Articles

Three-Dimensional Structure, Catalytic Properties, and Evolution of a Sigma Class Glutathione Transferase from Squid, a Progenitor of the Lens *S*-Crystallins of Cephalopods[†,‡]

Xinhua Ji,[§,||,⊥] Erik C. von Rosenvinge,[§] William W. Johnson,[§] Stanislav I. Tomarev,[#] Joram Piatigorsky,[#] Richard N. Armstrong,[*,§,○] and Gary L. Gilliland[*,||,⊥]

Department of Chemistry and Biochemistry, University of Maryland, College Park, Maryland 20742, Department of Biochemistry and Center in Molecular Toxicology, Vanderbilt University School of Medicine, Nashville, Tennessee 37232, Center for Advanced Research in Biotechnology of the Maryland Biotechnology Institute, University of Maryland, Shady Grove, and of the National Institute of Standards and Technology, 9600 Gudelsky Drive, Rockville, Maryland 20850, and Laboratory of Molecular and Developmental Biology, National Eye Institute, National Institutes of Health, Bethesda, Maryland 20892

Received January 13, 1995; Revised Manuscript Received February 24, 1995[⊗]

ABSTRACT: The glutathione transferase from squid digestive gland is unique in its very high catalytic activity toward 1-chloro-2,4-dinitrobenzene and in its ancestral relationship to the genes encoding the *S*-crystallins of the lens of cephalopod eye. The three-dimensional structure of this glutathione transferase in complex with the product 1-(*S*-glutathionyl)-2,4-dinitrobenzene (GSDNB) has been solved by multiple isomorphous replacement techniques at a resolution of 2.4 Å. Like the cytosolic enzymes from vertebrates, the squid protein is a dimer. The structure is similar in overall topology to the vertebrate enzymes but has a dimer interface that is unique when compared to all of the vertebrate and invertebrate structures thus far reported. The active site of the enzyme is very open, a fact that appears to correlate with the high turnover number (800 s^{-1} at pH 6.5) toward 1-chloro-2,4-dinitrobenzene. Both k_{cat} and k_{cat}/K_m^{CDNB} exhibit pH dependencies consistent with a pK_a for the thiol of enzyme-bound GSH of 6.3. The enzyme is not very efficient at catalyzing the addition of GSH to enones and epoxides. This particular characteristic appears to be due to the lack of an electrophilic residue at position 106, which is often found in other GSH transferases. The F106Y mutant enzyme is much improved in catalyzing these reactions. Comparisons of the primary structure, gene structure, and three-dimensional structure with class alpha, mu, and pi enzymes support placing the squid protein in a separate enzyme class, sigma. The unique dimer interface suggests that the class sigma enzyme diverged from the ancestral precursor prior to the divergence of the precursor gene for the alpha, mu, and pi classes.

The glutathione (GSH)[1] *S*-transferases (EC 2.5.1.18) are a group of proteins involved in the biotransformation of endogenous and xenobiotic electrophilic compounds. The enzymes are found in all eukaryotes and many bacteria. The soluble or cytosolic enzymes from vertebrates have been grouped into four distinct classes called alpha, mu, pi (Mannervik et al., 1985), and theta (Meyer et al., 1991). Candidates for other distinct classes of glutathione transferases include three recently characterized enzymes, a bacterial protein from *Proteus mirabilis* (Mignogna et al., 1993) which is 54% identical to an enzyme found in *Escherichia coli* (Nishida et al., 1994), and an enzyme from squid digestive gland (Harris et al., 1991; Tomarev et al., 1993). These proteins appear to be sufficiently different in primary structure from members of the other classes to warrant separate designations. Although the evolutionary relationships of the various isoenzyme classes are not entirely certain, a cogent argument has been made that the gene encoding the theta class protein is the precursor for genes encoding alpha, mu and pi isoenzymes (Pemble & Taylor, 1992). The class theta enzyme appears most closely related to the ancestral bacterial proteins. The three-dimensional structures of isoenzymes from the alpha (1GUH from human; Sinning et al., 1993), mu (1GST from rat, Ji et al., 1992; 1HNA from human, Raghunathan et al., 1994), and pi (1GSR from pig, Reinemer et al., 1991; 1GSS from human, Reinemer et al., 1992; 1GLP from mouse, Garcia-Saez et al., 1994) classes have been determined and support the

[†] Supported by National Institutes of Health Grant GM30910 and an Undergraduate Research Fellowship from the Howard Hughes Medical Institute to E.C.v.R.

[‡] The crystallographic coordinates for the structure have been deposited in the Brookhaven Protein Data Bank and assigned the file name 1GSQ.

* Address correspondence to these authors.
[§] University of Maryland, College Park.
[||] Center for Advanced Research in Biotechnology.
[⊥] National Institute of Standards and Technology.
[#] National Institutes of Health.
[○] Vanderbilt University School of Medicine.

[⊗] Abstract published in *Advance ACS Abstracts,* April 15, 1995.

[1] Abbreviations: GSH, glutathione; CDNB, 1-chloro-2,4-dinitrobenzene; GSDNB, 1-(*S*-glutathionyl)-2,4-dinitrobenzene; GSBzI, *S*-(3-iodobenzyl)glutathione; EMP, ethylmercuric phosphate; EDTA, ethylenediaminetetraacetic acid; IPTG, isopropyl β-D-thiogalactoside; PCR, polymerase chain reaction; PBO, 4-phenyl-3-buten-2-one; PO, phenanthrene 9,10-oxide; OCC, crystallographic occupancy factor; MIR, multiple isomorphous replacement.

0006-2960/95/0434-5317$09.00/0

unique classifications of the vertebrate proteins. In addition, the crystal structure of a GSH transferase from *Schistosoma japonicum* has recently been determined (Lim et al., 1994). The genetic, structural, and mechanistic aspects of the glutathione transferases have been reviewed recently (Armstrong, 1991, 1994; Rushmore & Pickett, 1993; Dirr et al., 1994; Wilce & Parker, 1994).

The enzyme from squid digestive gland occupies an interesting niche in the evolution of the GSH transferases in that the gene encoding the protein has been duplicated and evolved for another purpose, the production of soluble refractory proteins (*S*-crystallins) of the lens of cephalopod eye (Wistow & Piatigorsky, 1987; Tomarev & Zinovieva, 1988; Tomarev et al., 1991, 1992, 1993). In fact, the primary structure of the enzyme is more closely related to the crystallins of squid and octopus (42−44% identity) than it is to the class alpha, mu, and pi enzymes from vertebrates (19−34% identity) (Tomarev et al., 1993). Although the glutathione transferase-like crystallins perform primarily a refractive function in the lens, the SL11 crystallin has significant catalytic activity toward 1-chloro-2,4-dinitrobenzene (CDNB). The other *S*-crystallins appear devoid of catalytic activity toward this substrate. The parent enzyme from squid digestive gland is notable for its extremely high catalytic activity toward the same substrate.

In this paper we describe the three-dimensional structure of the glutathione transferase from squid digestive gland at 2.4 Å resolution and report the characterization of its catalytic properties toward three electrophilic substrates. The catalytic character of the enzyme is correlated with its structure and the participation of specific residues in the catalytic mechanism. Comparisons of the primary structure, gene structure, and three-dimensional structure with class alpha, mu, and pi enzymes support placing the squid protein in a separate enzyme class, sigma. In addition, the three-dimensional structure suggests that the class sigma enzyme diverged from the ancestral precursor prior to the divergence of the alpha/mu/pi precursor.

EXPERIMENTAL PROCEDURES

Construction of Expression Vectors. The cDNA insert of clone pGST5 (Tomarev et al., 1993) was amplified by the polymerase chain reaction (PCR). The primers created a *Nde*I restriction site immediately 5′ upstream to the translation initiation codon (5′-ccacacatATGCCTAAGTATACCCTACACTAT-3′) and a *Not*I site in the 3′ untranslated region (5′-ggagaagcggccgCTTGTTCTTGATTTCGGCTAGGA-3′) of the mRNA encoding the enzyme. After digestion of the amplified cDNA with *Nde*I and *Not*I a PCR product with the expected length of about 650 bp was isolated and purified from an agarose gel using the Geneclean[2] methodology (BIO 101, La Jolla, CA), ligated into the expression vector pET-17b (Novagen). The resulting vector was used to transform *E. coli* strain DH5α. Plasmids were isolated from several colonies and the cDNA inserts sequenced. A wild-type construct designated GST5/pET was used to transform the *E. coli* strain BL21(DE3). The construction of an expression vector encoding the F106Y mutant enzyme was accomplished using the PCR technique described by Jones and Howard (1990).

Expression and Purification of the Recombinant Enzymes. Bacteria harboring the plasmid GST5/pET or the F106Y mutant plasmid were grown to an optical density of 0.6−0.8 in LB medium supplemented with ampicillin. Induction was accomplished by adding IPTG to a concentration of 0.4 mM, and incubation was continued for another 3 h at 30 °C. Induction at 37 °C resulted in the accumulation of the expressed enzyme in inclusion bodies which was avoided by using the lower temperature. The glutathione transferase comprised about 20% of the total water-soluble proteins of the transformed bacteria. Cells from 6 L of culture were collected, washed once in 50 mM Tris-HCl (pH 8.0) containing 2 mM EDTA, resuspended in 120 mL 10 mM Tris buffer (pH 7.8) containing 1 mM EDTA, and disrupted with a Branson Sonifier II sonicator. The cell debris was removed by centrifugation (30000*g* for 30 min), and the supernatant was applied directly to a 1.7 cm × 6 cm affinity column of *S*-hexylglutathione coupled to Sepharose 6B previously equilibrated with 10 mM Tris (pH 7.8). The column was then washed with 900 mL of 10 mM Tris (pH 7.8) containing 1 mM EDTA and 0.2 M KCl. The enzyme was eluted with the same buffer containing 2.5 mM *S*-hexylglutathione. Fractions containing the enzyme were pooled and concentrated in an Amicon concentrator fitted with a PM-10 membrane. The protein was dialyzed overnight against 2 L of 10 mM potassium phosphate and 1 mM EDTA (pH 7.0).

Kinetic Characterization of the Enzyme. The kinetics of the enzyme-catalyzed addition of GSH to 1-chloro-2,4-dinitrobenzene (CDNB), 4-phenyl-3-buten-2-one (PBO), and phenanthrene 9,10-oxide (PO) were determined as previously described (Johnson et al., 1993). The pH dependence of the reaction with CDNB was established as described by Liu et al. (1992). Initial velocities and the pH dependencies of kinetic parameters were analyzed by the programs HYPER and HABELL, respectively (Cleland, 1979).

Crystallization of the Enzyme. Crystals of the recombinant protein were grown in sitting drops (15−30 μL) which initially consisted of between 6 and 8 mg/mL protein in 25 mM Tris (pH 7.0) containing 1 mM EDTA, 2 mM 1-(*S*-glutathionyl)-2,4-dinitrobenzene (GSDNB), and buffered (pH 7) ammonium sulfate at 40% saturation. The drops were equilibrated at 4 °C against wells containing between 60% and 70% saturated ammonium sulfate (pH 7). Single crystals grew in 5−7 days. Initial characterization by X-ray diffraction indicated that the protein crystallized in the space group $P3_121$ with unit cell dimensions of $a = b = 72.76$ Å, $c = 94.63$ Å, $\alpha = \beta = 90°$, and $\gamma = 120.0°$.

Heavy-Atom Derivatives. The first heavy-atom derivative (**1**) was obtained by replacement of the product inhibitor, GSDNB, which occupied the active site of the crystalline enzyme, with *S*-(3-iodobenzyl)glutathione (GSBzI). This was accomplished by diffusion of GSBzI into the crystals at 25 °C. A single crystal with the dimensions of 0.25 × 0.35 × 0.48 mm was transferred from the well in which it was grown to a well containing 100 μL of a 1.5 mM solution of 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

[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.

for 6 h. The yellow color of the crystal lightened, and the edges of the crystal were almost colorless. After sitting for additional 14 h, the soaking solution was again replaced every hour for an additional 8 h.

A second derivative (**2**) containing mercuric ions was prepared at room temperature by the following procedure. A single crystal with the dimensions of 0.30 × 0.30 × 0.40 mm was transferred to a well containing 50 μL of well solution [70% saturated ammonium sulfate (pH 7)] plus 50 μL of a 0.5 mM solution of ethylmercuric phosphate (EMP) in 70% saturated ammonium sulfate buffered with 25 mM Tris (pH 8). After 15 min, 50 μL of the solution was removed, and an additional 50 μL of the 0.5 mM EMP solution was added. This procedure was repeated three times. Cracks began to develop in the crystal after about five min and the crystal started losing color after 30 min. After a total soaking time of 1 h the crystal was mounted for data collection.

X-ray Data Collection and Data Processing. Crystals were mounted in thin-walled capillary tubes with diameters of 0.7−1.0 mm, depending upon the size of the crystal. X-ray diffraction data were collected using a Siemens electronic area detector. The area detector was mounted on a three-axis Supper oscillation camera controlled by a PCS microcomputer. Each electronic frame was composed of counts summed over a 0.2° range in ω with exposure times of 180 s. 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 Indigo2 computer for final processing. The data collection was carried out at controlled room temperature (19−20 °C).

The unit cell parameters and the space group were determined by using the automatic indexing facility in XDS (Kabsch, 1988). The refinement of the unit cell parameters and the integration of reflection intensities were performed with the XENGEN program system (Howard et al., 1987). A summary of the crystal and diffraction data and final results of data processing for the native and derivative data sets is presented in Table 1. The detailed statistics for the native data are given in Table 1 (supplementary material).

Structure Solution and Refinement. Attempts to solve the structure of the squid enzyme by molecular replacement using the alpha, mu, and pi structures as probe models were not successful. Therefore, the structure of the enzyme in complex with the product GSDNB was solved by using the multiple isomorphous replacement (MIR) technique, followed by iterative cycles of refinement and model building. Solvent flattening was also used in the early stages to improve the electron density maps. Two heavy-atom derivatives were used to produce the MIR map. Diffusion of the active site-directed heavy-atom-containing compound GSBzI into the native crystals gave the first isomorphous heavy-atom derivative **1** in much the same way that was successful in the structure solution of the mu class isoenzyme M1-1[3] from rat (Ji et al., 1992). The single isomorphous replacement procedure was carried out, and the packing of the molecules in the unit cell was revealed. However, the phasing power of the single derivative was not sufficient to discern secondary structural features of the protein. Derivative **2**, obtained by the reaction of the native crystals with ethylmercuric phosphate, was isomorphous with the native crystal. The heavy-atom positions were derived from the interpretation of the difference Patterson maps. A summary of heavy-atom positions and associated protein residues is presented in Table 2 (supplementary material). The two mercuric ion sites found were associated with the sulfur atoms of C123 and C174. The final mean figures of merit for the MIR phases and for phases after solvent flattening were 0.604 and 0.816, respectively, for 7060 phased reflections. 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).

Table 1: Results of X-ray Data Processing for Native and Two Derivative Data Sets and Statistics of MIR Phasing Calculations with Derivatives **1** and **2**

		native[a]	derivative **1**	derivative **2**
unit cell	a, b (Å)	72.76	73.29	73.01
	c (Å)	94.63	95.12	95.53
	α, β (°)	90.00	90.00	90.00
	γ (°)	120.0	120.0	120.0
data	D_{min} (Å)	2.40	2.56	3.16
	unique data obsd	11745	16648	8887
	data with $I \geq 2\sigma(I)$	8821	15739	7808
	R_w[b]	0.087	0.090	0.087
	R_{uw}[c]	0.080	0.082	0.081
	$R_{scaling}$[d]		0.187	0.193
MIR	data used (Å)		∞−2.6	∞−3.2
	no. of centric data		862	570
	no. ofheavy atom sites		3	2
	R_{Cullis}[e]		0.61	0.50
	phasing power[f]		1.46	2.17
	mean figure of merit		0.57	0.65

[a] Native, squid enzyme + GSDNB; derivative **1**, squid enzyme + GSBzI; derivative **2**, squid enzyme + GSDNB + EMP. [b] The weighted least-squares R factor on intensity for symmetry-related observations: $R_w = \sum[(I_{ij} - G_{ij}\langle I_j\rangle)/s_{ij}]^2/\sum(I_{ij}/s_{ij})^2$, where $G_{ij} = g_i + A_i s_j + B_i s_j^2$; $s = \sin\theta/\lambda$; g, A, and B are scaling parameters. [c] The unweighted absolute R factor on intensities: $R_{uw} = \sum(I_{ij} - G_{ij}\langle I\rangle_j/\sum|I_{ij}|)$. [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.

The solvent-flattened MIR map was readily interpretable. An initial model for most of the structure was built without particular difficulty. The entire model was completed after five cycles of simulated annealing with X-PLOR (Brünger et al., 1987; Brünger, 1992) and model building. The crystallographic R factor, which was 0.288 without B (temperature) factor refinement, decreased to 0.243 at a resolution of 2.40 Å after grouped B factor refinement (Powell, 1977; Brünger, 1992). Model building, the examination of MIR, $2F_o - F_c$, $F_o - F_c$ and omit maps, model adjustments, and the incorporation of water molecules during further refinement were done using O (Jones et al., 1991; Jones & Kjeldgaard, 1991) implemented on an Indigo2 computer with the GR3-XZ graphics system.

The final refinement procedure was performed with GPRLSA (Furey et al., 1982), the restrained least-squares refinement procedure of Hendrickson and Konnert (1980a,b) and Hendrickson (1985a,b). After seven cycles of least-squares refinement and model adjustment the model of the

[3] The nomenclature for the rat isoenzymes follows that recommended by Mannervik et al. (1992) for the human isoenzymes. Although the rat isoenzymes were not specifically named in this paper, it is clear that the class mu isoenzyme 3-3 from rat corresponds to M1-1 (Klinga-Levan et al., 1993).

Table 2: Least-Squares Refinement Parameters for the Enzyme−GSDNB Complex

	target values	final model
diffraction data from 6.0 to 2.4 Å with $I \geq 2\sigma(I)$		8302
crystallographic R^a factor		0.180
no. of residues		202
no. of water molecules		154
no. of non-H protein atoms/water molecules		10.6
rms deviations from ideal distances (Å)		
bond distances	0.025	0.018
angle distances	0.036	0.036
planar 1−4 distances	0.040	0.033
rms deviations from planarity (Å)	0.030	0.021
rms deviations from ideal chirality (Å^3)	0.200	0.221
thermal parameter correlation (mean/ΔB)		
main-chain bond	1.000	0.550
main-chain angle	1.500	0.942
side-chain bond	1.500	0.888
side-chain angle	2.000	1.371

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

glutathionyl moiety was built according to the $F_o - F_c$ map contoured at 3.0σ. The position of the dinitrophenyl moiety of the product was, however, not clear. During the next 28 cycles of refinement 154 water molecules were added, and the entire model was examined and adjusted using omit maps. At this stage, the R factor was 0.189 for a model that included only the glutathionyl portion of the product. The electron density showed obvious signs of disorder for the dinitrophenyl moiety of the product. After an additional eight cycles of least-squares refinement, the electron density for a dinitrophenyl group finally became very clear, assuming an occupancy of 70% for one possible conformer. The inclusion of this conformer in the refinement lowered the R to 0.180. At least one other conformer could be seen in the $F_o - F_c$ map. Further refinement including the second possible conformer of the dinitrophenyl group was attempted. However, since no improvement in the statistics or the electron density was evident, it was not included in the final model. The final coordinates of the squid glutathione transferase in complex with GSDNB had a crystallographic R value of 0.180 for 8302 diffraction data between 6.0 and 2.4 Å with $I \geq 2\sigma(I)$. A summary of the refinement statistics is given in Table 2. 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) under the file name 1GSQ.

RESULTS

Expression and Purification of the Enzyme. The squid GSH transferase could be expressed and purified in large quantities from *E. coli* by the standard affinity chromatographic procedure. The purified enzyme gave a single band upon SDS gel electrophoresis with an estimated molecular mass of 23 kDa. Gel filtration of the native enzyme on a calibrated Superdex 200 HR 10/30 column gave a molecular mass of 40 kDa, a result consistent with a dimeric holoenzyme. The catalytic properties of the enzyme as they relate to the three-dimensional structure are discussed in more detail below.

Three-Dimensional Structure. The crystals of the squid GSH transferase contain a single subunit in the asymmetric unit. A ribbon diagram of the single polypeptide chain in complex with the product, GSDNB, is illustrated in Figure 1a. The final model of the enzyme−GSDNB complex consists of 1784 atoms, including all non-hydrogen atoms of 202 amino acid residues (no. 1−202), one GSDNB molecule (no. 203), and the oxygen atoms of 154 water molecules (no. 301−454). The electron density map for the inhibitor GSDNB is shown in Figure 2. The water molecules are numbered sequentially from the largest to the smallest values of OCC^2/B (James & Sielecki, 1983). The electron density is in complete agreement with the amino acid sequence of the 202 residues deduced from the gene sequence, with the exception of the methionine encoded by the initiation codon which is removed by the bacteria during translation.

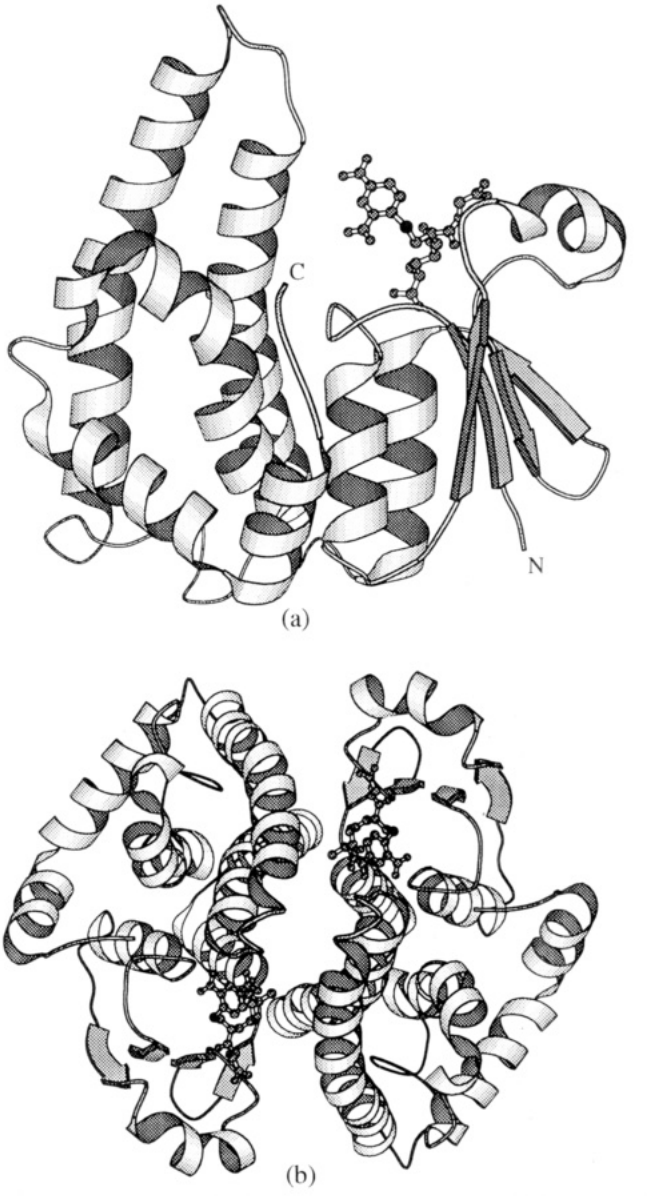

FIGURE 1: Ribbon diagram representations of squid GSH transferase in complex with 1-(*S*-glutathionyl)-2,4-dinitrobenzene: (a) one polypeptide chain (the asymmetric unit) viewed perpendicular to the crystallographic 2-fold axis and (b) a dimer as viewed down the crystallographic 2-fold axis. The product is illustrated as a ball-and stick model. The diagrams were generated using the program MOLSCRIPT (Kraulis, 1991).

The secondary structure of the squid enzyme is defined according to Kabsch and Sander (1983). The subunit contains four β-strands and eight α-helices arranged in a smaller α/β domain (domain I) and a larger α domain

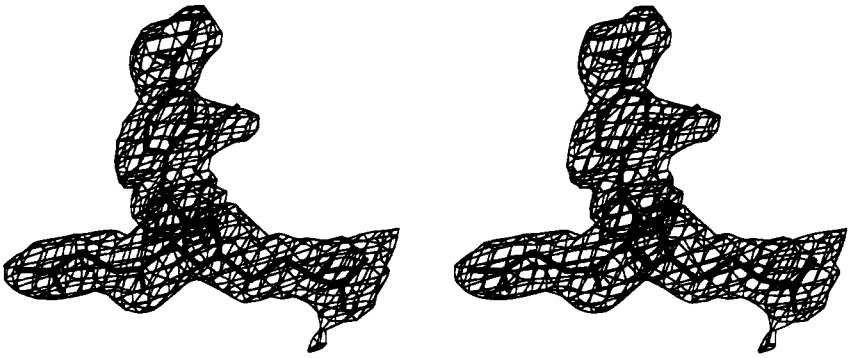

FIGURE 2: Stereoview of the $2F_o - F_c$ electron density map for GSDNB. The electron density map is contoured at 1σ except that the portion for the dinitrophenyl moiety is contoured at 0.7σ. The omit map for this region is quite similar to the $2F_o - F_c$ map.

(domain II). Domain I contains residues 1−74 and domain II 81−202. The two domains are connected via a short, six-residue, linker (residues 75−80). The major secondary structural elements of domain I are arranged in a $\beta\alpha\beta\alpha\beta\beta\alpha$ motif. The four β-strands form a central β-sheet and are arranged such that $\beta1$ and $\beta2$ are parallel and that $\beta1$ and $\beta3$ and $\beta3$ and $\beta4$ are antiparallel (Figure 1). Domain II is composed of five α-helices (α4, α5, α6, α7, and α8). One *cis*-proline is located in domain I (P51). Eleven hydrogen bonds and salt bridges (≤3.5 Å) are observed between the two domains (residues 9, 11, 12, 13, and 15 from domain I and residues 193, 197, 200, and 202 from domain II).

Dimer Interface. There is only one peptide chain in the asymmetric unit of the unit cell. The biologically active form of the enzyme is, however, dimeric. The second subunit of the dimeric product complex is generated by applying the symmetry operation (y, x, $-z$) to the fractional coordinates of the primary molecule. Figure 1b represents the constructed dimer looking down the crystallographic 2-fold axis.

The two subunits contact one another primarily by interactions between domain I of one subunit and domain II of the adjacent subunit (Figure 1b) with a hydrophilic channel that runs through the dimer and is coincident with the crystallographic 2-fold axis. Contacts between the two subunits involve several of the secondary structural elements including the $\beta4$, α3, α4, and α5. Among about 190 short contacts (≤4.0 Å) are about 30 salt bridges and hydrogen bonds (≤3.5 Å). These electrostatic interactions involve $\beta4$, α3, α4, and α5 of both subunits. Of particular note is the ratio of electrostatic interactions to the total number of short contacts between the two subunits. The ratio is higher (30/190) than that observed for the class mu M1-1 isoenzyme (17/170) (Ji et al., 1992). The hydrophobic "lock" between the subunits of M1-1, which involves the insertion of the side chain of F56 from domain I of one subunit into a pocket formed by I98, Q102, L136, Y137, and F140 from domain II of the other (Ji et al., 1992), is not observed in the structure of the squid enzyme. The increase of electrostatic interactions and the decrease of hydrophobic contacts in the interface are discussed in detail and compared to those found in vertebrate and *S. japonicum* glutathione transferases in the following sections.

Solvent Structure. A total of 154 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 distance criterion for hydrogen bonds to water is set at ≤3.5 Å. No angle restriction is assumed. Of the 154 water molecules, 121 form hydrogen bonds to protein atoms and 49 of these have more than one hydrogen bond to the protein. These 121 water molecules have temperature factors that are very similar to their hydrogen-bonding partners on the protein. There are 33 water molecules that do not form hydrogen bonds to protein atoms. However, they do exhibit hydrogen bonds to other water molecules.

Table 3: Electrostatic Interactions between 1-(*S*-Glutathionyl)-2,4-dinitrobenzene (GSDNB) and Glutathione *S*-Transferases from Mu (5GST; Ji et al., 1993) and Sigma Classes (1GSQ; This Work)[a]

GSDNB	5GST (mu)	1GSQ (sigma)
γ-Glu-N1	Q71-OE1	Q62-OE1
γ-Glu-N1	D105 (subunit b)-OD1	D96 (subunit b)-OD1 (3.90)
γ-Glu-N1	D105 (subunit b)-OD2	D96 (subunit b)-OD2 (3.68)
γ-Glu-O11	S72-N	S63-N
γ-Glu-O11	S72-OG (3.90)	S63-OG
γ-Glu-O12	Q71-OE1	Q62-OE1 (3.67)
γ-Glu-O12	S72-OG	S63-OG
Cys-N2	N58-OD1	
Cys-N2	L59-O	M50-O
Cys-O2	W7-NE1	
Cys-O2		M50-N
Gly-N3	N58-OD1	
Gly-O31		K42-NZ
Gly-O31		N48-O
Gly-O32	R42-NH2	W38-NE1
Gly-O32	W45-NE1	
Gly-O32		K42-NZ
Gly-O32	K49-NZ	

[a] Distance criterion for hydrogen bonds and salt bridges is 3.5 Å. Distances exceeding this criterion are given in parentheses.

Structure of the Active Site. The glutathionyl portion of the product is bound in an extended conformation with the involvement of a network of electrostatic interactions (Table 3) that is similar to those observed in the vertebrate enzymes. In contrast, other than a few van der Waals interactions with F106, the dinitrophenyl moiety of the product has little direct interaction with the protein. There are no hydrogen-bonding interactions between the nitro groups and the protein, although the hydroxyl group of Y7 is fairly close (3.58 Å) to one of the oxygen atoms of the *o*-nitro group. The *p*-nitro group appears not to be involved in hydrogen-bonding interactions. The lack of interactions with the dinitrophenyl group is almost certainly related to the less than full occupancy of this conformer.

Kinetic Characteristics of the Enzyme. The recombinant squid enzyme is very active toward CDNB (Table 4), a fact that was previously noted with enzyme isolated from the digestive gland (Harris et al., 1991; Tomarev et al., 1993).

Table 4: Kinetic Constants for Squid GSH Transferase with 1-Chloro-2,4-dinitrobenzene, 4-Phenyl-3-buten-2-one, and Phenanthrene 9,10-Oxide

substrate	enzyme	k_{cat} (s^{-1})	k_{cat}/K_m^S (M^{-1} s^{-1})	stereoselectivity (% isomer A)[a]
CDNB	native	$(8.00 \pm 0.64) \times 10^2$	$(1.05 \pm 0.07) \times 10^6$	
CDNB	F106Y	$(4.57 \pm 0.37) \times 10^2$	$(0.81 \pm 0.07) \times 10^6$	
PBO	native	$(8.00 \pm 0.88) \times 10^{-3}$	$(6.35 \pm 0.38) \times 10^1$	50 ± 2
PBO	F106Y	$(5.25 \pm 0.58) \times 10^{-1}$	$(1.47 \pm 0.04) \times 10^3$	71 ± 4
PO	native	$(2.60 \pm 0.20) \times 10^{-3}$		24 ± 1
PO	F106Y	$(1.60 \pm 0.16) \times 10^{-2}$		30 ± 1

[a] Isomer A is defined as the first product to elute on reversed phase HPLC (Kubo & Armstrong, 1989; Cobb et al., 1983). The absolute configuration of isomer A in the reaction with PBO has not been established. Isomer A in the reaction with PO is the (9*S*,10*S*)-diastereomer (Cobb et al., 1983; Ji et al., 1994).

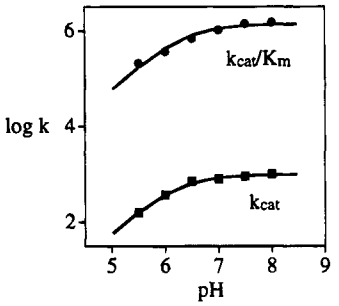

FIGURE 3: Dependence of k_{cat} (■) and k_{cat}/K_m^{CDNB} (●) on pH for squid glutathione transferase. The solid lines are computer fits of the experimental data to the equation $\log k = \log(C/(1 + [H^+]/K_a)$ using the program HABELL. The limiting values of k_{cat} and k_{cat}/K_m at high pH and the apparent pK_a's are 990 ± 30 s^{-1} ($pK_a = 6.2 \pm 0.1$) and $(1.38 \pm 0.14) \times 10^6$ M^{-1} s^{-1} ($pK_a = 6.4 \pm 0.3$), respectively.

In fact, it appears to be the most active GSH transferase yet described with CDNB as the electrophilic substrate. The k_{cat} for this reaction approaches 10^3 s^{-1}, and k_{cat}/K_m^{CDNB} is only about a factor of 10−100 less than the diffusion limit at pH 6.5. The influence of pH on both k_{cat} and k_{cat}/K_m^{CDNB} under conditions of saturating GSH is similar as shown in Figure 3 and suggests that the reaction is dependent on a single ionization with an apparent pK_a of between 6.2 and 6.4. The obvious interpretation of this result is that these two measures of kinetic pK_a's from k_{cat}/K_m^{CDNB} and k_{cat} reflect the ionization of GSH in the binary E·GSH and ternary E·GSH·CDNB complexes, respectively.

In marked contrast to the high efficiency toward CDNB, the squid enzyme is a very poor catalyst in the addition of GSH to enone and epoxide substrates such as 4-phenyl-3-buten-2-one or phenanthrene 9,10-oxide (Table 4). Moreover, the enzyme exhibits little stereoselectivity toward either of these substrates. It has been previously noted that Y115 of the M1-1 isoenzyme from rat appears to act as an electrophilic participant in both Michael additions and epoxide ring opening reactions (Johnson et al., 1993; Ji et al., 1994) and contributes significantly to the efficiency of the enzyme toward these substrates. Alignments of the structures of the class mu and the squid enzymes (see Figures 4 and 7) indicate that the cognate residue in the squid structure is F106, which of course lacks the relevant hydroxyl group. Mutation of this residue to tyrosine (F106Y) has little effect on the reaction with CDNB but increases the efficiency of the enzyme toward the enone and epoxide substrates as indicated in Table 4.

DISCUSSION

Folding Topology. The folding topology of the squid enzyme is similar to that of the other known GSH transferase structures. Although the rms deviation between Cα positions ranges from 1.56 to 2.19 Å for 362−376 best-aligned target pairs (Figure 5), significant differences are observed between the squid enzyme and other GSH transferases. The position of the helix−turn−helix (α4−turn−α5, residues 106−126) motif differs from that observed in all other GSH transferase structures in that the tip of this structural motif deviates about 5 Å from the position of the closest structural counterpart, the alpha class isoenzyme. The tips of the two α4−turn−α5 motifs in the dimer of the squid enzyme are at least 10 Å closer to each other than those in other enzyme classes (Figure 5a). Undoubtedly, this fact contributed to the failure of molecular replacement trials in the early stage of the crystal structure determination of the squid protein. As has been demonstrated (Sinning et al., 1993), molecular replacement solution of GSH transferase structures across gene classes can be difficult due to the different relative positions of subunits and domains for the different classes. The substantial deviation in the helix−turn−helix motif made it even more difficult in the case of the squid enzyme. After the structure was solved by using the multiple isomorphous replacement techniques, the use of a search model without this motif lead to a successful molecular replacement solution.

Dimer Interface. The biologically active forms of cytosolic GSH transferases are known to be either homo- or heterodimeric proteins. This is also the case for the enzyme from squid (this work). Both hydrophilic and hydrophobic interactions are involved in the dimerization. Common features of the subunit−subunit interface have been observed in all of the structures of cytosolic GSH transferases. In the middle of the interface, an intermonomer contact is created by the stacking of two symmetrically equivalent arginine guanidino groups (R69, alpha; R77, mu; R68, pi; R72, *S. japonicum*; R68, squid). Although R69 of the class alpha isoenzyme is not aligned with the arginine residues of other isoenzymes (Figure 4), it occupies a similar spatial position. The charges on the side chain of these arginyl residues are mitigated by their close proximity to aspartate or glutamate side chains. More than 10 (for example, 16 in the M1-1 isoenzyme; Ji et al., 1992) salt bridges and hydrogen bonds in addition to the arginine−arginine side-chain interactions are found in the middle of the dimer interface for all known structures. At each end of the interface, however, a different type of interaction has been observed. In alpha, mu, pi, and *S. japonicum*, a "lock-and-key" kind of hydrophobic interaction is established by wedging a hydrophobic side chain (F52, alpha; F56, mu; F47, pi; F51, *S. japonicum*) from one monomer into a hydrophobic pocket on the other side of the interface formed by the side

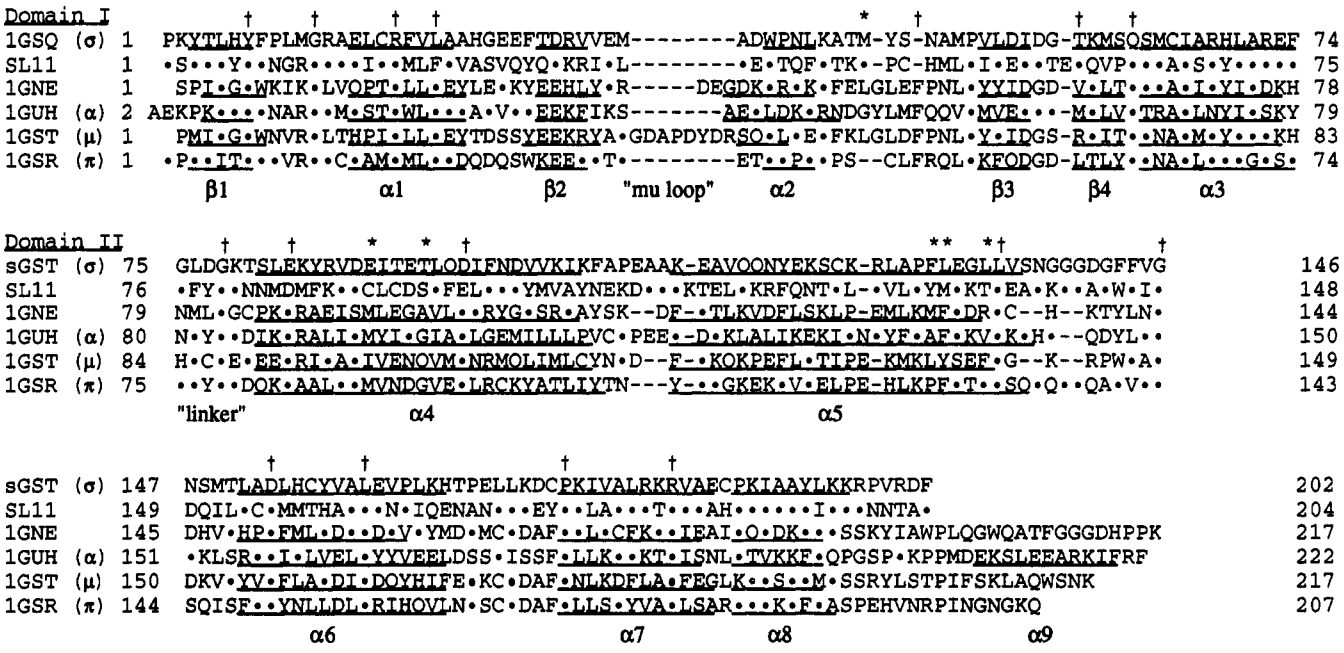

FIGURE 4: Sequence alignment based on the crystal structures of squid GSH transferase 1GSQ (from squid digestive gland, this work), 1GNE (from *S. japonicum*; Lim et al., 1994), 1GUH (class alpha from human liver; Sinning et al., 1993), 1GST (class mu from rat liver; Ji et al., 1992), and 1GSR (class pi from pig lung; Reinemer et al., 1991). Domains I and II refer to the smaller α/β domain and the larger α-helical domain, respectively. The elements of secondary structure are indicated by an α (α-helix) or a β (β-sheet). Nonloop residues are underlined. Residues that are identical to the squid enzyme are indicated by dots (•). Sixteen residues that are conserved in all five crystal structures are noted by crosses (†). Asterisks (*) highlight the residues involved in the "lock-and-key" type hydrophobic interaction observed at each end of the subunit−subunit interface in all, but the squid, GSH transferases. The helix α9 is unique for the alpha class enzyme, and the "mu loop" is unique to the mu class. The linker is the region that ties the two domains together. The sequence alignment of the SL11 crystallin is based solely on sequence similarity with the squid enzyme since the structure of SL11 is not known.

chains of five residues belonging to helices α4 and α5 (M94, G98, A135, F136, and V139, alpha; I98, Q102, L136, Y137, and F140, mu; M89, G93, P126, F127, and L130, pi; M93, A97, M131, F132, and R135, *S. japonicum*). Both "key" and "lock" residues are perfectly aligned in the sequences (Figure 4). A fundamental difference between the squid enzyme and all other cytosolic structures known is that this lock-and-key interaction is not present in the squid dimer. This is due to the fact that the key (phenylalanyl) residue and the loop on which it resides are absent (Figures 4−6), that two out of five lock residues form electrostatic rather than hydrophobic interactions across the dimer interface, and that the "hole" of the lock is blocked by the side chains of E89 and F129 (Figure 6a). The decrease in hydrophobic interactions appears to be compensated for by the increase of electrostatic interactions. There are 13 more salt bridges and hydrogen bonds formed between the two squid enzyme monomers when compared to isoenzyme M1-1 from the mu class (Figure 6a). Moreover, there are 18 water molecules that participate in electrostatic interactions between both monomers. Only 11 such water molecules are found between the two subunits of mu class isoenzyme M1-1 (Ji et al., 1992). It is concluded, therefore, that the dimer interface of the squid enzyme is more hydrophilic than that of enzymes of other classes.

The molecular recognition for dimerization of GSH transferase monomers is strictly dependent on the enzyme class, as evidenced by the fact that no heterodimer has been found with subunits from different classes. The two types of subunit−subunit interactions described above are quite different in nature. The squid enzyme illustrates one type of subunit−subunit interaction with a relatively hydrophilic interface, while all known structures of other classes of cytosolic enzymes share another. It is therefore proposed that the molecular recognition among GSH transferase subunits may be distinguished by the presence or absence of the hydrophobic lock between the loop of domain I and the α4- and α5-helices of domain II. Furthermore, the familial subunit recognition among enzyme classes that have the hydrophobic lock (e.g., alpha, mu, pi, and *S. japonicum*) is governed primarily by the relative orientation of the N- and C-terminal domains (Sinning et al., 1993) which alters the distance between the key of domain I and the lock located in domain II.

Active Site. The GSH transferase structures solved to date exhibit two distinct binding modes for the peptide. Although the γ-glutamyl side chain of GSH interacts with the enzyme in a very similar way in all known GSH transferase structures (Table 3; Sinning et al., 1993; Ji et al., 1992; Reinemer et al., 1991, 1992; Lim et al., 1994), the interactions between the cysteinyl and glycyl residues of the tripeptide show significant differences depending on the class of enzyme. For example, in the class mu enzyme, the carbonyl oxygen atom (O2) of the cysteinyl residue of GSH accepts a hydrogen bond from the indole ring NH (NE1) of W7. In contrast, the absence of a hydrogen bond donor at the equivalent position in the squid enzyme (occupied by F8) appears to favor an alternate conformation of the cysteinyl residue so that the carbonyl oxygen (O2) accepts a hydrogen bond with the backbone NH of M50 (Figure 7). The interactions of the glycyl residue of GSH with the enzyme are consequently different (Table 3 and Figure 7). Class alpha (Sinning et al., 1993) and pi (Reinemer et al., 1991, 1992) enzymes share the same GSH binding mode with the squid enzyme (this work), while the enzyme from *S. japonicum* (Lim et al., 1994) and the class mu enzyme share

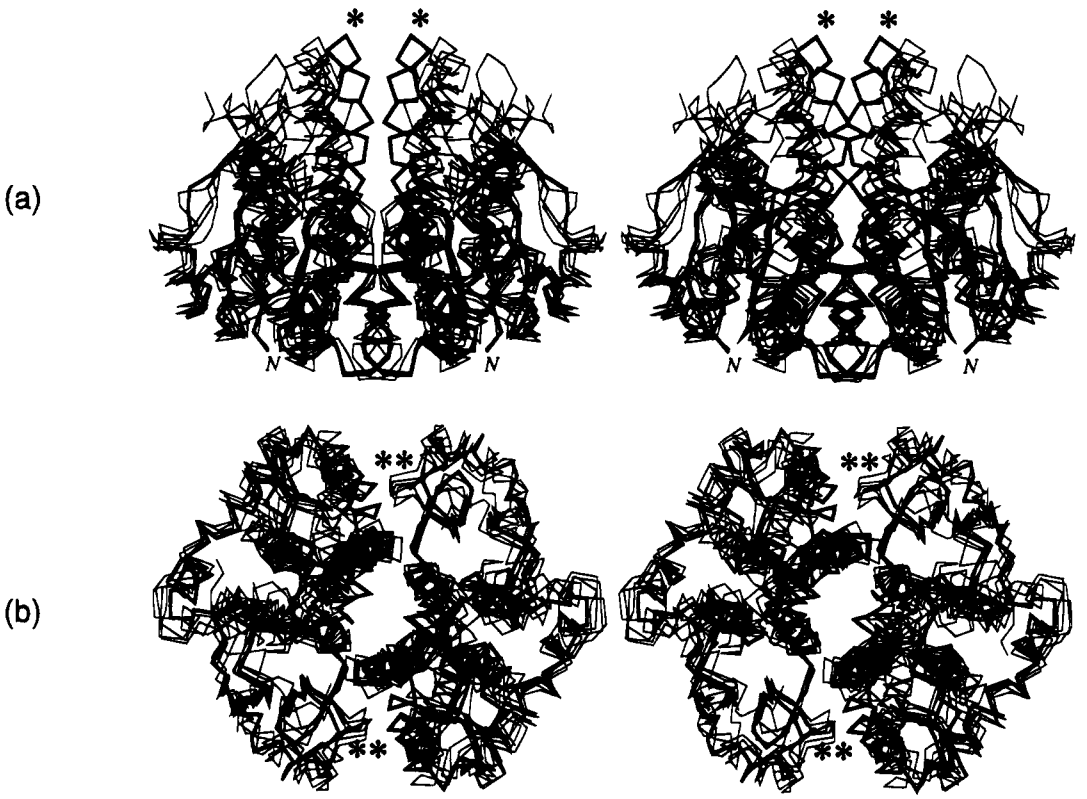

FIGURE 5: Stereoview of the Cα trace of the squid enzyme (this work; the second monomer is generated by symmetry operator $y, x, -z$) is aligned with class alpha (1GUH; Sinning et al., 1993), mu (1GST; Ji et al., 1992), pi (1GSR; Reinemer et al., 1991), and the enzyme from *S. japonicum* (1GNE; Lim et al., 1994). The squid enzyme is illustrated in thick solid lines, while others are represented by thin lines. The rms deviations for the α-carbons of the squid enzyme from the other four structures are alpha (1GUH, 2.19 Å, 372 target pairs), mu (1GST, 2.04 Å, 362 target pairs), pi (1GSR, 1.56 Å, 367 target pairs), and *S. japonicum* (1GNE, 1.80 Å, 376 target pairs). An asterisk (*) is used to highlight the displacement of the helix−turn−helix motif of the squid enzyme when viewed perpendicular to the 2-fold axis which relates the two subunits in (a). The double asterisk (**) in (b) highlights the position of the deletion of the phenylalanyl residue and loop in squid GSH transferase structure, which in other enzyme classes is involved in a key hydrophobic interaction at the dimer interface.

the other (Ji et al., 1992).

The addition of GSH to 1-chloro-2,4-dinitrobenzene is the most widely used spectrophotometric assay of the GSH transferases (Habig et al., 1974). However, it has proven difficult to obtain a well-defined structure of the enzyme in complex with the product of this reaction, GSDNB. For instance, electron density for the dinitrophenyl group was not observed in three different crystal forms of the human class mu isoenzyme (Raghunathan et al., 1994). The first crystal structure determination which revealed the complete model of GSDNB is the product complex with mu class M1−1 isoenzyme from rat (5GST; Ji et al., 1993), in which GSDNB assumes different conformations about the SG2−C1′ bond (180° rotation) in the two subunits. The crystal structure of the squid enzyme in complex with GSDNB (this work) locates a 70% populated conformer of the dinitrophenyl moiety which differs slightly from that found in the M1-1 complex. This conformation involves changes in the torsion angles about three bonds, CA2−CB2, CB2−SG2, and SG2−C1′ (Figure 7). In contrast to the peptidyl portion of the product, the dinitrophenyl group is not tethered to the protein with specific electrostatic interactions. As in the rat M1-1·GSDNB complex (Ji et al., 1993), the dinitrophenyl moiety points into the solvent channel between the two subunits and involves very few van der Waals contacts with the protein.

Structure and Catalysis. Some clear correlations can be made between the structure of the squid protein and its catalytic characteristics. The apparent pK_a of the enzyme-bound GSH (ca. 6.3) is about the same as has been measured for other isoenzymes, suggesting, not surprisingly, that one fundamental role of the protein is to lower the pK_a of the thiol of GSH. Although the hydroxyl group of Y7 is a bit further (4.9 Å) from the sulfur of GSDNB than it is in other structures (ca. 3.2−3.6 Å), it is likely that it plays a similar role in catalysis. The greater distance appears to be a consequence of the geometry imposed by the presence of the dinitrophenyl group and does not reflect any intrinsic difference in the spatial relationship of the catalytic hydroxyl group and the sulfur of the substrate GSH. In this regard, it should be noted that the refined structure of the squid enzyme in complex with GSBzI shows a normal O−H···S hydrogen-bonding distance of 3.2 Å. Furthermore, the Y7F mutant of the squid enzyme exhibits a 50−100-fold decrease in catalytic activity toward CDNB.[4]

The active site of the squid enzyme (Figure 1a) is relatively open, particularly when compared to the partially occluded active sites of the class alpha and mu enzymes. The openness is primarily a result of the shorter C-terminal tail and the absence of a "mu loop" in the squid protein. The high turnover number and the similar pH dependence of k_{cat} and k_{cat}/K_m^{CDNB} for the squid enzyme clearly suggest that product release is not the rate-limiting step in the addition of GSH to CDNB as it is with the M1-1 enzyme (Johnson et al., 1993). An open active site should help promote rapid

[4] S. I. Tomarev, S. Chung, and J. Piatigorsky, manuscript submitted for publication.

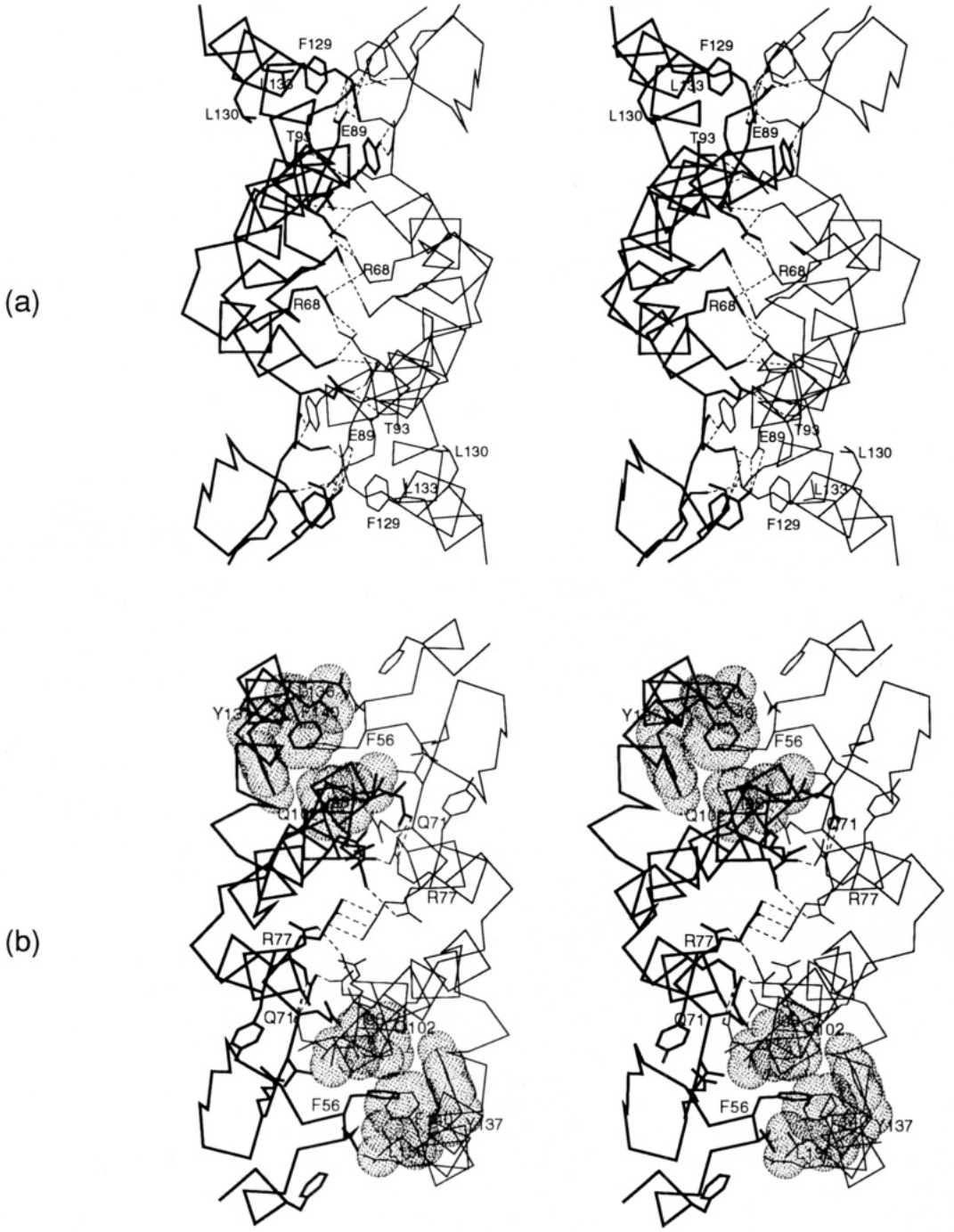

FIGURE 6: Stereoview of the hydrogen-bonding pattern and important van der Waals contacts between the two subunits of (a) the GSH transferase from squid (this work) and (b) class mu isoenzyme M1-1 from rat (Ji et al., 1992). The model of one subunit is shown in bold. Hydrogen bonds and electrostatic interactions are illustrated as dotted lines while side chains involved in the lock-and-key kind of hydrophobic interaction are highlighted as dot surfaces.

association of substrate and dissociation of product from the enzyme. The poor performance of the native enzyme in the addition of GSH to enones and epoxides is related, at least in part, to the absence of a residue at position 106 that is capable of proffering electrophilic assistance in these reactions. For example, both crystallographic and kinetic evidence suggests that the hydroxyl group of Y115 of the M1-1 isoenzyme protonates or hydrogen bonds to the oxirane oxygen of epoxides and the enolic oxygen of enones and thereby facilitates the addition of GS^- to these systems (Johnson et al., 1993; Ji et al., 1994). In the case of the squid enzyme this notion is buttressed by the fact that the F106Y mutant enzyme has k_{cat} and k_{cat}/K_m^{PBO} values that approach those of the M1-1 isoenzyme from rat where the equivalent position (115) is occupied by tyrosine. In fact, the structures and catalytic behavior of the cytosolic GSH transferases in general suggest that the presence of a tyrosyl residue at this position is a good predictor of whether a particular isoenzyme will catalyze Michael additions or epoxide ring openings efficiently.

Evolution of the Cephalopod GSH Transferases. The structural relatedness of the squid enzyme to other classes of GSH transferases may offer some clues as to the evolution of this group of enzymes. The squid enzyme has been variously suggested, on the basis of sequence comparisons, to be more closely related to the class alpha (Harris et al.,

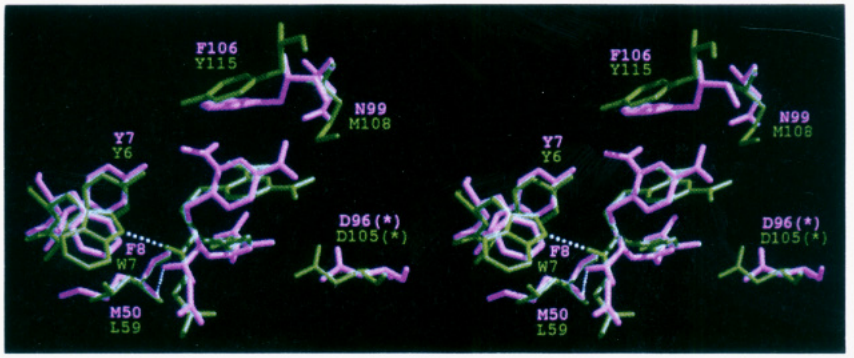

FIGURE 7: Stereoview of the alignment of squid GSH transferase (in pink, this work) and subunit A of class mu isoenzyme M1-1 (in green; 5GST, Ji et al., 1993) in complex with 1-(*S*-glutathionyl)-2,4-dinitrobenzene (GSDNB). The different hydrogen bond arrangements between the protein and the carbonyl oxygen of the cysteinyl residue of the bound GSDNB are illustrated by the dotted lines. The asterisks (*) indicate that the corresponding residue is from the other subunit of the dimeric enzymes.

1991) or pi (Tomarev et al., 1993) enzymes. However, the extent of sequence identities (19−34%) or similarities (about 40%; Wilce & Parker, 1994) with the four principal vertebrate classes is similar enough to make such distinctions questionable. A comparison of exon/intron boundaries of the squid gene with those from vertebrates provides little help in this regard. Pemble and Taylor (1992) have argued that the squid gene probably derived from a progenitor of the class alpha and pi genes subsequent to the divergence of mu from the alpha/mu/pi precursor. If this is true, then the squid gene must have undergone considerable evolution after divergence to account for the dramatic difference in the structure at the dimer interface compared to the class alpha, mu, and pi enzymes. An alternative evolutionary hypothesis is that the squid gene diverged from theta or an alpha/mu/pi precursor which originally encoded a protein that lacked the loop and hydrophobic lock for subunit dimerization. In this regard, it is interesting to note that, similar to the squid sequence, the class theta enzymes from rat do not have a phenylalanyl or tyrosyl residue in a position to participate in the intersubunit hydrophobic lock observed in the class alpha, mu, and pi structures. In addition, at least two of the cognate residues in the sequence of class theta enzymes that would be expected to form the hydrophobic pocket between the α4 and α5 helices are hydrophilic. Since it is difficult to argue these structural points with a high degree of certainty from sequence alignments alone, it remains to be seen if the class theta proteins lack the loop and the other structural elements necessary for this specific type of interfacial interaction. Nevertheless, this structural analysis suggests that the cephalopod GSH transferase may have diverged from its ancestral precursor before alpha/mu/pi diverged as illustrated in Figure 8.

Historically, the classification of GSH transferases has been based on the extent of sequence similarity between various isoenzymes. Although there are no established criteria for the extent of sequence similarity necessary to be a member of one class or another, it is fair to say that the squid enzyme is sufficiently different (<34% identity) from four of the principal vertebrate classes to warrant a separate classification. If the classification criteria are extended to include familial recognition through the formation of heterodimers within a particular enzyme class, a phenomenon that is a direct result of structural differences at the dimer interface, then the case for a separate classification of the squid GSH transferases is even more compelling. It is proposed, on the basis of the significant structural differences

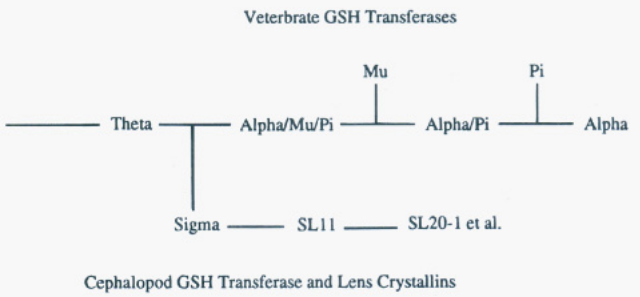

FIGURE 8: Proposed evolutionary scheme for cephalopod (class sigma) GSH transferases and the lens *S*-crystallins. The scheme for alpha, mu, and pi classes is after Pemble and Taylor (1992).

between the squid and vertebrate enzymes, that the GSH transferases from cephalopods be given the classification sigma, after the original suggestion by Buetler and Eaton (1992).

Relationship of the Enzyme to the Lens S-Crystallins. The gene for the squid GSH transferase is clearly closely related to those which encode the major soluble lens crystallins of the cephalopod eye. The primary structure of the enzyme more closely resembles those of the *S*-crystallins (42−44% identity) than the sequences of the vertebrate GSH transferases. In fact, the SL11 crystallin which most closely resembles the squid enzyme (Tomarev et al., 1993) has substantial GSH transferase activity.[4,5] This soluble refractory protein is about the same length as the squid enzyme, appears to be a dimer in dilute solution,[5] and does not bind to *S*-hexylglutathione affinity columns. Sequence alignments suggest that all but 3 of the 16 residues conserved in all of the soluble GSH transferases (Figure 4) are conserved in the SL11 crystallin. Of the six residues [Y7, W38, K42, Q62, S63, and D96(B)] of the squid enzyme that have side chains directly involved in the binding of GSH, five have identical counterparts in SL11 (Y7, W38, K42, Q63, and S64). The single exception is E97 of the lens protein which corresponds to D96(B), the aspartate that reaches over from the adjacent subunit to form an electrostatic contact with the α-amino group of the γ-glutamyl residue of GSH. Whether this or another difference in the GSH binding site is responsible for the lack of binding of SL11 to the *S*-hexylglutathione

[5] The SL11 crystallin has a k_{cat} of 0.23 s^{-1}, a k_{cat}/K_m^{CDNB} of 99 M^{-1} s^{-1} at pH 6.5, and an apparent molecular mass of 38 kDa, consistent with a dimeric protein in solution (E. C. von Rosenvinge, W. W. Johnson, and R. N. Armstrong, unpublished results.

affinity matrix is not clear.

The SL11 crystallin with its catalytic activity and close relationship to the squid GSH transferase may be an intermediate or "molecular fossil" in the evolution of the lens *S*-crystallins of cephalopods[4] as shown in Figure 8. Most *S*-crystallins have a substantial peptide insert of variable length at a position which corresponds to the junction between the α4 and α5 helices of the squid enzyme. Thus, in all *S*-crystallins except SL11, the helix−turn−helix motif appears to become a helix−insert−helix element by the incorporation of an additional exon precisely between exons 3 and 4 of the GSH transferase gene. For example, the most abundant crystallin (SL20-1) has an 18-residue peptide insert at the point between residues 112 and 113 of the squid GSH transferase (Tomarev et al., 1993). The insert occurs near the surface of the protein so that the incorporation of the additional peptide, whatever its structure, is less likely to result in a substantial disruption of the core (GSH transferase-like) structure of the crystallin.

Additional evidence that the SL20-1 and related crystallins have diverged further from the ancestral GSH transferase and the intermediate SL11 is their lack of catalytic activity. The SL20-1 crystallin has little or no catalytic activity toward CDNB,[4] yet it retains many of the conserved residues of the squid GSH transferase and SL11 proteins including four of the six residues whose side chains are involved in electrostatic contacts with GSH. The most notable exception is the residue E63 of SL20-1 which corresponds to Q62 and Q63 of the GSH transferase and SL11, respectively. This glutamine along with the subsequent serine or threonine residue is among the most highly conserved and crucial residues for the recognition of the γ-glutamyl moiety of GSH. Its replacement is convincing evidence of the continued divergence of the SL20-1 crystallin from the ancestral SL11 and GSH transferase and may be directly related to the further loss of catalytic activity of the crystallin used for refraction in the lens. It will be interesting to determine the extent to which the architecture of the enzyme has been altered to fulfill the physical role of a refractory protein.

One salient question is whether a "glutathione transferase-like" crystallin confers any biochemical advantage to the lens tissue through either residual catalytic activity or ability, even if diminished, to bind GSH. The concentration of GSH in squid lens is about 3 mM, which is similar to most vertebrate tissues including ocular tissue.[6] The very fact that the major lens crystallin SL20-1 has evolved so as to lose the ancestral enzymatic activity suggests, at the very least, that the catalytic activity is less important than the refractory properties of the protein. In fact, it can be argued that retention of the ability to bind GSH could be detrimental to the lens tissue since a high concentration of the peptide would be sequestered by the crystallin and less available for participation in cellular processes.

ACKNOWLEDGMENT

We thank Sambath Chung for help in constructing the F106Y mutant.

SUPPLEMENTARY MATERIAL AVAILABLE

Tables 1 and 2 describing the statistical details of the native diffraction data and the heavy-atom positions of derivatives **1** and **2** (2 pages). Ordering information is given on any current masthead page.

[6] S. I. Tomarev, unpublished results.

REFERENCES

Armstrong, R. N. (1991) *Chem. Res. Toxicol. 4*, 131−140.
Armstrong, R. N. (1994) *Adv. Enzymol. Relat. Areas Mol. Biol. 69*, 1−44.
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. (1992) *X-PLOR (Version 3.1) Manual*, pp 187−218, Yale University Press, New Haven, CT.
Brünger, A. T., Kuriyan, J., & Karplus, M. (1987) *Science 235*, 458−460.
Buetler, T. M., & Eaton, D. L. (1992) *Environ. Carcinogen. Ecotoxicol. Rev. C10*, 181−203.
Cleland, W. W. (1979) *Methods Enzymol. 63*, 103−138.
Cobb, D., Boehlert, C., Lewis, D., & Armstrong, R. N. (1983) *Biochemistry 22*, 805−812.
Dirr, H., Reinemer, P., & Huber, R. (1994) *Eur. J. Biochem. 220*, 645−661.
Furey, W. (1990) *Abstracts of the American Crystallographic Association Fortieth Anniversary Meeting*, New Orleans, PA33, LA.
Furey, W., Wang, B. C., & Sax, M. (1982) *J. Appl. Crystallogr. 15*, 160−166.
Garcia-Saez, I., Parraga, A., Phillips, M. F., Mantle, T. J., & Coll, M. (1994) *J. Mol. Biol. 237*, 298−314.
Graminski, G. F., Zhang, P., Sesay, M. A., Ammon, H. L., & Armstrong, R. N. (1989) *Biochemistry 28*, 6252−6258.
Habig, W. H., Pabst, M. J., & Jakoby, W. B. (1974) *J. Biol. Chem. 249*, 7130−7139.
Harris, J., Coles, B., Meyer, D. J., & Ketterer, B. (1991) *Comp. Biochem. Physiol. 98B*, 511−515.
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.
Howard, A. J., Gilliland, G. L., Finzel, B. C., Poulos, T. L., Ohlendorf, D. H., & Salemme, F. R. (1987) *J. Appl. Crystallogr. 20*, 383−387.
James, M. N. G., & Sielecki, A. R. (1983) *J. Mol. Biol. 163*, 299−361.
Ji, X., Zhang, P., Armstrong, R. N., & Gilliland, G. L. (1992) *Biochemistry 31*, 10169−10184.
Ji, X., Armstrong, R. N., & Gilliland, G. L. (1993) *Biochemistry 32*, 12949−12954.
Ji, X., Johnson, W. W., Sesay, M. A., Dickert, L., Prasad, S. M., Ammon, H. L., Armstrong, R. N., & Gilliland, G. L. (1994) *Biochemistry 33*, 1043−1052.
Johnson, W. W., Liu, S., Ji, X., Gilliland, G. L., & Armstrong, R. N. (1993) *J. Biol. Chem. 268*, 11508−11511.
Jones, D. H., & Howard, B. H. (1990) *BioTechniques 8*, 178−183.
Jones, T. A., & Kjeldgaard, M. (1993) *O Version 5.9.1*, 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. (1988) *J. Appl. Crystallogr. 21*, 67−71.
Kabsch, W., & Sander, C. (1983) *Biopolymers 22*, 2577−2637.
Klinga-Levan, K., Andersson, A., Hanson, C., Ridderström, M., Stenberg, G., Mannervik, B., Vajdy, M., Szpirer, J., Szpirer, C., & Levan, G. (1993) *Hereditas 119*, 285−296.

Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24,* 946−950.
Lim, K., Ho, J. X., Keeling, K., Gilliland, G. L., Ji, X., Rüker, F., & Carter, D. C. (1994) *Protein Sci. 3,* 2233−2244.
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.
Mannervik, B., Awasthi, Y. C., Board, P. G., Hayes, J. D., Di Ilio, C., Ketterer, B., Listowsky, I., Morgenstern, R., Muramatsu, M., Pearson, W. R., Pickett, C. B., Sato, K., Widersten, M., & Wolf, C. R. (1992) *Biochem. J. 282,* 305−308.
Meyer, D. J., Coles, B., Pemble, S. E., Gilmore, K. S., Fraser, G. M., & Ketterer, B. (1991) *Biochem. J. 274,* 409−414.
Mignogna, G., Allocatt, N., Aceto, A., Piccolomini, R., De Ilio, C., Barra, D., & Martini, F. (1993) *Eur. J. Biochem. 211,* 421−425.
Navaza, J. (1994) *Acta Crystallogr. A50,* 157−163.
Nishida, M., Kong, K.-H., Inoue, H., & Takahashi, K. (1994) *J. Biol. Chem. 269,* 32536−32541.
Pemble, S. E., & Taylor, J. B. (1992) *Biochem. J. 287,* 957−963.
Powell, M. J. D. (1977) *Math. Prog. 12,* 241−254.
Raghunathan, S., Chandross, R. J., Kretsinger R. H., Allison, T. J., Penington C. J., & Rule, G. S. (1994) *J. Mol. Biol. 238,* 815−832.
Reinemer, P., Dirr, H. W., Ladenstein, R., Schaffer, J., Gallay, O., & Huber, R. (1991) *EMBO J. 10,* 1997−2005.
Reinemer, P., Dirr, H. W., Ladenstein, R., Huber, R., Lo Bello,M., Federici, G., & Parker, M. W. (1992) *J. Mol. Biol. 227,* 214−226.
Rushmore, T. H., & Pickett, C. B. (1993) *J. Biol. Chem. 268,* 11475−11478.
Sinning, I., Kleywegt, G. J., Cowan, S. W., Reinemer, P., Dirr, H. W., Huber, R., Gilliland, G. L., Armstrong, R. N., Ji, X., Board, P. G., Olin, B., Mannervik, B., & Jones, T. A. (1993) *J. Mol. Biol. 232,* 192−212.
Tomarev, S. I., & Zinovieva, R. D. (1988) *Nature 336,* 86−88.
Tomarev, S. I., Zinovieva, R. D., & Piatigorsky, J. (1991) *J. Biol. Chem. 266,* 24226−24231.
Tomarev, S. I., Zinovieva, R. D., & Piatigorsky, J. (1992) *J. Biol. Chem. 267,* 8604−8612.
Tomarev, S. I., Zinovieva, R. D., Guo, K., & Piatigorsky, J. (1993) *J. Biol. Chem. 268,* 4534−4542.
Wilce, M. C. J., & Parker, M. W. (1994) *Biochim. Biophys. Acta 1205,* 1−18.
Wistow, G., & Piatigorsky, J. (1987) *Science 236,* 1554−1556.

BI950091Q