH, the H-site was the predicted binding site for ANS. Figure 4.14 shows the most probable binding mode of ANS. Interestingly, no binding site could be predicted for ANS in the presence of glutathione sulfonate that was pre-complexed to hGST A1-1. These results are in agreement with those obtained from the ANS displacement studies (Figure 4.13). Briefly, the saturation of the glutathione sulfonate binding site resulted in the displacement of ANS bound to hGST A1-1, whereas no ANS displacement occurred with saturating amounts of GSH bound to the G-site. Glutathione sulfonate is different to GSH in that it has a sulfonate moiety which, when bound to

hGST A1-1, protrudes into the H-site from the G-site (Adman *et al.*, 2001). The molecular docking results predict that the negatively charged sulfonate moiety of glutathione sulfonate is responsible for the displacement of ANS that is bound to the protein. This displacement is likely due to charge repulsion (as ANS itself has a sulfonate moiety) and/or steric clashes. The results in this study strongly suggest that the ANS binding site of hGST A1-1 is not at the amphipathic cleft of the dimeric protein (Sluis-Cremer *et al.*, 1996), but rather at the H-site.

4.7 Tertiary interactions predominate the stability of α-helix 9 in hGST A1-1

Peptide studies of the C-terminal region, equivalent to residues 207-222 of hGST A1-1, show that by removing the N-cap motif, where the Asp N-cap was substituted with a Gly, the helical content of the peptide decreased significantly favouring a more random coil conformer in water. Thus the local interactions at the N-cap motif contribute toward the stability, and hence the helical content of the C-terminal helix as a peptide. Interestingly, the difference in helical content of wild-type and the N-cap mutant peptides in the presence of 40% TFE were similar. It appears that TFE, being a more apolar solvent, may mimic an environment similar to that in proteins. It has been proposed that helical propensities of peptides can be compared to those of helices within the context of proteins. Thermodynamic and fluorescence binding studies of ANS to hGST A1-1 proved to be a sensitive means to elucidate the extent of the ordering/flexibility of this region in D209G relative to that in the wild-type. Even though the stability of the C-terminal region of D209G was reduced, and that a more solvent exposed active-site was evident in the presence of active-site ligands, this was however limited in that the decrease was only marginal. Nonetheless, the structural role of the Asp N-cap in contributing towards an increase in the stability and dynamics of α-helix 9, can give reason for its conservation amongst class Alpha GSTs. In conclusion, the importance of the N-cap residue contributing towards the stability of the

C-terminal region was more apparent in the peptide, than that observed in the protein. This provides evidence that even though local interactions, such as the hydrogen bonds at the N-cap motif do play a stabilising role towards helix stability, within context of the protein, protein-protein and/or ligand-protein interactions tend to predominate in providing for increased stability and induced docking of the C-terminal region of hGST A1-1.

Residues which have been identified in playing a role toward the C-terminal helix stability and dynamics, include Ile-219 (Mosebi *et al.*, 2003), Phe-220 in α-helix 9 (Nieslanik *et al.*, 1999a), active-site Tyr-9 (Nieslanik and Atkins, 2000), Met-208 (Gustafsson *et al.*, 1999; Widersten *et al.*, 1994) and Phe-10 (Ibarra *et al.*, 2001) and the ultimate C-terminal residues Arg-221 and Phe-222, which are not part of the α-helix 9. The benzyl group of Phe-222 is part of the hydrophobic wall of the H-site (Cameron *et al.*, 1995). Phe-220 and Phe-222 contribute toward an increased hydrophobic cavity at the H-site and form an aromatic cluster with Phe-10. Ile-219, in the presence of GSH, is buried in a hydrophobic pocket provided by the side-chain atoms of Phe-220 and G-site residues Leu-41 and Ala-38. The fact that no significant chemical exchange, observed by NMR, was detected for Asp-209, Glu-210, Lys-211, Ser-212, Ala-216, or Ile-219 in the absence of ligand bound at the active-site, is suggestive that these residues are forming extensive interactions with the surface of the protein, and hence contributing toward stability of the C-terminal region (Zhan and Rule, 2004). In the presence of GSH or ethacrynic acid, the ligand-protein interactions induce a well-ordered C-terminal helix over the active-site, rendering a more stable C-terminal region (Zhan and Rule, 2004; Cameron *et al.*, 1995). In saying this, the stability/dynamics of the C-terminal region of GST A1-1 appears to be governed predominantly by protein-ligand interactions, whereas protein-protein interactions appear to contribute to a lesser extent to a stabilised C-terminal region (Widersten *et al.*, 1996; Zhan and Rule, 2004).

The C-terminal region as a peptide has intrinsic helical properties in water, favouring a helical conformation, even in the absence of stabilising tertiary interactions. From this study it has been shown that the N-cap residue plays an important role in favouring a more helical conformation. Even though

NMR studies demonstrate that no significant chemical exchange was detected for Asp-209, Glu-210, Lys-211, Ser-212, Ala-216, or Ile-219 (millisecond time scale) in the absence of active-site ligands; Asp-209 and Phe-222 undergo motion on a short time scale, i.e. nanoseconds to picoseconds (Zhan and Rule, 2004). Residues, Met-208 and Asp-209 located at the N-terminal end of C-terminal region are visible in the crystal structures of hGST A1-1 in its apo form (Cameron *et al.*, 1995). These two residues occupy similar positions to those seen in a well-ordered α-helix 9, even though they have high temperature factors (Cameron *et al.*, 1995). This suggests that the C-terminal region of hGST A1-1 may occur as a structured helix, in its delocalised conformation in the absence of ligands, and unlikely to persist as a random coil conformer (Nieslanik *et al.*, 1999a). Work in this study supports the notion that the dynamics of the C-terminal region consists of a helical structure moving on a type of "hinge" at its N-terminal end (Dirr and Wallace, 1999; Nieslanik *et al.*, 2001). The marginal decrease of a well-ordered helical C-terminal region of hGST A1-1, attributed to the loss of the N-cap residue, is suggestive that N-cap residue does not play a predominant role in stabilising the α-helix 9. However, Met-208 when substituted with an Ala, has been shown to contribute significantly towards the stability of the C-terminal region. Even though Met-208 is known to be involved in the direct catalytic function of the enzyme (Gustafsson *et al.*, 1999; Widersten *et al.*, 1994), the role of Met-208 contributing to the stability of the C-terminal region as a staple motif identified at the N-terminal end of α-helix 9 (section 4.1.1, Table 4.1), remains to be elucidated.

Chapter 5

Role of α-helix 6 N-capping motif in hGST A1-1 folding and stability

A highly conserved N-capping motif (Ser/Thr-X-X-Asp) within the hydrophobic core of domain 2 in GSTs and GST-related proteins has been previously identified by Dragani *et al.*, 1997 (see section 2.3). The focus of this work is in investigating the role of this motif contributing toward the stability and folding of hGST A1-1.

5.1 Over-expression of α-helix 6 N-capping mutants

Single mutations were created by substituting the capping residues with an Ala to investigate the role of the N-cap (S154A) and N3 residues (D157A) of α-helix 6. The construction of mutant D157A and S154A pKHA1 plasmids were successful as judged by DNA sequencing, i.e. the expected sequences were in place and no undesired mutations were incorporated. The gene encoding for mutant proteins were over-expressed in BL21 *Escherichia coli* with 1 mM IPTG for 18 hours, at 37°C. As judged by SDS-PAGE, Figure 5.1 and 5.2, the

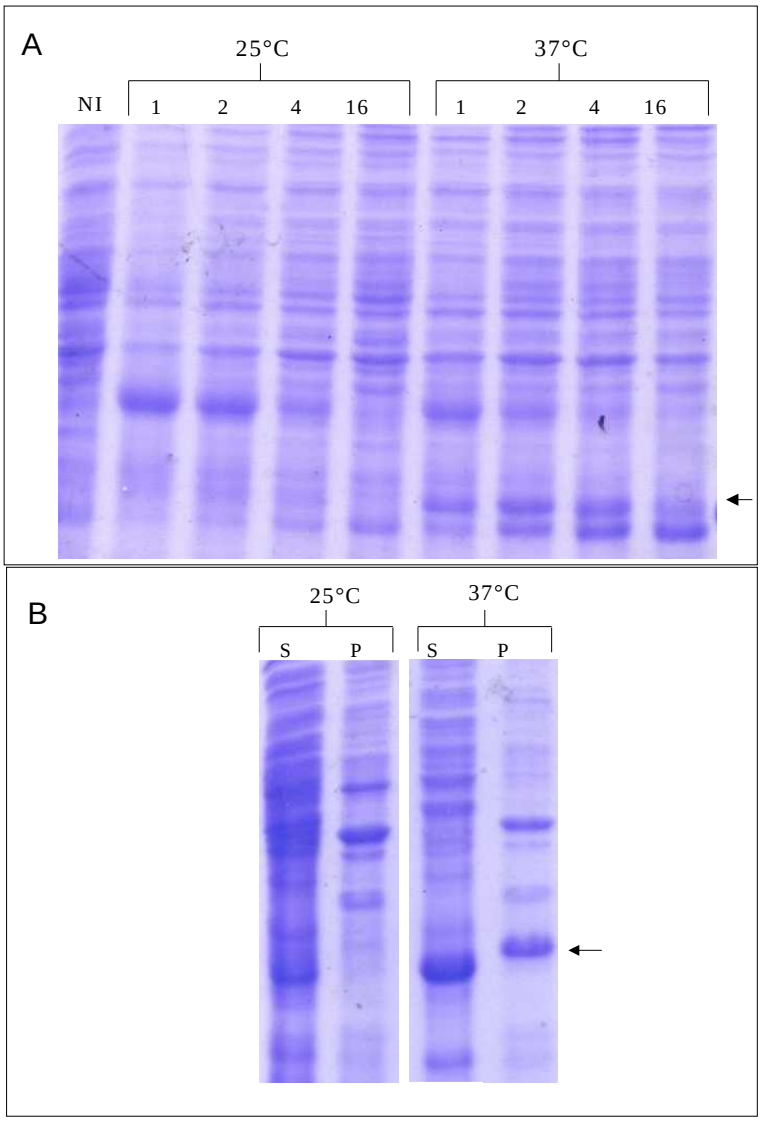

Figure 5.1: Panel A 15% SDS-PAGE stained with Coomassie Blue stain of over-expression levels of D157A in lysed BL21 (DE3) pLysS cells. NI denotes non-induced cells. The numbers indicate the hours of induction of cell cultures with 1 mM IPTG. The temperatures reported above are the growth temperatures for the cultures. The arrow indicates the position of wild-type hGST A1-1. Panel B shows the soluble and insoluble partitioning of over-expressed D157A and wild-type hGST A1-1 after 6 hour growth at the temperature indicated.

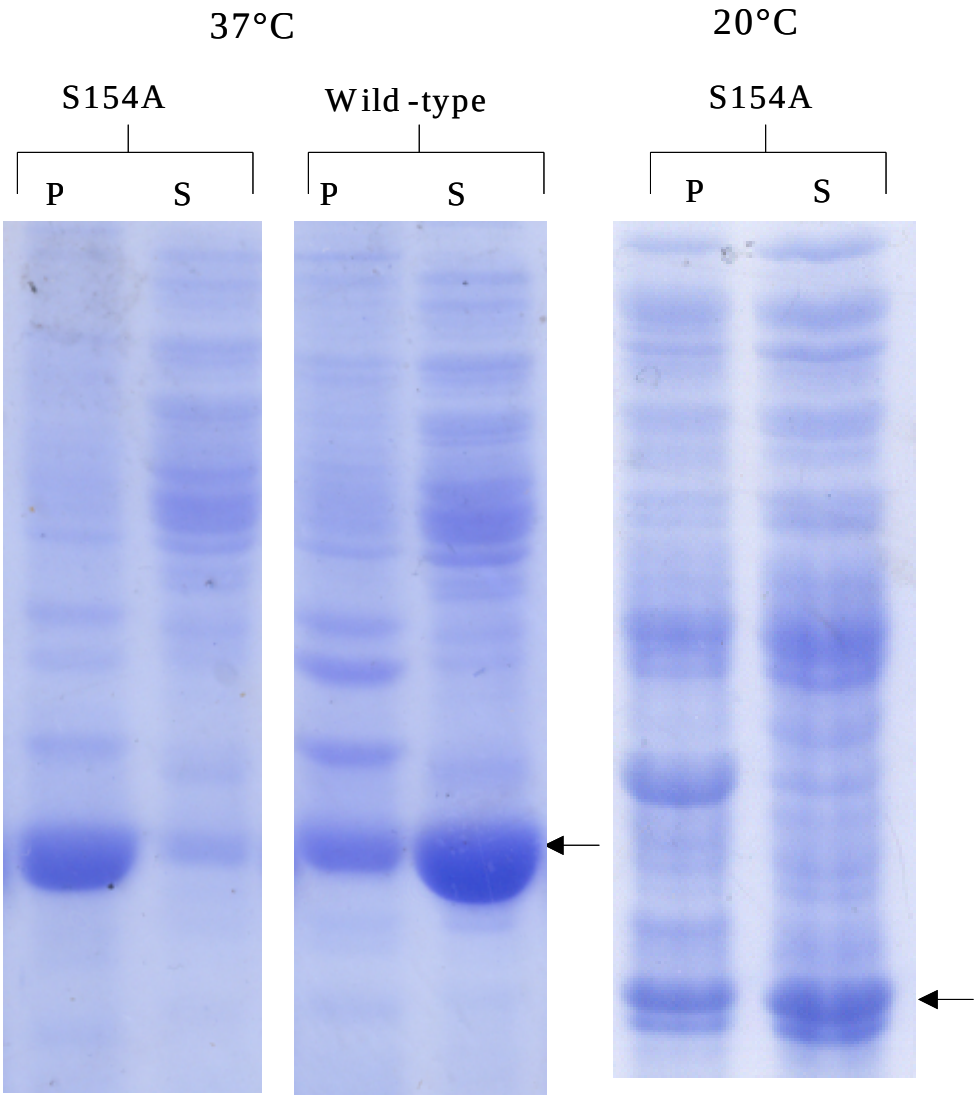

Figure 5.2: 15% SDS-PAGE stained with Coomassie Blue showing the soluble and insoluble fractions from lysed BL21 (DE3) pLysS *Escherichia coli* with over-expressed S154A and wild-type hGST A1-1. Cell cultures were induced at mid-log phase with 1 mM IPTG and grown for 16 hours. The growth temperatures are indicated. The arrow indicates the position of wild-type hGST A1-1.

expression levels of the N3 mutant (D157A) was considerably less when compared to that of the N-cap mutant (S154A), which in turn is less than wild-type. Both N-capping mutants were expressed in the insoluble fraction at 37°C. By lowering the growth temperatures to 20°C, there was no over-expression of D157A (Figure 5.1), whereas S154A was over-expressed with a sufficient amount of protein in the soluble fraction (Figure 5.2). Since no over-expression of D157A was obtained at 20°C, the N3 substitution appears more disruptive than the N-cap substitution. The more disruptive substitution of the N3 residue compared to the N-cap residue, is consistent with previous studies (Cocco *et al.*, 2001; Dragani *et al.*, 1997), and is explained in that the N3 substitution results in a complete loss a hydrogen bond network, which in turn results in an increase in mobility of the preceding loop and N-terminal end of α-helix 6 (Rossjohn *et al.*, 2000). Whereas upon removing the hydroxyl group of the Ser N-cap residue, the N-capping hydrogen bonding network is not entirely disrupted, where only 2 hydrogens bonds are lost (see Figure 2.5 for an illustration of the N-capping motif). From these results it is apparent that the hydrogen bond network provides an enthalpic contribution to the thermal stability of these proteins, located within the hydrophobic core of domain 2 of the protein.

5.2 Purification of S154A

As no over-expressed D157A could be obtained as a soluble protein, no further studies of this N3 mutant were carried out. The growth conditions to obtain soluble S154A were optimised. A 1 litre starting culture of S154A at 20°C was induced with 1 mM IPTG at mid-log phase, and harvested after 18 hours of growth. S154A was purified, as described for wild-type (see section 3.5). The purity and expected molecular mass of S154A, as judged by SDS-PAGE and SEC-chromatography, is greater than 98% pure with the expected molecular mass of 54 kDa. Due to the instability, and decreased expression-levels of S154A, lower yields of protein were obtained compared to that obtained for the wild-type. A typical purification of S154A yielded 40 mg of protein per 1 litre of culture.

5.3 Characterisation of S154A

The degree of ellipticity reports the extent of secondary structural content, while the selective excitation at 295 nm of Trp-21 fluorescence reports the tertiary structural conformation of hGST A1-1 (Wallace *et al.*, 1998b). The far-UV CD (Figure 5.4) and the intrinsic Trp-21 fluorescence emission spectra (Figure 5.3) of S154A coincide with those of the wild-type. The fluorescence emission maximum of S154A and wild-type protein was 330 nm. These data indicate that the α-helix 6 N-cap substitution did not alter the tertiary packing and secondary structural content of the protein. The specific activity of S154A (45 ± 6 μmol.min^{-1}.mg^{-1}) was similar to that of the wild-type (52 ± 5 μmol.min^{-1}.mg^{-1}). Therefore, S154A and wild-type share the same conformation at the active-site. The preservation of the structural integrity of the Ser to Ala N-cap mutant is consistent with that reported for the same mutant in class Pi GSTs (Rossjohn *et al.*, 2000). The crystal structure of the N-capping mutants in class Pi GST showed that even though the structure of the mutants did not vary from that of the wild-type, increased disordering of the loop preceding α-helix 6 was observed, as judged by the B-factors (Rossjohn *et al.*, 2000).

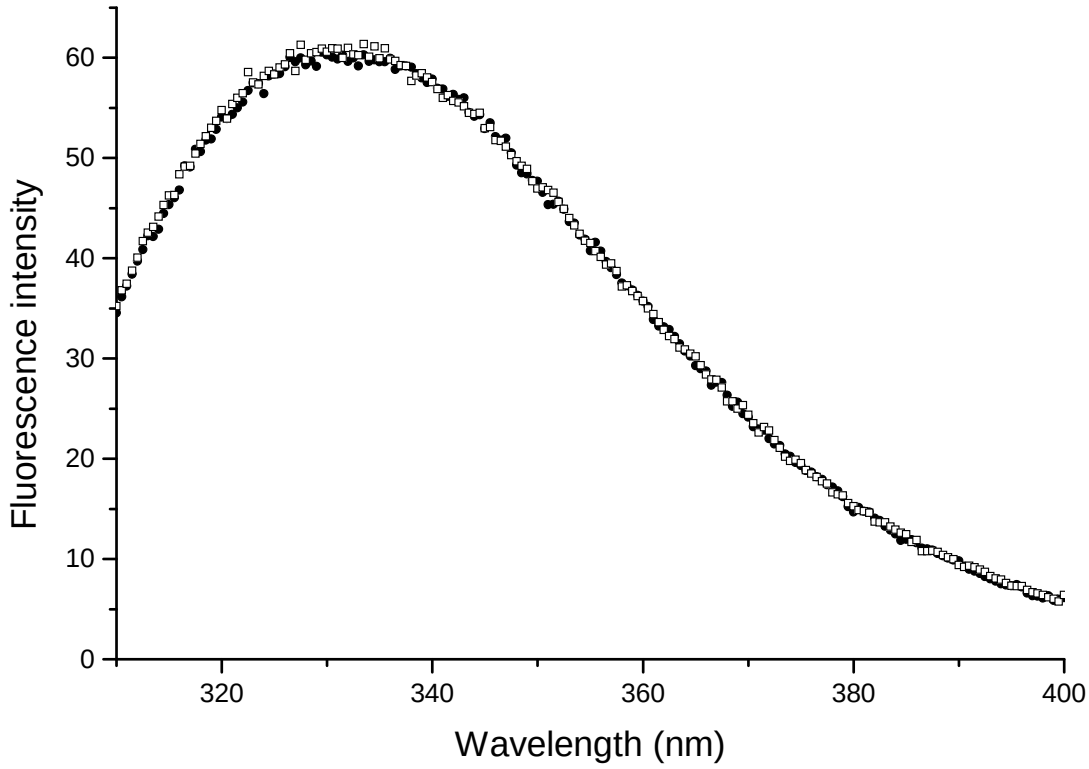

Figure 5.3: Trp-21 fluorescence emission spectra of S154A (□) and wild-type (•) hGST A1-1. Samples consisted of 1 μM protein in 20 mM sodium phosphate buffer, pH 6.5, 1 mM EDTA, 0.1 M NaCl, 0.02% sodium azide.

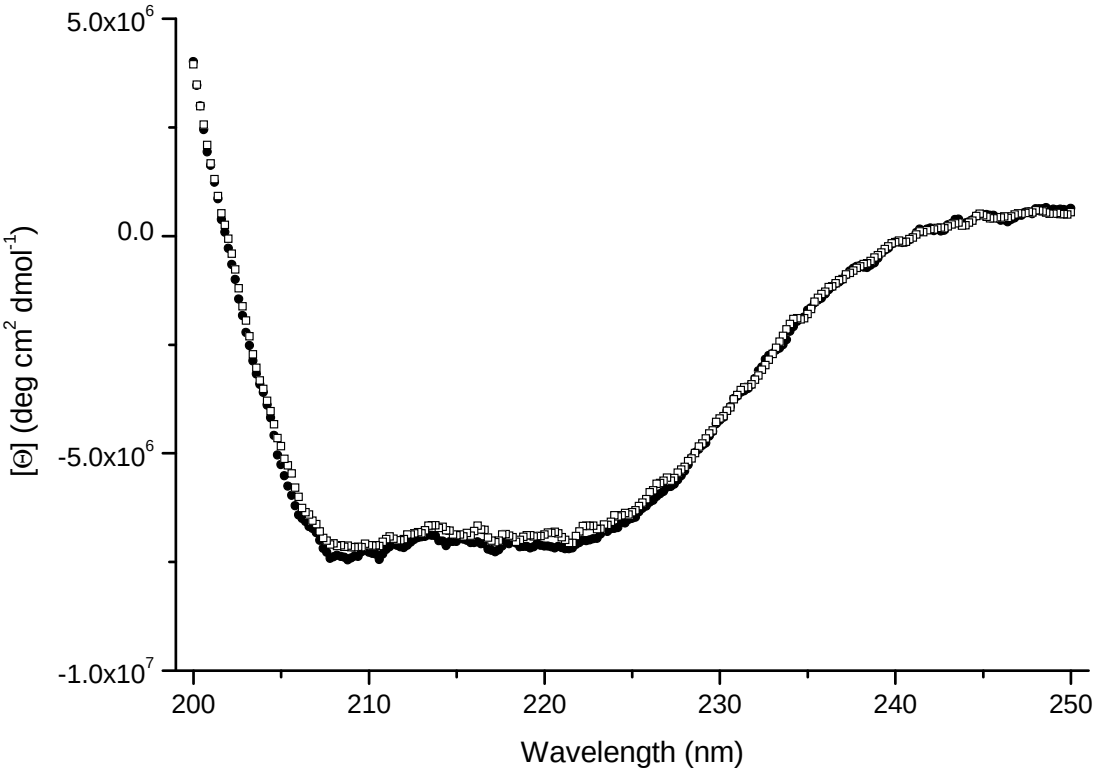

Figure 5.4: Far-UV CD spectra of S154A (□) and wild-type (•) hGST A1-1. Samples consisted of 2.5 μM protein in 5 mM sodium phosphate buffer, pH 6.5, 0.02% sodium azide.

5.4 Thermal stability of S154A and wild-type hGST A1-1

The thermal denaturation profile of wild-type and S154A were determined at different rates of heating (Figure 5.5). Thermal unfolding of wild-type and S154A were irreversible processes, as shown in Figure 5.5, and is a general feature of GSTs as a result of aggregation (Kaplan *et al.*, 1997; Wallace *et al.*, 1998b; Stenberg *et al.*, 2000). For wild-type hGST A1-1, the midpoint temperature (T_m) was 57°C, whereas the T_m for S154A was significantly reduced. For wild-type, the T_m was independent of the heating rate, whereas for S154A the T_m decreased with decreasing rate of heating (Figure 5.5). The most significant decrease in T_m of S154A was from 47°C and 43°C, when the heating rate was lowered from 2°C.min^{-1} to 1°C.min^{-1}. The lower T_m for S154A compared to wild-type indicates a less stable native conformation of S154A. The decrease in thermal stability of S154A can be explained in terms of the loss of the hydrogen bonding interactions, where for class Pi GSTs the removal of the side-chain hydroxyl group of Ser-154 (N-cap: numbering is according to class Alpha GSTs) results in a loss of two hydrogen bonds (section 2.3.1). Even though the crystal structure of the S154A variant indicates no increase in the conformational flexibility, a decrease in thermal stability does persist (Cocco *et al.*, 2001). Thus, this loss in conformational stability owed to the loss of two hydrogen bonds observed for class Pi GSTs (Rossjohn *et al.*, 2000; Cocco *et al.*, 2001), can be a general feature for GSTs, also giving reason for the decrease in thermal stability of S154A compared to that of the wild-type. This decrease in the T_m of S154A in relation to decreasing rate of heating, is suggestive that the transition of unfolding for this variant is under some kinetic control, i.e. most likely an unfolding step such as aggregation.

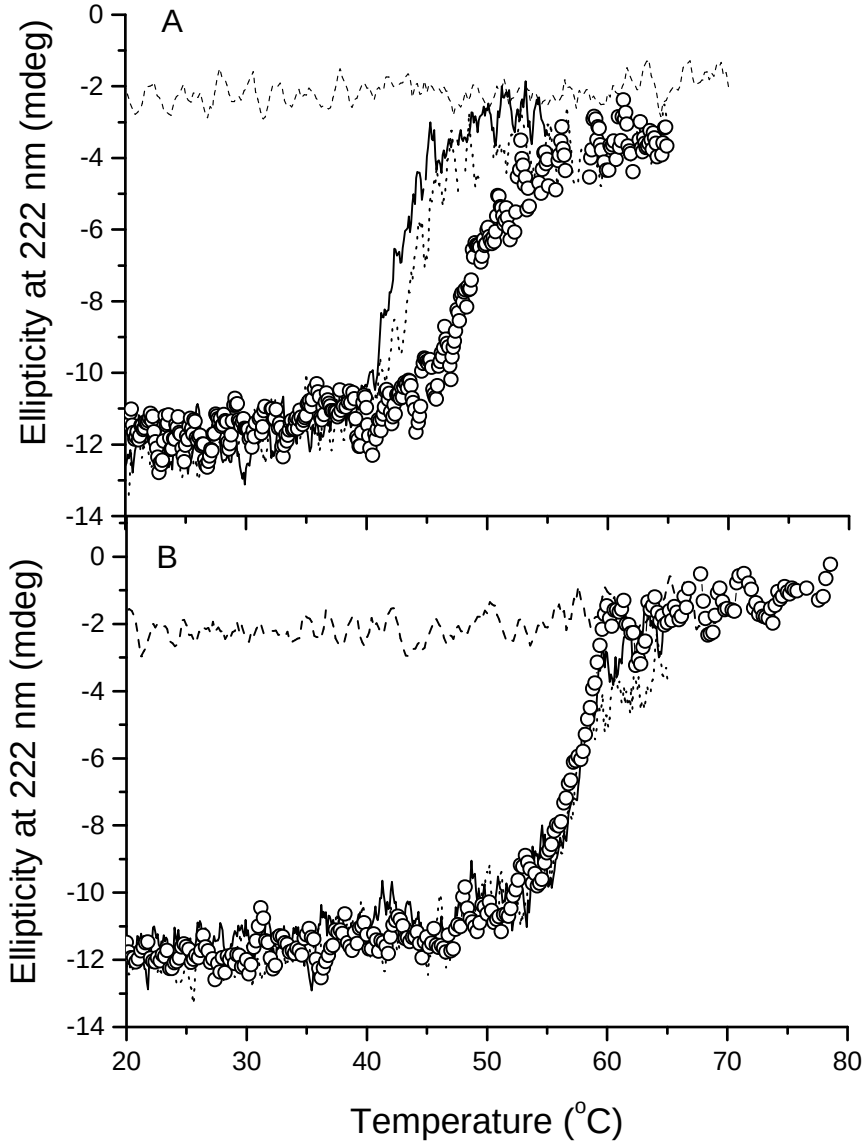

Figure 5.5: Thermal denaturation of (A) S154A and (B) wild-type monitored by circular dichroism spectroscopy at 222 nm. Sample of 1 μM protein were unfolded at different rates, namely 0.5°C.min^{-1} (dotted lines), 1°C.min^{-1} (solid lines) and 2°C.min^{-1} ($\circ$). The dashed lines indicate the re-cooling of the unfolded protein sample at 2°C.min^{-1} demonstrating irreversible heat denaturation of hGST A1-1.

5.5 Reversibility of urea-induced unfolding of S154A

To obtain thermodynamic parameters assigned to S154A, the unfolding transition of the protein is required to be a reversible process (Pace and Scholtz, 1998). Since the thermal denaturation of wild-type hGST A1-1 is an irreversible process, whereas urea-induced unfolding is reversible (Wallace *et al.*, 1998b), urea was therefore the preferred denaturant used for this study. The minimum time required for S154A to reach equilibrium in 6 M urea was 16 min. This is the same as for the wild-type (Barnwell, 2004). The reversibility of folding of S154A and wild-type was investigated. The protein was refolded in 6 M urea *via* a 6-fold dilution into renaturation buffer, i.e. without denaturant. The recovery of the native conformation was assessed by the extent to which the protein regained activity, the degree of ellipticity measured by far-UV CD and the intrinsic fluorescence of the protein. The yield of recovered protein in relation to the time the protein was incubated in 6 M urea is reported in Table 5.5. The reduced yield of recovered activity of both proteins, in a final concentration of 1 M urea, decreased with increased incubation periods of protein in 6 M urea, i.e. from 30 min and onwards. Between 16 - 30 min of incubating the protein in 6 M urea, the percentage in regained activity and far-UV CD signal of S154A and wild-type was greater than 95%. This is consistent to previous results (Wallace, 1998). The regain in the intrinsic fluorescence indicated a sufficient recovery, but slightly lower yield in refolded protein i.e., 88 - 95% is native protein. The recovery of the secondary and tertiary structure of both proteins indicate that the unfolding reaction under strong denaturing (6 M) and renaturing conditions (1 M urea) is reversible. Since a reduced yield of refolded protein is a common phenomena occurring most often at high concentrations of protein (Fink, 1998), this behaviour was investigated for S154A. Results in Table 5.5 show that the dependence of recovered S154A was independent of increasing concentrations of denatured protein in 6 M urea, i.e. between 3 to 10 μM. This is comparable to the wild-type (Wallace, 1998). To ensure the maximum yield of recovered protein was obtained after refolding, samples of protein in 6 M urea were incubated for no longer than 30 min.

Table 5.1: Percent recovery of native S154A and wild-type hGST A1-1 refolded from protein in 6 M urea. The percent recovered protein is shown with increasing times of incubation of protein in 6 M urea.

	Wild-type	S154A			
Incubation time (min)	% activity	% activity		% FI	% ellipticity
	3 μM	3 μM	10 μM	10 μM	
18	97	99	98	92	106
30	98	96	101	88	94
60	84	86	81	94	100
120	83	72	77	95	101
180	60	63	74	100	103
240	68	50	65	105	117

The activity, fluorescence intensity (FI) and ellipticity at 222 nm of hGST A1-1 in buffer was taken as 100%. Concentration of protein samples in 6 M urea are indicated. Protein samples were denatured in 6 M urea. The refolding reaction was initiated by diluting a sample of denatured protein by 6-fold in 20 mM phosphate buffer, pH 6.4. The minimum time required necessary for the refolding reaction to reach completion was 30 min.

5.6 Unfolding kinetics of S154A

The unfolding of S154A was monitored by fluorescence detection on a stopped-flow instrument. The unfolding of hGST A1-1 is biphasic (Wallace *et al.*, 1998b): $N_2 \rightarrow N_2^* \rightarrow 2U$, where N_2 is the native dimer, N_2^* is a transiently formed dimeric intermediate, and U is the unfolded monomer (see section 2.3.3, scheme 2.1). The two kinetic unfolding phases observed for S154A were similar to those of the wild-type (Figure 5.6). Within the first 200 ms of the reaction an increase in fluorescence was observed (fast phase), which was followed by a slower more gradual decrease in fluorescence, i.e. the slow phase (Figure 5.7). The fast unfolding event has been described as involving the partial dissociation of the domain 1-domain 2 interface (Wallace *et al.*, 1998b) which results in the destabilisation/unfolding of α-helix 9, represented as $N_2 \rightarrow N_2^*$. The slow unfolding event, represented as $N_2^* \rightarrow 2U$, involves the global unfolding of the dimeric protein. An unfolding temperature of 10°C was chosen to resolve the fast unfolding phase together with the slow phase (Wallace *et al.*, 1998b,a). The increased fluorescent amplitude for the fast phase of S154A and wild-type were similar (-0.6 $\pm$ 0.12 units), as well as for the decreased fluorescence amplitude of the slow phase (0.68 $\pm$ 0.5 units). It is apparent from the equivalent change in amplitudes for both folding phases of wild-type and S154A that the unfolding pathway for both proteins are similar, see also Figure 5.6 and 5.7. The order of the reaction is best described by the relationship of protein concentration to the rates of each unfolding phase. For wild-type, the rates of the unfolding phases are first-order, i.e. the unfolding reaction from the native dimer to unfolded monomer is unimolecular (Wallace *et al.*, 1998a). The rates of unfolding of S154A for both phases were independent of protein concentration (0.5 to 6 μM final concentration), and hence were first-order kinetic events.

The apparent rates for the fast unfolding phase of S154A, shown in Figure 5.8, displayed a slightly lower dependence upon increasing urea concentrations compared to that of the wild-type. The apparent rates for the slow phase of S154A resulted in an increased dependence with increasing urea concentrations

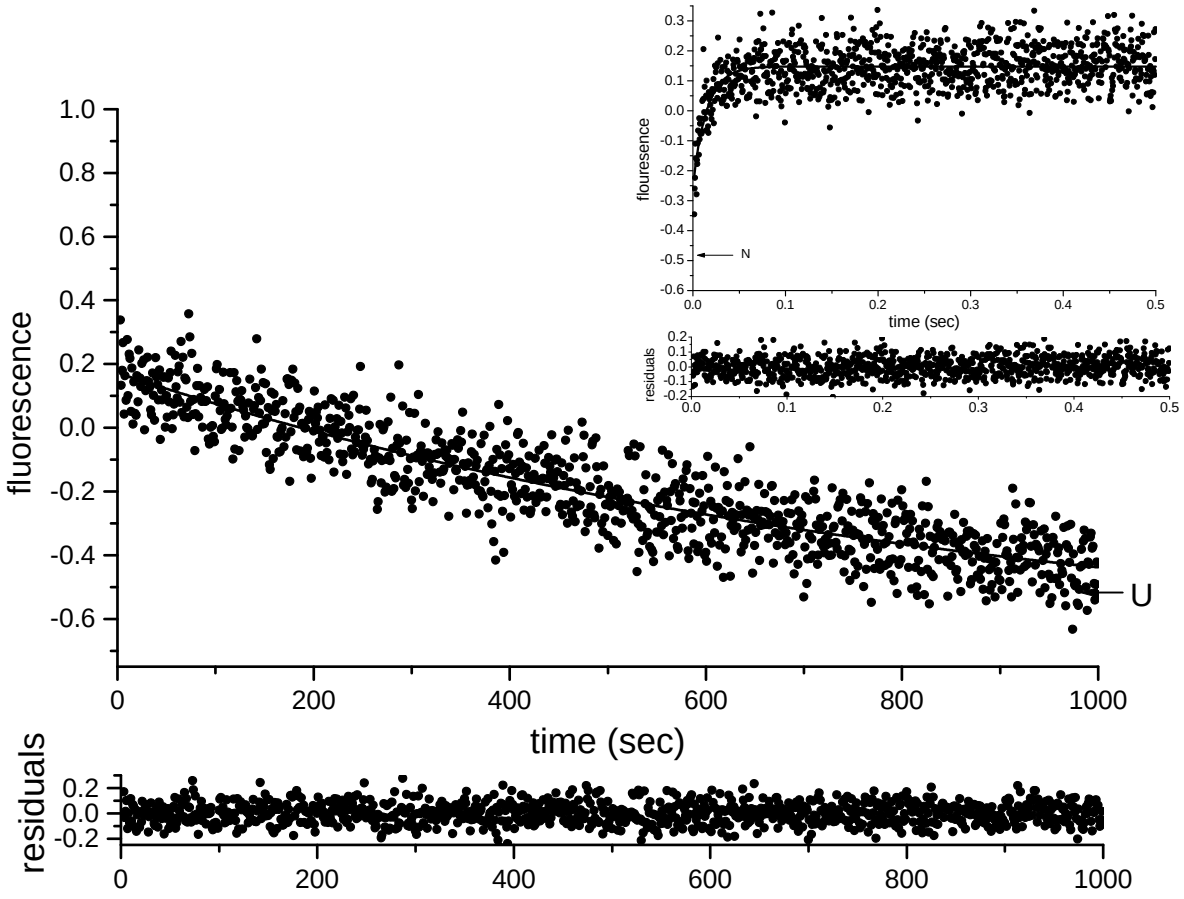

Figure 5.6: Kinetic unfolding trace of wild-type hGST A1-1 monitored by fluorescence (excitation at 280 nm and emission above 320 nm) using stopped-flow at 10°C. Two unfolding phases were resolved, being the fast (insert) and slow folding phases. Both phases are best fitted to a single exponential function (residuals are shown). Samples consisted of 1 μM final concentration of protein, pH 6.5. The arrow indicates the fluorescence signal corresponding to the native (N) and unfolded protein.

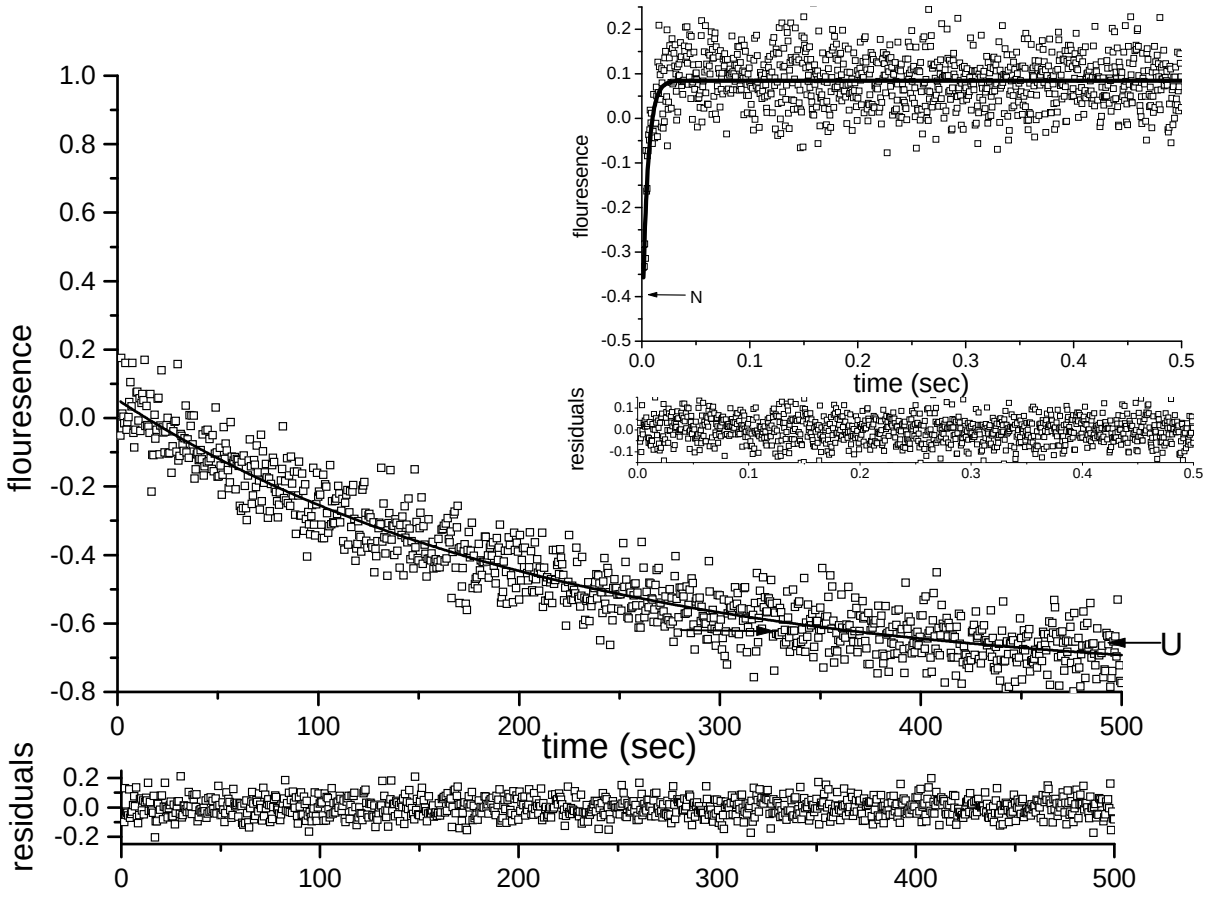

Figure 5.7: Kinetic unfolding trace of S154A monitored by fluorescence (excitation at 280 nm and emission above 320 nm) using the stopped-flow instrument. Two unfolding phases were followed at 10°C with a final protein concentration of 1 μM, pH 6.5. Both the fast (insert) and slow folding phases are shown, where both phases fit best to a single exponential function (residuals are shown). The arrow indicates the fluorescence signal corresponding to the native (N) and unfolded protein.

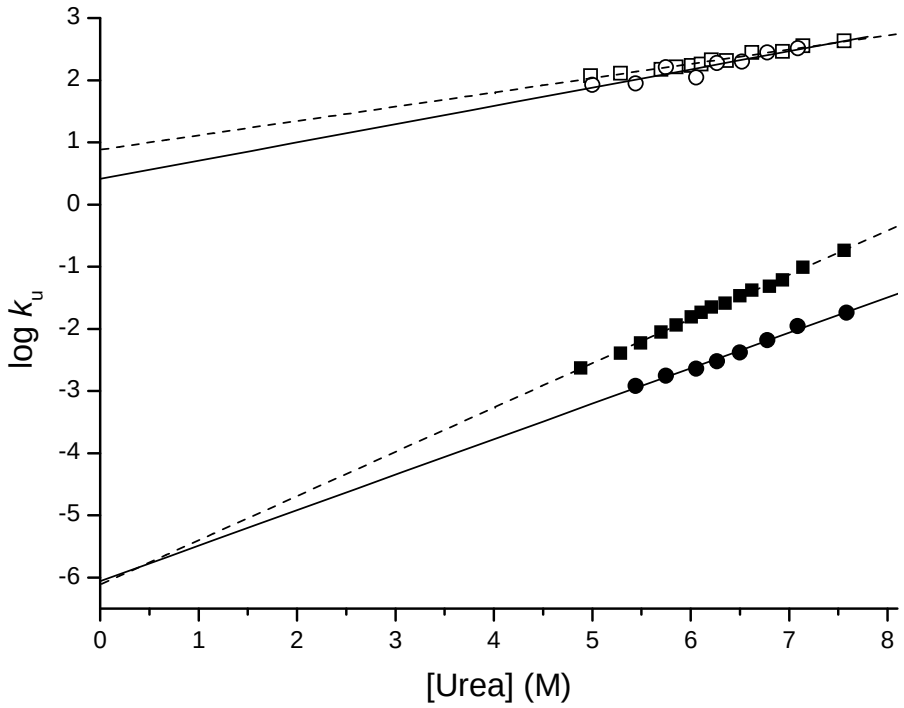

Figure 5.8: The dependence of the apparent unfolding rates of wild-type (circles) and S154A (squares) with increasing concentrations of urea at 10°C. The two unfolding phases are the fast (open symbols) and slow unfolding phases (closed symbols). Samples consisted of a final concentration of 1 μM protein, pH 6.5.

compared to that of the wild-type (Figure 5.8). The unfolding kinetic parameters obtained in Figure 5.8 are shown in Table 5.6. The $k_{\mathrm{u}}(\mathrm{H_20})$ of the fast unfolding phase of S154A was 3-fold greater compared to that of the wild-type, whereas for the slow phase the $k_{\mathrm{u}}(\mathrm{H_20})$ of both proteins were similar (Table 5.6). The m_u-value for the fast phase of S154A is less compared to that of the wild-type, whereas for the slow phase the m_u-value is greater for S154A compared to that of the wild-type (Table 5.6).

The m_u-values reflect changes in the solvent-accessible surface area of the transition state between the two species of a particular unfolding phase (Tanford, 1970). Thus, the increased m_u-value observed for the slow phase of S154A compared to that of the wild-type can be interpreted in that the transition state of S154A is more closely related to the solvent-accessible surface area of the unfolded state compared to that of the wild-type. Whereas, a decrease in the m_u-value of the fast folding event of S154A, is explained in terms of a less solvent exposed/more native-like transition state for S154A when compared to that of the wild-type (Shortle and Meeker, 1986; Shirley *et al.*, 1989). However, this seems unlikely since the slightly enhanced rate of unfolding of the fast phase of S154A compared to that of wild-type, suggests an increased susceptibility for the rapid dissociation at domain 1-domain 2 interface of this mutant. The possibility that the change in transition state may also reflect a change in the initial and final states, needs to be considered since the transition state is dependent upon the nature of these two conformers. Any differences in the unfolded species due to the N-cap mutation is unlikely. As mentioned previously, the structure of native S154A shares the same conformation to that of the wild-type. However, the decreased thermal stability of S154A compared to that of the wild-type, suggests that the dynamics of the structure of S154A could be greater than that of the wild-type. By removing the N-capping motif of α-helix 6, an increase in mobility of the loop preceding α-helix 6 in the native structure of class Pi GSTs has been reported. This resulted in increased motion of the entire structure, judged by the averaged B-values obtained from the crystal structures of the N-capping mutants when compared to that of the wild-type (Rossjohn *et al.*, 2000). Enhanced thermal instability of the native

Table 5.2: Unfolding kinetic parameters of S154A and wild-type hGST A1-1.

	$k_u(H_20)$ (s^{-1})	m_u-value (J.mol^{-1} M^{-1} urea)
	Fast	
Wild-type	2.6 ± 1.3 x 10^{-6}	686 ± 80
S154A 0.14	8.5 ± 1.8 x 10^{-6}	527 ± 50
	Slow	
Wild-type	1.2 ± 0.7 x 10^{-6}	1263 ± 16
S154A	0.67 ± 0.3 x 10^{-6}	1690 ± 16

structure of the N-capping mutants for class Pi GSTs is likely to be a general feature amongst GSTs. Therefore, in the case of hGST A1-1, an increase in the flexibility of the native structure could give reason for a less cooperative unfolding event for the fast phase of S154A compared to that of the wild-type.

As a general feature, the tight packing of a protein is crucial for stabilisation of its native three-dimensional conformation. It is well documented that the core of globular proteins are tightly packed and are stabilised by hydrogen bonds (Myers and Pace, 1996) and hydrophobic interactions (Dill, 1990). In the case of S154A, the loss of two hydrogen bonds at the N-motif of α-helix 6 resulted in the decrease stability of hGST A1-1. This decrease in thermal stability demonstrates the requirement of a tightly packed interior of the protein necessary for the overall structural stability of hGST A1-1. The hydrophobic core packing surrounding Leu-164 in α-helix 6 is reported to contribute toward stabilising domain 2 through van der Waals and hydrophobic contacts (Wallace *et al.*, 1998a). It has been suggested that the unfolding reaction of wild-type GST A1-1 is most likely attributed to the penetration of solvent into the hydrophobic core of domain 2 which in turn causes disruption of van der Waals forces and hydrophobic interactions at the domain 1-domain 2 interface (Wallace *et al.*, 1998a; Nathaniel *et al.*, 2003; Wallace *et al.*, 2000). This is reasoned since an increased rate in the partial unpacking and dissociation of the domain 1-domain 2 interface has been observed as a result of destabilised regions in both domain 2 (Wallace *et al.*, 1998a) and domain 1 (Nathaniel *et al.*, 2003) in hGST A1-1. The more loosely packed and destabilised hydrophobic core of domain 2 can give reason for the increased solvent-accessible surface area of the transition state ensemble along the unfolding reaction coordinate of S154A compared to that of the wild-type (Wallace *et al.*, 1998a; Nathaniel *et al.*, 2003). The increased rate of unfolding for the fast phase is best explained where a more loosely packed and destabilised core of domain 2 directly impacts on the stability of the domain 1-domain 2 interface rendering a less cooperative unfolding reaction with respect to the solvent environment (Wallace *et al.*, 1998a; Nathaniel *et al.*, 2003).

5.7 Implications of α-helix 6 in the folding of hGST A1-1

5.7.1 Folding kinetics of S154A

The folding pathway has been described for wild-type hGST A1-1 (Wallace and Dirr, 1999). The fast folding reaction of S154A was monitored by fluorescence spectroscopy on a stopped-flow instrument, described in section 3.13.1. The intrinsic Trp and Tyr fluorescence of hGST A1-1 are useful probes in providing sufficient structural information of this protein. Thus, an excitation wavelength of 280 nm with a 320 nm cut-off filter was used for this study. The folding reaction of hGST A1-1 was initiated where 6 μM of unfolded protein (in 6 M urea) was diluted 6-fold. This resulted in a final concentration of 1 μM protein in 1 M urea. All the refolding reactions were carried out at 15°C.

Figure 5.9 shows a refolding trace of both the wild-type and S154A in a final concentration of 1 M urea at 15°C. The refolding trace for both was triphasic. No burst-phase was observed for the folding reactions of both proteins, i.e. the total amplitude change in fluorescence signal of the unfolded and refolded protein was observed. The folding reaction of wild-type was followed by an initial increase in fluorescence signal (20 sec), and thereafter a decrease in signal was followed (200 - 1000 sec) reaching an end-point that corresponds to the fluorescence of native-folded hGST A1-1. For S154A, the refolding reaction was followed by an initial increase in fluorescence (0 - 100 sec), similar to that of the wild-type, however only an increase in fluorescence (200 - 1000 sec) was observed for S154A. Nonetheless the end-point of the reaction was equivalent in fluorescence signal to that of the native S154A. For both proteins the initial increase in signal was best fitted to a double-exponential growth, which describes the rates of the first two folding phases, i.e. the fast and intermediate phases of hGST A1-1 (Wallace and Dirr, 1999). The decrease in fluorescence signal observed for the wild-type (200 to 1000 sec) was best fitted to a single-exponential decay and has been described as the slow folding phase (Wallace *et al.*, 1998b).

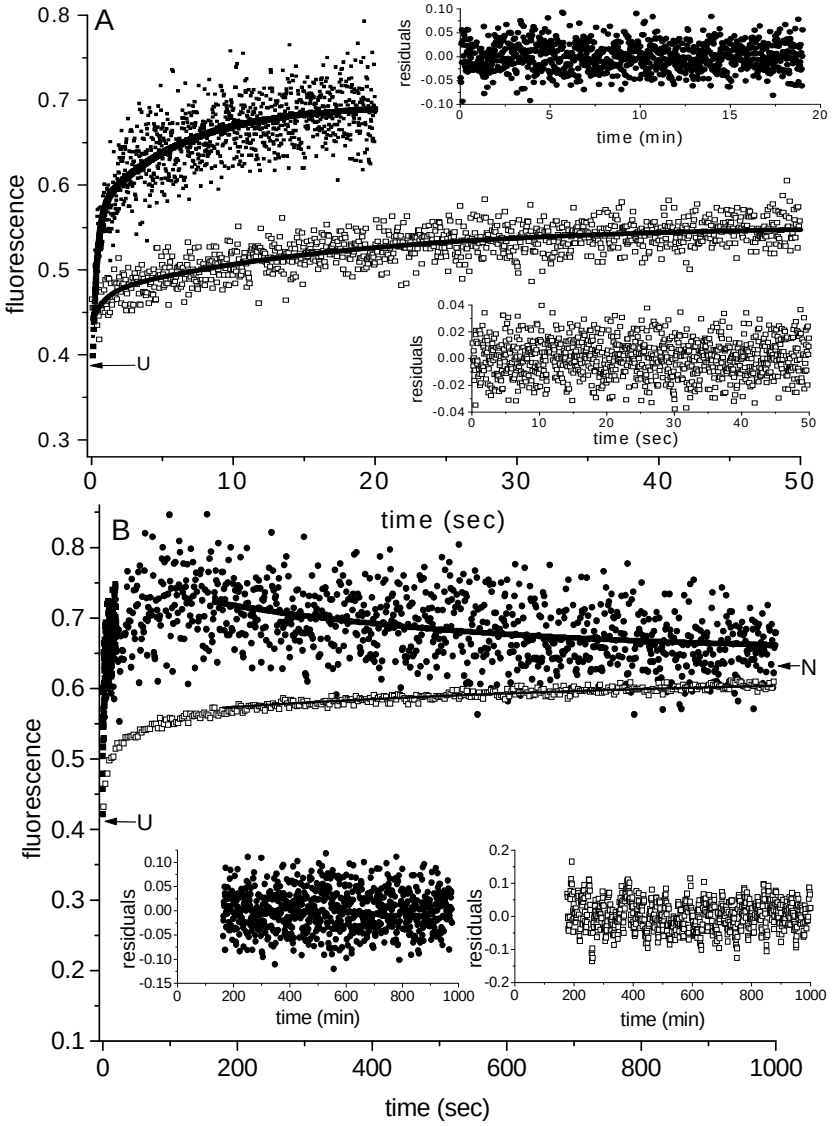

Figure 5.9: Folding kinetics of S154A and wild-type hGST A1-1 folded in a final concentration of 1 M urea and 1 μM protein at 15°C, pH 6.5. The folding of the proteins were monitored by fluorescence (excitation at 280 nm, monitored using a 320 nm cut-off filter). Panel A shows that the best fit to a double-exponential function (residuals are shown in the insets) for the fast and intermediate folding phases of wild-type (•) and S154A (□). Panel B shows the slow folding phase best fitted to a single-exponential decay (residuals are shown in the insets). S154A and the wild-type share the same fluorescence signal for native (N) and unfolded (U) protein as indicated by the arrows.

The slow folding phase involves the final tertiary packing and organisation of the dimeric protein leading to the native conformation (Wallace *et al.*, 1998b). Even though the reverse in polarity in fluorescence was not detected for the slow folding phase of S154A (Figure 5.9), this folding event was also best fitted to a single-exponential decay, as for the wild-type. The apparent folding rates for the respective folding phases are reported in Table 5.3.

Any occurrences of enhanced light scattering due to aggregate formation or mixing artifacts, for the respective refolding reactions, was monitored by the fluorescence emission intensity signal. An excitation wavelength of 340 nm was selected, using a 320 nm emission cut-off filter. For wild-type the absence of an increase in light scattering indicated that no aggregate formation and mixing artifacts was apparent, whereas the non-reproducible increase in light scattering observed during the refolding reaction of S154A suggested that aggregate formation was occurring to some extent (Figure 5.10). The kinetic traces of S154A monitored by fluorescence emission were reproducible for all the folding phases.

The standard deviation of the apparent refolding rates reported for the fast and intermediate folding phases of S154A were comparable to those of the wild-type. However a large standard deviation was apparent for the observed rate constant of the slow phase of S154A compared to that of the wild-type. This increase in standard deviation may be explained by the non-reproducible slow increase in light scattering observed for S154A, possibly as a result of some kinetic partition occurring during this phase, such as aggregate formation. The extent to which kinetic partitioning occurs is dependent upon the magnitude of the rates of the reactions. Intuitively, the partitioning of a reaction is unlikely to occur in the case where the two folding rates are very different, for instance a faster folding reaction would contribute predominantly to the observed rate of the reaction in the presence of a much slower reaction rate. Whereas the partitioning of two reactions of similar folding rates would contribute similarly to the observed rate. In the case of S154A, in the event of a slow rate of aggregation, this reaction is likely to contribute to the observed rate of the slow folding phase, moreover than to the observed rates of the fast and intermediate folding phases.

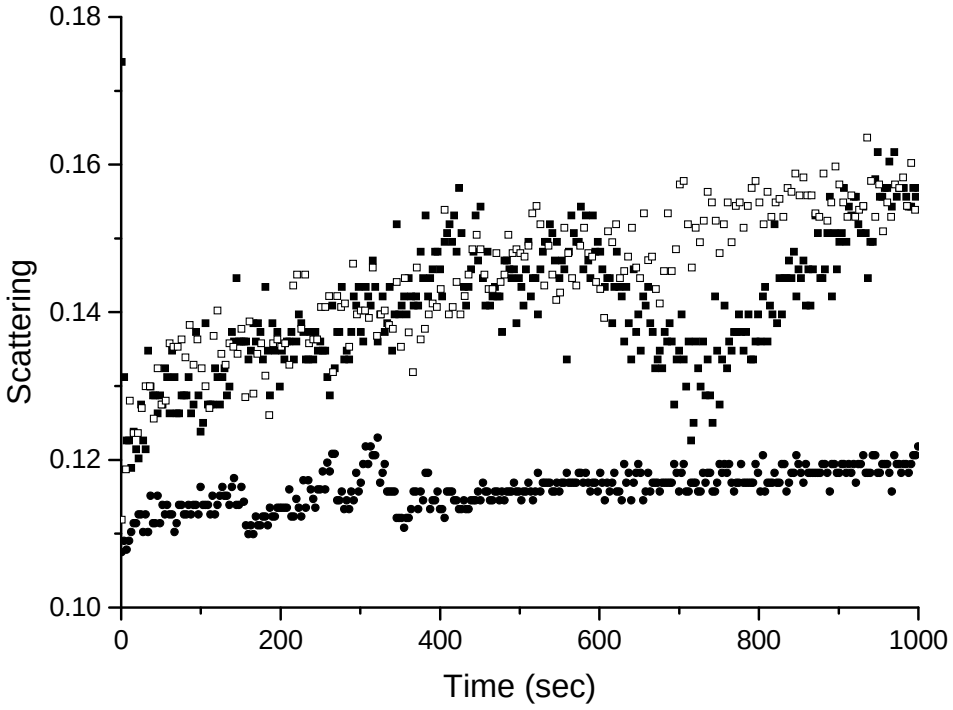

Figure 5.10: Fluorescence scattering monitored during the refolding of S154A (□,■) and wild-type (•) hGST A1-1. Protein was refolded in a final concentration of 1 M urea at 15°C, pH 6.5. The fluorescence scattering was monitored using a 320 nm emission cut-off filter, with an excitation of wavelength of 340 nm).

Table 5.3: Apparent refolding rates for the respective folding phases of wild-type and S154A hGST A1-1. A 6-fold dilution of unfolded protein was used to initiate the folding reaction. The refolding reaction consisted of a final concentration of 1 μM protein in 1 M urea at 15°C.

	Fast	Intermediate	slow
	k_{app} (sec^{-1})	k_{app} (sec^{-1})	k_{app} (sec^{-1})
wild-type	3.7 ± 0.2	0.4 ± 0.01	3.7 ± 0.3 x 10^{-4}
S154A	0.35 ± 0.05	0.029 ± 0.01	3.5 ± 2.2 x 10^{-4}

The increase in fluorescence amplitude of the fast (0.09 ± 0.02 units) and intermediate (0.14 ± 0.07 units) folding phases of S154A were significantly less compared to that of the fast (0.27 ± 0.04 units) and intermediate folding phases of the wild-type. The decrease in fluorescence during the slow folding phase of wild-type had an amplitude of 0.25 ± 0.05 units. Whereas the increase in fluorescence observed for the slow folding phase of S154A had an amplitude of 0.06 ± 0.06 units. The apparent refolding rate of the fast and intermediate folding phases of S154A were 18 and 14-fold lower compared to those of the wild-type, respectively. These reduced rates did not affect the fraction of recovered native protein observed for S154A (under strong refolding conditions, 0.5 - 1 M urea), as judged by the enzyme's specific activity. The reduced apparent rate and fluorescence amplitude of the fast and intermediate folding phases of S154A compared to that of the wild-type, suggests the importance of the N-cap residue playing a significant role during the fast and intermediate folding phases of hGST A1-1. The apparent rates of the slow folding phase of S154A and wild-type are similar. This indicates the insignificance of the stabilised N-cap motif of α-helix 6 during this folding phase of hGST A1-1.

Since the mechanism of the fast and intermediate refolding phases of wild-type hGST A1-1 involves the association of partially folded monomeric chains (Wallace and Dirr, 1999), the influence of protein concentration on the apparent refolding rate was investigated for both of these folding phases. Figure 5.11 shows the progressive linear dependence of the observed rates for the fast and intermediate folding phases with increasing concentration of S154A (0.8 - 2 μM final concentration protein in 1 M urea at 15°C). The value derived from this linear dependance is the second-order rate constant (k_f) of the rate of association of monomers forming dimeric hGST A1-1 (Wallace and Dirr, 1999). The k_f for the fast and intermediate phases of S154A were 4.0 x 10^4 $M^{-1}s^{-1}$ and 8.7 x 10^3 $M^{-1}s^{-1}$, respectively (Figure 5.11). For wild-type, the k_f for the fast and intermediate phases has been reported to be 5.7 x 10^5 $M^{-1}s^{-1}$ and 4.7 x 10^4 $M^{-1}s^{-1}$, respectively (Wallace and Dirr, 1999). In general the majority of the second-order rate constants, measured for subunit association on assembly folding pathways of proteins, are in the range of 10^3 to 10^6 $M^{-1}s^{-1}$ (Price, 1994; Jaenicke, 1987). The k_f for the intermediate phase for both proteins is

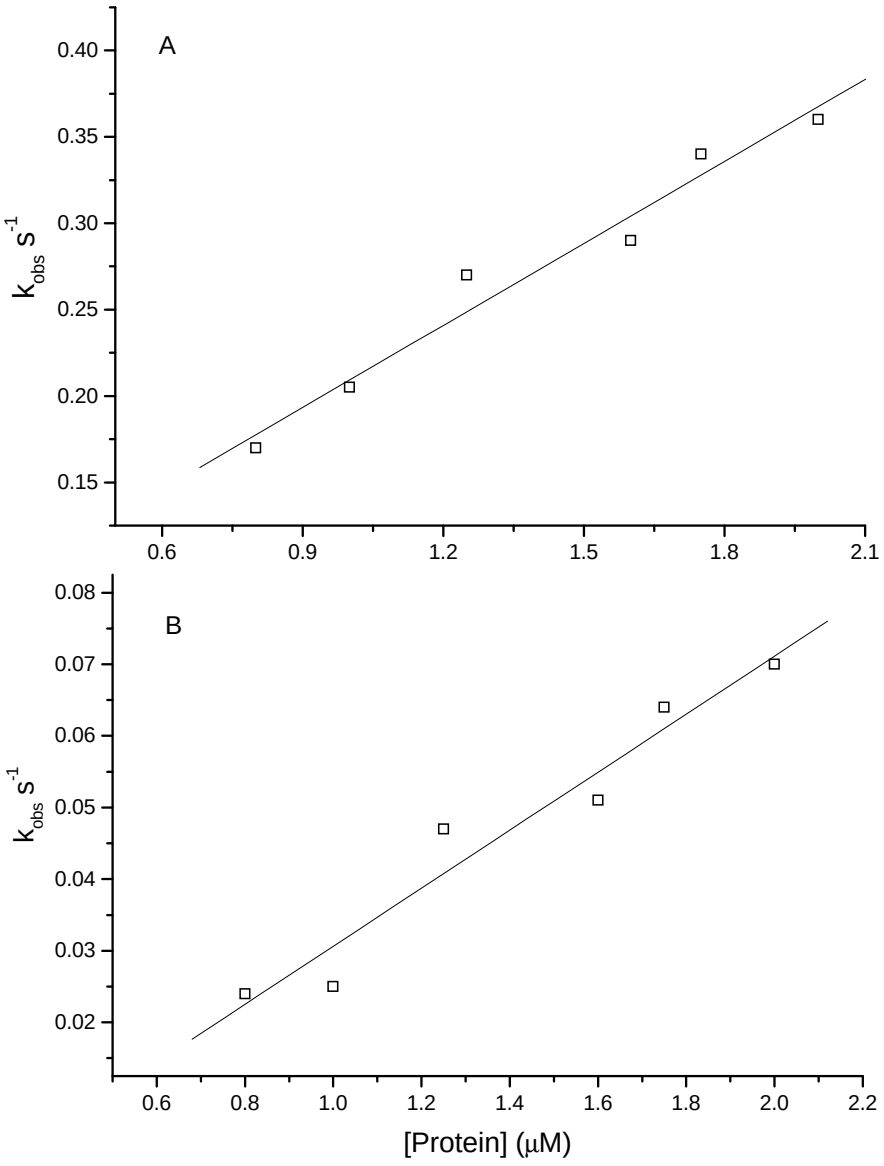

Figure 5.11: Protein concentration-dependence of the observed rate of refolding of S154A for the (A) fast and (B) intermediate phases. The straight line represents the best fit according to the equation, $(1/\tau) = k_{app} = 4[\text{Protein}]k_f$, where k_f is the second-order rate constant. The folding reactions were measured at 15°C in 1 M urea.

10-fold slower compared to that of the fast phase. The intermediate and fast folding phases for hGST A1-1 are parallel folding events involving the folding and dimerisation of the preassembled monomeric species. The intermediate phase is the kinetically slower event involving the slow isomerisation of X-Pro-56 peptide bond from a *trans* to *cis* configuration (Wallace *et al.*, 1998b).

The variation in second-order rate constants for different proteins has been noted to be tightly coupled with the association reaction and the extent to which the folding of the monomers occurs (Zitzewitz *et al.*, 1995). The rate of association approaches the diffusion limit in the presence of substantially folded monomers and the rates are 2-3 orders of magnitude smaller when the monomers are disordered (Chaffotte *et al.*, 1997). Thus, the larger the fraction of folded monomer or partially folded monomer the greater the observed bi-molecular rate constant. The 10-fold lower rate of association during the intermediate folding phase compared to that of the fast phase could be related to the tight coupling in the rate of association to the X-Pro-56 isomerisation step. The X-Pro-56 bond in the native *cis* configuration is suggested to allow for a rapid association, possibly due to a more native-like folded monomer structure than those folding conformers with the X-Pro-56 bond in the *trans* configuration (Wallace *et al.*, 1998b). Thus, an explanation for the reduced rate of association of monomers during the fast and intermediate folding phases of S154A compared to that of the wild-type, can be given in that there is a reduced fraction of folded monomeric species of S154A formed during the earlier phases of folding for hGST A1-1.

The rapid folding of S154A monitored by far-UV shows that 82% helical structure forms within 16 sec of the refolding reaction (Fig. 5.12), where the dead-time (unobserved reaction during the mixing of the experiment) is 10 sec. This is consistent with that observed for wild-type protein, where 85% helical structure forms within the burst-phase of the folding reaction, i.e. within the dead-time of the reaction being 10 ms (Wallace and Dirr, 1999). No burst-phase was observed for the folding of S154A and wild-type monitored by fluorescence, whereas a significant burst-phase was detected by far-UV. These results demonstrate the insensitivity of fluorescence as a structural probe for monitoring the formation of partially structured monomers (Wallace *et al.*,

1998b).

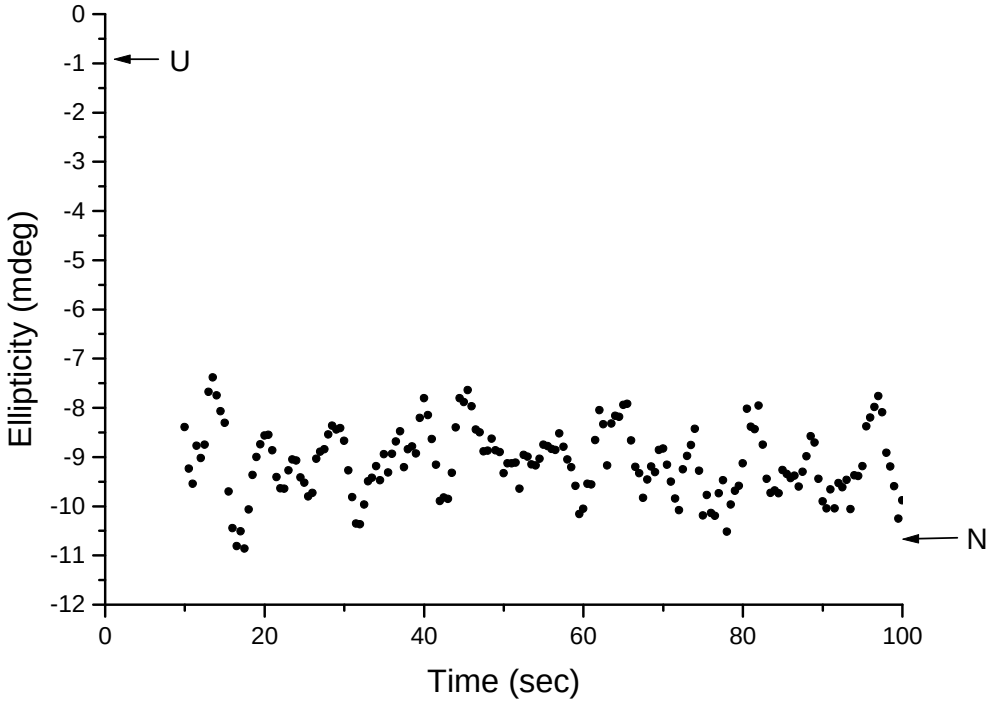

Figure 5.12: Folding of S154A monitored by far-UV CD at 15°C. The protein (1 μM final concentration) was refolded in a final concentration of 1 M urea at 15°C, pH 6.5. The dead-time of the folding experiment was 10 sec.

A more detailed study of the earliest folding events of hGST A1-1 was not possible, as in previous studies (Wallace *et al.*, 1998b), since the data obtained from far-UV CD had a low signal-to-noise ratio, as well as the fast mixing times were insufficient to observe the earlier more rapid formation of secondary structure. Whether the reduced rates of the fast and intermediate folding phases were attributed to the disruption or decrease in the formation of misfolded *vs* partially folded monomers remains to be resolved (Wallace *et al.*, 1998b). Nonetheless, since 85% of the secondary structure was formed during the burst phase of the reaction of S154A, which is comparable to the wild-type, this does provide some evidence that the N-cap motif is not an essential feature for the sufficient and rapid formation secondary structural elements. Thus, it is unlikely that this motif is a folding nucleation site for the formation of the α-helix 6.

The absence of a burst-phase monitored by fluorescence, and the similar structural content of the rapid formation of secondary structure reported by far-UV CD suggests no differences in the rapid folding of S154A compared to that of the wild-type, i.e. within the mixing time of the experiments. As mentioned previously, the rapid formation of well structured monomeric species for S154A was similar to that of the wild-type. Whereas the rate of dimerisation and tertiary packing of the subunits of S154A were significantly reduced compared to that of the wild-type. This is likely a result of an unstable loop preceding α-helix 6 in S154A, as suggested from thermal denaturation and unfolding kinetic studies mentioned in the previous sections. This structural motif appears to be an important feature for the tight packing of folded monomers which occurs during the earlier folding phases of hGST A1-1. Moreover, the unstabilised/misfolded monomers of S154A formed during folding, is likely a kinetic trap and hence gives rise to aggregate formation.

5.7.2 Folding kinetics of L164A

Even though the N-cap motif does play a significant role towards the thermal stability of the protein (section 5.4), the specific role of this motif in the folding

mechanism of hGST A1-1 remains unclear. Since α-helix 6 is an important structural feature for the core packing hGST A1-1, the question is whether the N-cap motif plays a role in the folding mechanism of hGST A1-1 as a specific local structural feature, that is exclusive to its role in stabilising α-helix 6 in domain 2. Leu-164 in α-helix 6 is known to play a role in the tight packing in the core of hGST A1-1 through hydrophobic contacts. The refolding of L164A, where the Leu was substituted with an Ala resulting in the removal of three methylene groups, was investigated to gain more insight as to the involvement of α-helix 6 in the folding of hGST A1-1.

The refolding kinetics L164A was carried out using the same conditions and fitting parameters as described for the wild-type and S154A, see section 5.7.1. The kinetic trace of L164A was the same as the wild-type, which was triphasic (Fig. 5.13), where the three folding phases are described as the fast, intermediate and slow phases. The m_f-value derived from the results in Figure 5.14 for the fast (891 $\pm$ 25 J.mol^{-1} M^{-1} urea) and intermediate (962 $\pm$ 75 J.mol^{-1} M^{-1} urea) phase of L164A are comparable to that of the wild-type, being 753 $\pm$ 25 J.mol^{-1} M^{-1} urea for the fast phase and 1054 $\pm$ 46 J.mol^{-1} M^{-1} urea for the intermediate phase. These comparable m_f-values gives evidence that the transition state of the respective folding phases are similar for both L164A and the wild-type. The derived $k_f(H_20)$ for the fast (7.8 $\pm$ 0.42 s^{-1}) and intermediate (1.2 $\pm$ 0.2 s^{-1}) folding phases of wild-type were 1.5-fold greater compared to that derived for L164A, being 6.2 $\pm$ 0.8 s^{-1} for the fast phase and 1.7 $\pm$ 0.1 s^{-1} for the intermediate phase. This fold difference between wild-type and L164A is insignificant. The $k_f(H_20)$ and m_f-values of the fast and intermediate folding events of L164A are similar to that of the wild-type. This suggests that the rate of dimerisation and the formation of structured monomers of L164A follows a similar folding mechanism to the wild-type. The $k_f(H_20)$ of the slow folding phase of L164A (5.0 $\pm$ 0.8 x 10$^-$3 s^{-1}) is similar to that of the wild-type (5.8 $\pm$ 0.7 x 10$^-$3 s^{-1}). However, the reported m_f-value of the slow phase of L164A (753 $\pm$ 87 J.mol^{-1} M^{-1}) is 1.8-fold greater compared to that of the wild-type. This increase in m_f-value suggests that the transition state ensemble of the slow folding phase of L164A has an increased solvent-accessible surface area compared to that of the wild-type. This is suggestive of a more

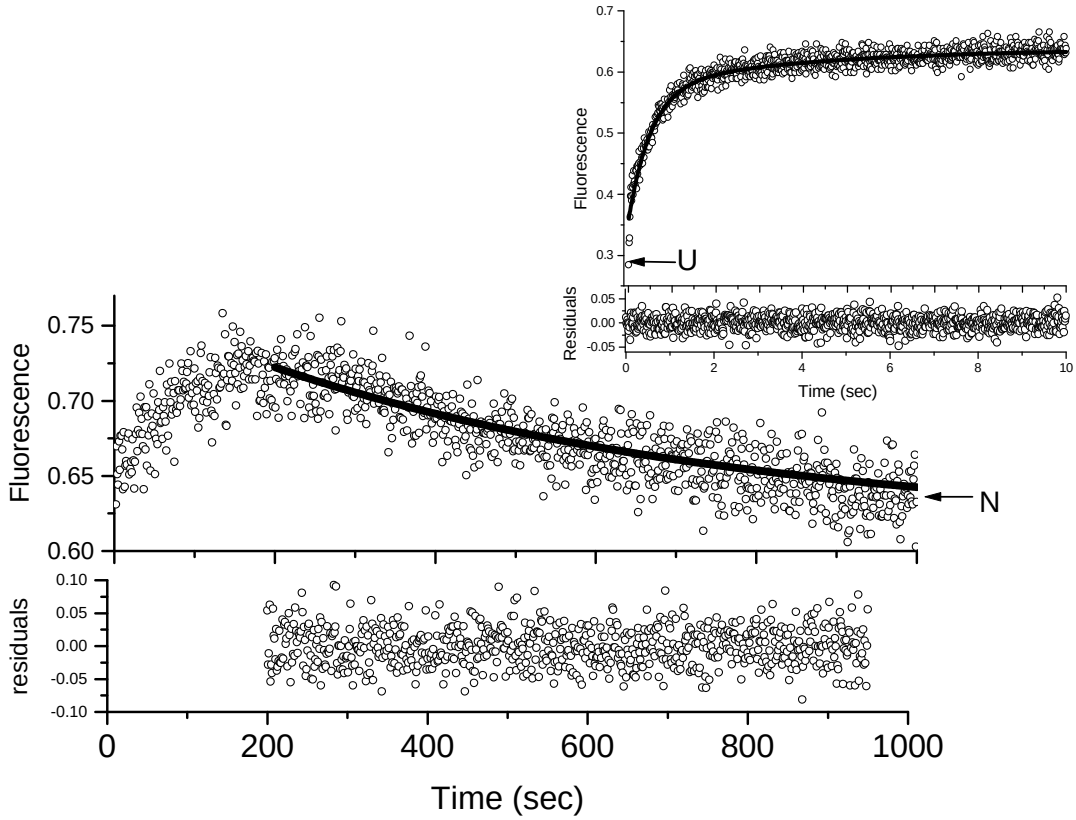

Figure 5.13: Kinetic trace of L164A showing the folding of L164A at 15°C. Samples consisted of final concentrations of 1 μM protein in 1 M urea, pH 6.5. The folding reaction were monitored by fluorescence (excitation 280 nm, with a cut-off 320 nm filter). The slow folding phase was fitted best to a single-exponential decay (residuals are shown), whereas the fast and intermediate folding phases were best fitted to a double-exponential function (inset). The arrows indicate the fluorescence signal for native (N) and unfolded (U) hGST A1-1.

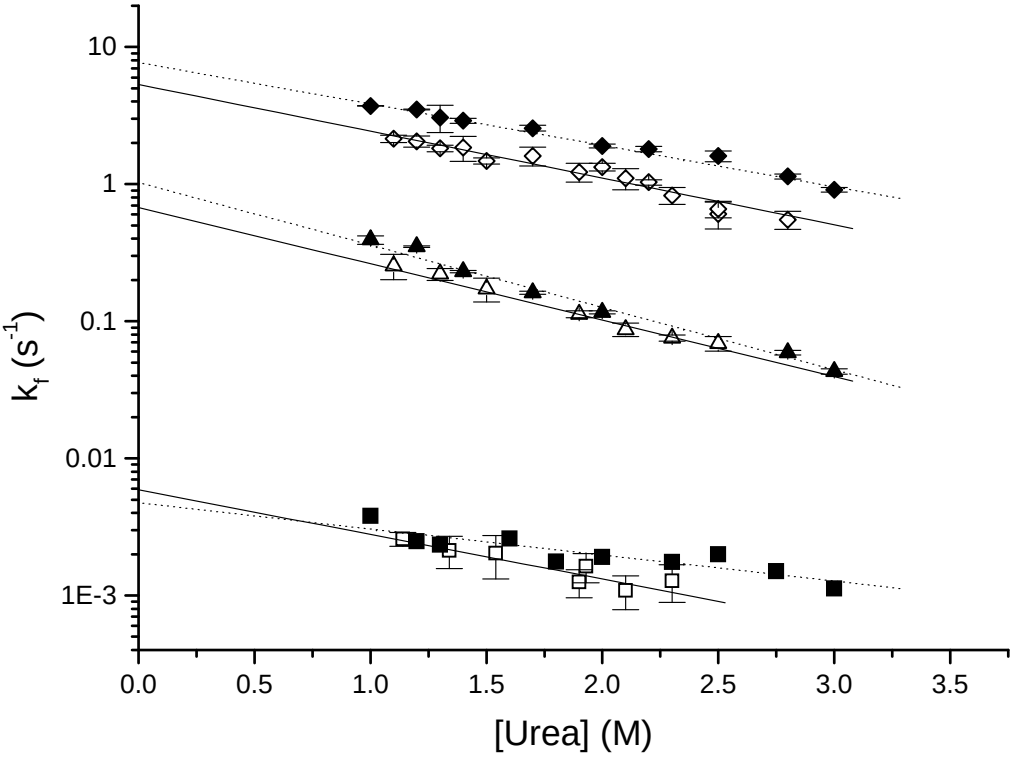

Figure 5.14: Chevron plot representing the linear dependence of the logarithmic rate of the fast (♦,◇), intermediate (▲,△) and slow folding phases (■,□) of L164A (open symbols) and wild-type (closed symbols) hGST A1-1 at 15°C. Samples consisted of 1 μM protein, pH 6.5.

loosely packed transition state, being the isomerisation reaction involving the final tight packing of the tertiary structure of the protein. From these results it can be said that the three methylene groups positioned at 164 in α-helix 6 participates in facilitating the final and tight packing of the tertiary structure of hGST A1-1 and less likely involved in the rapid formation and dimerisation of the structured monomers.

5.8 Role of hydrophobic core packing surrounding α-helix 6 in hGST A1-1 folding

From the kinetic refolding studies carried out in this study, the role of the three methylene groups of Leu-164 in α-helix 6 appears not to contribute to the rate of the folding of hGST A1-1. However, this residue appears to contribute significantly towards the tight packing of the hydrophobic core during the overall reorganisation of the tertiary interactions and tight packing of the hGST A1-1. This is reasoned since the transition state ensemble of the slow folding phase of L164A is more loosely packed compared to that of the wild-type. These results can shed further light into the mechanism of folding of hGST A1-1. From crystallographic data of hGST A1-1, the Leu-164 forms close van de Waals contacts with residues Leu-182 and Lys-185 (from α-helix 7) and Phe-133 (α-helix 5), as well as more distant contacts with Leu-181 (α-helix 7). α-Helix 7 is an amphipathic helix forming contacts with the interior of the protein, i.e. α-helix 6, α-helix 5 and bulk solvent (Figure 5.15). Including the above mentioned contacts, α-helix 5 also forms contacts with Phe-52 of the second subunit which forms the "lock-and-key" motif. This motif is located at the dimer interface where the Phe-52 forms the "key" to this motif, as shown in Figure 5.15, and is known to contribute to the stabilisation of the dimeric protein. Since the hydrophobic core packing surrounding the Leu-164 in α-helix 6 contributes to the overall tertiary packing of the dimeric protein, it seems reasonable to assume that this packing includes those structural features that Leu-164 forms close contacts with, being α-helix-7 and α-helix 5. However, since the interior packing encompassing α-helix 6 does not involve

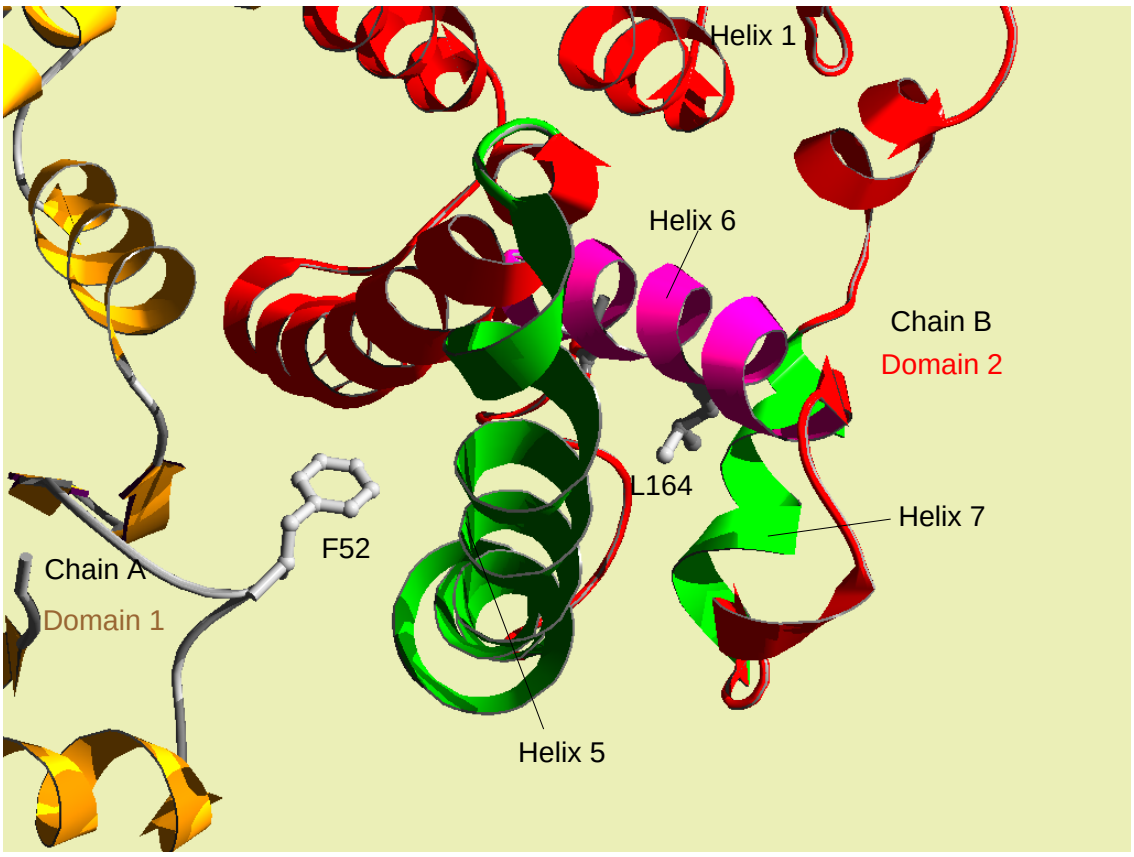

Figure 5.15: Ribbon representation of the structure of hGST A1-1. Viewing down the two-fold axis of the dimeric protein (Le Trong *et al.*, 2002), chain A is shown yellow, and chain B in red. Each subunit consists of an N-terminal domain (domain 1) connected to the C-terminal domain (domain 2). Leu-164 in helix 6 (in domain 2, chain B) and Phe-52 (in domain 1, chain A) are represented as ball-and-stick. The helices highlighted in green are α-helix 5 and 7 in domain 2, and α-helix 6 is shown in maroon. The figure was generated using Swiss-PdbViewer (Guex and Peitsch, 1997).

the "lock-and-key" motif, it is reasonable to assume that the role of Leu-164 is involved in the tight packing of α-helix-7, and to a lesser extent α-helix 5. This is based on the observation that the truncation of Leu-164 did not impact upon the folding and dimerisation of the monomers. α-Helix 5 forms contacts with the "lock-and-key" motif and α-helix 7 forms part of the exterior surface of the protein distant from the dimer interface, forming minimal contacts with α-helix 6 and α-helix 5.

5.9 Role of α-helix 6 N-capping motif in the folding of hGST A1-1

In this study and in previously published results, the conservation of the N-cap motif in the core of domain 2 has been shown to contribute to increased stability, and increased rate in folding of hGST A1-1. α-Helix 6 is within the core of domain 2 of the protein forming contacts with all neighbouring helices in domain 2 and forms interfacing interactions with α-helix 1 (Figure 5.15). It is therefore a central structural feature in the hydrophobic core of GSTs. It has been proposed that α-helix 1 in domain 1 is the least stabilised structural unit of GSTs (Aceto *et al.*, 1995; Dragani *et al.*, 1998), and that the structured α-helix 6 in domain 2, which forms contacts with the α-helix 1, participates in specifying the packing and stabilisation of the N-terminal end of domain 1 (Stenberg *et al.*, 2000). The hydrophobic core packing surrounding Leu-164 in α-helix 6 is shown to occur during the tight packing and reorganisation of tertiary interactions in hGST A1-1, whereas a stabilised N-cap motif in α-helix 6 appears to facilitate the rate of folding and dimerisation of the monomers, and does not contribute toward the final tight packing of the dimeric structure. The rate of folding and dimerisation of the monomers, as well as the Pro-56 isomerisation reaction are all tightly coupled folding events. The detailed mechanism into how the N-capping motif specifies the structuring of domain 2, and hence as a cooperative folding event with the isomerisation of X-Pro-56 and dimerisation of the monomers would require a more detailed study. Nonetheless, it is evident that the role of the N-capping region of α-helix 6

in the folding of hGST A1-1 is an important structural feature for the tertiary packing and dimerisation of the monomers of hGST A1-1, and to a much lesser extent, in contributing toward the final interior packing of the protein. Since the N-capping motif in α-helix 6 is not an essential structural feature in specifying the formation of helical structures, which are formed during the burst-phase of the reaction, this clearly indicates that this motif is not a folding nucleation site for hGST A1-1. Previous results have reported that the hydrophobic staple motif at the N-terminal end of α-helix 6, which excludes the involvement of hydrogen bonding, also contributes to increased rates of folding for both class Pi and Alpha GSTs (Stenberg *et al.*, 2000; Cocco *et al.*, 2001). Therefore, both the helical content and/or stability of the α-helix 6, defined by the N-capping and staple motif, contributes to the fast folding of both class Alpha and Pi GSTs. Importantly, it has also been shown that the conserved Gly-150 in the loop preceding α-helix 6, which also forms part of the same conserved folding module as the N-capping motif (Kong *et al.*, 2003), is also an important structural feature for the fast folding of class Pi GSTs. In conclusion it seems that the extensive hydrogen bond network described in section 2.3.1 and the N-capping and staple motif of the N-capping region of α-helix 6, all represent a universal motif that plays a critical role in the folding and stability of GST P1-1 and GST A1-1, and possibly for all other proteins homologous to the GST superfamily.

Chapter 6

Concluding remarks

In this study the helical content of the peptide, corresponding to the C-terminal region of hGST A1-1 (residues 208 - 222), and the stabilising role of the N-cap motif was correctly predicted by AGADIR. The significant stabilising effect of the N-cap motif in the α-helix 9 peptide, did not translate directly within context of the protein. In the absence of this N-capping motif, the stability and dynamics of the α-helix 9 was only marginally destabilised. An explanation for this is that the non-local stabilising tertiary interactions predominate the overall dynamics and stability of this amphipathic helix. These results are surprising and highlights the careful balance of local and non-local interactions play in contributing to the function and stability of proteins. Although one can carefully weigh out the context dependent nature of a single amino acid replacement by applying the helix-coil helical algorithm for protein design, the outcome remains largely context dependant. Thus, the specific ratio between local versus non-local interactions contributing to the stability of secondary structures and hence the overall conformation and stability of a protein appears to be inherently defined, as suggested by Munoz *et al.* (1996).

If the rate of protein folding and its stability is under some evolutionary pressure, it is possible to identify conserved regions representing key structural features that may be critical for the folding and structural stability of that particular fold. Work by Mirny and Shakhnovich (2001) have related folding nucleations site as being conserved features, however this does not mean that

all folding nucleation sites are conserved. Larger proteins may have multiple folding nuclei, i.e. possibly located separately from each domain. The conservation of Pro-56 in domain 2, i.e. the thioredoxin fold and an N-capping motif located in the GST-like fold, suggests for multiple folding nucleation sites of proteins homologous to GSTs. However, kinetic refolding studies in this work show that the N-capping motif is not an essential nor critical structural feature for the rapid formation of structured monomers. The same conclusion has been made for Pro-56 residue (Kelley and Richards, 1987; Kelley *et al.*, 1986; Galani *et al.*, 2002; Yu *et al.*, 2000), which represents the conserved Pro in the thioredoxin fold (Nathaniel *et al.*, 2003). It seems unlikely that the evolutionary conserved N-cap motif in domain 2 of GSTs serves as a folding nucleation site for GSTs, but rather it is an important structural feature in maintaining the conformational stability and the efficient folding of GSTs.

References

Abdalla, A. M., Bruns, C. M., Tainer, J. A., Mannervik, B. and Stenberg, G. (2002). Design of a monomeric human glutathione transferase GSTP1, a structurally stable but catalytically inactive protein. *Protein. Eng.* 15, 827–834.

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.

Aceto, A., Dragani, B., Melino, S., Allocati, N., Masulli, M., Ilio, C. D. and Petruzzelli, R. (1997). Identification of an N-capping box that affects the alpha 6-helix propensity in glutathione S-transferase superfamily proteins: a role for an invariant aspartic residue. *Biochem. J.* 322, 229–234.

Aceto, A., Sacchetta, P., Bucciarelli, T., Angelucci, S., Radatti, G. L. and Ilio, C. D. (1995). Structural and functional properties of the 34-kDa fragment produced by the N-terminal chymotryptic cleavage of glutathione transferase P1-1. *Arch. Biochem. Biophys.* 316, 873–878.

Adman, E. T., T Le, I., Stenkamp, R. E., Nieslanik, B. S., Dietze, E. C., Tai, G., Ibarra, C. and Atkins, W. M. (2001). Localization of the C-terminus of rat glutathione S-transferase A1-1: Crystal structure of mutants W21F and W21F/F220Y. *Proteins* 42, 192–200.

Alexander, P., Orban, J. and Bryan, P. (1992). Kinetic analysis of folding and unfolding the 56 amino acid IgG-binding domain of streptococcal protein G. *Biochemistry* 31, 7243–7248.

Allardyce, C. S., McDonagh, P. D., Lian, L. Y., Wolf, C. R. and Roberts, G. C. (1999). The role of tyrosine-9 and the C-terminal helix in the catalytic

mechanism of Alpha-class glutathione S-transferases. *Biochem. J.* 343, 525–531.

Allocati, N., Casalone, E., Masulli, M., Caccarelli, I., Carletti, E., Parker, M. W. and Ilio, C. D. (1999). Functional analysis of the evolutionary conserved proline 53 residue in Proteus mirabilis glutathione transferase B1-1. *FEBS Lett.* 445, 347–350.

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389–3402.

Arca, P., Reguera, G. and Hardisson, C. (1997). Plasmid-encoded fosfomycin resistance in bacteria isolated from the urinary tract in a multicentre survey. *J. Am. Chem. Soc.* 40, 393–399.

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

Armstrong, R. N. (1994). Glutathione S-transferases: structure and mechanism of an archetypical detoxication enzyme. *Adv. Enzymol. Relat. Areas Mol. Biol.* 69, 1–44.

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

Atkins, W. M., Dietze, E. C. and Ibarra, C. (1997). Pressure-dependent ionization of Tyr 9 in glutathione S-transferase A1-1: contribution of the C-terminal helix to a "soft" active site. *Protein Sci.* 6, 873–881.

Aurora, R. and Rose, G. D. (1998). Helix capping. *Protein Sci.* 7, 21–38.

Aurora, R., Srinivasan, R. and Rose, G. D. (1994). Rules for alpha-helix termination by glycine. *Science* 264, 1126–1130.

Baldwin, R. L. (1989). How does protein folding get started? *Trends Biochem. Sci.* 14, 291–294.

Baldwin, R. L. (1999). Protein folding from 1961-1982. *Nat. Struct. Biol.* 6, 814–817.

Baldwin, R. L. (2002). Making a network of hydrophobic clusters. *Science* 295, 1657–1658.

Baldwin, R. L. and Rose, G. D. (1999). Is protein folding hierarchic? I. Local structure and peptide folding. *Trends Biochem. Sci.* 24, 26–33.

Barnwell, D. C. (2004). Folding and aggregation of thioredoxin-like proteins in the presence of macromolecular crowders. Thesis (ph. d.), University of Witwatersrand.

Barycki, J. J. and Colman, R. F. (1997). Identification of the nonsubstrate steroid binding site of rat liver glutathione S-transferase, isozyme 1-1, by the steroid affinity label, 3beta-(iodoacetoxy)dehydroisoandrosterone. *Arch. Biochem. Biophys.* 345, 16–31.

Bell, J. A., Becktel, W. J., Sauer, U., Baase, W. A. and Matthews, B. W. (1992). Dissection of helix capping in T4 lysozyme by structural and thermodynamic analysis of six amino acid substitutions at Thr 59. *Biochemistry* 31, 3590–3596.

Bernat, B. A., Laughlin, L. T. and Armstrong, R. N. (1997). Fosfomycin resistance protein (FosA) is a manganese metalloglutathione transferase related to glyoxalase I and the extradiol dioxygenases. *Biochemistry* 36, 3050–3055.

Betz, S. F., Bryson, J. W. and DeGrado, W. F. (1995). Native-like and structurally characterized designed alpha-helical bundles. *Curr. Opin. Struct. Biol.* 5, 457–463.

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

Bilsel, O. and Matthews, C. R. (2000). Barriers in protein folding reactions. *Adv. Protein Chem.* 53, 153–207.

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

Bjornestedt, 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.

Blaber, M., Zhang, X. J. and Matthews, B. W. (1993). Structural basis of amino acid alpha helix propensity. *Science* 260, 1637–1640.

Blanco, F., Ramirez-Alvarado, M. and Serrano, L. (1998). Formation and stability of B-hairpin structures in polypeptides. *Curr. Opin. Struct. Biol.* 8, 107–111.

Board, P. G., Coggan, M., Chelvanayagam, G., Easteal, S., Jermiin, L. S., Schulte, G. K., Danley, D. E., Hoth, L. R., Griffor, M. C., Kamath, A. V., Rosner, M. H., Chrunyk, B. A., Perregaux, D. E., Gabel, C. A., Geoghegan, K. F. and Pandit, J. (2000). Identification, characterization, and crystal structure of the Omega class glutathione transferases. *J. Biol. Chem.* 275, 24798–24806.

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

Bonneau, R. and Baker, D. (2001). Ab initio protein structure prediction: progress and prospects. *Annu. Rev. Biophys. Biomol. Struct.* 30, 173–189.

Bousset, L., Belrhali, H., Melki, R. and Morera, S. (2001). Crystal structures of the yeast prion Ure2p functional region in complex with glutathione and related compounds. *Biochemistry* 40, 13564–13573.

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

Brandts, J. F. and Kaplan, L. J. (1973). Derivative spectroscopy applied to tyrosyl chromophores. Studies on ribonuclease, lima bean inhibitors, insulin, and pancreatic trypsin inhibitor. *Biochemistry* 12, 2011–2024.

Brokx, R. D., Lopez, M. M., Vogel, H. J. and Makhatadze, G. I. (2001). Energetics of target peptide binding by calmodulin reveals different modes of binding. *J. Biol. Chem.* 276, 14083–14091.

Brown, J. E. and Klee, W. A. (1971). Helix-coil transition of the isolated amino terminus of ribonuclease. *Biochemistry* 10, 470–476.

Bruns, C. M., Hubatsch, I., Ridderstrom, M., Mannervik, B. and Tainer, J. A. (1999). Human glutathione transferase A4-4 crystal structures and mutagenesis reveal the basis of high catalytic efficiency with toxic lipid peroxidation products. *J. Mol. Biol.* 288, 427–439.

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.

Burton, R. E., Myers, J. K. and Oas, T. G. (1998). Protein folding dynamics: quantitative comparison between theory and experiment. *Biochemistry* 37, 5337–5343.

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., Goldberg, M. E. and Tasayco, M. L. (1997). Recognition between disordered states: kinetics of the self-assembly of thioredoxin fragments. *Biochemistry* 36, 16040–16048.

Chakrabartty, A., Kortemme, T. and Baldwin, R. L. (1994). Helix propensities of the amino acids measured in alanine-based peptides without helix-stabilizing side-chain interactions. *Protein Sci.* 3, 843–852.

Chan, C. K., Hu, Y., Takahashi, S., Rousseau, D. L., Eaton, W. A. and Hofrichter, J. (1997). Submillisecond protein folding kinetics studied by ultrarapid mixing. *Proc. Natl. Acad. Sci. U. S. A* 94, 1779–1784.

Chan, H. S. and Dill, K. A. (1998). Protein folding in the landscape perspective: chevron plots and non-Arrhenius kinetics. *Proteins* 30, 2–33.

Chen, Y. W. and Fersht, A. R. (1994). Stability and solvation of Thr/Ser to Ala and Gly mutations at the N- cap of alpha-helices. *FEBS Lett.* 347, 304–309.

Chiti, F., Taddei, N., Webster, P., Hamada, D., Fiaschi, T., Ramponi, G. and Dobson, C. M. (1999). Acceleration of the folding of acylphosphatase by stabilization of local secondary structure. *Nat. Struct. Biol.* 6, 380–387.

Chou, P. Y. and Fasman, G. D. (1978). Empirical predictions of protein conformation. *Annu. Rev. Biochem.* 47, 251–276.

Chu, R.-A. and Bai, Y. (2002). Lack of definable nucleation sites in the rate-limiting transition state of barnase under native conditions1. *J. Mol. Biol.* 315, 759–770.

Clark, A. C., Raso, S. W., Sinclair, J. F., Ziegler, M. M., Chaffotte, A. F. and Baldwin, T. O. (1997). Kinetic mechanism of luciferase subunit folding and assembly. *Biochemistry* 36, 1891–1899.

Cocco, R., Stenberg, G., Dragani, B., Rossi, P. D., Paludi, D., Mannervik, B. and Aceto, A. (2001). The folding and stability of human alpha class glutathione transferase A1-1 depend on distinct roles of a conserved N-capping box and hydrophobic staple motif. *J. Biol. Chem.* 276, 32177–32183.

Cochran, D. A. and Doig, A. J. (2001). Effect of the N2 residue on the stability of the alpha-helix for all 20 amino acids. *Protein Sci.* 10, 1305–1311.

Cochran, D. A., Penel, S. and Doig, A. J. (2001). Effect of the N1 residue on the stability of the alpha-helix for all 20 amino acids. *Protein Sci.* 10, 463–470.

Creamer, T. P. and Rose, G. D. (1994). Alpha-helix-forming propensities in peptides and proteins. *Proteins* 19, 85–97.

Daggett, V. and Fersht, A. R. (2003). Is there a unifying mechanism for protein folding? *Trends Biochem. Sci.* 28, 18–25.

Danger, D. P., Baldwin, W. S. and LeBlanc, G. A. (1992). Photoaffinity labelling of steroid-hormone-binding glutathione S-transferases with [3H]methyltrienolone. Inhibition of steroid-binding activity by the anticarcinogen indole-3-carbinol. *Biochem. J.* 288, 361–367.

Dasgupta, S. and Bell, J. A. (1993). Design of helix ends. Amino acid preferences, hyrdogen bonding and electrostatic interactions. *Int. J. Pept. Protein Res.* 41, 499–511.

de Brevern, A. G., Valadie, H., Hazout, S. and Etchebest, C. (2002). Extension of a local backbone description using a structural alphabet: a new approach to the sequence-structure relationship. *Protein Sci.* 11, 2871–2886.

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

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

Dill, K. A., Fiebig, K. M. and Chan, H. S. (1993). Cooperativity in protein-folding kinetics. *Proc. Natl. Acad. Sci. U. S. A* 90, 1942–1946.

Dinner, A. R. and Karplus, M. (1998). A metastable state in folding simulations of a protein model. *Nature Struct. Biol.* 5, 236–241.

Dirr, H. W. (2001). Folding and assembly of glutathione transferases. *Chem. Biol. Interact.* 133, 19–23.

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

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

Dixon, D. P., Cummins, L., Cole, D. J. and Edwards, R. (1998). Glutathione-mediated detoxification systems in plants. *Curr. Opin. Plant. Biol.* 1, 258–266.

Dobson, C. M. (1992). Unfolded proteins, compact states and molten globules. *Curr. Opin. Struct. Biol.* 2, 6–12.

Doig, A. J., MacArthur, M. W., Stapley, B. J. and Thornton, J. M. (1997). Structures of N-termini of helices in proteins. *Protein Sci.* 6, 147–155.

Dragani, B., Iannarelli, V., Allocati, N., Masulli, M., Cicconetti, M. and Aceto, A. (1998). Irreversible thermal denaturation of glutathione transferase P1-1. Evidence for varying structural stability of different domains. *Int. J. Bioch. Cell Biol.* 30, 155–163.

Dragani, B., Stenberg, G., Melino, S., Petruzzelli, R., Mannervik, B. and Aceto, A. (1997). The conserved N-capping box in the hydrophobic core of glutathione S-transferase P1-1 is essential for refolding. *J. Biol. Chem.* 272, 25518–25523.

Eaton, W. A., Munoz, V., Hagen, S. J., Jas, G. S., Lapidus, L. J., Henry, E. R. and Hofrichter, J. (2000). Fast kinetics and mechanisms in protein folding. *Annu. Rev. Biophys. Biomol. Struct.* 29, 327–359.

Eaton, W. A., Muoz, V., Thompson, P. A., Chan, C. and Hofrichter, J. (1997). Submillisecond kinetics of protein folding. *Curr. Opin. Struct. Biol.* 7, 10–14.

Edwards, R., Dixon, D. P. and Walbot, V. (2000). Plant glutathione S-transferases: enzymes with multiple functions in sickness and in health. *Trends Plant Sci.* 5, 193–198.

elMasry, N. F. and Fersht, A. R. (1994). Mutational analysis of the N-capping box of the alpha-helix of chymotrypsin inhibitor 2. *Protein Eng.* 7, 777–782.

Fedorov, A. N. and Baldwin, T. O. (1997). GroE modulates kinetic partitioning of folding intermediates between alternative states to maximize the yield of biologically active protein. *J. Mol. Biol.* 268, 712–723.

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

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

Fersht, A. R. (2000). Transition-state structure as a unifying basis in protein-folding mechanisms: contact order, chain topology, stability, and the extended nucleus mechanism. *Proc. Natl. Acad. Sci. U. S. A* 97, 1525–1529.

Fink, A. L. (1998). Protein aggregation: folding aggregates, inclusion bodies and amyloid. *Fold. Des.* 3, R9–23.

Forood, B., Feliciano, E. J. and Nambiar, K. P. (1993). Stabilization of alpha-helical structures in short peptides via end capping. *Proc. Natl. Acad. Sci. U. S. A* 90, 838–842.

Fowler, S. B. and Clarke, J. (2001). Mapping the Folding Pathway of an Immunoglobulin Domain. Structural Detail from Phi Value Analysis and Movement of the Transition State. *Science* 9, 355–366.

Franks, F. (1975). *The hydrophobic effect. In water: A comprehensive treatise.* Vol. 4. Plenum Press, New York.

Galani, D., Fersht, A. R. and Perrett, S. (2002). Folding of the yeast prion protein Ure2: kinetic evidence for folding and unfolding intermediates. *J. Mol. Biol.* 315, 213–227.

Gans, P. J., P C Lyu Manning, M., Woody, R. W. and Kallenbach, N. R. (1991). The helix-coil transition in heterogeneous peptides with specific side-chain interactions: Theory and comparison with CD spectral data. *Biopolymers* 31, 1605–1614.

Gay, G. D. P., Ruiz-Sanz, J., Neira, J. L., Itzhaki, L. S. and Fersht, A. R. (1995). Folding of a nascent polypeptide chain in vitro: cooperative formation of structure in a protein module. *Proc. Natl. Acad. Sci. U. S. A* 92, 3683–3686.

Gloss, L. M. and Matthews, C. R. (1998). Mechanism of folding of the dimeric core domain of Escherichia coli trp repressor: a nearly diffusion-limited reaction leads to the formation of an on-pathway dimeric intermediate. *Biochemistry* 37, 15990–15999.

Gloss, L. M., Simler, B. R. and Matthews, C. R. (2001). Rough energy landscapes in protein folding: dimeric E. coli Trp repressor folds through three parallel channels. *J. Mol. Biol.* 312, 1121–1134.

Greene, Jr. R. F. and Pace, C. N. (1974). Urea and guanidine hydrochloride denaturation of ribonuclease, lysozyme, alpha-chymotrypsin, and beta-lactoglobulin. *J. Biol. Chem.* 249, 5388–5393.

Gu, Y., Singh, S. V. and Ji, X. (2000). Residue R216 and catalytic efficiency of a murine class alpha glutathione S-transferase toward benzo[a]pyrene 7(R),8(S)-diol 9(S), 10(R)-epoxide. *Biochemistry* 39, 12552–12557.

Gu, Y., Xiao, B., Wargo, H. L., Bucher, M. H., Singh, S. V. and Ji, X. (2003). Residues 207, 216, and 221 and the catalytic activity of mGSTA1-1 and mGSTA2-2 toward benzo[a]pyrene-(7R,8S)-diol-(9S,10R)-epoxide. *Biochemistry* 42, 917–921.

Guex, N. and Peitsch, M. C. (1997). SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* 18, 2714–2723.

Gustafsson, A., Etahadieh, M., Jemth, P. and Mannervik, B. (1999). The C-terminal region of human glutathione transferase A1-1 affects the rate of glutathione binding and the ionization of the active-site Tyr9. *Biochemistry* 38, 16268–16275.

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

Hagen, S. J., Hofrichter, J., Szabo, A. and Eaton, W. A. (1996). Diffusion-limited contact formation in unfolded cytochrome c: estimating the maximum rate of protein folding. *Proc. Natl. Acad. Sci. U. S. A* 93, 11615–11617.

Harper, E. T. and Rose, G. D. (1993). Helix stop signals in proteins and peptide: the capping box. *Biochemistry* 32, 7605–7609.

Harrop, S. J., DeMaere, M. Z., Fairlie, W. D., Reztsova, T., Valenzuela, S. M., Mazzanti, M., Tonini, R., Qiu, M. R., Jankova, L., Warton, K., Bauskin, A. R., Wu, W. M., Pankhurst, S., Campbell, T. J., Breit, S. N. and Curmi, P. M. (2001). Crystal structure of a soluble form of the intracellular chloride ion channel CLIC1 (NCC27) at 1.4-A resolution. *J. Biol. Chem.* 276, 44993–45000.

Hayes, J. D. and Pulford, D. J. (1995). The glutathione S-transferase supergene family: regulation of GST and the contribution of the isoenzymes to cancer chemoprotection and drug resistance. *Crit. Rev. Biochem. Mol. Biol.* 30, 445–600.

Higo, J., Ito, N., Kuroda, M., Ono, S., Nakajima, N. and Nakamura, H. (2001). Energy landscape of a peptide consisting of alpha-helix, 3(10)-helix, beta-turn, beta-hairpin, and other disordered conformations. *Protein Sci.* 10, 1160–1171.

Horovitz, A., Matthews, J. M. and Fersht, A. R. (1992). Alpha-helix stability in proteins II Factors that influence stability at an internal position. *J. Mol. Biol.* 227, 560–568.

Hummer, G., Garcia, A. E. and Garde, S. (2001). Helix nucleation kinetics from molecular simulations in explicit solvent. *Proteins* 42, 77–84.

Ibarra, C., Nieslanik, B. S. and Atkins, W. M. (2001). Contribution of aromatic-aromatic interactions to the anomalous pK(a) of tyrosine-9 and the C-terminal dynamics of glutathione S-transferase A1-1. *Biochemistry* 40, 10614–10624.

Iqbalsyah, T. M. and Doig, A. J. (2004). Effect of the N3 residue on the stability of the alpha-helix. *Protein Sci.* 13, 32–39.

Itzhaki, L. S., Neira, J. L., Ruiz-Sanz, J., de Prat Gay, G. and Fersht, A. R. (1995a). Search for nucleation sites in smaller fragments of chymotrypsin inhibitor 2. *J. Mol. Biol.* 254, 289–304.

Itzhaki, L. S., Otzen, D. E. and Fersht, A. R. (1995b). The structure of the transition state for folding of chymotrypsin inhibitor 2 analysed by protein engineering methods: evidence for a nucleation-condensation mechanism for protein folding. *J. Mol. Biol.* 254, 260–288.

Jackson, S. E. (1998). How do small single-domain proteins fold? *Fold. Des.* 3, R81–91.

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

Jennings, P. A., Finn, B. E., Jones, B. E. and Matthews, C. R. (1993). A reexamination of the folding mechanism of dihydrofolate reductase from Escherichia coli: verification and refinement of a four-channel model. *Biochemistry* 32, 3783–3789.

Jimenez, M. A., Munoz, V. and Serrano, L. (1994). Helix stop and start signals in peptides and proteins. The capping box does not necessarily prevent helix elongation. *J. Mol. Biol.* 242, 487–496.

Jones, B. E., Jennings, P. A., Pierre, R. A. and Matthews, C. R. (1994). Development of nonpolar surfaces in the folding of Escherichia coli dihydrofolate reductase detected by 1-anilinonaphthalene-8-sulfonate binding. *Biochemistry* 33, 15250–15258.

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

Kauzmann, W. (1959). *Adv. Protein Chem.* 14, 763.

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

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

Ketley, J. N., Habig, W. H. and Jakoby, W. B. (1976). Binding of nonsubstrate ligands to the glutathione S-transferases. *J. Biol. Chem.* 250, 8670–8673.

Ketterer, B. (2001). A birds eye view of the glutathione transferase field. *Chem. Biol. Interact.* 138, 27–42.

Kim, M. K. and Kang, Y. K. (1999). Positional preference of proline in alpha-helices. *Protein Sci.* 8, 1492–1499.

Kirk, W. R., Kurian, E. and Prendergast, F. G. (1996). Characterization of the sources of protein-ligand affinity: 1-sulfonato-8-(1')anilinonaphthalene binding to intestinal fatty acid binding protein. *Biophys. J.* 70, 69–83.

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.

Kolobe, D., Sayed, Y. and Dirr, H. W. (2004). Characterisation of bromosulfophthalein binding to human glutathione transferase A1-1: thermodynamics and inhibition kinetics. *Biochem. J.* in print.

Kong, G. K., Polekhina, G., McKinstry, W. J., Parker, M. W., Dragani, B., Aceto, A., Paludi, D., Principe, D. R., Mannervik, B. and Stenberg, G. (2003). Contribution of glycine 146 to a conserved folding module affecting stability and refolding of human glutathione transferase P1-1. *J. Biol. Chem.* 278, 1291–1302.

Kraulis, P. J. (1991). MOLSCRIPT: A Program to Produce Both Detailed and Schematic Plots of Protein Structures. *J. App. Crystallogr.* 24, 946–950.

Krengel, U., Schroter, K. H., Hoier, H., Arkema, A., Kalk, K. H., Zimniak, P. and Dijkstra, B. W. (1998). Crystal structure of a murine alpha-class glutathione S-transferase involved in cellular defense against oxidative stress. *FEBS Lett.* 422, 285–290.

Kumaran, S. and Roy, R. P. (1999). Helix-enhancing propensity of fluoro and alkyl alcohols: influence of pH, temperature and cosolvent concentration on the helical conformation of peptides. *J. Pept. Res.* 53, 284–293.

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

Lacroix, E., Viguera, A. R. and Serrano, L. (1998). Elucidating the folding problem of alpha-helices: local motifs, long-range electrostatics, ionic-strength dependence and prediction of NMR parameters. *J. Mol. Biol.* 284, 173–191.

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

Larson, S. M., Ruczinski, I., Davidson, A. R., Baker, D. and Plaxco, K. W. (2002). Residues participating in the protein folding nucleus do not exhibit preferential evolutionary conservation. *J. Mol. Biol.* 316, 225–233.

Li, L., Mirny, L. A. and Shakhnovich, E. I. (2000). Kinetics, thermodynamics and evolution of non-native interactions in a protein folding nucleus. *Nat. Struct. Biol.* 7, 336–342.

Li, L. and Shakhnovich, E. I. (2001). Different circular permutations produced different folding nuclei in proteins: a computational study. *J. Mol. Biol.* 306, 121–132.

Lian, L. Y. (1998). NMR structural studies of glutathione S-transferase. *Cell Mol. Life Sci.* 54, 359–362.

Listowsky, I., Abramovitz, M., Homma, H. and Niitsu, Y. (1988). Intracellular binding and transport of hormones and xenobiotics by glutathione-S-transferases. *Drug. Metab. Rev.* 19, 305–318.

Lockhart, D. J. and Kim, P. S. (1992). Internal stark effect measurement of the electric field at the amino terminus of an alpha helix. *Science* 257, 947–951.

Lockhart, D. J. and Kim, P. S. (1993). Electrostatic screening of charge and dipole interactions with the helix backbone. *Science* 260, 198–202.

Loladze, V. V., Ermolenko, D. N. and Makhatadze, G. I. (2001). Heat capacity changes upon burial of polar and nonpolar groups in proteins. *Protein Sci.* 10, 1343–1352.

Lopez-Hernandez, E., Cronet, P., Serrano, L. and Munoz, V. (1997). Kinetics of CheY mutants with enhanced native alpha-helix propensities. *J. Mol. Biol.* 266, 610–620.

Lopez-Hernandez, E. and Serrano, L. (1996). Structure of the transition state for folding of the 129 aa protein CheY resembles that of a smaller protein, CI-2. *Fold. Des.* 1, 43–55.

Luo, J. K., Hornby, J. A., Wallace, L. A., CHEN, J., Armstrong, R. N. and Dirr, H. (2002). Impact of domain interchange on conformational stability and equilibrium folding of chimeric class glutathione transferases. *Protein Sci.* 11, 2208–2217.

Luque, I., Mayorga, O. L. and Freire, E. (1996). Structure-based thermodynamic scale of alpha-helix propensities in amino acids. *Biochemistry* 35, 13681–13688.

Lyu, P. C., Liff, M. I., Marky, L. A. and Kallenbach, N. R. (1990). Side chain contributions to the stability of alpha-helical structure in peptides. *Science* 250, 669–673.

Lyu, P. C., Wemmer, D. E., Hongxing, X. Z., Pinker, P. J. and Kallenbach, N. R. (1993). Capping interactions in isolated α-helices: position-dependant substitution effects and structure of a serine-capped peptide helix. *Biochemistry* 32, 421–425.

Maness, S. J., Franzen, S., Gibbs, A. C., Causgrove, T. P. and Dyer, R. B. (2003). Nanosecond temperature jump relaxation dynamics of cyclic beta-hairpin peptides. *Biophys. J.* 84, 3874–3882.

Mannervik, B. (1985). The isoenzymes of glutathione transferase. *Adv. Enzymol. Relat. Areas Mol. Biol.* 57, 357–417.

Mannervik, B., Awasthi, Y. C., Board, P. G., Hayes, J. D., Ilio, C. D., Ketterer, B., Listowsky, I., Morgenstern, R., Muramatsu, M., Pearson, W. R. and et al. (1992). Nomenclature for human glutathione transferases. *Biochem. J.* 282, 305–306.

Mannervik, B. and Danielson, U. H. (1988). Glutathione transferases structure and catalytic activity. *CRC Crit. Rev. Biochem.* 23, 283337.

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

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

Matouschek, A., Kellis, Jr. J. T., Serrano, L., Bycroft, M. and Fersht, A. R. (1990). Transient folding intermediates characterized by protein engineering. *Nature* 346, 440–445.

Matthews, B. W. (1993). Structural and Genetic Analysis of Protein Stability. *Annual Review of Biochemistry* 62, 139–160.

McLellan, L. I. and Wolf, C. R. (1999). Glutathione and glutathione-dependent enzymes in cancer drug resistance. *Drug. Resist. Updat.* 2, 153–164.

McTigue, M. A., 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.

Mirny, L. and Shakhnovich, E. (2001). Evolutionary conservation of the folding nucleus. *J. Mol. Biol.* 308, 123–129.

Mirny, L. A., Abkevich, V. I. and Shakhnovich, E. I. (1998). How evolution makes proteins fold quickly. *Proc. Natl. Acad. Sci. U. S. A* 95, 4976–4981.

Mirny, L. A. and Shakhnovich, E. I. (1999). Universally conserved positions in protein folds: reading evolutionary signals about stability, folding kinetics and function. *J. Mol. Biol.* 291, 177–196.

Moran, L. B., Schneider, J. P., Kentsis, A., Reddy, G. A. and Sosnick, T. R. (1999). Transition state heterogeneity in GCN4 coiled coil folding studied by using multisite mutations and crosslinking. *Proc. Natl. Acad. Sci. U. S. A* 96, 10699–10704.

Morozova-Roche, L. A., Jones, J. A., Noppe, W. and Dobson, C. M. (1999). Independent nucleation and heterogeneous assembly of structure during folding of equine lysozyme. *J. Mol. Biol.* 289, 1055–1073.

Mosebi, S., Sayed, Y., Burke, J. and Dirr, H. W. (2003). Residue 219 impacts on the dynamics of the C-terminal region in glutathione transferase A1-1: implications for stability and catalytic and ligandin functions. *Biochemistry* 42, 15326–15332.

Munoz, V., Blanco, F. J. and Serrano, L. (1995). The hydrophobic-staple motif and a role for loop-residues in α-helix stability and protein folding. *Nat. Struct. Biol.* 2, 380–385.

Munoz, V., Cronet, P., Lopez-Hernandez, E. and Serrano, L. (1996). Analysis of the effect of local interactions on protein stability. *Fold. Des.* 1, 167–178.

Munoz, V., Henry, E. R., Hofrichter, J. and Eaton, W. A. (1998). A statistical mechanical model for beta-hairpin kinetics. *Proc. Natl. Acad. Sci. U. S. A* 95, 5872–5879.

Munoz, V. and Serrano, L. (1995a). Analysis of i,i+5 and i,i+8 hydrophobic interactions in a helical model peptide bearing the hydrophobic staple motif. *Biochemistry* 34, 15301–15306.

Munoz, V. and Serrano, L. (1995b). Helix design, prediction and stability. *Curr. Opin. Biotech.* 6, 382–386.

Munoz, V. and Serrano, L. (1996). Local versus nonlocal interactions in protein folding and stability–an experimentalist's point of view. *Fold. Des.* 1, R71–77.

Munoz, V., Thompson, P. A., Hofrichter, J. and Eaton, W. A. (1997). Folding dynamics and mechanism of beta-hairpin formation. *Nature* 390, 196–199.

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

Myers, J. K., Pace, C. N. and Scholtz, J. M. (1995). Denaturant m values and heat cpacaity changes: relation to changes in accessible surface areas of protein. *Protein Sci.* 4, 2138–2148.

Myers, J. K., Pace, C. N. and Scholtz, J. M. (1998). Trifluoroethanol effects on helix propensity and electrostatic interactions in the helical peptide from ribonuclease T1. *Protein Sci.* 7, 383–388.

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

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

Nieslanik, B. S. and Atkins, W. M. (2000). The catalytic Tyr-9 of Glutathione S-transferase A1-1 controls the dynamica of the c terminus. *J. Biol. Chem.* 275, 17447–17451.

Nieslanik, B. S., Dabrowski, M. J., Lyon, R. P. and Atkins, W. M. (1999a). Stopped-flow kinetic analysis of the ligand-induced coil-helix transition in glutathione S-transferase A1-1: evidence for a persistent denatured state. *Biochemistry* 38, 6971–6980.

Nieslanik, B. S., Dietze, E. C., Atkins, W. M., Trong, I. L. and Adman, E. (1999b). The locally denatured state of glutathione S-transferase A1-1: transition state analysis of ligand-dependent formation of the C- terminal helix. *Pac. Symp. Biocomput.* pp. 554–565.

Nieslanik, B. S., Ibarra, C. and Atkins, W. M. (2001). The C-terminus of glutathione S-transferase A1-1 is required for entropically-driven ligand binding. *Biochemistry* 40, 3536–3543.

Oakley, A. J., Bello, M. L., Nuccetelli, M., Mazzetti, A. P. and Parker, M. W. (1999). The ligandin (non-substrate) binding site of human Pi class glutathione transferase is located in the electrophile binding site (H-site). *J. Mol. Biol.* 291, 913–926.

O'Niel, K. T. and DeGrado, W. F. (1990). A thermodynamic scale for the helix-forming tendencies of the commonly occuring amino acids. *Science* 250, 646–651.

Onuchic, J. N., Socci, N. D., Luthey-Schulten, Z. and Wolynes, P. G. (1996). Protein folding funnels: the nature of the transition state ensemble. *Fold. Des.* 1, 441–450.

Onuchic, J. N., Wolynes, P. G., Luthey-Schulten, Z. and Socci, N. D. (1995). Toward an outline of the topography of a realistic protein-folding funnel. *Proc. Natl. Acad. Sci. U. S. A* 92, 3626–3630.

Pace, C. N. and Scholtz, J. M. (1998). A helix propensity scale based on experimental studies of peptides and proteins. *Biophys. J.* 75, 422–427.

Padmanabhan, S. and Baldwin, R. L. (1994). Helix-stabilizing interaction between tyrosine and leucine or valine when the spacing is i, i + 4. *J. Mol. Biol.* 241, 706–713.

Pal, L., Chakrabarti, P. and Basu, G. (2003). Sequence and structure patterns in proteins from an analysis of the shortest helices: implications for helix nucleation. *J. Mol. Biol.* 326, 273–291.

Parker, M. H. and Hefford, M. A. (1998). Introduction of potential helix-capping residues into an engineered helical protein. *Biotechnol. Appl. Biochem.* 28, 69–76.

Perkins, S. J. (1986). Protein volumes and hydration effects. The calculations of partial specific volumes, neutron scattering matchpoints and 280-nm absorption coefficients for proteins and glycoproteins from amino acid sequences. *Eur. J. Biochem.* 157, 169–180.

Perl, D., Welker, C., Schindler, T., Schroder, K., Marahiel, M. A., Jaenicke, R. and Schmid, F. X. (1998). Conservation of rapid two-state folding in mesophilic, thermophilic and hyperthermophilic cold shock proteins. *Nat. Struct. Biol.* 5, 229–235.

Peterson, R. W., Nicholson, E. M., Thapar, R., Klevit, R. E. and Scholtz, J. M. (1999). Increased helix and protein stability through the introduction of a new tertiary hydrogen bond. *J. Mol. Biol.* 286, 1609–1619.

Petukhov, M., Uegaki, K., Yumoto, N., Yoshikawa, S. and Serrano, L. (1999). Position dependence of amino acid intrinsic helical propensities II: non-charged polar residues: Ser, Thr, Asn, and Gln. *Protein Sci.* 8, 2144–2150.

Plaxco, K. W., Millett, I. S., Segel, D. J., Doniach, S. and Baker, D. (1999). Chain collapse can occur concomitantly with the rate-limiting step in protein folding. *Nat. Struct. Biol.* 6, 554–556.

Presta, L. G. and Rose, G. D. (1988). Helix signals in proteins. *Science* 240, 1632–1641.

Price, N. C. (1994). *Assembly of multisubunit structures.* Mechanisms of protein folding. Oxford University Press, Oxford.

Ptitsyn, O. B. (1998). Protein folding and protein evolution: common folding nucleus in different subfamilies of c-type cytochromes? *J. Mol. Biol.* 278, 655–666.

Ptitsyn, O. B. and Ting, K. L. (1999). Non-functional conserved residues in globins and their possible role as a folding nucleus. *J. Mol. Biol.* 291, 671–682.

Ranson, H., Prapanthadara, L. and Hemingway, J. (1997). Cloning and characterization of two glutathione S-transferases from a DDT-resistant strain of Anopheles gambiae. *Biochem. J.* 324, 97–102.

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

Richardson, J. S. and Richardson, D. C. (1988). Amino acid preferences for specific locations at the ends of a helices. *Science* 240, 1648–1652.

Robinson, C. R. and Sauer, R. T. (1996). Equilibrium stability and submillisecond refolding of a designed single-chain Arc repressor. *Biochemistry* 35, 13878–13884.

Rohl, C. A., Chakrabartty, A. and Baldwin, R. L. (1996). Helix propagation and N-cap propensities of the amino acids measured in alanine-based peptides in 40 volume percent trifluoroethanol. *Protein Sci.* 5, 2623–2637.

Rossjohn, J., McKinstry, W. J., Oakley, A. J., Parker, M. W., Stenberg, G., Mannervik, B., Dragani, B., Cocco, R. and Aceto, A. (2000). Structure of thermolabile mutants of human glutathione transferase P1-1. *J. Mol. Biol.* 302, 295–302.

Rowe, J. D., Patskovsky, Y. V., Patskovska, L. N., Novikova, E. and Listowsky, I. (1998). Rationale for reclassification of a distinctive subdivision of mammalian class Mu glutathione S-transferases that are primarily expressed in testis. *J. Biol. Chem.* 273, 9593–9601.

Sadqi, M., Lapidus, L. J. and Munoz, V. (2003). How fast is protein hydrophobic collapse? *Proc. Natl. Acad. Sci. U. S. A* 100, 12117–12122.

Sauder, J. M. and Roder, H. (1998). Amide protection in an early folding intermediate of cytochrome c. *Fold. Des.* 3, 293–301.

Sayed, Y. (2001). Biochemical and thermodynamic chracterisation of ligand-binding to human class alpha glutathione transferase A1-1. Thesis (ph. d.), University of Witwatersrand.

Sayed, Y., Hornby, J. A., Lopez, M. and Dirr, H. (2002). Thermodynamics of the ligandin function of human class Alpha glutathione transferase A1-1: energetics of organic anion ligand binding. *Biochem. J.* 363, 341–346.

Sayed, Y., Wallace, L. A. and Dirr, H. W. (2000). The hydrophobic lock-and-key intersubunit motif of glutathione transferase A1-1: implications for catalysis, ligandin function and stability. *FEBS Lett.* 465, 169–172.

Schindler, T. and Schmid, F. X. (1996). Thermodynamic properties of an extremely rapid protein folding reaction. *Biochemistry* 35, 16833–16842.

Scholtz, J. M. and Baldwin, R. L. (1992). The mechanism of alpha-helix formation by peptides. *Annu. Rev. Biophys. Biomol. Struct.* 21, 95–118.

Scholtz, J. M., Marqusee, S., Baldwin, R. L., York, E. J., Stewart, J. M., Santoro, M. and Bolen, D. W. (1991). Calorimetric determination of the enthalpy change for the alpha-helix to coil transition of an alanine peptide in water. *Proc. Natl. Acad. Sci. U. S. A* 88, 2854–2858.

Semisotnov, G. V., Rodionova, N. A., Razgulyaev, O. I., Uversky, V. N., Gripas, A. F. and Gilmanshin, R. I. (1991). Study of the "molten globule" intermediate state in protein folding by a hydrophobic fluorescent probe. *Biopolymers* 31, 119–128.

Serrano, L., Sancho, J., Hirshberg, M. and Fersht, A. R. (1992). α-helix stability in proteins I. Empirical correlations concerning substitution of side-chain at the N- and C-cap and the reaplacement of alanine by glycine or serine at solvent-exposed surfaces. *J. Mol. Biol.* 227, 544–559.

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

Shimada, J., Kussell, E. L. and Shakhnovich, E. I. (2001). The folding thermodynamics and kinetics of crambin using an all-atom monte carlo simulation1. *J. Mol. Biol.* 308, 79–95.

Shirley, B. A., Stanssens, P., Steyaert, J. and Pace, C. N. (1989). *J. Biol. Chem.* 264, 11621–11625.

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

Shortle, D. and Meeker, A. K. (1986). *Proteins Struct. Funct. Genet.* 1, 81–98.

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

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

Sobolev, V., Wade, R. C., Vriend, G. and Edelman, M. (1996). Molecular docking using surface complementarity. *Proteins* 25, 120–129.

Sosnick, T. R., Shtilerman, M. D., Mayne, L. and Englander, S. W. (1997). Ultrafast signals in protein folding and the polypeptide contracted state. *Proc. Natl. Acad. Sci. U. S. A* 94, 8545–8550.

Stark, G. R. (1965). Reactions of cyanate with functional groups of proteins. IV. Inertness of aliphatic hydroxyl groups. Formation of carbamyl- and acyl-hydantoins. *Biochemistry* 4, 2363–2367.

Stenberg, G., Bjrnestedt, R. and Mannervik, B. (1992). Heterologous expression of recombinant human glutathione transferase A1-1 from a hepatoma cell line. *Protein Expr. Puri.* 3, 80–84.

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

Stenberg, G., Dragani, B., Cocco, R., Mannervik, B. and Aceto, A. (2000). A conserved "hydrophobic staple motif" plays a crucial role in the refolding of human glutathione transferase P1-1. *J. Biol. Chem.* 275, 10421–10428.

Stevens, J. M., Hornby, J. A., Armstrong, R. N. and Dirr, H. W. (1998). Class sigma glutathione transferase unfolds via a dimeric and a monomeric intermediate: impact of subunit interface on conformational stability in the superfamily. *Biochemistry* 37, 15534–15541.

Stryer, L. (1965). The interaction of a naphthalene dye with apomyoglobin and apohemoglobin. A fluorescent probe of non-polar binding sites. *J. Mol. Biol.* 13, 482–495.

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

Tang, A. H. and Tu, C. P. (1994). Biochemical characterization of Drosophila glutathione S-transferases D1 and D21. *J. Biol. Chem.* 269, 27876–27884.

Tasayco, M. L., Fuchs, J., Yang, X. M., Dyalram, D. and Georgescu, R. E. (2000). Interaction between two discontiguous chain segments from the beta-sheet of Escherichia coli thioredoxin suggests an initiation site for folding. *Biochemistry* 39, 10613–10618.

Tew, K. D. (1994). Glutathione-associated enzymes in anticancer drug resistance. *Cancer. Res.* 54, 4313–4320.

Thompson, P. A., Eaton, W. A. and Hofrichter, J. (1997). Laser temperature jump study of the helix¡==¿coil kinetics of an alanine peptide interpreted with a 'kinetic zipper' model. *Biochemistry* 36, 9200–9210.

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

Le Trong, I., Stenkamp, R. E., Ibarra, C., Atkins, W. M. and Adman, E. T. (2002). 1.3-A resolution structure of human glutathione S-transferase with S-hexyl glutathione bound reveals possible extended ligandin binding site. *Proteins* 48, 618–627.

Tseng, Y. Y. and Liang, J. (2004). Are residues in a protein folding nucleus evolutionarily conserved? *J. Mol. Biol.* 335, 869–880.

Uversky, V. N. and Fink, A. L. (2002). The chicken-egg scenario of protein folding revisited. *FEBS Lett.* 515, 79–83.

van Gunsteren, W. F. and Berendsen, H. J. C. (1990). Computer Simulation of Molecular Dynamics: Methodology, Applications and Perspectives in Chemistry. *Angew. Chem. Int.* 29, 992–1023.

Velazquez-Campoy, A., Todd, M. J. and Freire, E. (2000). HIV-1 protease inhibitors: enthalpic versus entropic optimization of the binding affinity. *Biochemistry* 39, 2201–2207.

Viguera, A. R., Martinez, J. C., Filimonov, V. V., Mateo, P. L. and Serrano, L. (1994). Thermodynamic and kinetic analysis of the SH3 domain of spectrin shows a two-state folding transition. *Biochemistry* 33, 2142–2150.

Viguera, A. R. and Serrano, L. (1995). Experimental analysis of the Schellman motif. *J. Mol. Biol.* 251, 150–160.

Viguera, A. R. and Serrano, L. (1999). Stable proline box motif at the N-terminal end of alpha-helices. *Protein Sci.* 8, 1733–1742.

Viguera, A. R., Serrano, L. and Wilmanns, M. (1996a). Different folding transition states may result in the same native structure. *Nat. Struct. Biol.* 3, 874–880.

Viguera, A. R., Virtudes, V., Aviles, F. X. and Serrano, L. (1996b). Favourable native-like helical local interactions can accelerate protein folding. *Fold. Des.* 2, 23–33.

Vriend, G. (1990). WHAT IF: a molecular modeling and drug design program. *J. Mol. Graph.* 8, 52–56.

Wallace, A. C., Laskowski, R. A. and Thornton, J. M. (1995). LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Protein Eng.* 8, 127–134.

Wallace, L. A. (1998). The unfolding and refolding of human glutathione transferase A1-1. Thesis (ph. d.), University of the Witwatersrand.

Wallace, L. A., Blatch, G. L. and Dirr, H. W. (1998a). A topologically conserved aliphatic residue in alpha-helix 6 stabilizes the hydrophobic core in domain II of glutathione transferases and is a structural determinant for the unfolding pathway. *Biochem. J.* 336, 413–418.

Wallace, L. A., Burke, J. and Dirr, H. W. (2000). Domain-domain interface packing at conserved Trp-20 in class alpha glutathione transferase impacts on protein stability. *Biochim. Biophys. Acta* 1478, 325–332.

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

Wallace, L. A., Gildenhys, S. and Dirr, H. (in preparation). Dimeric Glutathione Transferases and Monomeric Glutaredoxin-2 share common folding mechanisms .

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

Wan, W. Y. and Milner-White, E. J. (1999). A natural grouping of motifs with an aspartate or asparagine residue forming two hydrogen bonds to residues ahead in sequence: their occurrence at alpha-helical N termini and in other situations. *J. Mol. Biol.* 286, 1633–1649.

Warren, J. R. and Gordon, J. A. (1966). On the refractive Indices of aqueous solutions of urea. *J. Phys. Chem.* 70, 297–300.

Werner, J. H., Dyer, R. B., Fesinmeyer, R. M. and Andersen, N. H. (2002). Dynamics of the primary processes of protein folding: helix nucleation. *J. Phys. Chem. B.* 106, 487494.

Widersten, M., Bjornestedt, 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., Bjornestedt, 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, J. A., Zeng, W., Rose, K., Craik, D. J. and Jackson, D. C. (1996). 1H NMR structural study of free and template-linked antigenic peptide representing the C-terminal region of the heavy chain of influenza virus hemagglutinin. *Biomed. Pept. Proteins Nuc. Acids* 2, 51–58.

Wilcock, D., Pisabarro, M. T., Lopez-Hernandez, E., Serrano, L. and Coll, M. (1998). Structure analysis of two CheY mutants: importance of the hydrogen-bond contribution to protein stability. *Acta Crystallogr. D. Biol Crystallogr.* 54, 378–385.

Williams, S., Causgrove, T. P., Gilmanshin, R., Fang, K. S., Callender, R. H., Woodruff, W. H. and Dyer, R. B. (1996). Fast events in protein folding: helix melting and formation in a small peptide. *Biochemistry* 35, 691–697.

Winder, A. F. and Gent, W. L. (1971). Correction of light-scattering errors in spectrophotometric protein determinations. *Biopolymers* 10, 1243–1251.

Wolynes, P., Luthey-Schulten, Z. and Onuchic, J. (1996). Fast-folding experiments and the topography of protein folding energy landscapes. *Chem. Biol.* 3, 425–432.

Wong, K. B., Clarke, J., Bond, C. J., Neira, J. L., Freund, S. M., Fersht, A. R. and Daggett, V. (2000). Towards a complete description of the structural and dynamic properties of the denatured state of barnase and the role of residual structure in folding. *J. Mol. Biol.* 296, 1257–1282.

Xia, B., Vlamis-Gardikas, A., Holmgren, A., Wright, P. E. and Dyson, H. J. (2001). Solution structure of Escherichia coli glutaredoxin-2 shows similarity to mammalian glutathione-S-transferases. *J. Mol. Biol.* 310, 907–918.

Yu, W. F., Tung, C. S., Wang, H. and Tasayco, M. L. (2000). NMR analysis of cleaved Escherichia coli thioredoxin (1-73/74-108) and its P76A variant: cis/trans peptide isomerization. *Protein Sci.* 9, 20–28.

Zhan, Y. and Rule, G. S. (2004). Glutathione induces helical formation in the carboxy terminus of human glutathione transferase A1-1. *Biochemistry* 43, 7244–7254.

Zhou, H. X., Lyu, P., Wemmer, D. E. and Kallenbach, N. R. (1994). alpha-helix capping in synthetic model peptides by reciprocal side chain-main-chain interactions:evidence for an N-terminal 'capping box'. *Proteins* 18, 1–7.

Zhukovsky, E. A., Mulkerrin, M. G. and Presta, L. G. (1994). Contribution to global protein stabilization of the N-capping box in human growth hormone. *Biochemistry* 33, 9856–9864.

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