A Conserved “Hydrophobic Staple Motif” Plays a Crucial Role in the Refolding of Human Glutathione Transferase P1-1*

(Received for publication, September 20, 1999, and in revised form, January 11, 2000)

Gun Stenberg‡, Beatrice Dragani§, Roberta Cocco§, Bengt Mannervik‡, and Antonio Aceto§¶

From the ‡Dipartimento di Scienze Biomediche, Università “G. D’Annunzio,” Via dei Vestini, 31, 66100 Chieti, Italy and the §Department of Biochemistry, Uppsala University, Biomedical Center, Box 576, S-751 23 Uppsala, Sweden

The specific (i, i+5) hydrophobic staple interaction involving a helix residue and a second residue located in the turn preceding the helix is a recurrent motif at the N terminus of α-helices. This motif is strictly conserved in the core of all soluble glutathione transferases (GSTs) as well as in other protein structures. Human GSTP1-1 variants mutated in amino acid Ile149 and Tyr154 of the hydrophobic staple motif of the α6-helix were analyzed. In particular, a double mutant cycle analysis has been performed to evaluate the role of the hydrophobic staple motif in the refolding process. The results show that this local interaction, by restricting the number of conformations of the α6-helix relative to the α1-helix, favors the formation of essential interdomain interactions and thereby accelerates the folding process. Thus, for the first time it is shown that the hydrophobic staple interaction has a role in the folding process of an intact protein. In Π class GSTs, Tyr154 appears to be of particular structural importance, since it interacts with conserved residues Leu21, Asp24, and Gln25 of the adjacent α1-helix which contributes to the active site. Human GSTP1-1 variants L21A and Y154F have also been analyzed in order to distinguish the role of interdomain interactions from that of the hydrophobic staple. The experimental results reported here suggest that the strict conservation of the hydrophobic staple motif reflects an evolutionary pressure for proteins to fold rapidly.

It is known that the folded conformation adopted by a protein is governed by its amino acid sequence, suggesting an underlying stereoechemical code for protein folding. Current protein folding models attribute a critical role to local interactions that reduce the conformational freedom at certain points of the polypeptide chain and thus define the folding pathways. Consequently, significant attention has recently been directed to identifying local motifs responsible for the formation of secondary structure element as well as of supersecondary regions including combination of such elements. At the N terminus of helices two local motifs have been described. The first one, named the N-capping box (1, 2), involves the reciprocal main chain-side chain hydrogen bonds between the N-cap (Ser/Thr) and the N-3 (Glu/Asp) residues of a helix. (We follow the nomenclature proposed by Richardson and Richardson (1): N’-N- Ncap-N1-N2-N3-N4; where N1-N4 belong to the helix and N- cap is the boundary residue.) The second one, termed the hydrophobic staple motif (3–5), involves a specific hydrophobic interaction between residues located at the N’ and N-4 positions. A global survey of helix capping in proteins of known structure has recently been reported (6). When present in combination with a capping box (expanded box (3)), the hydrophobic staple motif gives a cooperative effect in defining the α-helix starting point. In a statistical analysis (4) its presence in the absence of a capping box has also been shown, suggesting that it may represent an independent motif. Analysis of protein structures in data bases as well as studies involving short peptides have shown that residues at positions N’ and N-cap adopt dihedral angles characteristic of the β-region of Ramachandran plots (4). It has therefore been hypothesized that the hydrophobic staple motif may serve to define the direction of a polypeptide chain as it forms the α-helix (4, 5). However, the function of the hydrophobic staple motif still remains to be clarified. Notably, its suggested contribution to folding and stability has never been studied experimentally in a protein, despite of its frequent occurrence at the N terminus of helices.

Glutathione transferases (GSTs) are a superfamily of detoxication enzymes (7) that catalyze the nucleophilic attack of glutathione (GSH) on the electrophilic groups of a wide range of hydrophobic toxic compounds (8). GSTs are dimeric proteins in which each 25-kDa subunit is characterized by two distinct domains and an active site that acts independently of the other subunit. The smaller N-terminal domain (domain I) adopts an α/β topology and contributes most of the contacts to GSH. The C-terminal domain (domain II) is all-α-helical and provides most of the amino acid residues of the hydrophobic substrate-binding site which lies adjacent to the GSH-binding site (9–13).

We have previously shown that amino acid residues forming an N-capping box and a hydrophobic staple motif at the N terminus of the α6-helix in domain II, are strictly conserved in all known sequences of soluble GSTs (more than 50) (14). While the N-capping box, being involved in the α6-helix formation, is essential to the folding of GSTs (15), the possible role of the hydrophobic staple motif in the refolding of GST is still unknown. It should be noted that it is located in the hydrophobic core of the GSTs and that, in a very similar structural environment, it is also conserved in other protein superfamilies (16). Thus, GSTs serve as good model proteins for investigations of the specific role played by this supersecondary interaction in protein folding.

Site-directed mutagenesis was employed to change the amino acid residues forming the hydrophobic staple motif in

* This work was supported by grants of the Ministero dell’Università e della Ricerca Scientifica e Tecnologica (60%), the Carl Trygger Foundation, and the Swedish Natural Science Research Council. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† To whom correspondence should be addressed: Dipartimento di Scienze Biomediche, Università degli Studi “G. D’Annunzio,” Via dei Vestini, 31, I-66100 Chieti, Italy. Tel.: 39-0871-3555272; Fax: 39-0871-3555356; E-mail: a.aceto@dsb.unich.it.

‡ The abbreviations used are: GST, glutathione S-transferase; GSH, glutathione; CDNB, 1-chloro-2,4-dinitrobenzene.

This paper is available on line at http://www.jbc.org

10421
the human enzyme GSTP1-1. The results indicate that the interaction between the side chains of the loop residue N and the α6-helix residue N-4 favors the appropriate conformation of the α6-helix with respect to the preceding structural element thereby enhancing the rate of protein folding and defining the folding pathway.

**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 described previously (17). GSH and 1-chloro-2,4-dinitrobenzene (CDNB) were purchased from Sigma.

**Construction of GSTP1-1 Mutants**—Site-directed mutagenesis was employed to generate the hydrophobic staple mutants. The following oligonucleotides were used for the mutations: I149A, 5’-GACGACGGTTCTCTCGTAGTCTACAACTCGTGTG-3’; 1149G, 5’-GACGACGGCTCTTAGTCTACAACTCGTGTG-3’; 154A, 5’-GACACGATCTCTCTCCGTCGTGGACCTCAACCTCGTGTG-3’; 154F, 5’-GACACGATCTCTCTCCGTCGTGGACCTCAACCTCGTGTG-3’; 1149Y154A, 5’-GACACGAGCCCTCTCTCCGTCGTGGACCTCAACCTCGTGTG-3’; 1149Y154F, 5’-GACACGAGCCCTCTCTCCGTCGTGGACCTCAACCTCGTGTG-3’; 1149A154Y, 5’-GACGACGGCTTCGCTGACTACAACCTGCTG-3’; 1149G154Y, 5’-GACGACGGCTTCGCTGACTACAACCTGCTG-3’; 1149A154F, 5’-GACGACGGCTTCGCTGACTACAACCTGCTG-3’; 1149G154F, 5’-GACGACGGCTTCGCTGACTACAACCTGCTG-3’; 1149A/154Y, 5’-GGCTTCCTTCGCTGACTACAACCTGCTG-3’; I149G/Y154F, 5’-GACCAGGGCTC-9 CTTCGCTGACTTCAACCTGCTG-3’; I149G/Y154A, 5’-GACCAGGGCTC-9 CTTCGCTGACTTCAACCTGCTG-3’; Y154A, 5’-GACACGCGGCTC-9 CTTCGCTGACTTCAACCTGCTG-3’; Y154F, 5’-GACACGCGGCTC-9 CTTCGCTGACTTCAACCTGCTG-3’; I149G/Y154A, 5’-GACCAGGGCTC-9 CTTCGCTGACTTCAACCTGCTG-3’; I149G/Y154F, 5’-GACCAGGGCTC-9 CTTCGCTGACTTCAACCTGCTG-3’; I149G/Y154A, 5’-GACCAGGGCTC-9 CTTCGCTGACTTCAACCTGCTG-3’; I149G/Y154F, 5’-GACCAGGGCTC-9 CTTCGCTGACTTCAACCTGCTG-3’; I149G/Y154A, 5’-GACCAGGGCTC-9 CTTCGCTGACTTCAACCTGCTG-3’; I149G/Y154F, 5’-GACCAGGGCTC-9 CTTCGCTGACTTCAACCTGCTG-3’; I149G/Y154A, 5’-GACCAGGGCTC-9 CTTCGCTGACTTCAACCTGCTG-3’; I149G/Y154F, 5’-GACCAGGGCTC-9 CTTCGCTGACTTCAACCTGCTG-3’. The reverse primer had the sequence: 5’-TCCCAATAGGAGGGTGCTCTGGCTCGTGTG-3’. For the L21A mutant on the mutagenic primer 5’-CTGATGGGACCCGAC-3’ and the reverse primer 5’-CAGTGGGACCCGAC-3’. Alteration nucleotides were phosphorylated and then used in pairs of one mutagenic and the reverse primer in inverted polymerase chain reaction. The expression clone pKH1(17) was used as a template. The polymerase chain reaction mixture contained 0.8 μM of each primer, 0.2 mM dNTPs, 2.5 μl of 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 polymerase chain reaction product was recovered from the agarose gel. The DNA was ligated and used to transform *E. coli* XL-1 Blue cells. The cDNA encoding the isolated GSTP1-1 mutants was sequenced in its entirety to verify that no undesired mutations had been introduced in the polymerase chain reaction.

**Protein Expression and Stability in the Host Cells** as a Function of the Growth Medium Temperature—Cultures of *E. coli* XL-1 Blue containing plasmids were grown in 500 ml of LB broth in a 2-liter Erlenmeyer flask at 37 °C. At an A600 of 0.35, isopropyl-β-D-thiogalactopyranoside was added to a final concentration of 0.2 mM. From the time of addition the culture was grown for about 20 h at two different temperatures of incubation, i.e. 25 and 37 °C. The subsequent purification of wild-type and mutant proteins, expressed at different growth temperatures, was performed as described previously (18) with the only modification that the enzymes were purified on a GSH-Sepharose affinity column. The intracellular protein yields of GST in bacteria were expressed as percent of the enzymes was purified on a GSH-Sepharose affinity column.

**Kinetic and Structural Studies on Wild-type and Hydrophobic Staple Motif in GST Folding**

**Hydrophobic Staple Motif in GST Folding**

**RESULTS**

**Identification of a Conserved Hydrophobic Staple Motif and Molecular Graphics Analysis**—At the N terminus of the α6-helix in domain II, the sequence (Ser/Thr)XXAasp is conserved in the alpha, mu, Pi, theta, and sigma classes of GSTs. We have recently demonstrated that this motif corresponds to a “capping box” signal (1, 2, 6) which exerts a strong influence on the α6-helix nucleation (14) as well as on GST folding (15). The sequence analysis has also revealed the presence of a second structural motif conserved at the N terminus of the α6-helix in GSTs of the same classes (14). This motif is specific to a group of GSTs that contains an α-helix (5) and is designated (defining time 0) 1.40 into renaturation buffer (0.2 mM phosphate, 1 mM EDTA, 5 mM dithiothreitol, pH 7.0) at the same temperature. The final guainichloride concentration was 0.1 mM during refolding. All refolding experiments were carried out by rapid addition of the denatured enzyme to the renaturation buffer. Recovered activity was monitored by either by the rate of the reactivation value at time 0. The effect of mutation on the energy of the transition state of folding can be calculated using transition state theory in a similar manner to that reported by Jackson et al. (20).

The stability of the transition state of a mutant protein relative to that of wild-type is calculated from, where ΔGf is the difference in the energy of the transition state of folding relative to the unfolded state between wild-type and mutant proteins; kf and k- are the respective rate constants of folding.

Double mutant cycles have been used to calculate ΔGf, and the interaction between two residues in a protein, under conditions in which they can make additional interactions. ΔGf is obtained by subtraction of the ΔGf values of the corresponding single mutants from the ΔGf of the double mutant, calculated as above. The general theory of double mutant cycle analysis has been discussed extensively elsewhere (21). It is assumed that the ΔGf is 0 in the unfolded state and this is taken as the reference state. The reactivation behavior of all GSTP1-1 variants can be described by a single exponential equation. Thus, if none of the mutants have any significant effect on the folding pathway, at least not when analyzed at subphysiological temperatures.

**Identification of a Conserved Hydrophobic Staple Motif and Molecular Graphics Analysis**—Coordinates of GST x-ray structures were derived from the Brookhaven Protein Data Bank via the Anonymous file-transfer protocol. The crystal structures were analyzed by using Hyperchem (22) (Autodesk, Sausalito, CA) and MolView 1.4.6 (Purdue University) programs. Figures were generated by using the RasMol (version 2.6) program.

**RESULTS**

**Identification of a Conserved Hydrophobic Staple Motif and Molecular Graphics Analysis**—At the N terminus of the α6-helix in domain II, the sequence (Ser/Thr)XXAasp is conserved in the alpha, mu, Pi, theta, and sigma classes of GSTs. We have recently demonstrated that this motif corresponds to a “capping box” signal (1, 2, 6) which exerts a strong influence on the α6-helix nucleation (14) as well as on GST folding (15). The sequence analysis has also revealed the presence of a second structural motif conserved at the N terminus of the α6-helix in GSTs of the same classes (14). This motif is specific to a group of GSTs that contains an α-helix (5) and is designated (defining time 0) 1.40 into renaturation buffer (0.2 mM phosphate, 1 mM EDTA, 5 mM dithiothreitol, pH 7.0) at the same temperature. The final guainichloride concentration was 0.1 mM during refolding. All refolding experiments were carried out by rapid addition of the denatured enzyme to the renaturation buffer. Recovered activity was monitored by either by the rate of the reactivation value at time 0. The effect of mutation on the energy of the transition state of folding can be calculated using transition state theory in a similar manner to that reported by Jackson et al. (20).

The stability of the transition state of a mutant protein relative to that of wild-type is calculated from, where ΔGf is the difference in the energy of the transition state of folding relative to the unfolded state between wild-type and mutant proteins; kf and k- are the respective rate constants of folding.

Double mutant cycles have been used to calculate ΔGf, and the interaction between two residues in a protein, under conditions in which they can make additional interactions. ΔGf is obtained by subtraction of the ΔGf values of the corresponding single mutants from the ΔGf of the double mutant, calculated as above. The general theory of double mutant cycle analysis has been discussed extensively elsewhere (21). It is assumed that the ΔGf is 0 in the unfolded state and this is taken as the reference state. The reactivation behavior of all GSTP1-1 variants can be described by a single exponential equation. Thus, if none of the mutants have any significant effect on the folding pathway, at least not when analyzed at subphysiological temperatures.

**Identification of a Conserved Hydrophobic Staple Motif and Molecular Graphics Analysis**—Coordinates of GST x-ray structures were derived from the Brookhaven Protein Data Bank via the Anonymous file-transfer protocol. The crystal structures were analyzed by using Hyperchem (22) (Autodesk, Sausalito, CA) and MolView 1.4.6 (Purdue University) programs. Figures were generated by using the RasMol (version 2.6) program.
serves to cause a reversal of the polypeptide chain direction as it forms the \( \alpha \)-helix. Three-dimensional structure analysis of GSTs shows that the side chains of the hydrophobic residues at the N’ and N-4 positions are located within a distance of 4 Å from each other in all known GST structures and that the N’ and N-cap residues always adopt dihedral angles in the \( \beta \)-region of the Ramachandran plot. Thus, this supersecondary interaction possesses all structural characteristics of a hydrophobic staple motif and, despite the low overall sequence identity (in some cases less than 5%), is well conserved among GSTs and GST-related proteins (14).

As shown in Fig. 1, residue Tyr154, in addition to the hydrophobic staple interaction with Ile149, forms very favorable contacts with amino acid residues Leu21, Asp24, and Gln25, all belonging to an important structural element (the \( \alpha \)-helix) which contributes to the active site. Moreover, the hydrophobic bond as well as the hydrophobic interactions made by Tyr154 with Asp24 and Leu21, respectively, are conserved in class Pi GSTs. It is interesting to note that these polar and apolar interdomain interactions seem to be specific for \( \Pi \) class GSTs, since they are maintained in the bovine, murine, porcine, and rat GSTP1-1 but not in GST structures from other classes (not shown).

**Design, Expression, and Purification of Hydrophobic Staple Mutants of Human GSTP1-1**—To determine the role of the hydrophobic staple motif in the folding process, site-directed mutagenesis was used to produce the two single mutations I149A and Y154A as well as the double mutation I149A/Y154A. However, considering that an alanine residue still potentially can form, during the folding process, a weak hydrophobic staple and a glycine cannot, the I149G variant and the corresponding I149G/Y154F and I149G/Y154A double mutants were also generated. Furthermore, in order to distinguish the contribution of the hydrophobic staple versus the other contributions of Tyr154 with the \( \alpha \)-helix, the variants Y154F and L21A were also analyzed. These two mutants make it possible to separate the contribution of the hydrogen bonds (Y154F mutant) from the hydrophobic interaction (L21A variant) made by Tyr154 with domain I. The mutants and the wild-type GSTP1–1 were expressed in *E. coli* and purified in a single chromatographic step using a GSH affinity matrix. No more than 2–3% of the total enzymatic activity remained unbound for any of the different GST variants. Apparently, the mutations do not affect the affinity for GSH-Sepharose. The purified proteins were homogeneous as judged by SDS-polyacrylamide gel electrophoresis (results not shown).

The possibility that the mutations could have resulted in temperature-sensitive species was investigated by expressing the proteins at different growth temperatures of the host cells. As shown in Table I the intracellular yields of wild-type GSTP1-1 and mutants Y154A and I149A, expressed in bacteria grown at 25 °C, were essentially the same. With increasing temperature, the yields of these mutants slightly decreased, unlike the yield of the wild-type, L21A, Y154F, and I149G variants, which remained unchanged. A much lower intracellular yield, strongly dependent on the temperature of host cell growth, was found for the I149A/Y154A, I149G/Y154F, and I149G/Y154A double mutants. At 37 °C the protein expression levels of all double mutants were more than 15 times lower than those of the wild-type and single-point mutants. These results probably have the explanation that the double mutants fail to attain the native conformation at physiological temperatures. The activity in the crude bacterial lysate of all double mutants (not shown) and the specific activity of the purified I149A/Y154A and I149G/Y154A variants (Table I) were markedly lower when the growth temperature of the host cell was 37 °C.

**Kinetic and Structural Properties of Wild-type and Hydrophobic Staple Mutant Enzymes Expressed at 25 °C**—The level of protein expression of some mutants was more or less dependent on the growth temperature of the bacterial host cells. Moreover, some mutants displayed a lower specific activity when expressed at 37 °C. Thus, kinetic and structural characterizations were performed using enzymes expressed at the more permissive temperature of 25 °C.

Table II summarizes the kinetic parameters for the conjugation of CDNB with GSH catalyzed by wild-type GSTP1-1 and mutant variants. The \( k_{cat}^{\text{GSH}} \), \( K_m^{\text{GSH}} \), and \( k_{cat}/K_m \) values of all mutants, although to various extents, were all higher than those of the wild-type enzyme. In particular, mutants I149A/ Y154A, L21A, and I149G/Y154A displayed \( k_{cat} \) values about 2-fold higher as compared with the parent enzyme. The \( K_m \) values for CDNB were also higher than that of the wild-type enzyme. The \( k_{cat}^{\text{GSH}} \) values for CDNB were also higher than that of the wild-type enzyme.

*Table I*

| Protein yield | 25 °C | 37 °C |
|---------------|-------|-------|
| Wild-type     | 39 ± 5| 34 ± 5|
| L21A          | 40 ± 6| 50 ± 7|
| Y154F         | 52 ± 7| 27 ± 4|
| I149G         | 43 ± 8| 76 ± 8|
| I149G/Y154F   | 53 ± 8| 62 ± 9|
| I149G/Y154A   | 42 ± 6| 23 ± 3|

**Table II**

| Unit/mg | 25 °C | 37 °C |
|---------|-------|-------|
| Wild-type | 6.0 | 7.5 |
| I149A | 8.2 | 5.9 |
| Y154A | 5.2 | 2.9 |
| I149A/Y154A | 0.8 | 0.1 |
| L21A | 6.4 | 17.4 |
| Y154F | 4.3 | 22.3 |
| I149G | 4.1 | 8.7 |
| I149G/Y154F | 3.3 | 5.0 |
| I149G/Y154A | 0.2 | 0.1 |

---

**Fig. 1.** The hydrophobic staple motif in human GSTP1-1 represented by the loop residue Ile149 (N’ position) and \( \alpha \)-helix residue Tyr154 (N-4 position). Tyr154 can form favorable hydrogen bonds with Gln25 and Asp24 and a hydrophobic interaction with Leu21. These last three residues belong to the \( \alpha \)-helix which is an important element of the active site. Interactions made by Tyr154 with Asp24 and Gln25 form, during the folding process, a weak hydrophobic staple motif in the folding process, site-directed mutagenesis was used to produce the two single mutations I149A and Y154A as well as the double mutation I149A/Y154A. However, considering that an alanine residue still potentially can form, during the folding process, a weak hydrophobic staple and a glycine cannot, the I149G variant and the corresponding I149G/Y154F and I149G/Y154A double mutants were also generated. Furthermore, in order to distinguish the contribution of the hydrophobic staple versus the other contributions of Tyr154 with the \( \alpha \)-helix, the variants Y154F and L21A were also analyzed.
enzyme for the Y154A, I149A/Y154A, L21A, and I149G/Y154A mutants.

Expression at 25 °C yielded wild-type and mutant proteins with very similar structural properties. The similar far-UV CD spectra of all variants as well as the same gel filtration retention times (not shown) indicate that the secondary structure content and the dimeric arrangement of the proteins are essentially unaffected by the mutations. Fig. 2 shows the intrinsic fluorescence spectra of the proteins. The \( \lambda_{\text{max}} \) values were the same, indicating that a similar polarity characterizes the environment of the tryptophanyl residues of all enzyme variants. However, the normalized intensities of fluorescence of most mutants were lower than that of the wild-type enzyme. This finding suggests that significant conformational changes distinguish the final structure of the mutants from that of the parent enzyme. These differences imply shorter distances between some fluorescence quenching groups and one or both tryptophanyl residues located in the GST domain I, far from the mutation site.

**Thermal Stability of the Wild-type and Mutant Enzymes—**
The thermal stability of wild-type and hydrophobic staple motif mutant proteins, heterologously expressed at 25 °C, was investigated. Residual activity was monitored after 10 min of exposure to different temperatures. As shown in Fig. 3, A and B, all mutants were destabilized compared with the wild-type GSTP1-1. The I149A (Fig. 3A) mutant was completely inactivated at 50 °C, whereas the wild-type enzyme, under the same conditions, was inactivated to less than 10%. Furthermore, mutants Y154A and I149A/Y154A were significantly more unstable than both wild-type GSTP1-1 and the I149A mutant, being almost completely inactivated by incubation at 40 °C for 10 min. For the Y154A and the I149A/Y154A mutants about 60% of the activity remained after the same time at 37 °C. The fact that the thermal inactivation curves for the single Y154A and the double I149A/Y154A mutants were essentially the same suggests that interactions made by the Tyr154 residue with other regions of the molecule play a major role in the global stability of the protein. At 40 °C, the L21A and Y154F mutants were 40 and 10% inactivated, respectively, suggesting that the hydrophobic interaction made by Tyr154 with the \( \alpha \)-helix is essential for enzyme stability (Fig. 3B). The results also show that about 40% of the activity of mutant I149G is lost when this variant is incubated at 40 °C. Moreover, the I149G/Y154A variant, being inactivated to 70% at 37 °C (Fig. 3B), is more unstable than the I149A/Y154A double mutant (Fig. 3A).

Thus, an Ala to Gly substitution at position 149 in GSTP1-1 clearly affects the overall stability of the enzyme.

**Reactivation Yields at Different Temperatures—**
The proposed role of the hydrophobic staple motif involves determination of the direction of the polypeptide chain during folding. One consequence of substituting residues that contribute to this motif could be the generation of temperature-sensitive folding mutants. To test this possibility, reactivation yields of GSTP1-1 and its mutants at different temperatures were determined. Fig. 4 shows that when the reactivation was performed at 25 °C the double mutant I149A/Y154A possessed a higher specific activity (35 units/mg) than that observed for the wild-type and the single mutants I149A and Y154A. Still, at 25 °C all enzymes show the same (70–80%) reactivation yield. Fig. 4 also shows that while the wild-type enzyme was essentially unaffected by temperature in the 25–40 °C range, the reactivation yields of mutants, even if to different extents, decreased with temperatures increasing toward physiological values. In particular, the I149A/Y154A double mutant, already at 32 °C, displayed a substantial reduction in the yield of refolded active form. On the contrary, at the same temperature, the reactivation yield of the I149A single mutant was complete with a specific activity (25 units/mg), slightly higher than that of the wild-type. At 32 and 40 °C lower refolding yields were observed for the Y154A variant than for the I149A mutant, suggesting that a role in the refolding could be played by the specific interdomain interactions between the \( \alpha \)-6- and \( \alpha \)-1-helices identified in the crystal structure of the native enzyme. To address this aspect, the refolding yields of L21A and Y154F

---

**TABLE II**

| Protein           | \( k_{\text{cat}}^{\text{Wild-type}} \) | \( K_{\text{m}}^{\text{Wild-type}} \) | \( K_{\text{m}}^{\text{GSTP1-1}} \) | \( k_{\text{cat}}^{\text{Wild-type}}/K_{\text{m}}^{\text{Wild-type}} \) |
|-------------------|---------------------------------------|-----------------------------------|----------------------------------|--------------------------------------------------|
| Wild-type         | 0.15 ± 0.02                           | 0.94 ± 0.10                       | 36                               |
| I149A             | 0.16 ± 0.02                           | 0.87 ± 0.10                       | 48                               |
| Y154A             | 0.26 ± 0.04                           | 1.18 ± 0.18                       | 38                               |
| I149A/Y154A       | 0.17 ± 0.04                           | 1.16 ± 0.17                       | 73                               |
| L21A              | 0.60 ± 0.09                           | 1.94 ± 0.19                       | 38                               |
| Y154F             | 0.29 ± 0.03                           | 0.90 ± 0.09                       | 63                               |
| I149G             | 0.20 ± 0.01                           | 0.97 ± 0.11                       | 63                               |
| I149G/Y154F       | 0.41 ± 0.02                           | 0.64 ± 0.06                       | 77                               |
| I149G/Y154A       | 0.26 ± 0.01                           | 1.25 ± 0.11                       | 61                               |

---

**FIG. 2. Intrinsic fluorescence spectra of human GSTP1-1 and hydrophobic staple mutants.** Uncorrected intrinsic fluorescence spectra at the normalized protein concentration of 0.1 mg/ml. The spectra correspond to: A, wild-type (---); I149A (----), Y154A (-----); I149A/Y154A (-- ---); I149A/Y154A (-- - ---); B, wild-type (---), I149G (-----), Y154F (-----), L21A (----), I149G/Y154F (-----), I149G/Y154A (--- ---). The excitation wavelength was 280 nm; the standard \( N \)-acetyl-l-tryptophanamide gave a \( \lambda_{\text{max}} \) at 355 ± 3 nm.
mutants were also analyzed. As shown in Fig. 5 much lower refolding yields were observed for the L21A mutant as compared with those of mutant Y154F.

The I149G single mutant is completely unable to refold at 40 °C and even at 32 °C it displayed a refolding yield significantly lower than that attained at 25 °C (Fig. 5). Moreover, for the corresponding I149G/Y154A (Fig. 6) and I149G/Y154F (not shown) double mutants complete recovery of enzymatic activity could not be obtained even at 25 °C.

Analysis of Reactivation at 25 ºC—A single exponential equation could be fitted to all data sets for the refolding kinetics (not shown). However, as shown in Table III, while the refolding rates of both I149A and Y154A single mutants were only slightly different from that of the wild-type enzyme, the I149A/Y154A double mutant, in which the staple motif has been significantly weakened, was characterized by a reactivation velocity about 6 times lower than that of the wild-type. These differences in the refolding rates reflect changes in the free energy of activation of folding upon mutation. Double mutant cycles analysis (20, 21) has been used to estimate the energy contribution ($\Delta G_{int}$) of the staple interaction at the transition state of folding. As reported in Table III and Fig. 7, an estimated value significantly different from 0 ($\pm 0.53$ kcal/mol) indicates that the hydrophobic staple interaction is present in the transition state of the reactivation process. The results show that also the I149G single mutant refolds at a lower rate than does the wild-type enzyme and the estimated $\Delta G_{int}$ was $-0.25$ kcal/mol. The I149G/Y154F and I149G/Y154A double mutants, in distinction from the I149A/Y154A variant, were unstable even at temperatures as low as 25 °C suggesting that the substitution of Ile$^{149}$ to Gly destabilizes the transition state of folding more dramatically than does the I149A muta-
tion. In contrast to the other GSTP1-1 variants, the reactivation rate of the L21A mutant was not significantly different from that of the wild-type, suggesting that the hydrophobic contact formed by Tyr154 with domain I of the protein probably forms later on the folding pathway. For all variants, the reactivation rate was independent of the protein concentration (not shown).

**DISCUSSION**

Although the (i, i+5) hydrophobic staple interaction is a recurrent motif at the N terminus of α-helices, its possible function in folding and stability has not previously been tested experimentally in a complete protein. In the present work we have found that this supersecondary interaction plays an essential role in the folding and stability of a GST protein. In the GST superfamily the amino acid residues that make up this motif represent, together with Gly146, the only residues strictly conserved in domain II in all GSTs, but their role has not been investigated until now.

To study folding and stability in vitro, GSTP1-1 and its corresponding mutants were expressed at 25 °C. At this permissive temperature the kinetic and structural properties of all mutants were only slightly different from those of the wild-type protein, suggesting that the global folding of all variant enzymes were very similar. However, all mutants were less stable, although to various degrees, than the wild-type enzyme (Fig. 3). The N-4 mutant (Y154A), in particular, was significantly more destabilized than the N' mutant (I149A). A simple explanation for this difference is provided by the analysis of the three-dimensional structure of human GSTP1-1. As shown in Fig. 1, the residue Tyr154, in addition to the hydrophobic staple motif.
interactions with Ile\textsuperscript{149}, forms conserved interactions with amino acid residues Leu\textsuperscript{21}, Asp\textsuperscript{24}, and Gln\textsuperscript{25}, belonging to an important structural element (the \(\alpha\)-helix) of the active site. The almost identical thermal inactivation curves for the single mutant Y154A and the double mutant I149A/Y154A support the interpretation that these interdomain interactions are critical for stability of class \(P\) GSTs. Furthermore, the L21A mutant is much more unstable than the Y154F variant. This indicates that among all interdomain contacts made by Tyr\textsuperscript{154} and the \(\alpha\)-helix, the conserved hydrophobic interaction of Tyr\textsuperscript{154} and Leu\textsuperscript{21} plays an essential role for the overall stability of the protein. The results also show that the I149G/Y154A variant is more unstable than the I149A/Y154A double mutant. Thus, although the complex unfolding behavior of GSTP1-1 does not allow accurate determinations of thermodynamic parameters (18), the present results demonstrate that a glycine residue at position 149 significantly destabilizes the structure of GSTP1-1.

Refolding \textit{in vitro} of all mutants was thermosensitive. The results show that the reactivation yield of the I149A/Y154A double mutant was much lower than that found for either of the two I149A and Y154A single mutants and provide an explanation for the very low yield observed for the heterologous expression of this mutant in \textit{E. coli}. In particular, the double mutant I149A/Y154A is completely unable to refold at 32 °C, although the final structure of this variant is stable at this temperature. This means that during refolding of GSTP1-1 the contribution from the staple motif is significant also in a situation where other interactions between Tyr\textsuperscript{154} and the \(\alpha\)-helix occur. The results also show that refolding yields of the Y154A variant were lower than those of the I149A mutant. Considering the interactions made by Tyr\textsuperscript{154} with other residues of the \(\alpha\)-helix, these results suggest that the appropriate docking of the \(\alpha\)-helix with the \(\alpha\)-helix in domain I is also required for a productive folding pathway. In order to better clarify the significance of different interactions formed by Tyr\textsuperscript{154} and domain I, the reactivation of L21A and Y154F mutants have also been analyzed. While the refolding yield of this last variant was almost identical to that of wild-type, the very low refolding yield observed for the L21A mutant suggest that the interdomain hydrophobic interaction made by Tyr\textsuperscript{154} with Leu\textsuperscript{21} plays a major role in the juxtapositioning of the domains and in establishing the proper conformation of class \(P\) GSTs.

The analysis of the reactivation rates at 25 °C indicates that the removal of the hydrophobic staple motif destabilizes the transition state of folding. Upon mutation of GSTP1-1, the rate constants decrease for nearly all mutants (Table III). This is because side chain deletions, to various extents, destabilize the transition state and thereby increase the activation energy for folding. In particular the \(\Delta G_f\) value for the Y154A variant is higher than that for the I149A mutant (Fig. 7). As discussed below, some additional interactions made by Tyr\textsuperscript{154} with its environment further contribute to the stabilization of the transition state. The results also show that the reactivation rate of the I149A/Y154A double mutant is lower than that of either the I149A and Y154A single mutants as could be expected (Table III). It should be noted that a single exponential equation can be fitted to the refolding data for all GSTP1-1 variants suggesting that no significant amount of intermediate is accumulated during the reactivation of wild-type and variants and that none of the mutations have a dramatic effect on the refolding pathway. Thus, double mutant cycle analysis was used to estimate quantitatively free-energy contribution of the staple interaction between Ile\textsuperscript{149} and Tyr\textsuperscript{154} to the activation energy of the folding process. As shown in Fig. 7, the estimated value of \(\Delta G_f\) for the native is 0.17 kcal/mol, for I149A single mutant 0.35 kcal/mol, for I149A/Y154A double mutant 0.70 kcal/mol, and for Y154A single mutant 0.88 kcal/mol.
Hydrophobic Staple Motif in GST Folding

(−0.70 kcal/mol) for the substitution of Ile149 by Ala in the Y154A single variant was higher than that determined (−0.17 kcal/mol) for the same substitution in the wild-type enzyme. The difference (−0.53 kcal/mol, Table III and Fig. 7) represents the contribution (ΔΔG) of the staple interaction to the activation barrier of folding and indicates that this motif affects the transition state.

Similarly, the analysis of reactivation rate of the Y154F mutant also suggests that the interdomain hydrogen bond between Tyr154 and Asp24 is formed early during GSTP1-1 refolding. On the contrary, the refolding rate of the L21A mutant was similar to that of the wild-type. This indicates that the interaction between Tyr154 and Leu21 does not influence the activation energy of the folding process. It is conceivable that this interdomain contact is formed very late on the folding pathway, after the transition state, and maybe is not realized until the approach of the final state.

Since Ala residues still can form a weak hydrophobic staple during folding (4, 5), the reactivation analysis of the I149G variant as well as of the I149G/Y154A and I149G/Y154F double mutants was also performed. The I149G mutant is unable to refold at physiological temperatures. It should be noted that at 37 °C the native state at this variant (expressed at 25 °C) is stable for several hours (not shown) confirming that the mutation affects an essential step during GSTP1-1 refolding. The corresponding I149G/Y154A double mutant, in contrast to the I149A/Y154A variant, was characterized by a very low refolding yield even at 25 °C. Thus, the presence of a glycine at position 149 in the core of GSTs instead of a hydrophobic residue is even more destabilizing and precludes accurate energy determinations of the possible Ala-Ala interaction during refolding.

The present results indicate that the removal of the hydrophobic staple motif does not necessarily prevent folding but, at least at permissive temperatures, decreases the rate of the process. At 25 °C, variants of GSTP1-1 are generated that are characterized by a higher catalytic efficiency but a lower structural stability. Given the lower solubility of folding intermediates (23), the slower refolding rate could favor, at more restrictive temperatures, competing aggregation reactions that lead to the complete loss of reactivation at physiological temperatures. The reactivation rate was independent of the protein concentration indicating that the elimination of the hydrophobic staple motif gives rise to a new limiting step of refolding that does not involve dimerization. Considering the proposed role of this motif in determining a reversal of the polypeptide chain direction, the results suggest that the appropriate conformation of the α-helix relative to the preceding long loop represents an important feature of a transition state occurring during monomer formation. The conserved hydrophobic staple motif could stabilize one out of several closely related conformations of the α-helix that favors essential interdomain interactions that in some cases are formed later in the refolding process.

It has been hypothesized (24, 25) that, for monomeric single domain proteins, buried and highly conserved residues, making a definite set of native like contacts, determine the formation of a specific folding nucleus that serves as a necessary template onto which other parts of the molecule can form. The present work provides evidence also that in a more complex multidomain enzyme, such as specific nucleus, contacts do exist and, in GSTs, they are at least in part determined by conserved local motifs.

Different theoretical studies suggest that there is an evolutionary pressure for proteins to fold rapidly (23, 25–28). Amino acid replacements that dramatically retard the folding process might be selected against, because competing aggregation reactions could reduce the yield of the native form. The present work identifies a rapid-folding determinant of the polypeptide chain that, at least in GSTs, is represented by a conserved hydrophobic staple motif.

REFERENCES

1. Richardson, J. S., and Richardson, D. C. (1988) Science 240, 1648–1652
2. Presta, L. G., and Rose G. D. (1988) Science 240, 1632–1641
3. Seale, J. W., Srinivasan, R., and Rose, D. G. (1994) Protein Sci. 3, 174–1–1745
4. Munoz, V., Blanco, F. L., and Serrano, L. (1995) Nat. Struct. Biol. 2, 380–385
5. Munoz, V., and Serrano, L. (1995) Biochemistry 34, 15301–15306
6. Aurora, R., and Rose, D. G. (1996) Protein Sci. 7, 21–38
7. Hayes, J. D., and Pulford, D. J. (1995) Crit. Rev. Biochem. Mol. Biol. 30, 445
8. Mannervik, B., and Danielsson, U. H. (1988) CRC Crit. Rev. Biochem. 23, 283–337
9. Armstrong, R. N. (1997) Chem. Res. Toxicol. 10, 2–18
10. Reinemer, P., Dirr, H. W., Ladenstein, R., Schaffer, J., Gallay, O., and Huber, R. (1991) EMBO J. 10, 1997–2005
11. Ji, X., Zhang, P., Armstrong, R. N., and Gilliland, G. L. (1992) Biochemistry 31, 10169–10184
12. Sinning, I., Kierweg, G. J., Cowan, S. W., Reinemer, P., Dirr, H. W., Huber, R., Gilliland, G., Armstrong, R. N., Ji, X., Board, P. G., Olin, B., Mannervik, B., and Jones, T. A. (1995) J. Mol. Biol. 252, 192–212
13. Wilke, M. C. J., Board, P. G., Feil, S. C., and Parker, M. W. (1995) EMBO J. 14, 2133–2143
14. Aetao, A., Dragani, B., Melino, S., Allocaii, N., Masulli, M., Di Ilio, C., and Petruzzelli, R. (1997) Biochem. J. 322, 229–234
15. Dragani, B., Stenberg, G., Melino, S., Petruzzelli R., Mannervik, B., and Aetao, A. (1997) J. Biol. Chem 272, 25518–25523
16. Koonin, E. V., Mushogbui, A. R., Tatunov, R. L., Altshul, S. F., Bryant, S. H., Bork, P., and Valencia, A. (1994) Protein Sci. 3, 2045–2054
17. Kolm, R. H., Stenberg, G., Widersten, M., and Mannervik, B. (1995) Protein Exp. Purif. 6, 265–271
18. Aetao, A., Caccuri, A. M., Sacchetta, P., Bucciarelli, T., Dragani, B., Rosato, N., Federici, G., and Di Ilio, C. (1992) Biochem. J. 285, 241–245
19. Aetao, A., Di Ilio, C., L Bello, M., Bucciarelli, T., Angelucci, S., and Federici, G. (1990) Carcinogenesis 11, 2267–2269
20. Jackson, S. E., elMasry, N., and Fersht, A. R. (1993) Biochemistry 32, 11270–11278
21. Horowitz, A., and Fersht, A. R. (1990) J. Mol. Biol. 214, 613–617
22. Weiner, S. J., Kollman, P. A., Case, D. A., Singh, U. C., Ghio, C., Alagona, G., Profeta, J. S., and Weiner, P. (1984) J. Am. Chem. Soc. 106, 765
23. Stackhouse, T. M., Onduffer, J. J., Matthews, R., Ahmed, S. A., and Miles, K. W. (1988) Biochemistry 27, 824–832
24. Pitsyn, O. B. (1996) FEBS Lett. 10, 3–4
25. Shakhnovich, E., Akhrevich, V., and Pitsyn, O. (1996) Nature 378, 96–98
26. Fersht, A. R. (1997) Curr. Opin. Struct. Biol. 7, 3–9
27. Shakhnovich, E. I. (1997) Curr. Opin. Struct. Biol. 7, 29–40
28. Pitsyn, O. B. (1995) Adv. Protein Chem. 47, 83–229
A Conserved "Hydrophobic Staple Motif" Plays a Crucial Role in the Refolding of Human Glutathione Transferase P1-1

Gun Stenberg, Beatrice Dragani, Roberta Cocco, Bengt Mannervik and Antonio Aceto

J. Biol. Chem. 2000, 275:10421-10428.
doi: 10.1074/jbc.275.14.10421

Access the most updated version of this article at http://www.jbc.org/content/275/14/10421

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 28 references, 5 of which can be accessed free at http://www.jbc.org/content/275/14/10421.full.html#ref-list-1