The Three-Dimensional Structure of a Glutathione *S*-Transferase from the Mu Gene Class. Structural Analysis of the Binary Complex of Isoenzyme 3–3 and Glutathione at 2.2-Å Resolution[†,‡]

Xinhua Ji,[§,∥] Pinghui Zhang,[∥] Richard N. Armstrong,[*,∥] and Gary L. Gilliland[*,§,⊥]

Center for Advanced Research in Biotechnology of the Maryland Biotechnology Institute, University of Maryland, Shady Grove, and National Institute of Standards and Technology, 9600 Gudelsky Drive, Rockville, Maryland 20850, and Department of Chemistry and Biochemistry, University of Maryland, College Park, Maryland 20742

Received May 26, 1992; Revised Manuscript Received July 22, 1992

ABSTRACT: The crystal structure of a mu class glutathione *S*-transferase (EC 2.5.1.18) from rat liver (isoenzyme 3–3) in complex with the physiological substrate glutathione (GSH) has been solved at 2.2-Å resolution by multiple isomorphous replacement methods. The enzyme crystallized in the monoclinic space group *C*2 with unit cell dimensions of $a = 87.98$ Å, $b = 69.41$ Å, $c = 81.34$ Å, and $\beta = 106.07°$. Oligonucleotide-directed site-specific mutagenesis played an important role in the solution of the structure in that the cysteine mutants C86S, C114S, and C173S were used to help locate the positions of mercuric ion sites in nonisomorphous derivatives with ethylmercuric phosphate and to align the sequence with the model derived from MIR phases. A complete model for the protein was not obtained until part of the solvent structure was interpreted. The dimer in the asymmetric unit refined to a crystallographic $R = 0.171$ for 19 298 data and $I \geq 1.5\sigma(I)$. The final model consists of 4150 atoms, including all non-hydrogen atoms of 434 amino acid residues, two GSH molecules, and oxygen atoms of 474 water molecules. The dimeric enzyme is globular in shape with dimensions of 53 × 62 × 56 Å. Crystal contacts are primarily responsible for conformational differences between the two subunits which are related by a noncrystallographic 2-fold axis. The structure of the type 3 subunit can be divided into two domains separated by a short linker, a smaller α/β domain (domain I, residues 1–82), and a larger α domain (domain II, residues 90–217). Domain I contains four β-strands which form a central mixed β-sheet and three α-helices which are arranged in a $\beta\alpha\beta\alpha\beta\beta\alpha$ motif. Domain II is composed of five α-helices. Domain I can be considered the glutathione binding domain, while domain II seems to be primarily responsible for xenobiotic substrate binding. The active site is located in a deep (19-Å) cavity which is composed of three relatively mobile structural elements: the long loop (residues 33–42) of domain I, the $\alpha4/\alpha5$ helix–turn–helix segment, and the C-terminal tail. GSH is bound at the active site in an extended conformation at one end of the β-sheet of domain I with its backbone facing the cavity and the sulfur pointing toward the subunit to which it is bound. Fifteen hydrogen bond or salt-bridge contacts are involved in the interaction between GSH and the enzyme, among which the hydrogen bond (3.2 Å) between the sulfur of GSH and the side-chain hydroxyl group of Y6 is of special interest. This interaction appears to lower the pK_a of the sulfhydryl group [Liu, S., Zhang, P., Ji, X., Johnson, W. W., Gilliland, G. L., & Armstrong, R. N. (1992) *J. Biol. Chem. 267*, 4296–4299] and appears to help orient the thiolate anion in the active site such that half of its surface is shielded by hydrophobic residues and the other half is exposed for reaction with the electrophilic substrate. The nucleophilic reactivity of the GSH appears to be enhanced by both destabilization of the thiol and desolvation of the thiolate anion. The structure of the class mu isoenzyme is compared to that of a pi class isoenzyme which was recently reported at a preliminary stage of refinement [Reinemer, P., Dirr, H. W., Ladenstein, R., Schaffer, J., Gallay, O., & Huber, R. (1991) *EMBO J. 10*, 1997–2005].

The glutathione *S*-transferases (EC 2.5.1.18) are a group of enzymes that catalyze the nucleophilic addition of the tripeptide glutathione (GSH)[1] to substrates bearing electrophilic functional groups. As such they are probably the single, most important class of enzymes involved in the detoxication of endogenous and xenobiotic electrophilic substances. The enzymes are ubiquitous and quite abundant in mammalian tissues comprising, for example, about 5% of the total cytosolic protein in hepatic tissue. Their very abundance and ability to initiate the metabolism of a broad range of alkylating agents suggest that this family of proteins plays a central role in the detoxication of many carcinogens as well as anticancer chemotherapeutic agents. For a historical perspective as well as more recent reviews the reader may consult Jakoby (1978), Jakoby et al. (1984), Mannervik (1985), Mannervik and Danielson (1988), Pickett and Lu (1989), and Armstrong (1991).

The cytosolic GSH transferases are dimeric molecules composed of subunits with molecular masses of about 25 kDa. The various subunit types that make up the several isoenzymic species are derived from four distinct gene families which have been designated alpha, mu, and pi (Mannervik et al.,

[†] This work was supported by grants from the National Institutes of Health (GM30910) and the American Cancer Society (BC 632).

[‡] The crystallographic coordinates have been deposited in the Brookhaven Protein Data Bank under the file name 1GST.

* Address correspondence to these authors.

§ Center for Advanced Research in Biotechnology.

∥ University of Maryland, College Park.

⊥ National Institute of Standards and Technology.

[1] Abbreviations: GSH, glutathione; GSBzI, *S*-(3-iodobenzyl)glutathione; EMP, ethylmercuric phosphate; GSPhen, (9*R*,10*R*)-9-glutathionyl-10-hydroxy-9,10-dihydrophenanthrene; DTNB, 5,5′-dithiobis(2-nitrobenzoic acid); EDTA, ethylenediaminetetraacetic acid.

0006-2960/92/0431-10169$03.00/0

1985) and theta (Meyer et al., 1991). Sequence identities among subunits within the same gene class are quite high, typically in the range of 60–80%, while inter-gene-class identities are considerably less, about 25–35% sequence identity. Within a given gene class the holoenzymes can exist as either homodimers or heterodimers. However, inter-gene-class heterodimers have not been observed, suggesting that there are specific structural requirements for familial subunit–subunit interactions. The comparative structural details for subunit recognition are not known.

Each isoenzyme of the GSH transferases tends to have characteristic catalytic properties that are flexible enough to accommodate several types of electrophilic substrates. The remarkably broad substrate spectrum of this group of enzymes is manifest in two levels of molecular design: namely, the existence of structurally and catalytically distinct subunits and a relative lack of constraints on substrate recognition by a given subunit type. The structural determinants for substrate recognition by any GSH transferase have not been elucidated. However, some early hints that the amino- and carboxy-terminal regions of the polypeptides may be involved in the architecture of the active site, which were obtained by regional mutagenesis of isoenzyme 3–3 of the mu gene class (Zhang & Armstrong, 1990) and photoaffinity labeling of isoenzymes 1–1 and 2–2 of the alpha class (Hoesch & Boyer, 1989), are consistent with the recently published three-dimensional structure of a porcine enzyme in the pi family (Reinemer et al., 1991). Nevertheless, the details of the structural basis for the catalytic specificity of various subunits, even within a particular gene class, have not been elucidated.

Even though considerable progress has been made in elucidating the nature of the catalytic mechanism of the GSH transferases, many of the fundamental details remain poorly understood. For example, chemical intuition and, more importantly, considerable spectroscopic and chemical evidence indicate that GSH exists predominantly as the thiolate anion in the binary enzyme–GSH complex (e.g., E·GS$^-$) (Graminski et al., 1989a). Yet the nature of the electrostatic field in the active site which is necessary to stabilize the anion has not been fully defined. Site-specific mutagenesis of conserved and positively charged residues has been a bit informative in this regard. For example, the direct involvement of an imidazolium side chain in catalysis has been unambiguously ruled out (Wang et al., 1991; Zhang et al., 1991). In contrast, Stenberg et al. (1991a) have reported evidence, again by site-directed mutagenesis, that three conserved arginines in an isoenzyme from the alpha gene class are involved either in the binding of GSH or in stabilization of productive conformational states of the enzyme. The preliminary analysis of the structure of a pi class enzyme from pig complexed with the inhibitor glutathione sulfonate (GSO_3^-) (Reinemer et al., 1991) is consistent with the proposed involvement of secondary structural elements (e.g., dipoles of α-helices) in the stabilization of bound GS$^-$ (Chen et al., 1988; Zhang et al., 1991). In addition, this structure provided the first evidence that a tyrosine residue might be involved in catalysis, a conclusion drawn from the observation that the hydroxyl group of Y7 was within hydrogen-bonding distance of one of the sulfonate oxygens of the inhibitor. The case for the involvement of a tyrosine residue in catalysis has been corroborated both by X-ray crystallography of the binary complex of isoenzyme 3–3 from the mu gene class with the actual substrate GSH (the full details of which are reported in this paper) and by spectroscopic and kinetic analysis of the Y6F mutant of the same isoenzyme (Liu et al., 1992) and a related human isoenzyme (Penington & Rule, 1992). Stenberg et al. (1991b) have also confirmed that this conserved tyrosine is also important in an isoenzyme from the alpha class.

Significant progress has been made toward the determination of the three-dimensional structures of some GSH transferases. Crystals suitable for X-ray diffraction studies have been obtained for at least one example of a GSH transferase from three of the principal gene classes (Sesay et al., 1987; Schaffer et al., 1988; Cowan et al., 1989; Parker et al., 1990; Dirr et al., 1991). Furthermore, the crystals of the mu class isoenzyme 3–3 (Sesay et al., 1987) have been shown to be chemically competent to form the σ-complex between GSH and 1,3,5-trinitrobenzene (Graminski et al., 1989b). However, only one structure, that of a pig enzyme of the pi gene class, has been recently reported at a preliminary stage of refinement (Reinemer et al., 1991).

In this paper we report the first determination of the three-dimensional structure of an isoenzyme of GSH transferase from the mu gene class. The structure of the binary complex of isoenzyme 3–3 from rat with the physiologic substrate GSH is analyzed to reveal a number of the structural features of the active site that may be involved in substrate recognition and catalysis. The overall architecture and the details of the subunit–subunit interactions in the dimeric holoenzyme are discussed and compared to the recently reported preliminary structure of the class pi isoenzyme from porcine lung (Reinemer et al., 1991).

EXPERIMENTAL PROCEDURES

General Materials and Methods. Enzymes were assayed with 1-chloro-2,4-dinitrobenzene at pH 6.5 and 25 °C as previously described (Graminski et al., 1989b). The binding constant for GSBzI was determined from the competitive inhibition kinetics observed with GSH as the variable substrate as described by Graminski et al. (1989a). Sulfhydryl group titrations were performed by the method of Ellman (1959). Buffers and chemical reagents were of the highest quality commercially available.

Enzymes. Recombinant isoenzyme 3–3 of glutathione *S*-transferase was produced in *Escherichia coli* (M5219) harboring the expression plasmid pGT33MX (Zhang et al., 1991) and purified as previously described (Zhang & Armstrong, 1990). Genes encoding the three mutants C86S, C114S, and C173S in which the codon for each cysteine residue was replaced with that for serine were constructed by oligonucleotide-directed site-specific mutagenesis. The 563-base-pair fragment between the *Xho*I and *Csp*45I restriction sites of pGT33MX was asymmetrically cloned into the same two restriction sites of phagemid pGEM-7Zf(–). The region was then removed using the *Xba*I and *Hin*dIII sites of the phagemid which flank the *Xho*I and *Csp*45I sites, respectively, and directionally subcloned into M13mp19. Mutagenesis was carried out by using the oligonucleotide-directed in vitro mutagenesis system (version 2) from Amersham[2] (Sayers et al., 1988) following the manufacturer's instructions. This involved codon changes of TGT to TCT for the C86S and C114S mutants and TGC to TCC for the C173S mutant. The appropriate oligonucleotide primers were all 19 bases long with the mismatched base in the middle. Plaques obtained

[2] Certain commercial equipment, instruments, and materials are identified in this paper in order to specify the experimental procedure as completely as possible. In no case does such identification imply a recommendation or endorsement by the National Institute of Standards and Technology nor does it imply that the material, instrument, or equipment identified is the best available for the purpose.

from the Amersham system were screened by DNA sequencing. The regions bearing the mutation to be cloned back into the expression plasmid pGT33MX were completely sequenced to ensure that no secondary mutation occurred. The mutant expression plasmids were designated pGT33MXC86S, pGT33MXC114S, and pGT33MXC173S to indicate the C86S, C114S, and C173S mutations, respectively. The mutant enzymes were expressed and isolated in the same way as was the native protein. Expression of the three mutant proteins in *E. coli* was unremarkable with typical isolated yields of between 0.1 and 0.2 g of protein per liter of cell culture. Sulfhydryl group titers of the mutant proteins obtained with DTNB (Ellman, 1959) were 1.9, 1.8, and 1.9 per subunit for C86S, C114S, and C173S, respectively, as compared to 2.9 for the native enzyme, a result consistent with the loss of one cysteine per subunit. None of the mutations affected the specific activity of the enzyme to any significant extent. Turnover numbers for the C86S, C114S, and C173S mutants with 1-chloro-2,4-dinitrobenzene were 19 ± 4, 24 ± 3, and 24 ± 2 s^{-1}, respectively, and were comparable to that of the native enzyme (19 ± 1 s^{-1}).

S-(3-Iodobenzyl)glutathione. This active site-directed heavy-atom-containing compound was prepared by the reaction of glutathione (1 mmol, dissolved in 2 mL of H_2O containing 2 mmol of NaOH) with 3-iodobenzyl bromide (Lancaster Synthesis, Inc.) (1 mmol, dissolved in 2 mL of ethanol) under nitrogen at room temperature. The reaction was judged to be complete after 4 h by sulfhydryl group titration with DTNB (Ellman, 1959). The crude product was precipitated by adjusting the pH of the reaction mixture to 2.1 with 6 N HCl and storing overnight at 4 °C. The crude material was crystallized from a 1/1 mixture of ethanol and water to give 0.69 mmol of *S*-(3-iodobenzyl)glutathione (GSBzI) as a white powder with a mp 196–7 °C (dec) which had ^1H and ^{13}C NMR spectra consistent with the expected structure.

Crystal Growth. Crystals of the recombinant protein were grown in sitting drops which initially consisted of between 8 and 12 mg/mL protein in 25 mM Tris buffer (pH 8.0) containing 1 mM EDTA, 0.2% octyl β-D-glucopyranoside, 1 mM product inhibitor (9*R*,10*R*)-9-*S*-glutathionyl-10-hydroxy-9,10-dihydrophenanthrene (GSPhen), and buffered (pH 8) ammonium sulfate at 40–50% saturation. The drops were equilibrated at room temperature against wells containing between 60% and 72% saturated ammonium sulfate in 25 mM Tris buffer (pH 8.0). This procedure, which is a modification of the original one reported by Sesay et al. (1987), gives crystals of the same habit that grow in 5–10 days at between 65% and 72% ammonium sulfate. Crystallization of the cysteine mutants was accomplished by the same technique.

Replacement of the Product Inhibitor in Crystals. Replacement of the product inhibitor GSPhen, which occupies the active site of the crystalline enzyme, with either GSH or GSBzI was accomplished through diffusion by mass action at 25 °C. Single crystals were transferred from the wells in which they were grown to wells containing 100 μL of a 1.0 mM solution of either GSH or GSBzI in 70% saturated ammonium sulfate buffered with 25 mM Tris (pH 8.0). The solution surrounding the crystal was removed and refreshed with 100-μL aliquots every hour for 4–6 h.

X-ray Data Collection and Data Processing. X-ray diffraction data were collected using a Siemens electronic area detector mounted 10 cm from the crystal. The volume of the crystals used for data collection was no smaller than 0.05 mm^3. The area detector was mounted on a three-axis Supper oscillation camera controlled by a PCS microcomputer. X-rays were generated with a Rigaku RU200B rotating anode operated at 40 kV and 60–80 mA. A graphite monochromator was used, followed by a 0.3–0.5-mm collimator, depending upon the size of the crystal. Each electronic frame was composed of counts summed over a 0.1–0.25° range in ω with exposure times of 60 s or longer. The individual frames were contiguous in that the beginning of each angular range in ω coincided with the end of the previous range. The raw data frames were transferred from the PCS microcomputer to a Silicon Graphics Personal IRIS computer for final processing. All data collection was carried out at controlled room temperature (19–20 °C).

The space group was determined from the systematic absences on X-ray precession photographs, and the unit cell parameters were first obtained from X-ray diffraction measurements (Sesay et al., 1987) on a CAD4 diffractometer. The determination of crystal orientation and the integration of reflection intensities were performed with the XENGEN program system (Howard et al., 1987). A summary of the crystal data and final results of data processing for the native and five derivative data sets is presented in Table I, and detailed statistics for the native data are given in Table II. The unit cell parameters listed in Table I were obtained from the area detector data and refined by using the XENGEN data processing package (Howard et al., 1987).

Heavy-Atom Derivatives. The structure of the binary complex of mu gene class isoenzyme 3–3 with the physiologic substrate GSH was solved by using the multiple isomorphous replacement technique (MIR), followed by iterative cycles of refinement and model building. Solvent flattening and electron density averaging were also used in the early stages to improve the electron density maps. A large variety of heavy-atom derivatives were surveyed, but only three were used to produce the final MIR map. Diffusion of the active site-directed heavy-atom-containing compound GSBzI, which is a potent competitive inhibitor (K_i = 0.2 μM) of isoenzyme 3–3, into the native crystals gave the first isomorphous heavy-atom derivative, **2**, that also facilitated the location of the active site. Heavy-atom derivatives **2** and **3** (CdI_2) (Table I) were isomorphous with the native crystals, and the heavy-atom positions were derived from the interpretation of the difference Patterson maps. A third derivative was sought by taking advantage of the fact that the type 3 subunit of GSH transferase contains three cysteine residues (C86, C114, and C173). However, reaction of the native crystals with a variety of mercuric ions led to derivatives that were not isomorphous with the native crystals. For example, reaction of native crystals (containing GSPhen) with ethylmercuric phosphate produced a derivative (**6**) that was not isomorphous with the native crystals as evidenced by the change in unit cell dimensions, a high scaling *R* factor, and a cluttered difference Patterson map.

The three cysteine mutants, C86S, C114S, and C173S, were prepared in an effort to reduce the number of sites of reaction for mercurials and increase the changes of obtaining isomorphous heavy-atom derivatives (Dao-Pin et al., 1987; Veerapandian et al., 1992). Single crystals of the C86S and C114S mutants having the same morphology as crystals of the native protein were obtained using crystallization conditions identical to those used for the native protein. Suitable crystals of the C173S mutant were not obtained. Data sets collected to 3.2 Å with crystals of C86S and C114S were found to index with unit cell dimensions very similar to those of the native crystals and to have scaling *R* factors of ≤0.10 with the native

Table I: Crystal Data and Results of X-ray Data Processing for the Native Crystals (**1**) and Five Derivative Data Sets (**2–6**) and Statistics of MIR Phasing Calculations with Derivatives **2–4**

		derivatives				
	native **1**[a]	**2**	**3**	**4**	**5**	**6**
unit cell						
a (Å)	87.98	88.12	87.80	88.77	88.06	88.37
b (Å)	69.41	69.02	68.61	69.08	69.46	69.67
c (Å)	81.34	82.29	81.43	80.40	80.65	80.57
β (deg)	106.07	106.09	105.57	107.01	107.23	107.48
data						
D_{min}(Å)	2.2	3.2	2.5	3.2	3.2	3.2
unique data obsd	22810	7576	12704	7649	7553	7647
data with $I \geq 2\sigma(I)$	19920	7318	7911	7294	7229	7333
R_ω[b]	0.076	0.073	0.053	0.061	0.059	0.074
$R_{u\omega}$[c]	0.064	0.059	0.042	0.052	0.049	0.063
$R_{scaling}$[d]		0.16	0.20	0.39	0.42	0.41
MIR						
data used (Å)		10.0–3.2	10.0–3.2	10.0–5.0	not used in MIR phasing	
centric data		357	322	180		
R_{Cullis}[d]		0.54	0.52	0.52		
phasing power[f]		2.06	1.85	1.55		
mean figure of merit		0.69	0.61	0.50		

[a] Native **1**: isoenzyme 3–3 + GSH. Derivatives: **2**, isoenzyme 3–3 + GSBzI; **3**, isoenzyme 3–3 + GSH + CdI_2; **4**, C114S mutant of isoenzyme 3–3 + GSPhen + EMP; **5**, C86S mutant of isoenzyme 3–3 + GSPhen + EMP; **6**, isoenzyme 3–3 + GSPhen + EMP. [b] The weighted least-squares R factor on intensity for symmetry-related observations: $R_\omega = \sum[(I_{ij} - G_{ij}\langle I\rangle_j)/\sigma_{ij}]^2/\sum(I_{ij}/\sigma_{ij})^2$, where $G_{ij} = g_i + A_i s_j + B_i s_j^2$; $s = \sin\theta/\lambda$, and g, A, and B are scaling parameters. [c] The unweighted absolute R factor on intensities: $R_{u\omega} = \sum(I_{ij} - G_{ij}\langle i\rangle_j)/\sum|I_{ij}|$. [d] Scaling R between derivative and native data. [e] $R_{Cullis} = \sum\|FPH_{obs}| \pm |FP_{obs}\| - |FH_{cal}\|/\sum\|FPH_{obs}| \pm |FP_{obs}\|$. [f] Phasing power = FH_{cal}/E.

Table II: Statistical Details of Native Data (**1**) for Isoenzyme 3–3 + GSH

				no. of unique reflections				
shell lower limit (Å)	av intensity I	av $I/\sigma(I)$	no. of measurements	expected	obsd	$I \geq 2\sigma(I)$	R_ω[a]	$R_{u\omega}$[a]
3.96	1162	64.2	19 853	4 193	4 191	4 148	0.060	0.044
3.15	654	41.1	17 769	4 098	4 098	4 001	0.075	0.061
2.75	282	20.6	16 960	4 124	4 124	3 824	0.112	0.101
2.50	165	12.3	15 686	4 100	4 100	3 530	0.147	0.138
2.32	111	7.8	13 078	4 039	4 036	3 104	0.183	0.173
2.18	84	4.2	4 124	4 120	2 260	1 313	0.193	0.185
totals	440	26.9	87 470	24 674	22 810	19 920	0.076	0.064

[a] The definitions for R_ω and $R_{u\omega}$ are given in footnotes *b* and *c* of Table I.

data. However, derivatives **4** and **5** obtained by the reaction of crystals of C86S and C114S with EMP had the same unit cell dimensions of **6** as well as the same nonisomorphous character with respect to the native crystals. Derivatives **4–6**, however, were isomorphous with respect to each other, a fact which was exploited to locate the positions of C86 and C114 in both monomers. The positions of the two C86 and the two C114 residues were located from the solutions of the difference Patterson maps between **5** and **6** and between **4** and **6**, respectively. The difference Patterson map between **4** and **5** was consistent with the above solutions. Guided by the known positions of the two C86 residues, the difference Patterson map between **4** and **1** was solved at lower resolution (5 Å) by means of $\Delta\Delta F$ (cross difference Fourier syntheses) using the SIR phases from **2**, resulting in the positions of two additional heavy atoms which bind to the two C173 residues. The best MIR phases were obtained, at 3.2 Å, with the use of the isomorphous contributions of three derivatives, **2–4**, and the anomalous contribution of derivative **2**. The mean figure of merit of the resulting phases was 0.692 for 7341 phased reflections (Table I). The program package PHASES (Furey, 1990) was used to carry out all of the calculations. The correct handedness was established by using the technique described by Blundell and Johnson (1976).

Structure Solution and Refinement. FRODO (Jones, 1978) was used on an Evans & Sutherland PS390 graphics system for the examination of MIR, $2F_o - F_c$, $F_o - F_c$, and omit maps, the adjustment of the model, and the deletion and addition of water molecules. In the later stage of model building, the "Lego" facilities embedded in O (Jones et al., 1991; Jones & Kjeldgaard, 1991) and run on a Silicon Graphics Personal IRIS were used to check the loops against a protein database of best refined protein structures.

The original MIR map was not readily interpretable, although it did reveal some secondary structural features of the molecule. In 70% of the molecule, however, the secondary structure was not apparent. A portion of the MIR map for residues 62–66 of a β-strand along with the final model is shown in Figure 1a, while the corresponding final electron density ($2F_o - F_c$) map is shown in Figure 1c. The noncrystallographic symmetry, a 2-fold axis between the two subunits (Sesay et al., 1987), was used to average the electron density using the program package SARASA developed by J. K. Mohana Rao (Jaskolski et al., 1990). The correlation coefficient between the two monomers was 0.28. The electron density averaged map was then subjected to solvent flattening (Wang, 1985; Furey, 1990). The density averaging and the solvent flattening improved the mean figure of merit from 0.692 to 0.785. Figure 1b shows the improvement of the resulting map. Such improvement, however, was achieved mainly in the regions with a good MIR map for at least one monomer. When the MIR map was not interpretable in either monomer, the electron density averaging and solvent flattening did not improve the quality of the map significantly. The overall quality of the resulting map was marginal so traditional minimaps were used to assist in the tracing of the polypeptide

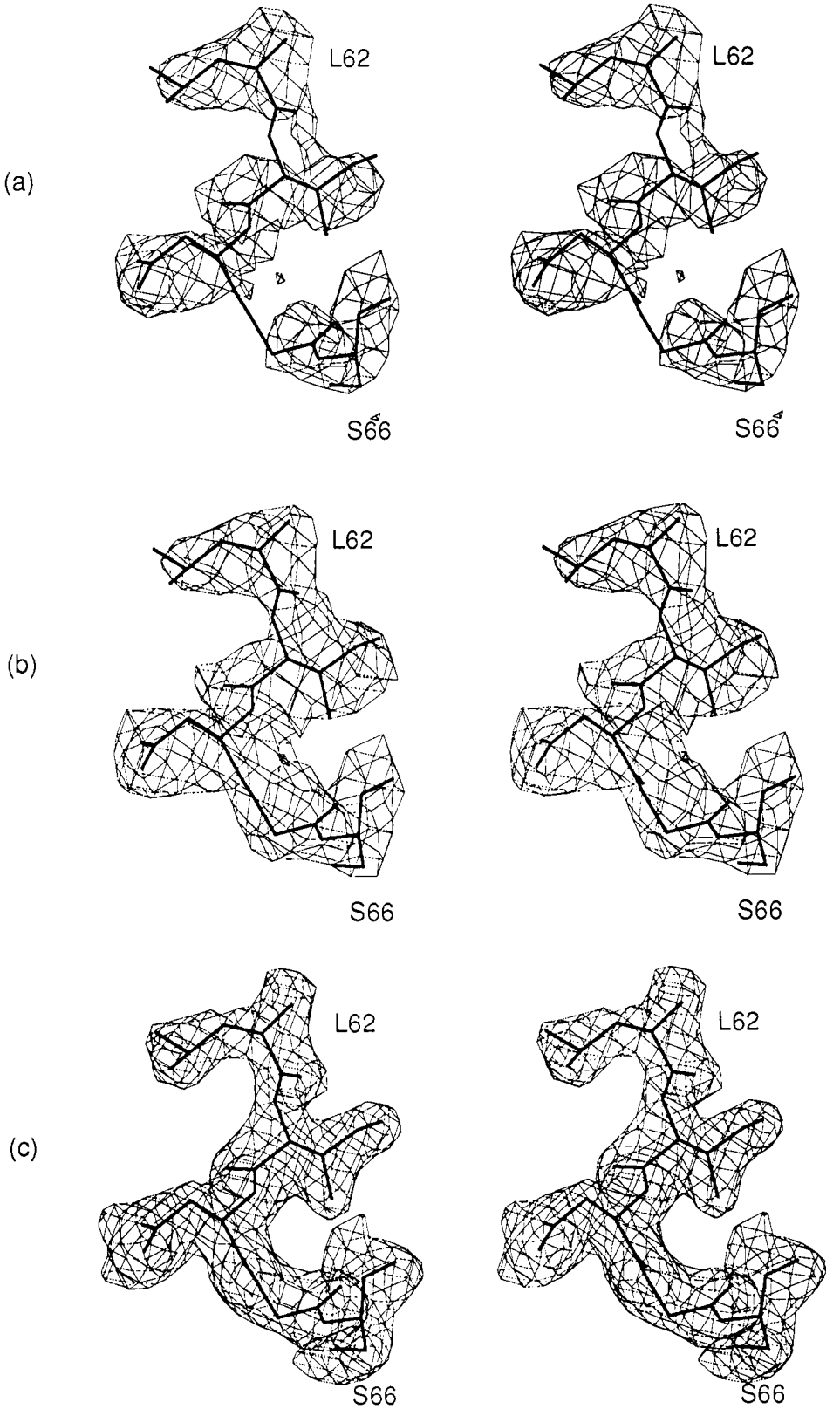

FIGURE 1: An example of the electron density maps (a) for original MIR phases, (b) for the phases after electron density averaging and solvent flattening, and (c) for the final phases ($2F_o - F_c$). All electron density is contoured at 1σ.

chain. The first model was composed of two polypeptide chains, each containing 150 serine or alanine residues. The coordinates of all free atoms were then used as variables, and the total energy was minimized with the use of the Powell method conjugate gradient minimizer embedded in X-PLOR (Powell, 1977; Brünger, 1990). The calculated phases of the energy-minimized model were combined with the original MIR phases with no phase extension. Electron density averaging and solvent flattening were applied to the combined phases, and the model was rebuilt according to the new map with more side chains assigned. After one more cycle of energy minimization and phase combination and map fitting, the adjusted model was subjected to simulated annealing (Brünger et al., 1987; Brünger, 1990) with the initial temperature set at 3000 °C. The combined phases thereafter gave rise to the recognition of both N-termini in the electron density map

Table III: Summary of Least-Squares Refinement Parameters for the Binary Enzyme–GSH Complex at the Resolution of 2.2 Å

	target values	final model
diffraction data from 6.0 to 2.2 Å with $I \geq 1.5\sigma(I)$		19298
crystallographic R^a factor		0.171
rms deviations from ideal distances (Å)		
bond distances	0.025	0.019
angle distances	0.036	0.032
planar 1–4 distances	0.040	0.033
rms deviations from planarity (Å)	0.030	0.019
rms deviations from ideal chirality (Å^3)	0.200	0.199
thermal parameter correlations (mean/$\Delta\beta$)		
main-chain bond	1.000	0.739
main-chain angle	1.500	1.254
side-chain bond	1.500	1.189
side-chain angle	2.000	1.798

^{a} $R = \sum_{hkl}||F_o| - |F_c||/\sum_{hkl}|F_o|$.

which, when combined with the known positions of the three cysteine residues in each subunit, made the sequence alignment possible in some regions. Two continuous chains were built, starting with the N-termini, each containing 200 residues. The R value after simulated annealing was 0.246 at a resolution of 3.2 Å. Thereafter, more data, up to 2.2 Å, were included in the refinement. This was followed by a long process of model rebuilding and refinement during which the electron density map improved in small increments. During this process a number of tracing errors were gradually discovered and corrected, and the substrate (GSH) molecules were built into very well defined electron density in both subunits. No solvent molecules were included although some were clearly defined in the electron density map. The overall temperature factor was refined, followed by grouped temperature factor refinement (Brünger, 1990). After 26 cycles of rebuilding, the R value was 0.250 at 2.2-Å resolution. Further refinement was performed with GPRLSA (Furey et al., 1982), the restrained least-squares refinement procedure of Hendrickson and Konnert (1980a,b) and Hendrickson (1985a,b). The entire model was checked, and rebuilt when necessary, against a series of omit maps (Furey, 1990) with 30 residues omitted each time. However, the electron density in the regions from residues 37 to 43, from 115 to 122, and from 201 to 217 in both subunits did not improve sufficiently to build an unambiguous model of these areas. In order to improve the phasing power of the model, about 150 water molecules, which were found as the peaks of the $F_o - F_c$ map greater than 3σ and which either formed more than one hydrogen bond with the protein or formed hydrogen bonds to the backbone amides, were gradually added to the model. With the interpretation of solvent structure, the density in the problem areas mentioned above slowly improved. After an additional 20 cycles of least-squares refinement and model adjustment, the final coordinates of the two independent monomers of isoenzyme 3–3 of glutathione S-transferase had a crystallographic R value of 0.171 for 19 298 diffraction data between 6.0 and 2.2 Å with $I \geq 1.5\sigma(I)$. A summary of the refinement statistics is given in Table III. Low-resolution data excluded from the refinement were included in all map calculations. The final coordinates have been deposited in the Brookhaven Protein Data Bank (Bernstein et al., 1977).

RESULTS

Final Model of the Structure. A stereoview of the dimeric enzyme in complex with the substrate GSH is presented as ribbon diagrams in panels a and b of Figure 2. The final model consists of 4150 atoms, including all non-hydrogen atoms of 434 amino acid residues (A1–A217 and B1–B217), two GSH molecules (I218–I220 and J218–J220), and oxygen atoms of 474 water molecules (1–474). The water molecules are numbered sequentially from the largest to the smallest values of OCC^2/B (James & Sielecki, 1983).[3] The polypeptide torsional angles (ϕ and ψ) are presented in the Ramachandran plot (Ramachandran & Sasisekharan, 1968) of Figure 3 (supplementary material). Among the ϕ, ψ values of non-glycine residues, only those of the two Q71 residues are located outside the energetically preferred area. This is the only residue between the two secondary structural elements, β-strand β4 and α-helix α3 (see following sections). It is very well defined by the electron density map. The Q71 residues are involved in subunit interactions and in substrate binding (see corresponding sections below) and, therefore, may have structural significance related to the function of the enzyme. The electron density is in complete agreement with the amino acid sequence of the 217-residue type 3 subunit. Three of the twelve proline residues (38, 60, and 206) are observed as having the cis conformation. Of the three P60 and P206 are conserved in the cytosolic glutathione S-transferases.

Secondary Structure. The secondary structure of isoenzyme 3–3 is defined according to Kabsch and Sander (1983) in Figure 4 (supplementary material). The main secondary structural features of the two subunits, A and B, are summarized and compared in Table IV. The secondary structure of subunit A and of the full asymmetric unit is shown in a RASTER 3D representation (Bacon & Anderson, 1988) in Figure 5. Each subunit contains four β-strands and eight α-helices. The fifth α-helix is actually composed of two segments, α5a and α5b, by virtue of a proline kink near the middle which gives α5 the overall appearance of a long, crescent-shaped helix.[4] The most obvious difference in secondary structure between subunits A and B of isoenzyme 3–3 is in the region of residues 49–51. Residues A49–A51 participate in the formation of a β-turn (A48–A51) while B49, B50, and B51 form a 3_{10} helical turn [Figure 4 (supplementary material) and Table IV]. One intermolecular hydrogen bond is observed between atoms NZ of K51 of the B subunit and the symmetry-related main-chain O of $M197_{1/2-x,1/2+y,-z}$ with an N–O distance of 3.13 Å, which may be the force that stabilizes the helical conformation. This distance is 3.85 Å in subunit A. Otherwise, the two tripeptide segments are subject to very similar intermolecular interactions. Other differences in the secondary structure between the two subunits include the different lengths of four helices, α3, α4, α5a, and α7. α3 is longer in subunit B, while α4, α5, and α7 are longer in subunit A. The differences involve only one or two residues and can be considered minor since they are the consequence of differences in ϕ or ψ angles of $<10°$.

Folding Topology and Domain Structure. The structure of the type 3 subunit can be divided into two domains separated by a short linker, a smaller α/β domain (domain I, residues 1–82), and a larger α domain (domain II, residues 90–217) as illustrated in Figure 2c. Domain I contains four β-strands (β1, β2, β3, and β4) which form a central β-sheet, three α-helices (α1, α2, and α3), two *cis*-proline bends, and two β-turns (one β-turn and a 3_{10} helical turn in subunit B; see Table V). The major secondary structural elements of domain

[3] The quality index OCC^2/B is the ratio of the square of the fractional occupancy (OCC) of the oxygen atom position and the crystallographic temperature factor B.

[4] Technically, helix α5 is two helices due to the interruption in the hydrogen-bonding pattern by the proline residue at position 131. The two helices are referred to as α5a and α5b in Table IV to reflect this fact.

Table IV: Summary and Comparison of the Secondary Structures of the Two Subunits in Isoenzyme 3-3 of Glutathione S-Transferase from Rat Liver and Comparison of Secondary Structures between the Mu Class Isoenzyme from Rat (This Work) and the Pi Class Isoenzyme from Pig (Reinemer et al., 1991)

rat mu class isoenzyme 3-3[a]			pig pi class[b]	
	no. of residues and range[d]			
structural element[c]	subunit A	subunit B	no. of residues and range[d]	structural element[c]
		Domain I		
β1	6 (2–7)	6 (2–7)	**5** (3–7)	β1
β-turn	(10–13)	(10–13)	(11–14)	β-turn
α1	9 (14–22)	9 (14–22)	9 (15–23)	αA
β2	6 (27–32)	6 (27–32)	**4** (29–32)	β2
cis-Pro bend	(36–39)	(36–39)	4 (34–37)	β-turn
α2	4 (43–46)	4 (43–46)	4 (38–41)	αB
3_{10}		(49–51)	(41–43)	3_{10}
β-turn	(48–51)			
bend	(54–57)	(54–57)	(45–48)	β-turn
cis-Pro bend	(58–61)	(58–61)	(49–52)	*cis*-Pro bend
β3	4 (61–64)	4 (61–64)	4(52–55)	β3
β4	4 (67–70)	4 (61–70)	4 (58–61)	β4
α3	**11** (72–82)	12 (72–83)	12 (63–74)	αC
		Domain II		
α4	**25** (90–114)	**24** (90–113)	**27** (81–107)	αD
β-turn	(116–119)	(116–119)		
α5a	**9** (119–127)	**8** (119–126)	**24** (109–132)	αE
α5b	12 (130–141)	12 (130–141)		
β-turn	(141–144)	(141–144)	(135–137)	3_{10}
bend	(146–149)	(146–149)	(140–143)	β-turn
α6	16 (154–169)	16 (154–169)	16 (148–163)	αF
β-turn	(170–173)	(170–173)	(164–167)	β-turn
β-turn	(174–177)	(174–177)	(168–171)	β-turn
α7	**12** (178–189)	**10** (178–187)	**11** (172–182)	αG
α8	7 (191–197)	7 (191–197)	**8** (185–192)	αH
β-turn	(199–202)	(199–202)	(196–199)	β-turn
cis-Pro turn	(204–207)	(204–207)		
β-turn	(209–212)	(209–212)		

[a] Rat class mu glutathione S-transferase (this work). [b] Porcine class pi glutathione S-transferase (Reinemer et al., 1991). [c] α1, α-helix 1; αA, α-helix A; β1, β-stand 1; etc. [d] Number of residues in the element of secondary structure. Boldfaced numbers indicate differences. Numbers in parentheses are sequence numbers in corresponding structures of which the sequence alignment can be found in Figure 11.

I are arranged in a βαβαββα motif. The four β-strands in the β-sheet are arranged such that β1 and β2 are parallel and that β1 and β3 and β3 and β4 are antiparallel (Table IV, Figure 5). Domain II is composed of five α-helices [α4, α5 (a and b), α6, α7, and α8], six β-turns, and a *cis*-proline bend (Table IV). The two domains are linked via a short polypeptide segment, residues 83–89. Hydrogen bonds and electrostatic interactions appear to play an important role in stabilizing the relative positions of the two domains (Table V). Fifteen hydrogen bonds and salt bridges are found between the two domains in subunit A and nineteen in subunit B. In addition, 12 water molecules (1, 3, 28, 30, 33, 51, 70, 75, 85, 119, 131, and 204) form hydrogen bonds to both domains in subunit A and 10 (2, 7, 16, 27, 49, 84, 113, 135, 204, and 416) in subunit B. Water molecule 204 forms hydrogen bonds to the four domains at the same time.

Interactions between Subunits. The two subunits of isoenzyme 3-3 are related by a noncrystallographic 2-fold axis corresponding to a rotation of 179.8° around an axis defined by the direction cosines (0.954, –0.048, 0.297) with a translation of 0.057 Å along the rotation axis. The equivalent Eulerean rotation is $\phi = 87.467°$, $\theta = 145.498°$, and $\psi = 93.231°$. The two subunits contact one another primarily by interactions between domain I of one subunit and domain II of the adjacent subunit (Figures 2b and 5b) with a relatively open hydrophilic channel which runs through the dimer and is coincident with the 2-fold axis. Contacts between the two subunits involve several of the secondary structural elements (Table IV), including the β-turn A48–A51 (3_{10} helix B49–B51 in subunit B), the bend at residues 54–57, β4, α3, α4, and α5b. Among 173 short contacts (≤4.0 Å) are 17 salt bridges and hydrogen bonds. These electrostatic interactions involve Q71 and the two helices, α3 and α4, of both subunits as is illustrated in Figure 6. Of particular note is the interaction between the guanidino groups of R77 of each subunit. The two guanidino groups are actually stacked on top of one another at the dimer interface, as can be seen in the center of Figure 6. The charge on the side-chain R77 is mitigated by close intrasubunit contact (<3 Å) with D97 and E100 (Table V). Eleven water molecules (5, 8, 13, 21, 31, 32, 48, 58, 75, 113, and 204) form hydrogen bonds to both subunits. The burial of solvent-accessible surface upon dimerization of the two subunits is a modest 13% (2820 Å^2) with a 10% (970 Å^2) loss in hydrophilic surface and 15% (1850 Å^2) loss of accessible hydrophobic surface. One particularly obvious hydrophobic contact between domain IA and domain IIB is the intercalation of the side chain of F56 of one subunit between α4 and α5b of the other subunit where it is wedged between L136, Y137, F140, I98, and Q102 as illustrated in Figure 6.

Crystal Packing and Comparison of the Two Subunit Structures. The dimer in the asymmetric unit (*x*, *y*, *z*) is in contact with nine symmetry-related molecules with 21 intermolecular salt bridges and hydrogen bonds and 224 van der Waals contacts (Table VI, supplementary material). The locations of crystal contacts in the primary structure of subunits A and B are illustrated in Figure 7 (panels c and d). Sixty-six surface residues (31 in subunit A and 35 in subunit B) are involved in intermolecular contacts. In addition, 11 water molecules (38, 49, 57, 59, 61, 132, 136, 193, 233, 260, and 380) form hydrogen bonds to both *x*, *y*, *z* dimer and symmetry-related molecules. Crystal packing is the main reason for the differences found in the least-squares fitting of the Cα backbones of the two subunits.

Despite the small differences in the secondary structures of the two subunits, the three-dimensional structures are nearly identical. A comparison of the two subunits is shown in Figure 8 as a least-squares fit of Cα backbones of the two polypeptide chains. The rms deviation for 217 pairs of Cα's is 0.340 Å. Regions of maximum deviation for the main-chain atoms include amino acids 35–39, residue 65, and residue 151 shown in Figure 8 as regions a, b, and c, respectively. The rms deviation for 210 residue pairs excluding the above is 0.271 Å. Two of the regions of maximum deviation appear to be the result of intermolecular interactions due to crystal packing. Residues A37–A40 are subject to 18 intermolecular interactions with distances shorter than 4.0 Å (Table VI, supplementary material) while the corresponding region of the B subunit (B35–B40) has no contact with symmetry-related molecules (Figure 7, compare panels c and d). In addition, the region near B65 (residues B64–B67) has 12 short intermolecular contacts while no such contacts are observed between A64–A67 and symmetry-related molecules (Table VI). The third major difference (residues A151 and B151) is a result of two intramolecular hydrogen bonds between NZ of K151 of the B subunit and the main-chain carbonyls of H84 (3.13 Å) and C86 (2.86 Å) of the same subunit which are not formed between the corresponding residues of subunit A. It is expected that in solution the two subunits should have the same overall conformation.

Local Mobility. The local mobility in the structure as represented by the mean crystallographic temperature factors is instructive in highlighting the similarities and differences

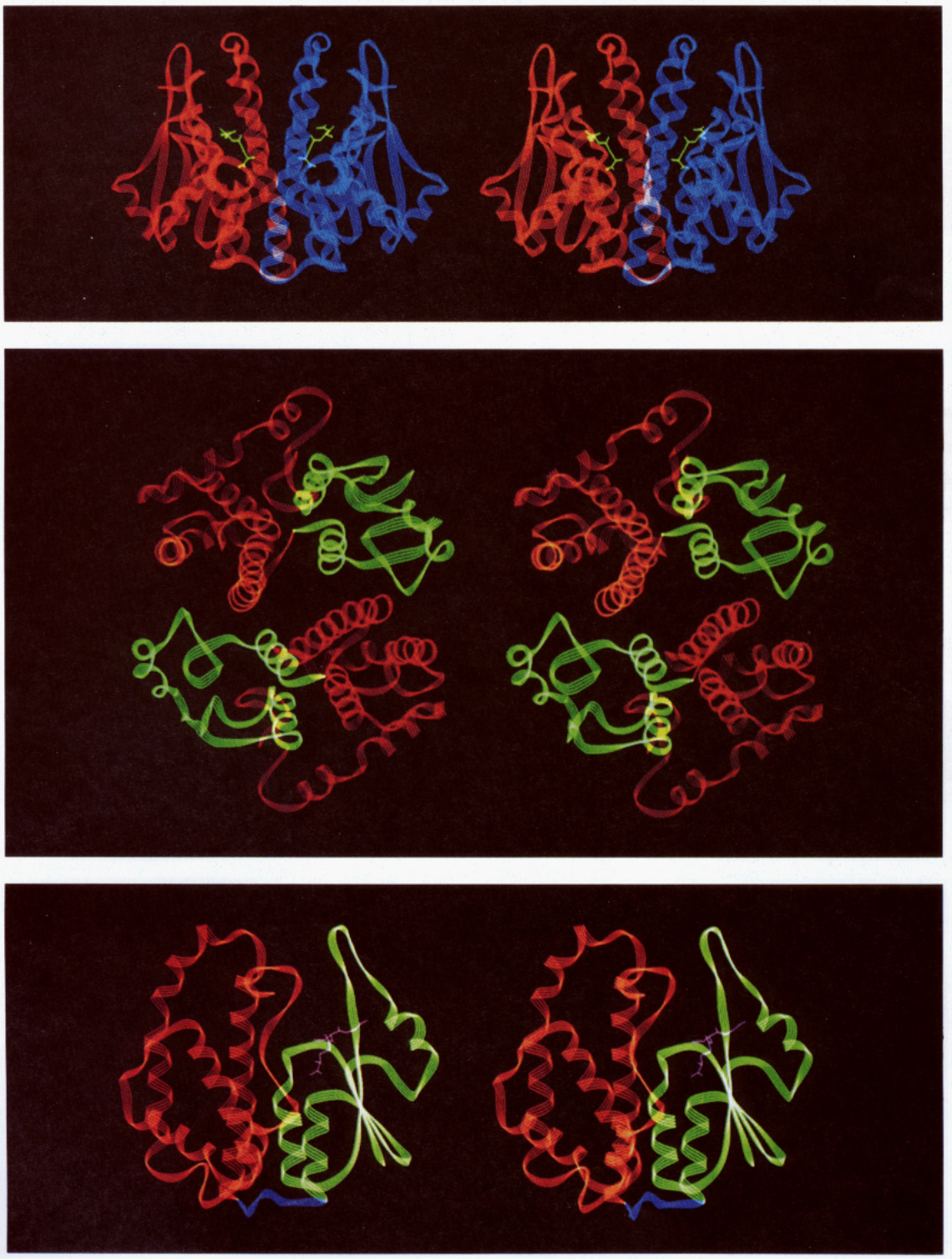

FIGURE 2: (a, top) Stereoview of the dimer of isoenzyme 3–3 of rat glutathione *S*-transferase (ribbon diagram). Subunits A and B are shown in blue and red, respectively. Glutathione is shown as a stick model in green. (b, middle) Stereoview of the dimer looking down the 2-fold axis. Domains I and II of each subunit are shown in green and red, respectively. (c, bottom) Stereoview of monomer A of isoenzyme 3–3. The two domains, I (residues 1–82) and II (residues 90–217), are shown in green and red, respectively, while the heptapeptide link between the domains is shown in blue.

between the subunits as well as the molecular symmetry of the subunit–subunit interactions in the dimer. The overall similarity of the folding of the two subunits is reflected in the mean group temperature factors for each subunit as shown in Figure 7 (panels c and d). In general, the regions of high mobility are associated with external loops and turns, whereas low *B* values involve regions with extensive hydrogen bonding such as α-helices and β-strands. One notable exception is residues 110–130 which comprise the N-terminal half of the α4 helix and α5a which are connected by a tight two-residue turn. This section of structure as well as the loop formed by residues 33–43 and the C-terminal tail (residues 205–217) has higher than average temperature factors as shown in the stereoview of Figure 9 and together constitute a channel to the active site. The structural elements that define the approach to and constitute part of the architecture of the binding site for the xenobiotic substrate have a higher than average local mobility.

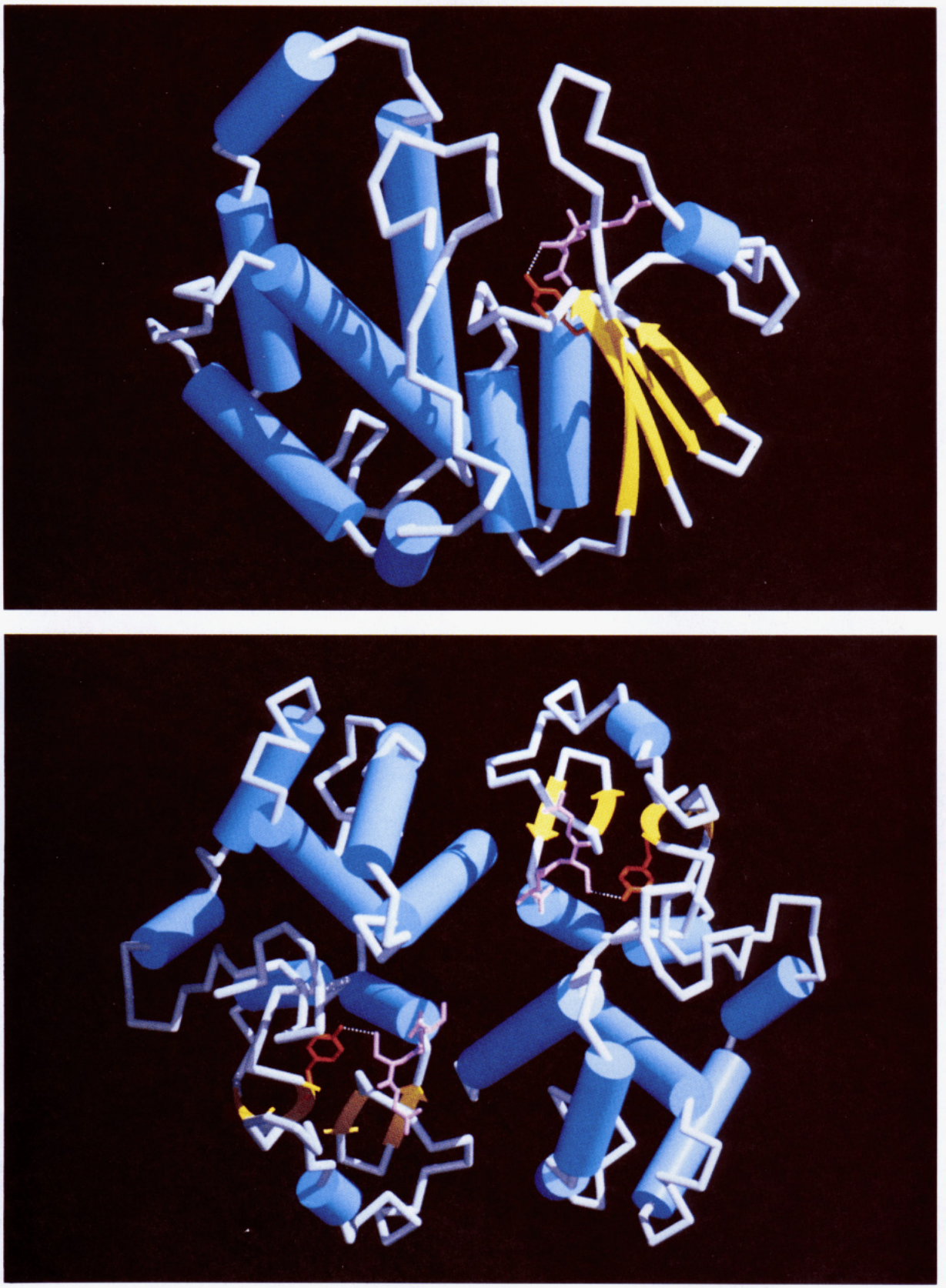

FIGURE 5: RASTER 3D images (Bacon & Anderson, 1988) for (a, top) subunit A of a mu class glutathione *S*-transferase 3–3 from rat liver and (b, bottom) the dimer of the asymmetric unit. β-Strands are indicated by yellow arrows and α-helices by blue barrels. Loops are represented by grey "stove pipes". The active site tyrosine is shown in red and GSH is in purple. The hydrogen bond between the sulfur of GSH and the side-chain hydroxyl group of Y6 is shown by the dotted line. The access channel to the active site from the top is clearly shown in (a) and is defined by the C-terminal tail, the C-terminal end of the α4 helix, and the long mu loop contributed by domain I. The spatial relationship between the two active sites in the dimer is emphasized in (b).

The differences in temperature factors (ΔB) between the two subunits are also interesting, particularly with regard to the relationship of the domain structure to the molecular symmetry of the subunit–subunit interactions. For the purpose

Table V: Hydrogen Bonds and Salt Bridges Observed between Domains I and II

hydrogen bond or salt bridge	distance (Å) subunit A	distance (Å) subunit B
NE(R10)···OD1(D161)	2.92	2.97
NH2(R10)···OD1(D161)	2.88	2.75
NH1(R10)···OD2(D164)	2.76	2.38
NH2(R10)···OD2(D164)	2.34	2.68
NH1(R10)···O(S204)	3.07	2.84
NH1(R19)···OG(S204)		2.79
N(R10)···O(P206)		3.02
N(G11)···O(I207)	2.78	2.66
O(G11)···NE2(Q165)		3.18
ND1(H14)···O(L203)		3.33
OE2(E21)···NE(R201)		2.69
OE1(E21)···NH1(R201)	3.26	
OE2(E21)···NH1(R201)	2.82	
OE1(E21)···NH2(R201)		3.17
OE2(E21)···NH2(R201)		3.35
O(Y22)···NZ(K192)		2.88
OD2(D24)···NZ(K192)	3.01	3.17
NH2(R31)···O(R201)	2.75	2.72
OD1(N73)···OE2(E100)	3.26	
NH2(R77)···OD1(D97)	2.72	3.01
NE(R77)···OE1(E100)	2.89	2.77
NH2(R77)···OE1(E100)	3.14	2.69
NE(R77)···OH(Y154)	3.37	
total hydrogen bonds and salt bridges	15	19

of comparison, ΔB is plotted as a function of sequence number for side-chain atoms and for main-chain atoms separately (Figure 7, panels a and b). There is a relatively clear change (positive to negative) in the average sign of ΔB that coincides with the break between domains I and II of the subunits as might be expected from the molecular symmetry of the dimer. That is, the higher than average B value for regions in domain IA is also manifest in its partner (domain IIB) in the dimer interface. The IB–IIA domain pair has just the opposite property. That the IA–IIB domain pair has higher mobility than does the IB–IIA pair appears to be another consequence of the intermolecular interactions of crystal packing. Domains IA and IIB have 54 and 65 short van der Waals (≤4 Å) and electrostatic (≤3.4 Å) contacts with neighboring molecules, respectively. The less mobile domains, IB and IIA, have 66 and 76 intermolecular interactions, respectively.

Solvent Structure. A total of 474 water molecules are included as solvent structure. They are numbered sequentially from the largest to the smallest values of a reliability factor which is defined as OCC^2/B (James & Sielecki, 1983). The average occupancy factor OCC, temperature factor B, and *reliability* factor OCC^2/B (and ranges) for the 474 solvent molecules are 0.77 (1.0–0.37), 30.26 (5.33–35.52 Å^2), and 0.0236 (0.180–0.004), respectively. The distance criterion for hydrogen bonds to water is set at ≤3.5 Å. No angle restriction is assumed. Of the 474 water molecules 225 form hydrogen bonds to protein atoms, and 154 of these can be considered an integral part of the structure, due to the fact that each forms either two or more hydrogen bonds to protein atoms or has hydrogen bonds to the backbone amides. These water molecules have temperature factors that are very similar to their hydrogen-bonding partners on the protein. The average OCC, B, and OCC^2/B (and ranges) for the 154 protein-bound water molecules are 0.90 (1.0–0.50), 26.75 (5.33–34.80 Å^2), and 0.0369 (0.180–0.007), respectively. Among the 154 are 43 solvent molecules involved in the interaction between the two subunits (11), between the two domains (21), and between the substrate and the protein molecules (11; see corresponding sections). There are 249 water molecules that do not form hydrogen bonds to protein atoms. However, 233 of these do exhibit hydrogen bonds to other water molecules. Water clusters are located not only in the cavities of the dimeric molecule but on the surface of the protein as well. Only 16 of the solvent molecules included in the final model appear not to be hydrogen bonded to any other atom. These 16 waters (62, 88, 99, 101, 161, 168, 186, 210, 299, 305, 313, 349, 379, 435, and 446) all have well-defined electron density and OCC^2/B values that span a wide range (0.035–0.007) of values included in the final cutoff. In some areas, the electron density is more than enough for a water molecule and may be attributable to a larger species such as sulfate ions. However, they are temporarily assigned as water molecules (101, 126, 161, 168, 186, 206, 214, 246, 299, 323, 435) and will certainly be more accurately interpreted at higher resolution.

Glutathione Binding Site. Glutathione is bound (K_d = 20 μM) at the active site of isoenzyme 3–3 in an extended conformation at one end of the β-sheet of domain I with its backbone facing the cavity between the two subunits and the sulfur pointing toward the subunit to which it is bound (Figure 5). The sulfur atoms of the two GSH molecules are 21.6 Å apart. The interaction of GSH with the active site of the enzyme involves up to 15 hydrogen bond or salt-bridge contacts to the protein as illustrated for subunit A in Figure 10. Of particular note are the interactions between the α-amino group of the γ-glutamyl residue of GSH and the protein. These include a hydrogen bond between OE1 of Q71, the residue which displays unusual polypeptide torsional angles, and a salt bridge to D105 of the B subunit, which is the only interaction between GSH and the other subunit in the dimer. Each oxygen of the α-carboxylate of the γ-glutamyl residue is hydrogen bonded to both the main-chain NH and the side-chain hydroxyl of S72. It is also notable that the α-carboxylate of the γ-glutamyl group is on the axis of and within 4 Å of the N-terminal end of the α3 helix (Figure 5b). The α3 helix dipole may therefore make a significant electrostatic contribution to the binding of GSH. There are several interactions between the amides of the substrate and the protein, including hydrogen bonds between the carbonyl of the central cysteine residue and NE1 of W7 and between the NH group and OD1 of A58 of O of L59. The amide carbonyl of the γ-glutamyl group does not interact directly with the protein. The glycine carboxylate of GSH exhibits hydrogen-bonding or salt-bridge interactions with NE1 of W45, NH2 of R42, and NZ of K49.

The interactions of the sulfhydryl group of GSH with the enzyme are of special interest. The primary interaction with the protein is a relatively short hydrogen bond (3.2 Å) between the side-chain hydroxyl group of Y6 and the sulfur. The sulfur atom is also in van der Waals contact with CE1 of Y6 as well as CD and CG2 of L12, interactions which effectively shield one hemisphere of the sulfur from contact with solvent. The remainder of the coordination sphere of the sulfur is occupied by water molecules. The sulfur of GSH in the active site of the A chain shows a hydrogen bond to a single water molecule at 3.4 Å while the GSH associated with the B chain has hydrogen bonds to two waters at distances of 3.3 and 3.4 Å.[5] No evidence for oxidation of the sulfur was seen in the electron density maps. Although the sulfur of GSH is located near the N-terminal end of the α1 helix (Figure 5), the influence of the helix dipole on the electrostatic environment is not clear since the sulfur is located about 7.5 Å from the end of the helix and is off to one side of the helix axis.

FIGURE 6: Stereoview of the hydrogen-bonding pattern and important van der Waals contacts between the two subunits of isoenzyme 3–3. The model of the A subunit is shown in bold lines. Hydrogen bonds and electrostatic interactions are illustrated as dotted lines while side chains involved in the hydrophobic contacts are highlighted as dot surfaces.

Regions of the active site that are probably important in the binding of the xenobiotic substrate are readily apparent from model-building exercises in which various substrate molecules are positioned within the active site and from preliminary structures of product complexes. The active site is located in a deep (19 Å) cavity which is composed of three relatively mobile structural elements: the long loop (residues 33–42) of domain I, the $\alpha 4/\alpha 5a$ helix–turn–helix segment, and the C-terminal tail (Figure 9). Three regions of the polypeptide that lie near enough to the active site to influence the approach of the xenobiotic substrate to the sulfur are the loop (residues 7–13) which connects strand $\beta 1$ to $\alpha 1$, a portion (residues 104–115) of the hydrophobic face of the $\alpha 4$ helix, and part of the C-terminal tail. Specific side chains in these three regions that are near (≤8.5 Å) the sulfur of GSH are W7, V9, L12, M104, M108, I111, M112, Y115, F208, and S209. The nearest residue from the opposite subunit is D105 (mentioned above), which is about 9 Å from the sulfur. The relative independence of the two active sites is apparent in Figure 5b.

DISCUSSION

Comparison of the Structures of the Mu and Pi Class Isoenzymes. The availability of the structure of the pi class enzyme from pig, which has been reported at a preliminary stage of refinement, allows the first three-dimensional structural comparison of isoenzymes from different gene classes. The rat class mu and porcine class pi glutathione *S*-transferases display a relative low degree of sequence identity (31%) as illustrated in Figure 11. Nevertheless, they adopt very similar folding topology. Both proteins have a smaller α/β domain (domain I) containing a similar four-strand β-sheet and a larger all-helical domain II. The secondary structures of the mu and pi enzymes are compared in Table IV. The four β-strands in the β-sheets of both are arranged in the same manner. However, two obvious differences are observed between rat and pig isoenzymes in domain I. One, as discussed in more detail below, is the different lengths of β-strands (Table IV). The other is the long loop between $\beta 2$ and $\alpha 2$, which is present in isoenzyme 3–3 (residues 35–41) but not in pig enzyme. This is obviously a consequence of the difference in the amino acid sequences found in domain I (Figure 11). This loop constitutes part of the principal difference in the overall shape of the two enzyme molecules. In particular, the extra loop (the mu loop) together with the C-terminus of the polypeptide chain and the "loop" formed by $\alpha 4$ and $\alpha 5a$ form a cleft which leads to the active site of the mu isoenzyme that is much deeper than that seen in the pi isoenzyme.

Reinemer et al. (1991) pointed out that domain I of the pig enzyme has same folding pattern ($\beta\alpha\beta\alpha\beta\beta\alpha$) that has been found in thioredoxin from bacteriophage T4 (Söderberg et al., 1978) and that related folds have also been observed in thioredoxin from *E. coli* (Holmgren et al., 1975) and glutathione peroxidase (Ladenstein et al., 1979). The same, of course, is true of domain I of isoenzyme 3–3. Each of these three molecules has a very different sequence from domain I of the GSH transferases. The folding of domain II of each enzyme is very similar even in the details of the secondary structure compared below.

The overall features of the subunit contacts in both dimers are quite similar although fewer secondary structural elements are found to be involved in the intersubunit contacts of porcine enzyme. The secondary structural elements that are involved in subunit contacts in the pi isoenzyme from pig are the β-turn (residues 45–48), strand $\beta 4$, and helices αC, αD, and αE (Reinemer et al., 1991; also see Tables IV and V). The 3_{10} helix 41–43, which is the counterpart of β-turn A48–A51 in

[5] The arrangement of the sulfur atom of GS^- and the two oxygen atoms of waters 153 and 164 approximates an equilateral triangle with S---O distances of 3.26 and 3.42 Å, and O---O distance of 3.16 Å, S–O–O angles of 59.2° and 64.5°, and an O–S–O angle of 56.3°. Since it is not possible to distinguish a water molecule from an ammonium ion, an alternative interpretation could involve an ion pair between an ammonium ion and the thiolate of GS^- with the third corner of the triangle occupied by a water. This arrangement could easily accommodate three linear hydrogen bonds.

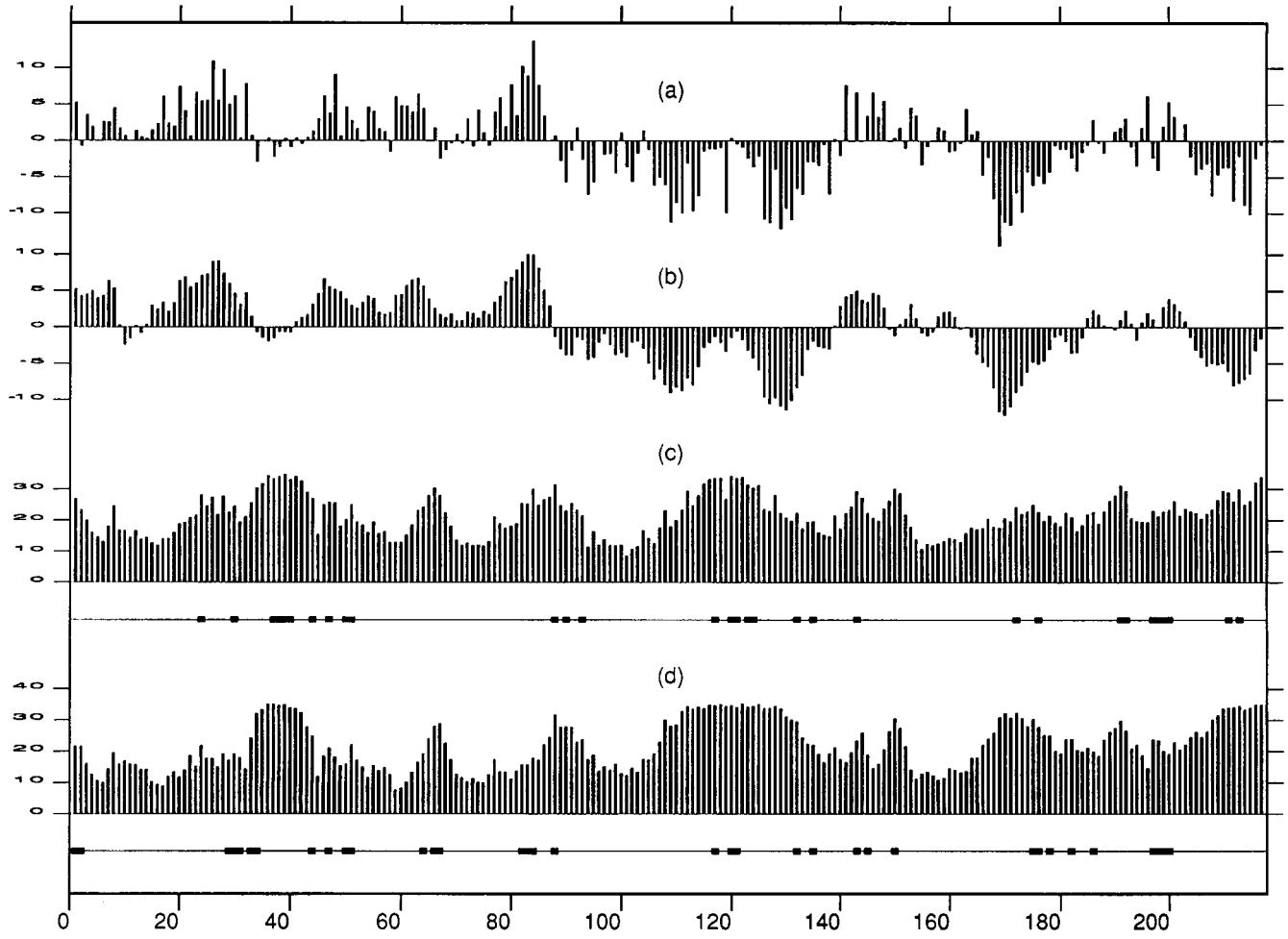

FIGURE 7: Plots of temperature factors, B, and ΔB ($B_A - B_B$) between the two subunits for the final model of the structure of isoenzyme 3–3: (a) ΔB for side-chain atoms and (b) ΔB for the main-chain atoms. Mean group crystallographic temperature factors for the A and B subunits as a function of residue number are shown in panels c and d, respectively. The regions of sequence involved in crystal contacts are illustrated by the bold bars beneath plots c and d of the mean temperature factors.

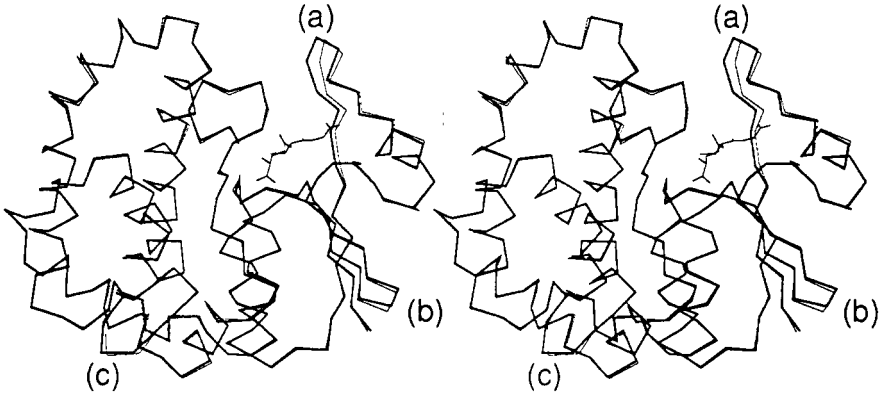

FIGURE 8: Stereoview of the least-squares fit of the Cα positions of the two subunits in the asymmetric unit of isoenzyme 3–3 of rat liver glutathione *S*-transferase. Subunit A is shown in bold lines. The position of GSH is included. The three regions of maximum deviation are shown as (a) residues 35–39, (b) residue 65, and (c) residue 151.

subunit A and the 3_{10} helix B49–B51 in subunit B of rat enzyme, is not reported to be involved in intersubunit interactions.

Although the folding topology of the class mu isoenzyme is quite similar to that of the pi isoenzyme, they do differ in a number of details of the secondary structure as shown in Table IV. For example, three *cis*-proline bends (Table IV) are observed in rat isoenzyme, while only one is described in the pi structure even though two of the three *cis*-prolines are conserved in the primary structures (Figure 11). Only P60 (P51 in Reinemer et al., 1991) assumes the cis conformation in both structures. Two β-turns (45–48 and 140–143) of pig isoenzyme are bends (54–57 and 146–149) in structure of isoenzyme 3–3. The change of R48 (pig) to P57 (rat) prevents the formation of the hydrogen bond necessary to define the first of these β-turns, while the second is probably prevented by the large side chains of two consecutive residues W146 and F147 (rat). A third β-turn (209–212) is not found in pig isoenzyme. Furthermore, the lengths of two β-strands (β1 and β2) and four α-helices (α4, α5, α7, and α8) in the class mu isoenzyme are different from the corresponding structural elements in pig isoenzyme from the pi class. For example, β1

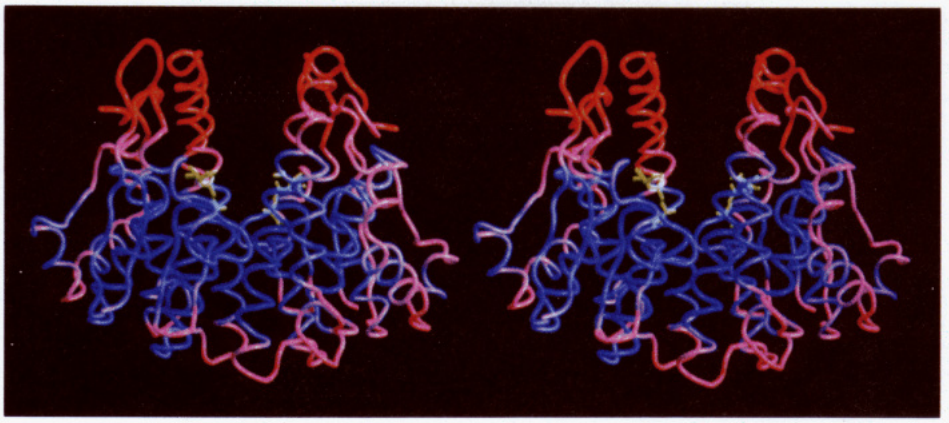

FIGURE 9: Stereoview of a RASTER 3D representation of the backbones of the A and B chains colored as a function of the crystallographic temperature factors. Red corresponds to a high *B* factor (high local mobility) and blue to low *B* factor and mobility.

and β2 with six residues each in rat have an unequal number of residues (five and four) in the porcine protein.

The glutathione binding sites of isoenzymes from the mu and pi class are remarkable for both their similarities and differences. Of the eleven residues that are involved in polar interactions with GSH in the active site of isoenzyme 3–3, seven [Y6, W45, K49, L59, Q71, S72, and D105 (from the opposite subunit)] are conserved in the primary structure of the mu and pi class enzymes and appear to serve precisely the same function.[6] The most remarkable aspects of the conservation are the involvement of Q71, the only residue that has unusual main-chain torsional values in both structures, and the participation of the aspartate residue (D105) from the opposite subunit. In the isoenzyme 3–3 structure Q71 and D105 clearly interact with the free α-NH_2 group of the γ-glutamyl residue of GSH (Figure 10). In contrast, the equivalent side chains Q62 and D96 (opposite subunit) of the pi isoenzyme are reported (Reinemer et al., 1991) to interact with the α-CO_2^- of the γ-glutamyl residue of the GSH analogue, GSO_3^-. This conclusion, drawn from the structure of the pi enzyme at a preliminary stage of refinement (Reinemer et al., 1991), is probably incorrect inasmuch as it is unlikely that the carboxylate of GSH and D96 would constitute a favorable interaction. Moreover, it is difficult to believe that such a remarkable conservation in the structure of the GSH binding site would lead to recognition of a different functional group on the substrate.

There are several outstanding differences in the interactions between GSH and the mu and pi class isoenzymes. For example, the indole nitrogen of W7 of isoenzyme 3–3 forms a hydrogen bond to the carbonyl of the cysteine of GSH. The equivalent residue in the porcine pi isoenzyme is F8, which obviously cannot interact in such a manner. However, the carbonyl of the cysteic acid residue of GSO_3^- is hydrogen bonded to the main-chain amide nitrogen of L50 in the pi structure. The pi isoenzyme does not have a residue equivalent to R42 of isoenzyme 3–3 that interacts with the carboxylate of the glycine of GSH nor does the mu isoenzyme have the equivalent of R13 which is reported to interact with the carboxylate of the γ-glutamyl residue of GSO_3^-. These differences as well as others clearly derive from changes in the amino acid sequences of the two proteins and not from any difference between GSH and GSO_3^-. Although both crystal structures support the notion that arginine residues may play some role in binding GSH, the role is not a conserved one so that its fundamental importance in either enzyme remains uncertain.

Catalytic Mechanism. The two general issues that are crucial to the understanding of how the GSH transferases catalyze their reactions are, first, and perhaps foremost, how the enzyme activates the sulfhydryl group of GSH for nucleophilic attack and, second, what, if any, devices are used by a particular isoenzyme to stabilize transition states for reactions with specific xenobiotic substrates. The structure of the enzyme in complex with GSH directly addresses the first issue. Increasing the reactivity of a thiol can be accomplished by promoting its ionization and by desolvation of the resulting thiolate anion. Consideration of the structural, spectroscopic, and kinetic data suggests that isoenzyme 3–3 does just that in the activation of GSH for nucleophilic attack. The structures of both isoenzyme 3–3 and the porcine enzyme clearly rule out the often proposed direct participation of histidine or cysteine residues in catalysis.

A considerable amount spectroscopic and kinetic evidence has accumulated that the predominant ionization state of the sulfur of GSH in complex with the enzyme is the thiolate (GS^-) (Chen et al., 1988; Graminski et al., 1989a,b; Liu et al., 1992). It is interesting to note that no positively charged residues are near enough to the sulfur of GSH to help stabilize the thiolate anion. The only residue that makes a direct polar contact with the sulfur of GSH is the hydroxyl group of Y6, suggesting that the hydrogen bond (Y6–O–H···$^-$SG) assists in destabilizing (lowering its pK_a) the thiol in the active site. The pK_a of GSH in the binary complex with isoenzyme 3–3 has been measured to be between 6.2 and 6.9 (Liu et al., 1992), which is 2–3 pK_a units below that in aqueous solution. Although the equilibrium position of the proton in the hydrogen-bonded complex has been inferred from its UV spectroscopic properties (Liu et al., 1992), the assignment of this position is not unambiguous. Other possibilities such as (Y6–O^-···H–S–G) where the barrier to proton transfer is very low or a very tight hydrogen bond in which the proton resides in a single-well potential (Y6–O^-···H^+···$^-$SG) (Cleland et al., 1992) cannot be strictly ruled out. The later two possibilities require a sufficiently short O---S distance and that the pK_a values of Y6–OH and GSH be closely matched. It could be that the active site of the enzyme provides an environment in which Y6–OH and GSH have very similar pK_as. This situation provides a mechanism for lowering the effective pK_a of the conjugate acid of the nucleophile without stabilizing it so much as to mitigate its reactivity.

[6] The fact that the sequence alignment of Figure 11 does not indicate that K49 is a conserved residue in the two structures is misleading. Comparison of the secondary structural arrangements of the two proteins in this region as well as the function of K49 suggests that it is a conserved residue with its counterpart in the porcine enzyme being K42.

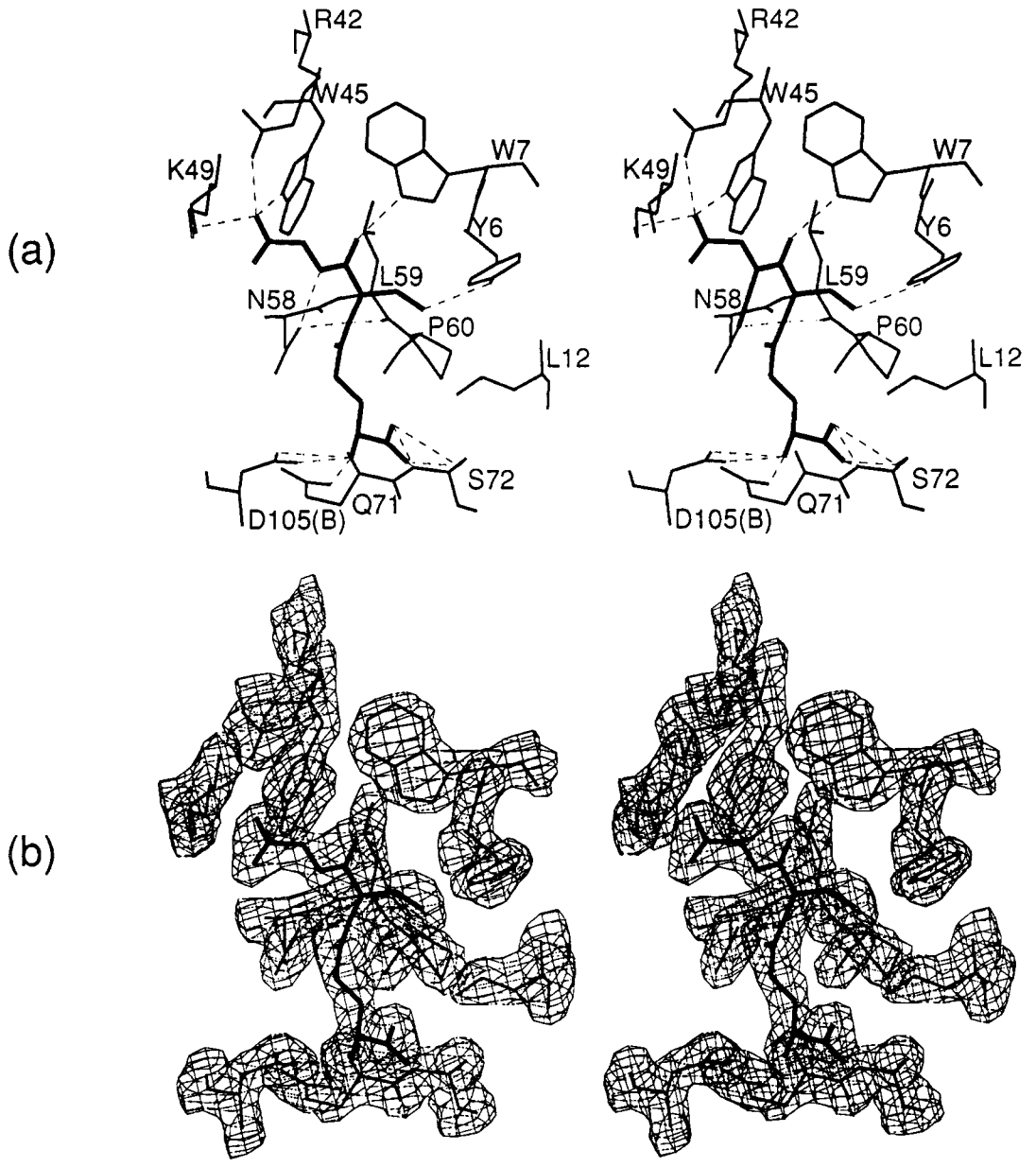

FIGURE 10: Stereoview of the close contacts of GSH bound in the active site of subunit A. (a) Hydrogen bond and salt-bridge contacts are shown as dashed lines. (b) The electron density for GSH and the surrounding residues contoured at 1.0σ is shown to illustrate the quality of the map from which the contacts were derived.

Removal of the hydroxyl group as in the Y6F mutant of isoenzyme 3–3 (Liu et al., 1992) is very instructive with respect to the role of the hydrogen bond in catalysis. Both spectroscopic and kinetic evidence indicates that the pK_a of the sulfhydryl group in complex with the Y6F mutant (E^{Y6F}·GSH) is ≥7.8. The shift in pK_a suggests that the hydrogen bond to Y6 contributes about a factor of 10 to this aspect of the rate acceleration at neutral pH. Proper Brønsted behavior of the nucleophile would predict that E^{Y6F}·GS^- would be as reactive as if not more reactive than E·GS^-. It is not. In fact, pH–rate profiles show that E^{Y6F}·GS^- is about 10-fold less active than native isoenzyme 3–3 at high pH, suggesting that another factor is involved in catalysis. The influence of other direct electrostatic effects on the stability and reactivity of the thiolate of GSH is not obvious. Even though the helix dipole of the N-terminal end of the α1 helix points in the vicinity of the sulfur of GSH and the hydroxyl group of Y6, it is probably too far away to have a significant influence on the ionization of either.

The enzyme also appears to provide for the partial desolvation of the thiolate anion in the active site. One face of the sulfur atom is completely shielded from solvent being pushed up against the side chain of L12. The solvent structure in the active site shows only one or two water molecules hydrogen bonded to the thiolate where at least two or three others would be expected if the anion was fully exposed to solvent. A preliminary examination of the crystal structure of the Y6F mutant in complex with GSH indicates that the geometry of the active site is essentially unchanged with respect to the native enzyme but that an additional water molecule occupies the solvation shell of the sulfur. This change in the hydration state of the nucleophile needs further investigation but may explain why the mutant enzyme is not as active as the native enzyme at high pH.

Xenobiotic Substrate Binding Site. The xenobiotic substrate binding site of isoenzyme 3–3 is composed of regions of the N- and C-terminus as well as the middle of the polypeptide chain. With the exception of Y6 and the neighboring residues which form one wall of the active site, the segments of structure that comprise the electrophilic substrate binding site are located in the large C-terminal domain (domain II) of the molecule. The possible conse-

FIGURE 11: Sequence alignment of the rat class mu isoenzyme 3–3 (this work) and the porcine class pi isoenzyme described by Reinemer et al. (1991). The individual numbering systems for the two isoenzymes are shown at the top and bottom, respectively. Sequence identities (31%) are highlighted by the vertical bars between the two sequences.

quences of this to the evolution of catalytic diversity in the mu class isoenzymes are discussed briefly below and in some detail in the accompanying paper (Zhang et al., 1992). What is not readily apparent from the crystal structure of isoenzyme 3–3 is possible residues that may be involved in the stabilization of transition states for specific reactions. For example, this particular isoenzyme is particularly good at catalyzing nucleophilic aromatic substitution reactions (Graminski et al., 1989b), yet there are no obvious clues as to why it has the ability to stabilize the transition state for this reaction. It is possible that the helix dipole contributed by the α1 helix helps to stabilize anionic transition states or intermediates; however, why one isoenzyme differs from another in this regard is not clear.

Significance of the Domain Structure of the Enzyme. It probably has not escaped the attention of most readers that the two structural domains of the GSH transferases appear to have distinct functional roles. All but one of the polar interactions between GSH and the enzyme involve residues in domain I so it is not unreasonable to consider domain I as a GSH binding domain. In addition, domain II provides a great deal (though certainly not all) of the hydrophobic surface that interacts with the xenobiotic substrates and the products of the reaction. Comparison of the primary structures of the class mu isoenzymes from rat (Zhang et al., 1992) indicates that most of the sequence variation presumably associated with the different catalytic properties of the isoenzymes is located in domain II. Although it may be stretching a point to call domain II the xenobiotic substrate binding domain, that does appear to be the major contribution of this portion of the protein to the function of the enzyme. Efforts are in progress to test these ideas through the construction of appropriate vectors encoding each domain so that they may be expressed and studied independently.

ACKNOWLEDGMENT

The authors thank Gerard Lacouciere and Dr. Gerard Graminski for their assistance with various aspects of this work and Drs. Osnat Herzburg and Thomas Poulos for many helpful discussions.

SUPPLEMENTARY MATERIAL AVAILABLE

The Ramachandran plot (Figure 3) along with Figure 4 and Table VI describing the details of the secondary structure of each subunit and the subunit interactions in the dimer (6 pages). Ordering information is given on any current masthead page.

REFERENCES

Armstrong, R. N. (1991) *Chem. Res. Toxicol. 4*, 131–140.

Bacon, D., & Anderson, W. A. (1988) *J. Mol. Graphics 6*, 219–220.

Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Jr., Brice, M. D., Rogers, J. R., Kennard, O., Shimanouchi, T., & Tasumi, M. (1977) *J. Mol. Biol. 112*, 535–547.

Blundell, T. L., & Johnson, L. N. (1976) *Protein Crystallography*, pp 374–375, Academic Press, New York.

Brünger, A. T. (1990) *X-PLOR (Version 2.1) Manual*, pp 92–186, Yale University, New Haven, CT.

Brünger, A. T., Kuriyan, J., & Karplus, M. (1987) *Science 235*, 458–460.

Chen, W.-J., Graminski, G. F., & Armstrong, R. N. (1988) *Biochemistry 27*, 647–654.

Cleland, W. W. (1992) *Biochemistry 31*, 317–319.

Cowan, S. W., Bergfors, T., Jones, T. A., Tibbelin, G., Olin, B., Board, P. G., & Mannervik, B. (1989) *J. Mol. Biol. 208*, 369–370.

Dao-Pin, S., Alber, T., Bell, J. A., Weaver, L. H., & Matthews, B. W. (1987) *Protein Eng. 1*, 115–123.
Dirr, H. W., Mann, K., Huber, R., Ladenstein, R., & Reinemer, P. (1991) *Eur. J. Biochem. 196*, 693–698.
Ellman, G. L. (1959) *Arch. Biochem. Biophys. 82*, 70–77.
Furey, W. (1990) Abstracts of the American Crystallographic Association Fortieth Anniversary Meeting, New Orleans, LA, PA33.
Furey, W., Wang, B. C., & Sax, M. (1982) *J. Appl. Crystallogr. 15*, 160–166.
Graminski, G. F., Kubo, Y., & Armstrong, R. N. (1989a) *Biochemistry 28*, 3562–3568.
Graminski, G. F., Zhang, P., Sesay, M. A., Ammon, H. L., & Armstrong, R. N. (1989b) *Biochemistry 28*, 6252–6258.
Hendrickson, W. (1985a) *Methods Enzymol. 115*, 252–270.
Hendrickson, W. (1985b) *Crystallographic Computing 3: Data Collection, Structure Determination, Proteins, and Databases* (Sheldrick, G., Kruger, C., & Goddard, R., Eds.) pp 306–311, Clarendon Press, Oxford.
Hendrickson, W., & Konnert, J. (1980a) *Computing in Crystallography* (Diamond, R., Ramaseshan, S., & Venkatesan, K., Eds.) pp 1301–1323, Indian Academy of Sciences, Bangalore.
Hendrickson, W., & Konnert, J. (1980b) *Biomolecular Structure, Function Conformation and Evolution* (Srinivasan, R., Ed.) Vol. 1, pp 43–57, Pergamon, Oxford.
Hoesch, R. M., & Boyer, T. D. (1989) *J. Biol. Chem. 264*, 17712–17717.
Holmgren, A., Söderberg, B.-O., Eklund, H., & Bränden, C.-I. (1975) *Proc. Natl. Acad. Sci. U.S.A. 72*, 2305–2309.
Howard, A. J., Gilliland, G. L., Finzel, B. C., Poulos, T. L., Ohlendorf, D. H., & Salemme, F. R. (1987) *J. Appl. Crystallog. 20*, 383–387.
Jakoby, W. B. (1978) *Adv. Enzymol. Relat. Areas Mol. Biol. 46*, 383–414.
Jakoby, W. B., Ketterer, B., & Mannervik, B. (1984) *Biochem. Pharmacol. 33*, 2539–2540.
James, M. N. G. & Sielecki, A. R. (1983) *J. Mol. Biol. 163*, 299–361.
Jaskolski, M., Miller, M., Rao, J. K. M., Leis, J., & Woldawer, A. (1990) *Biochemistry 29*, 5889–5898.
Jones, T. A. (1978) *J. Appl. Crystallogr. 11*, 268–272.
Jones, T. A., & Kjeldgaard, M. (1991) O Version 5.6, Department of Molecular Biology, BMC, Uppsala University, Sweden, and Department of Chemistry, Aarhus University, Denmark.
Jones, T. A., Zou, J.-Y., Cowan, S. W., & Kjeldgaard, M. (1991) *Acta Crystallogr. A47*, 110–119.
Kabsch, W., & Sander, C. (1983) *Biopolymers 22*, 2577–2637.
Ladenstein, R., Epp, O., Bartels, K., Jones, A., Huber, R., & Wendel, A. (1979) *J. Mol. Biol. 134*, 199–218.
Liu, S., Zhang, P., Ji, X., Johnson, W. W., Gilliland, G. L., & Armstrong, R. N. (1992) *J. Biol. Chem. 267*, 4296–4299.
Mannervik, B. (1985) *Adv. Enzymol. Relat. Areas Mol. Biol. 57*, 357–417.
Mannervik, B., & Danielson, U. H. (1988) *CRC Crit. Rev. Biochem. 23*, 283–337.
Mannervik, B., Alin, P., Guthenberg, C., Jensson, H., Tahir, M. K., Warholm, M., & Jornvall, H. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 7202–7206.
Meyer, D. J., Coles, B., Pemble, S. E., Gilmore, K. S., Fraser, G. M., & Ketterer, B. (1991) *Biochem. J. 274*, 409–414.
Parker, M. W., Lo Bello, M., & Federici, G. (1990) *J. Mol. Biol. 213*, 221–222.
Penington, C. J., & Rule, G. S. (1992) *Biochemistry 31*, 2912–2920.
Pickett, C. B., & Lu, A. Y. H. (1989) *Annu. Rev. Biochim. 58*, 743–764.
Powell, M. J. D. (1977) *Math. Program. 12*, 241–254.
Ramachandran, G. N., & Sasisekharan, V. (1968) *Adv. Protein Chem. 23*, 283–437.
Reinemer, P., Dirr, H. W., Ladenstein, R., Schaffer, J., Gallay, O., & Huber, R. (1991) *EMBO J. 10*, 1997–2005.
Sayers, J. R., Schmidt, W., & Eckstein, F. (1988) *Nucleic Acids Res. 16*, 791–802.
Schaffer, J., Gallay, O., & Ladenstein, R. (1988) *J. Biol. Chem. 263*, 17405–17411.
Sesay, M. A., Ammon, H. L., & Armstrong, R. N. (1987) *J. Mol. Biol. 197*, 377–378.
Söderberg, B.-O., Sjöberg, B.-M., Sonnerstam, U., & Bränden, C.-I. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 5827–5830.
Stenberg, G., Board, P. G., Carlberg, I., & Mannervik, B. (1991a) *Biochem. J. 274*, 549–555.
Stenberg, G., Board, P. G., & Mannervik, B. (1991b) *FEBS Lett. 293*, 153–155.
Veerapandian, B., Gilliland, G. L., Ragg, R., Svensson, A. L., Masui, Y., Hirai, Y., & Poulos, T. L. (1992) *Proteins: Struct., Funct., Genet. 12*, 10–23.
Wang, B. C. (1985) *Methods Enzymol. 115*, 90–111.
Wang, R. W., Newton, D. J., Pickett, C. B., & Lu, A. Y. H. (1991) *Arch. Biochem. Biophys. 286*, 574–578.
Zhang, P., & Armstrong, R. N. (1990) *Biopolymers 29*, 159–169.
Zhang, P., Graminski, G. F., & Armstrong, R. N. (1991) *J. Biol. Chem. 266*, 19475–19479.
Zhang, P., Liu, S., Shan, S., Ji, X., Gilliland, G. L., & Armstrong, R. N. (1992) *Biochemistry* (following paper in this issue).