Contribution of Glycine 146 to a Conserved Folding Module Affecting Stability and Refolding of Human Glutathione Transferase P1-1*

Received for publication, September 18, 2002, and in revised form, October 29, 2002

Geoffrey K-W. Kong‡‡‡, Galina Polekhina‡, William J. McKinstry‡‡, Beatrice Dragani‡, Antonio Aceto‡, Domenico Paludi‡, Daniela Rossi Principe‡, Bengt Mannervik‡‡‡, and Gun Stenberg‡‡‡‡§§§§

From the ‡Biota Structural Biology Laboratory, St. Vincent's Institute of Medical Research, Fitzroy, Victoria 3065, Australia, the §Russell Grimwade School of Biochemistry and Molecular Biology, The University of Melbourne, Victoria 3010, Australia, the ¶Dipartimento di Scienze Biomediche and the **Dipartimento di Scienze del Farmaco Università, G. D'Annunzio, Via dei Vestini, 31, I-66100 Chieti, Italy, and the ‡‡Department of Biochemistry, Uppsala University, Biomedical Center, Box 576, SE-751 23 Uppsala, Sweden

In human glutathione transferase P1-1 (hGSTP1-1) position 146 is occupied by a glycine residue, which is located in a bend of a long loop that together with the α6-helix forms a substructure (GST motif II) maintained in all soluble GSTs. In the present study G146A and G146V mutants were generated by site-directed mutagenesis in order to investigate the function played by this conserved residue in folding and stability of hGSTP1-1. Crystallographic analysis of the G146V variant, expressed at the permissive temperature of 25 °C, indicates that the mutation causes a substantial change of the backbone conformation because of steric hindrance. Stability measurements indicate that this mutant is inactivated at a temperature as low as 32 °C. The structure of the G146A mutant is identical to that of the wild type with the mutated residue having main-chain bond angles in a high energy region of the Ramachandran plot. However even this Gly → Ala substitution inactivates the enzyme at 37 °C. Thermodynamic analysis of all variants confirms, together with previous findings, the critical role played by GST motif II for overall protein stability. Analysis of reactivation in vitro indicates that any mutation of Gly-146 alters the folding pathway by favoring aggregation at 37 °C. It is hypothesized that the GST motif II is involved in the nucleation mechanism of the protein and that the substitution of Gly-146 alters this transient substructure. Gly-146 is part of the buried local sequence GXXh(T/S)XXDh (X is any residue and h is a hydrophobic residue), conserved in all GSTs and related proteins that seems to behave as a characteristic structural module important for protein folding and stability.

Numerous investigations over the last 40 years have focused on the folding and structural determinants that govern how a polypeptide adopts its native structure. Most of our knowledge on protein folding derives from studies on small monomeric proteins. However, the majority of native proteins are more complex structures, composed of several subunits that in turn consist of domains. The extrapolation of results obtained in the study of small proteins to larger ones is not always appropriate, and it is therefore important to investigate proteins composed of more than one subunit.

Human glutathione transferase P1-1 (hGSTP1-1),1 a homodimeric enzyme and thus representing the simplest type of oligomeric structure, has been the subject of recent studies on protein folding (1–3). The aim of these investigations was to identify sequence patterns of importance to folding kinetics and/or structure of the final native state. hGSTP1-1 is a Pi class member of a multifunctional superfamily of enzymes, the GSTs. The role of GSTs is considered to be the detoxication of a large number of hydrophobic compounds by catalyzing the conjugation to glutathione (GSH) and thus increase their water solubility (4). The cytosolic GSTs have been grouped into a number of different evolutionary classes denoted by the names of Greek letters such as Alpha, Mu, Pi, and so on (4, 5). The classification was originally based on primary structure similarities, substrate specificities, and immunologic properties (5), but sequence similarities are currently the overriding criterion. The mass of the dimeric GSTs is ∼50 kDa, and subunits within the same class can combine to form either homo- or heterodimers (6). Each subunit contains an active site, and two domains form the subunit. The smaller N-terminal domain (domain I) adopts an α/β topology and provides most of the contacts with GSH. The larger C-terminal domain (domain II) is completely helical and contains most of the residues that form the hydrophobic binding site (10–13). Alignment of all known GST structures (more than 100) shows that only 6–7 residues, which is less than 5% of the entire polypeptide chain, are conserved in all GSTs and related proteins. Consequently, the folding of GSTs is regarded as a characteristic example of protein folding.

1 The abbreviations used are: hGSTP1-1, human glutathione transferase P1-1; CDNB, 1-chloro-2,4-dinitrobenzene; GST, glutathione transferase; MES, 2-(N-morpholino)ethanesulfonic acid; Ncap, first residue of the N-capping box motif located at the beginning of an α-helix; NCS, non-crystallographic symmetry; r.m.s., root mean-square.
are strictly conserved. Despite this limited sequence similarity, all GSTs adopt the same native fold. GST structure analysis with special emphasis on the few conserved residues led to the identification of two local structural motifs in a characteristic position, the N-terminal region of the α-helix (1). The N-capping box [(S/T)/X(XD), and the hydrophobic staple motif (14–19) in which two hydrophobic residues flanking the N-capping box are present in all GSTs known today (1–3, 20). Previous investigations demonstrated that single mutations of residues forming the capping box [(S/T)/X(D) and the hydrophobic staple motifs have a dramatic effect on the protein stability (1–3). The same amino acid substitutions also have significant effects on protein folding, generating temperature-sensitive folding mutants unable to refold at the physiological temperature 37 °C (1–3). Mutations corresponding to those made in hGSTP1-1 have also been constructed in hGSTA1-1, a member of the Alpha class showing 31.6% sequence identity with hGSTP1-1 (21). The results obtained were in accordance with those found for hGSTP1-1, further emphasizing that the highly conserved N-capping box and hydrophobic staple motifs play critical and universal roles in GST folding and stability. The α-helix and preceding loop form a substructure, named GST motif II, which is conserved in the core of all GSTs and related proteins. The analysis of the protein crystal structure indicates that GST motif II is stabilized by a buried network of eight hydrogen bonds, half of which involve water-mediated contacts (23). Crystalllographic studies of the capping mutants, expressed at the permissive temperature of 25 °C, indicated that these amino acid substitutions locally destabilize GST motif II through a partial or complete loss of the hydrogen bond network (23). All these results indicate that a local destabilization of GST motif II has a critical effect on the overall protein stability and strongly support the hypothesis that this buried region might be involved in the nucleation mechanism of protein folding.

A strictly conserved glycine residue, Gly-146, in hGSTP1-1 is located four residues before Ser or Thr of the Ncap motif in all known GSTs (1). In particular Gly-146 is part of a buried local sequence GXaXh/T/S/XaXdh (X is any residue and h is a hydrophobic residue), which is maintained in all GSTs and, as a more general folding module, in other proteins such as EF1α and URE2. Its role has until now remained unexplored. This amino acid residue is located in a bend of the long loop preceding the α-helix and does not make any specific contacts with other structural parts of the molecule but only with neighboring residues in the polypeptide. Its strict conservation through evolution implicates the absolute necessity of a small amino acid residue in this position. The present investigation addresses the function of the strictly conserved Gly-146 in folding, and its significance for structural stability. By a combination of protein engineering and X-ray structure-function analysis we have obtained evidence that Gly-146 is part of a conserved folding module. As a universal motif it plays a critical role in the refolding and stability of all GSTs and, probably, of other structurally related proteins such as EF1α and URE2.

EXPERIMENTAL PROCEDURES

Materials—Wild-type human GSTP1-1 was obtained by expression of a cloned cDNA in Escherichia coli XL-1 Blue (Stratagene, La Jolla, CA) as previously described (24). GSH and 1-chloro-2,4-dinitrobenzene (CDNB) were purchased from Sigma. Oligonucleotides and dNTPs were obtained from Amersham Biosciences.

Construction of GSTP1-1 Mutants—Site-directed mutagenesis was employed to generate single-point mutants. The following oligonucleotides were used: 5’-CC ATT GTG AGC/TC/G CCC CAG-3’ and 5’-TCT TGG CTC CCT GCT TGT G-3’. The altered codon is underlined. The oligonucleotides were phosphorylated and then used in inverted polymerase chain reaction. The expression clone pKH1 (24) was used as a template. The polymerase chain reaction mixture contained 0.8 μM each primer, 0.2 mM dNTPs, 2.5 units of Pfu DNA polymerase (Stratagene, La Jolla, CA); the buffer supplied with the enzyme and various amounts of DNA template. The temperature program started at 94 °C for 10 min and was followed by 25 cycles of 94 °C for 1 min, 70 °C for 1 min, and 72 °C for 9 min. The program terminated with a reaction at 72 °C for 30 min. After electrophoresis the DNA product from the reaction was recovered from the agarose gel. The DNA was ligated and used for transformation of competent E. coli XL-1 Blue cells. The cDNAs encoding the isolated GSTP1-1 mutants were sequenced in their entirety to verify that no undesired mutations had been introduced in the polymerase chain reaction.

Protein Expression and Stability As a Function of Growth Temperature of the Host Cells— Cultures of E. coli XL-1 Blue-containing plasmids were grown in 300 ml of LB broth in a 1-liter Erlenmeyer flask at 37 °C. At an OD_{600} of 0.35, isoprpyl-1-thio-β-galactopyranoside was added to a final concentration of 0.2 mM. From the time of addition the culture was grown for about 20 h at 25 or 37 °C. The subsequent purification of wild-type and mutant proteins was performed as described previously (24) with the only modification that the enzymes were purified on a GSH-Sepharose affinity column (25).

Kinetic and Structural Studies of Wild-type hGSTP1-1 and Gly-146 Mutants Expressed at 25 °C—The kinetic parameters, k_{cat} and K_{m}, were determined at 25 °C as previously described (26). The concentrations of GSH and CDNB were 2 and 1 mM, respectively. Spectroscopic properties of the mutants and the wild-type enzyme were also studied. A Jasco 600 spectrophotometer was used for CD measurements in the far-ultraviolet region from 200 to 250 nm. Spectra were recorded using a protein concentration of 0.3 mg/ml with cuvettes of 0.1 cm path length in a thermostat-controlled cell holder. Intrinsic fluorescence emission spectra were measured with a Spex (model Fluoromax) spectrophotometer. The excitation wavelength was 280 nm and the λ_{max} and all emission spectra were analyzed at the same protein concentration (0.1 mg/ml).

Heat Inactivation Assays—Enzyme was incubated at each temperature for 10 min at a protein concentration of 0.05 mg/ml in 0.01 M potassium phosphate (pH 7.0) containing 1 mM EDTA and 5 mM dithiothreitol to prevent oxidative inactivation. The enzyme was heat-inactivated in sealed Eppendorf tubes, and the temperature was monitored with a Cryson telethermometer. The inactivation time courses were determined by withdrawing suitable aliquots at different time points from the denaturation mixture for assay of remaining activity. The activity was assayed in 0.1 M potassium phosphate (pH 6.5) with 2 mM GSH and 1 mM CDNB at 25 °C. Lowering the temperature of incubation could not reverse the thermal inactivation for any of the proteins.

Kinetics of Thermal Denaturation—The denaturation of the wild-type and glycine mutants was monitored at different temperatures. The enzymes were incubated in 10 mM potassium phosphate, pH 7.0, 1 mM EDTA, 5 mM dithiothreitol, and their activity was monitored for 120 min taking the first value as 100% native protein. An equation describing a single exponential decay with a rate constant of thermal unfolding k_u was fitted to the data according to Equation 1.

\[-ln(\% \text{ native}/100%) = k_u t \]  

(Eq. 1)

The free energy of activation of thermal unfolding (ΔG_u) was calculated according to Eyring theory (27) as Equation 2.

\[ \frac{ln k_u}{R} = \frac{ΔG_u}{RT} \]  

(Eq. 2)

where k_u is the Boltzmann constant; T, the absolute temperature in Kelvin; h, Planck’s constant; R, the gas constant; and K is the transmission factor, which was set to unity. The difference of free energy of activation of thermal denaturation between wild-type and each mutant protein (ΔG_u) was calculated according to Equation 3.

\[ ΔG_u = ΔG_u,\text{wt} - ΔG_u,\text{mut} = -RT \ln(k_u/wt/k_u/\text{mut}) \]  

(Eq. 3)

Substitution of Equation 4,

\[ ΔG_u = ΔH_u - TΔS_u \]  

(Eq. 4)

into Equation 2 yields Equation 5.

\[ ln\frac{k_u}{k_u/\text{WT}} = \frac{ΔH_u}{R} - \frac{T}{R} + \frac{ΔS_u}{R} \]  

(Eq. 5)

Both activation enthalpy ΔH_u and entropy ΔS_u were determined from the temperature dependence of k_u.
Temperature Dependence of Refolding in Vitro for Wild-type and Gly-146 Mutants—When the refolding of human GSTP1-1 and its mutants was to be monitored, 10 μl enzyme was first denatured in 4 M guanidinium chloride (0.2 M potassium phosphate, 1 mM EDTA, 5 mM dithiothreitol, pH 7.0) at 25, 33, and 40 °C for 30 min and then diluted (defining time 0) 1:40 into renaturation buffer (0.2 M potassium phosphate, 1 mM EDTA, 5 mM dithiothreitol, pH 7.0) at the same temperature. The final guanidinium chloride concentration was 0.1 M during refolding. All refolding experiments were carried out by rapid addition of the denatured enzyme to renaturation buffer. Activity recovered as a function of time was monitored by withdrawal of appropriate aliquots of the renaturation mixture followed immediately by dilution into 2.0 ml of assay buffer. The kinetic parameters of refolding were determined by non-linear regression analysis by fitting equations with one or two exponentials to the experimental data using the KaleidaGraph 3.0.5 program (Abelbek Software). The values reported in this study represent the means of at least three different experimental data sets.

Effect of a Second Dilution on Refolding Yields—A second dilution during refolding was performed at different time points after the initiation of reactivation. Enzymes at an optimized concentration (1 mg/ml) were denatured and allowed to start as described above. After 1 min (wild-type) and after 5 and 20 min (all enzymes), the refolding mixture was diluted 10-fold and analyzed for reactivation by measuring increase of activity with time.

Molecular Graphics Analysis and Computer Search for Structural Motifs—Coordinates of GST x-ray structures were derived from the Protein Data Bank (www.rcsb.org/pdb/) via the anonymous file-transfer protocol. The crystal structures were analyzed by using Hyperchem (Autodesk, Sausalito, CA) and Molview 1.4.6 (Purdue University) programs. The figures were generated using the RasMol (version 2.6) program. A PHI-BLAST (28) (www.ncbi.nlm.nih.gov/blast/psiblast.cgi) search was performed using the BLOSUM62 matrix. The pattern used as a query was GX(2)−[LIVY]−[ST]−X(2)−[D]−[LYIVA] and the threshold value was set to 0.001.

Crystallization—Crystallization was performed by the hanging drop vapor diffusion method as described elsewhere (29). Briefly, a 2-μl drop of a protein solution containing the GST mutant (4.2 mg/ml for G146V, 2 mg/ml for G146A) in 1 mM EDTA, 1 mM dithiothreitol, and 10 mM HEPES buffer (pH 7.0) was mixed with an equal volume of reservoir solution, which consisted of 15–25% (w/v) polyethylene glycol (PEG) 8000, 20 mM CaCl2, 1 mM GSH, 10 mM dithiothreitol, and 100 mM MES buffer (pH range 5.2–5.8). All trials were carried out at a constant temperature of 22 °C. Crystals took between 3 and 5 days to appear and grew to their final size within 2 weeks. The x-ray diffraction data were collected using a MARResearch area detector with CuKα x-rays generated by a Rigaku RU-200 rotating anode x-ray generator. The data were collected at 100 K. The diffraction data were processed and analyzed using programs in the HKL (30) and CCP4 suites (31). The G146A and G146V mutants crystallized in the same space group and cell as wild type.

Structure Determination—For both mutants the refinement began with wild-type Pi class GST in the C2 space group (10GS; Ref. 29) that had inhibitor and water molecules omitted. In addition residues between 144 and 150 in each monomer were removed from the starting model. Rigid body refinement in CNS (32) was used to compensate for any possible changes in crystal packing. As the asymmetric unit of the crystal contains one GST monomer, the restraint made of the nanocrystallographic symmetry (NCS) restraints on all non-hydrogen atoms throughout the course of the positional refinements. The 2Fo − Fo map of the G146A mutant showed clear and continuous density for residues between 144 and 150 that had been omitted from the search model. These residues were then built into the map. Among the significant features in the subsequent Fobs − Fcalc maps were positive peaks (close to 4σ) for the r.m.s. error of the map) within 2 Å of the Ca of Gly-146, thus confirming the mutation. The residue at position 146 was changed from glycine to alanine and the model refined. The 2Fo − Fobs and Fobs − Fcalc maps of the G146V mutant clearly indicated a change in the backbone conformation of the region between residues 144 and 146. After the backbone alterations were made, there was additional density next to the Ca position of residue 146 sufficient to accommodate a valine side-chain (close to 3σ the r.m.s. error of the Fobs − Fcalc map), and hence the glycine residue was changed to a valine residue. These residues were then built into the map of the G146V mutant. For both mutants, a number of rounds of positional refinement were performed followed by model building and then by rounds of positional and individual NCS-restrained B-factor refinement. In the final stages of refinement a bulk solvent correction was employed. The correctness of the final structures in the regions around the mutation were confirmed by calculating omit maps between residues 144 and 148. A stereochemical analysis of the refined structure with the program PROCHECK (33) gave values either similar or better than expected for structures refined at similar resolutions. The coordinates for the G146A and G146V models have been deposited in the Protein Data Bank with accession numbers 1MD3 and 1MD4, respectively.

RESULTS

Expression and Purification of Wild-type and Gly-146 Mutants of Human GSTP1-1—To investigate the role of the buried and conserved glycine residue (Gly-146, in hGSTP1-1) this amino acid was replaced with alanine and valine by oligonucleotide-directed mutagenesis producing the mutants G146A and G146V. Mutant and wild-type enzymes were expressed in E. coli XL-1 Blue and purified by affinity chromatography on immobilized GSH. The purified enzymes gave a single band on SDS-PAGE (not shown). 65–90% of the total activity was recovered showing that the affinity for GSH-Sepharose was essentially unaffected by the mutations. Considering that the above substitutions could represent temperature-sensitive mutants, the activity could be observed (not shown). The yields of the G146A and G146V mutants, in percentage of total cytosolic proteins, were 2.4 and 1.7% respectively after purification, both slightly lower than that (4.0%) of the wild-type enzyme.

Kinetic and Structural Characterization of Wild-type Human GSTP1-1 and Gly-146 Mutants—Table I summarizes the kinetic parameters for the conjugation of CDNB to GSH catalyzed by the wild-type hGSTP1-1 and Gly-146 mutants. The replacement of Gly-146 with either alanine or valine significantly increases the kcat values of the corresponding mutants as compared with that of the parental enzyme. The Km values of the mutants were 1.5–2-fold higher for both substrates than those of the wild type. This suggests that the above substitutions slightly but significantly decrease the affinity for GSH and CDNB in both mutants. The physical properties of the Gly-146 mutants and the wild-type enzyme were very similar. The far-UV CD spectra as well as their gel filtration retention times (not shown) were the same, suggesting that all enzyme variants are, in terms of secondary structure and dimeric state of the molecule, essentially identical. The λmax values of the intrinsic fluorescence spectra were 338, 338, and 339 nm for the wild-type enzyme and the G146V and G146A mutants, respec-
tively (not shown), suggesting that a similar polarity characterizes the environment of the tryptophan residues of all enzyme variants. The normalized intensities of the fluorescence of the mutants were slightly higher than that of the wild-type enzyme. This indicates that limited conformational changes distinguish the final structure of the mutants from that of the parent enzyme. These differences involve the environment of one or both tryptophan residues located in GST domain I, far from the mutation site.

**Structure of Mutant G146A**—The structure of helix α6 and the preceding loop is well defined in the wild-type 1.9 Å resolution electron density maps (29). The region encompassing the loop (residues 141–149) and the N-terminal end of the helix is characterized by a network of eight hydrogen bonds between the two, half of which involving a water-mediated contact (Fig. 1A). In addition to these contacts, other significant polar contacts involving the loop include the backbone carbonyl moiety of Gly-146 forming hydrogen-bonding interactions with the side-chains of Asn-137 and Thr-142 and the side-chain of Gln-148 forming water-mediated contacts with the backbone of Ile-149, Gly-78, and Gly-81. Gly-146 adopts $\psi/\phi$ angles of 70°, −163° lying within an allowed region for glycine residues in the Ramachandran plot (33). Statistical data generated from the structure determination of the G146A mutant are presented in Tables II and III. The G146A structure is essentially identical to the wild-type structure with an r.m.s. deviation on superposition of $\alpha$-carbon atoms of 0.2 Å and no deviations greater than 0.6 Å. Ala-146 adopts $\psi/\phi$ angles of 56°, −143° that lie within the additionally allowed region of the Ramachandran plot (33). The extensive network of interactions seen in the wild-type
structure in the region of the loop is maintained in the mutant structure (Fig. 1B). Hence the only obvious difference between the two structures is that the mutated residue lies in a higher energy conformation than wild type as judged by its location in the Ramachandran plot.

**Structure of Mutant G146V**—Statistical data generated from the structure determination of the G146V mutant are presented in Tables II and III. Most of the G146V structure superimposes closely on the wild-type structure (r.m.s. deviation on Cα positions of about 0.3 Å). However, within the region of the mutation (residues 144–147) the r.m.s. deviation rises to 0.6 Å (Fig. 1C). In particular there are shifts of about 1 Å for residues 145 and 146. Also deviations between 0.5 and 0.8 Å occur in the region encompassing the first turn of helix α4 (residues 82–86), which abuts onto the long loop of GST motif II. The origin of these shifts can be traced to changes in the backbone conformation of the mutated residue. In the mutant structure the backbone conformation of residue 146 has changed considerably with the introduction of the valine sidechain. The ϕ/ψ angles for Val-146 are −124°, −177° (compared to wild type of 70°, −163°) which falls in the additionally allowed region of the Ramachandran plot (33). The other significant change occurs at Val-145 where the angles in the mutant structure are −144°, −25° compared with the wild-type structure where they are −141°, 124°. These changes have led...
to some changes in protein contacts. The critical water molecule in the loop region has lost three potential hydrogen-bonding contacts with Val-145 and Ser-150 although it has gained a new hydrogen-bonding interaction with the main-chain of Gly-148. The loss of contact with the side-chain of Ser-150 appears to be due to a movement of the water molecule away from this residue caused by loss of contacts with Val-145. Surprisingly, the backbone carbonyl moiety of Val-146 is still in hydrogen-bonding contact with Asn-137 (but not with Thr-142) despite the change in backbone conformation. Despite all these changes the temperature factors of the loop are not significantly higher than for the wild-type protein.

Thermal Stability of the Wild-type and Mutant Enzymes—Heat inactivation of wild-type and Gly-146 mutants, expressed at 25 °C, was investigated. The results indicate that single Gly-146 substitutions with hydrophobic residues cause very large effects on the catalytic competence even though the active site is situated far from the location of the mutations. Both mutants were significantly destabilized compared with the wild-type enzyme, by being partially inactivated already at subphysiological temperatures (Fig. 2). Furthermore, the G146V mutant was much more unstable than the wild type and the G146A mutant (Fig. 2). It is important to note that the mutants became unstable at a lower temperature than the wild-type; at 32 and 37 °C the mutants were unstable, and their inactivation rates were dependent on the protein concentration (not shown). Different experiments, performed by gel-filtration and electrophoresis techniques, made to clarify the change in backbone conformation. Despite all these changes the temperature factors of the loop are not significantly higher than for the wild-type protein.

Effect of Protein Concentration on the Final Recovery of Native GST after Refolding from 4 M Guanidinium Chloride—To identify the optimal protein concentration for reversible refolding studies, a series of different concentrations were tested (Fig. 4). At low protein concentration (< 5 μg/ml), both wild-type and mutant enzymes appeared to be unstable since the final yield was significantly reduced. At protein concentration above 25 μg/ml, the yield was compromised, presumably due to aggregation, a phenomenon that has been reported for other

![Fig. 3. Eyring plots of wild-type (■), G146A (●), and G146V (○). The lines represent linear fits according to Equation 5, as reported under "Experimental Procedures."](http://www.jbc.org/)

![Fig. 4. Effect of hGSTP1-1 concentration on the final yield of active enzyme following dilution from 4 M guanidinium chloride. hGSTP1-1 was permitted to refold at the concentration indicated for 2 h at 25 °C after rapid 40-fold dilution from 4 M guanidinium chloride into renaturation buffer (0.2 M potassium phosphate, 5 mM dithiothreitol, pH 7.0). The different symbols represent the mean of three different experiments and correspond to wild-type (■), G146A (●), and G146V (○). Percent recovery is expressed relative to the activity of the native control sample at each concentration diluted into the same renaturation buffer with 0.1 M guanidinium chloride added, and incubated for the same period of time.)

### Table IV

|           | $k_u$ | $k_u/k_{u,w}$ | $\Delta G_{u}$ | $\Delta S_{u}$ | $\Delta H_{u}$ | $\Delta M_{u}$ | $\Delta S_{u}$ |
|-----------|-------|---------------|----------------|---------------|---------------|---------------|---------------|
| WT        | 0.0004 | 1             | 135.27         | 0             | 228 ± 10.3    | 0             | 0.304 ± 0.03  |
| G146A     | 0.0112 | 28            | 126.46         | 8.81          | 133 ± 10.6    | 95            | 0.025 ± 0.03  | 0.279         |
| G146V     | 0.0710 | 177           | 121.58         | 13.69         | 226 ± 5.9     | 2             | 0.341 ± 0.02  | −0.037        |

* From the denaturation kinetics at 40 °C according to Equations 1 and 3.

* From the temperature dependence of thermal unfolding (Eyring plot) according to Equation 5.
Conserved Folding Module in hGSTP1-1

proteins. Maximal refolding yields of 75–90% were observed at protein concentrations in the range of 5–15 μg/ml.

Reactivation Yields at Different Temperatures—Replacing Gly-146 with amino acids containing larger hydrophobic side-chains could result in temperature-sensitive mutants with impaired folding properties. To test this hypothesis, we have investigated the reactivation yields of hGSTP1-1 and its mutants at different temperatures of refolding. 10 μM enzyme, expressed at 25 °C, was denatured in 4 M guanidinium chloride for 30 min. This denaturant concentration was sufficient to completely unfold the protein as indicated by the loss of their CD signal at 222 nm (data not shown). The unfolded enzyme was diluted 40-fold with phosphate buffer at pH 7.0 to a final guanidinium chloride concentration of 0.1 M. Appropriate aliquots from this incubation mixture were immediately assayed for activity at 25 °C. Fig. 5 shows that the reactivation yields of the Gly-146 mutants were very different from that of the wild-type protein at increasing temperatures of refolding. While the reactivation yield of the wild-type enzyme was essentially unaffected by temperatures in the 25–40 °C range, the yields of the mutants, albeit to different extents, decreased markedly with temperatures increasing toward physiological values. At 25 °C, when the specific activity recovered by both mutants was significantly higher than that of the wild type, all enzyme variants were characterized by a similar reactivation yield (80–90% of initial activity). Both mutants showed a lower refolding yield than the wild-type enzyme already at 33 °C. At 40 °C the mutants, in contrast to the wild-type enzyme, displayed very poor or complete lack of refolding capacity even at the very beginning of the refolding reaction. It should be noted that during the time course of reactivation competing inactivation reactions also take place. However, most of the activity is regained in a few minutes when the inactivation of both mutants is negligible. Thus, these results seem to suggest that single substitutions of Gly-146 also destabilize the folding of GSTP1-1. This possibility is further supported by the analysis of the formation of aggregates (Fig. 6), as followed by measuring the apparent absorbance (turbidity) at 360 nm. During the reactivation at 37 °C of the G146V mutant, detrimental aggregation reactions are observed even at the very beginning of the refolding reaction. In contrast, the thermal inactivation of this variant, at the same temperature, generated large aggregates only at extended times of incubation (30 min). This means that the destabilization of a folding intermediate causes faster aggregation than during the unfolding of the final structure. Thus, the results could suggest that substitution of the conserved Gly-146 not only affects the stability of the final folded protein but, to a larger extent, also destabilizes a productive intermediate of folding.

Kinetics of Reactivation at 25 °C—A single-exponential equation could be fitted to the kinetic reactivation curve of the wild-type enzyme at 25 °C (Fig. 7). In contrast, as shown in Table V, the refolding kinetics for the G146A and G146V mutants were better described by a double-exponential equation. The attempt to fit a single-exponential equation to the reactivation data of the mutants resulted in higher chi-square values of at least one order of magnitude. The initial part of the refolding curve for the mutants corresponds to an exponential phase with rate constants of 0.29 min⁻¹ and 0.31 min⁻¹ for the G146V and G146A mutants, respectively, which are lower than the single-exponential rate constant of the wild-type enzyme (0.43 min⁻¹). Their relative amplitudes were 73 and 75% for the G146V and G146A variants, respectively. Rate constants of 0.015 min⁻¹ and 0.010 min⁻¹ and relative amplitudes of 27 and 25% characterized the slower phase of the reactivation for the G146V and G146A mutants, respectively. The reactivation ki-
tant quenching characterized the spectra of both variants at the beginning of refolding. However, while no changes in the spectra of the wild-type protein could be observed after 10 min of refolding, significant changes in the intrinsic fluorescence of the G146V mutant was seen even after 120 min. These results are in accordance with the reactivation kinetics and clearly indicate that the substitution of Gly-146 by valine causes the formation of a second refolding step. In order to better identify the molecular species characterizing this slower phase of re-folding another approach was tried. 20 min after the initiation of refolding, when the contribution of the first rapid phase was negligible, a second 10-fold dilution of the protein was performed. Rate and yield of refolding were essentially identical to that obtained during refolding without secondary dilution for all protein variants. In contrast to the results with the wild-type enzyme, the final refolding yield of the mutants was significantly reduced when the second dilution was performed after 5 min (Fig. 7). This is in accordance with the results of the concentration-dependence analysis (Fig. 4) in which a lower protein concentration results in a lower overall yield of reactivation of mutants as compared with the wild type. In addition, because of the higher refolding rate constant, more than 90% refolding is reached by the wild type at 5 min of reactivation. The results also show that the refolding rates were unaffected by the secondary dilution.

Molecular Graphics Analysis and Computer Searches for Local Motifs—Previous sequence analysis showed that only two residues are strictly conserved in domain II (C-terminal domain) in all GSTs. These are Asp-153 and Gly-146, both located in the hydrophobic core of the molecule distant from the active site. Structure analysis showed that Asp-153 belongs to a local motif (S/T)XXD named the N-capping box that is strictly conserved at the N terminus of the α6-helix in GSTs. Studies performed by the peptide approach (1) and site-directed mutagenesis (2) verified that this motif has a crucial role in folding and stability of hGSTP1-1. The sequence analysis also revealed the presence of a second motif at the N terminus of the α6-helix (1). This consists of a specific interaction made by two hydrophobic residues flanking the residues that constitute the capping box (hydrophobic staple motif). Recent work shows that this specific hydrophobic interaction serves to enhance the rate of protein folding and to define the folding pathway (3). The fourth residue preceding the Ncap, a glycine residue, is conserved in all GST sequences (Gly-146 in hGSTP1-1). This residue is possibly playing a joint role with the above described structural motifs in forming a folding nucleus that initiates the folding process. Gly-146 is partially buried (solvent accessible surface, 9.0 Å²), and it is located in a bend of the long loop preceding the α6-helix (Fig. 9). A PHI-BLAST search (28) was performed by using the conserved structural motif GXXh(S/
domain II in all GSTs, create a specific pattern that serves to generate a characteristic alpha-loop-alpha supersecondary structure strictly maintained only in the core of GSTs.

**DISCUSSION**

Gly-146 is one of few residues that are strictly conserved in the GST superfamily. More specifically, Gly-146 is part of a buried local sequence GXXh(T/S)XXDh, which is maintained in all GSTs and, as a more general folding module, in other proteins such as EF1 and URE2. Previous investigations indicated that single mutations of capping box (T/S)XXD and hydrophobic staple (h) residues cause similar dramatic effects on the folding and stability of hGSTP1-1 and hGSTA1-1 (2, 3, 21, 23). Gly-146 is always located four positions before the first Ncap residue (S/T), and its role has not previously been investigated. In the present work the G146A and G146V mutants were generated by site-directed mutagenesis in order to investigate the possible function played by this strictly conserved amino acid residue for folding and stability of human GSTP1-1.

Gly-146 is located in a hydrophobic region, and the choice of alanine was considered the most conservative replacement. The introduction of valine was expected to require more space and therefore could be more disruptive. The steady-state kinetic parameters of both GST variants were very similar at 25 °C to those of the wild-type enzyme (Table I). This similarity was in accordance with the data obtained by CD, gel filtration, and intrinsic fluorescence experiments. These results indicate that the overall structure and dimeric state were the same, suggesting that the global fold of mutants and wild-type GST were similar. However, both mutants, as compared with the wild-type enzyme, were characterized by a higher sensitivity toward heat inactivation. The mutants, expressed at the permissive temperature of 25 °C, have been crystallized and their

### Table V

|          | k_{ref} | k_{2ref} | A_1   | A_2   | χ² |
|----------|---------|----------|-------|-------|----|
| Wild-type| 0.43 ± 0.02 | 100      | 0.010 ± 0.001 | 75    | 2.04 |
| G146A    | 0.31 ± 0.02 | 0.015 ± 0.002 | 73    | 27    | 2.65 |
| G146V    | 0.29 ± 0.01 | 0.001    | 75    | 25    | 0.76 |

**Fig. 9. Ribbon picture of a monomer of human Pi class GST.** The location of GST motif II (the conserved sequence motif consisting of helix α6 and the preceding long loop) is rendered in dark yellow. GSH and the site of mutation are shown in ball-and-stick. This figure was produced using BOBSCRIPT (41).
three-dimensional structures determined by x-ray crystallography. No structural change to the protein was caused by the G146A mutation and the principal interactions formed between the main-chain carbonyl group with Asn-137 and Thr-142 were preserved. Nevertheless, the Ala-146 residue lies in a higher energy conformation than wild type as judged by its location in the Ramachandran plot. This is sufficient to destabilize the active site at 37 °C, even though it is located far from the mutation site. Due to a large decrease in activation enthalpy (∆H 95 kJ/mol), the destabilizing effect of the G146A mutation was less important at higher temperatures but clearly significant at subphysiological and physiological ones. In addition, the magnitude of T∆S at 40 °C (87 kJ/mol) indicates that major changes in the conformational freedom characterize the denaturation of the wild type compared with those occurring during the thermal inactivation of the G146A mutant. This means that the wild-type hGSTP1-1, because it is more rigid than the G146A mutant, may tolerate larger perturbations of its structure before the unfolding transition state is reached. The crystallographic data also indicate that the mutation of Gly-146 to valine, unlike the change to alanine, had to result in a substantial change of the backbone conformation because otherwise the valine side-chain would have sterically clashed with Glu-86. Thermal inactivation measurements indicate that these local structural changes, observed at 25 °C, had destabilizing effects at temperatures as low as 32 °C. The decrease in the unfolding free energy (14 kJ/mol) for the G146V mutant is higher than that estimated for the G146A variant. It is important to note that the entropic contribution of thermal inactivation of the G146V mutant is higher than it is to the wild type and G146A mutant and is the driving force for the denaturation of this variant. Thus the thermal inactivation of this mutant causes perturbations of the active site structure, which are larger than those experienced in similar unfolding conditions by the other variants and more readily leads to a more denatured state of the protein.

The α6-helix and preceding loop form a substructure (GST motif II), which is conserved in the core of all GSTs. It is stabilized by a network of eight hydrogen bonds with half involving a water-mediated contact and makes interdomain contacts with the α1-helix which is an important structural element of the active site. Previous investigations indicated that single mutations of residues forming the capping box ((S/T)XXD) and hydrophobic staple motif, located at the N-terminal of the α6-helix, have dramatic effect on the stability of hGSTP1-1 and hGSTA1-1. X-ray structure analysis of the capping mutants, expressed at the permissive temperature of 25 °C, indicated that these amino acid substitutions locally destabilize GST motif II through a partial or complete loss of the buried hydrogen bond network. Glycine 146 is located in the bend of the long loop (residues 141–149) preceding the α6-helix (Fig. 9). The crystallographic and thermodynamic analyses of the G146V mutant confirm that alterations of the buried hydrogen bond network inactivate the enzyme already at temperature as low as 32 °C. In this region, a glycine residue must be maintained at position 146 principally because of lack

![Conserved folding module of GST](http://www.jbc.org)
of space. However, even the more conservative Gly/Ala substitution, which does not alter the protein structure at 25 °C, inactivates the enzyme at 37 °C. It is important to note that the formation of a hydrogen bond network in the interior of the protein poses geometric constraints that have to be reconciled with the geometric constraints of the dense packing. A major flexibility of the main-chain, due to the presence of a glycine residue, permits the maintenance of the internal hydrogen bond network even at higher temperatures, despite larger motions of the side-chains. Thus the present results, in accordance to previous findings, confirm the critical role of GST motif II substructure for the overall protein stability. In addition they strongly support the hypothesis that GST motif II, stabilized by a characteristic hydrogen bond network, behaves as a general structural motif whose precise conformation serves to stabilize GSTs and, likely also other proteins, in the physiological range of temperature.

The in vitro refolding experiments suggest that the thermal lability of the final structure of the mutants reflects differences in the conformational properties of a productive intermediate of folding. Reactivation in vitro of both mutants, in fact, was thermosensitive with the most pronounced effect for the G146V variant. It should be noted that during the time course of reactivation competing inactivation reactions also take place. Thus, although most of the catalytic activity is regained in a few minutes, a time period in which the inactivation of the enzyme is negligible, it is difficult to decide whether single substitutions of Gly-146 generate temperature sensitive folding mutants. However this possibility is further supported by the analysis of the formation of aggregates (Fig. 5), as followed by measuring the apparent absorbance at 360 nm. During the reactivation at 37 °C of the G146V mutant, large aggregates are formed faster than those that can be observed during the thermal inactivation of this variant at the same temperature. Thus, the results indicate that substitution of the conserved Gly-146 not only affects the stability of the final folded protein, but also favors detrimental aggregation events during refolding. This also means that the fourth position before the Ncap motif must be maintained as a glycine residue because of its essential contribution to a productive folding pathway. Its substitution by another residue likely alters the conformation of an essential intermediate of folding, favoring protein aggregation at 37 °C. The fact that the Gly-146 mutations cause lesser effects on thermal stability (Fig. 2) than on the refolding kinetics (Fig. 5) at comparable temperatures suggests that the conformation of this intermediate is more labile than that of the fully folded state. This interpretation is confirmed by the analysis performed at a permissive temperature (25 °C) of the re- forming of mutants, as followed by regain of activity and fluorescence. The results show that a slower refolding step, subsequent to dimerization, becomes rate-limiting for reactivation. It is conceivable that an altered conformation of GST motif II during the refolding of mutants makes the formation of the necessary interdomain interactions made by the α6-helix with the α1-helix more difficult (1–3). It has already been shown, that the proper conformation of GST motif II is important for protein stability and, to an even higher extent, for a productive folding pathway (2, 3). From a more general point of view, the present results together with previous studies suggest that a well defined stereochemical code underlies the refolding of GSTs. All conserved residues forming the local sequence G146XHx(S/T)XXDh, albeit through different mechanisms, are critical for folding (Fig. 7). The important role played by Gly-146 for the GST refolding is reported here. Present work supports the hypothesis that substitution of Gly-146 destabilizes a conserved loop-helix substructure that is essential during the first refolding events. However, it cannot be completely ruled out that substitution of Gly-146 promotes the formation of off-pathway intermediates under the in vitro refolding conditions used here. We have previously shown, in fact, that the other conserved residues belonging to the above local pattern form an N-capping box and a hydrophobic staple motif are important for the α6-helix nucleation and stabilization (1–3, 21). These few conserved residues are all crucial for the formation of GST motif II that might represent a nucleation site for the refolding of GST. It has been hypothesized that for monomeric single domain proteins, the two-state kinetics of folding follows a molecular growth mechanism (34–36). Buried and highly conserved residues, making a definite set of native like contacts, determine the formation of a specific folding nucleus that represents the transition state of refolding (37–40). Conserved non-functional amino acid residues in globins and in the cytochrome c family have been implicated in the folding process of these proteins (39–40). Certain amino acid residues are favored in a few positions that are located far apart in the primary structure of the polypeptide chain. Folding is initiated when these residues interact to form a folding nucleus, which is thought to promote α-helix formation in these particular regions from which folded motifs then proceed (39–40). Thus the present work, together with our previous studies, also support the hypothesis that even in more complex multi-domain proteins such a specific nucleation mechanism exists. However, in GSTs, the strictly conserved non-functional residues that play different and well defined roles for GST folding are clustered in a linear local sequence (Fig. 10). We propose that this conserved and buried pattern represents a general folding module that plays an important role in the refolding and stability of GSTs and, likely, of other proteins such as EF1γ and URE2.

REFERENCES

1. Aceto, A., Dragani, B., Melino, S., Allocci, N., Manelli, M., Di Ilio, C., and Petruzzi, R. (1997) Biochem. J. 329, 229–234
2. Dragani, B., Stenberg, G., Melino, S., Petruzzi, R., Mannervik, B., and Aceto, A. (1997) J. Biol. Chem. 272, 25518–25523
3. Stenberg, G., Dragani, B.,occo, R., Mannervik, B., and Aceto, A. (2000) J. Biol. Chem. 275, 10421–10428
4. Hayes, J. D., and Pulford, D. J. (1995) Crit. Rev. Biochem. Mol. Biol. 30, 445–460
5. Mannervik, B., Alín, P., Gutenberg, C., Jansson, H., Tahir, M. K., Warholm, M., and Jornvall, H. (1985) Proc. Natl. Acad. Sci. U. S. A. 82, 7202–7206
6. Mannervik, B., and Jansson, H. (1982) J. Biol. Chem. 257, 9909–9912
7. Danielson, U. H., and Mannervik, B. (1985) Biochemistry 24, 231, 263–267
8. Tahir, M. K., and Mannervik, B. (1986) J. Biol. Chem. 261, 1044–1051
9. Gustafsson, A., and Mannervik, B. (1999) J. Mol. Biol. 286, 787–800
10. Bontemmer, P., Durr, H. W., Schultz, G., Wences, A., and Huber, R. (1991) EMBO J. 10, 1997–2005
11. Ji, X., Zhang, P., Armstrong, R. N., and Gililand, G. L. (1992) Biochemistry 31, 10169–10184
12. Sinning, I., Kleywegt, G. J., Cowan, S. W., Reinemer, P., Durr, H. W., Huber, R., Gililand, G., Armstrong, R. N., Ji, X., Board, P. G., Olin, B., Mannervik, B., and Jones, T. A. (1993) J. Mol. Biol. 232, 192–212
13. Wilce, M. C. J., Board, P. G., Pelf, S. C., and Parker, M. W. (1995) EMBO J. 14, 2133–2143
14. Richardson, J. S., and Richardson, D. C. (1988) Science 240, 1648–1652
15. Presta, L. G., and Rose G. D. (1988) Science 240, 1632–1641
16. Seale, J. W., Srinivasan, R., and Rose, D. G. (1994) Protein Sci. 3, 1741–1745
17. Munoz, V., Blanco, F. L., and Serrano, L. (1995) Nat. Struct. Biol. 2, 380–385
18. Munoz, V., and Serrano, L. (1995) Biochemistry 34, 15301–15306
19. Aurora, R., and Rose, D. G. (1998) Protein Sci. 7, 21–38
20. Snyder, M. J., and Maddison, D. R. (1997) DNA Cell Biol. 16, 1373–1384
21. Coco, R., Stenberg, G., Dragani, B., Rossi Principi, D., Fulabu, D., Mannervik, B., and Aceto, A. (2001) J. Biol. Chem. 276, 32177–32183
22. Koonin, E. V., Mushegian, A. R., Tatusov, R. L., Altschul, S. F., Bryant, S. H., Bork, P., and Valencia, A. (1994) Protein Sci. 3, 2045–2054
23. Rossjohn, J., Oakley, M. J., Parker, M. W., Stenberg, G., Mannervik, B., Dragani, B., Coco, R., and Aceto, A. (2000) J. Mol. Biol. 302, 295–302
24. Klein, R. H., Stenberg, G., Wisterden, M., and Mannervik, B. (1995) Protein Expression Purif. 6, 265–271
25. Aceto, A., Caccuri, A. M., Sacchetta, P., Bucciarelli, T., Dragani, B., Rosato, N., Federic, G., and Di Ilio, C. (1993) Biochem. J. 295, 241–245
26. Aceto, A., Di Ilio, C., Lo Bello, M., Bucciarelli, T., Angelucci, S., and Federici, G. (1990) Carcinogenesis 11, 2267–2269
27. Johnson, F. H., Eyring, H., and Polissar, M. J. (1954) The Kinetic Basis of
Conserved Folding Module in hGSTP1-1

Molecular Biology, John Wiley and Sons, New York

28. Zhang, Z., Schaffner, A. A., Miller, W., Madden, T. L., Lipman, D. J., Koonin, E. V., and Altschul, S. F. (1998) Nucleic Acids Res. 26, 3986–3990
29. Oakley, A. J., Lo Bello, M., Battistoni, A., Rieci, G., Rossjohn, J., Villar, H. O., and Parker, M. W. (1997) J. Mol. Biol. 274, 84–100
30. Otwinowski, Z. (1993) in Data Collection and Processing, (Sawyer, L., Isaacs, N. & Bailey, S., eds), SERC Daresbury Laboratory, Warrington, UK
31. CCP4 (1994) Acta Crystallogr. Sect. D 50, 750–763
32. Brünger, A. T., Adams, P. D., Cloare, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Acta Crystallogr. Sect. D Biol. Crystallogr. 54, 905–921
33. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) J. Appl. Crystallogr. 26, 283–291
34. Pitsyn, O. B. (1996) FASEB J. 10, 3–4
35. Shakhnovich, E., Abkevich, V., and Pitsyn, O. (1996) Nature 379, 96–98
36. Fersht, A. R. (1997) Curr. Opin. Struct. Biol. 7, 3–9
37. Itzhaki, L. S., Neira, J. L., Ruiz-Sanz, J., de Prat Gay, G., and Fersht, A. R. (1998) J. Mol. Biol. 284, 289–304
38. Mirny, L. A., Abkevich, V. I., and Shakhnovich, E. I. (1999) Proc. Natl. Acad. Sci. U. S. A. 95, 4976–4981
39. Pitsyn, O. B., (1999) J. Mol. Biol. 278, 655–666
40. Ptitsyn, O. B., and Ting, K. L. (1999) J. Mol. Biol. 291, 671–682
41. Esnouf, R. M. (1999) Acta Crystallogr. Sect. A 55, 938–940
Contribution of Glycine 146 to a Conserved Folding Module Affecting Stability and Refolding of Human Glutathione Transferase P1-1
Geoffrey K-W. Kong, Galina Polekhina, William J. McKinstry, Michael W. Parker, Beatrice Dragani, Antonio Aceto, Domenico Paludi, Daniela Rossi Principe, Bengt Mannervik and Gun Stenberg

J. Biol. Chem. 2003, 278:1291-1302.
doi: 10.1074/jbc.M209581200 originally published online October 31, 2002

Access the most updated version of this article at doi: 10.1074/jbc.M209581200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 39 references, 12 of which can be accessed free at http://www.jbc.org/content/278/2/1291.full.html#ref-list-1