Kim, P. S. (1986) *Methods Enzymol. 131*, 136-156.
Kim, P. S., & Baldwin, R. L. (1982) *Biochemistry 21*, 1-5.
Krimm, S., & Bandekar, J. (1986) *Adv. Protein Chem. 38*, 181-364.
Kummer, H., & Rüdiger, H. (1988) *Biol. Chem. Hoppe-Seyler 369*, 639-646.
Lis, H., & Sharon, N. (1984) in *Biology of Carbohydrates* (Ginsburg, V., & Robbins, P., Eds.) Vol. 2, pp 1-86, J. Wiley, New York.
Lotta, T. I., Salonen, I. S., Virtanen, J. A., Eklund, K. K., & Kinnunen, P. K. J. (1988) *Biochemistry 27*, 8158-8169.
Naik, V. M., & Krimm, S. (1986) *Biophys. J. 49*, 1131-1145.
Pace, C. N. (1986) *Adv. Enzymol. 131*, 266-280.
Perchard, C., Baron, M.-H., & de Lozē, C. (1984) *J. Mol. Struct. 112*, 247-262.
Pêre, M., Bourrillon, R., & Jirgensons, B. (1975) *Biochim. Biophys. Acta 393*, 31-36.
Poole, P. L., & Barlow, D. J. (1986) *Biopolymers 25*, 317-335.
Schellman, J. A. (1987) *Biopolymers 26*, 549-559.
Segawa, S. I., & Kume, K. (1986) *Biopolymers 25*, 1981-1996.
Sharif, A., Brochier, J., & Bourrillon, R. (1977) *Cell. Immunol. 31*, 302-310.
Strosberg, A. D., Buffard, D., Lauwereys, M., & Foriers, A. (1986) in *The Lectins: Properties, Functions and Applications in Biology and Medicine* (Liener, I. E., Sharon, N., & Goldstein, I. J., Eds.) pp 249-264, Academic Press, New York.
Surewicz, W. K., & Mantsch (1988) *Biochim. Biophys. Acta 952*, 115-130.
Surewicz, W. K., Moscarello, M. A., & Mantsch, H. H. (1987) *Biochemistry 26*, 3881-3886.
Susi, H., & Byler, D. M. (1987) *Arch. Biochem. Biophys. 258*, 465-469.
Trewhella, J., Liddle, W. K., Heidorn, D. B., & Strynadka, N. (1989) *Biochemistry 28*, 1294-1301.
Von Hippel, P. H., & Schleich, T. (1969) in *Structure and Stability of Biological Macromolecules* (Timasheff, S. N., & Fasman, G. D., Eds.) pp 417-574, M. Dekker, New York.
Wantyghem, J., Goulut, C., Frênoy, J.-P., Turpin, E., & Goussault, Y. (1986) *Biochem. J. 237*, 483-489.

Structure of Yeast Triosephosphate Isomerase at 1.9-Å Resolution[†,‡]

Elias Lolis, Tom Alber,[§] Robert C. Davenport,[∥] David Rose,[⊥] Fred C. Hartman,[#] and Gregory A. Petsko*
Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139
Received June 30, 1989; Revised Manuscript Received December 11, 1989

ABSTRACT: The structure of yeast triosephosphate isomerase (TIM) has been solved at 3.0-Å resolution and refined at 1.9-Å resolution to an *R* factor of 21.0%. The final model consists of all non-hydrogen atoms in the polypeptide chain and 119 water molecules, a number of which are found in the interior of the protein. The structure of the active site clearly indicates that the carboxylate of the catalytic base, Glu 165, is involved in a hydrogen-bonding interaction with the hydroxyl of Ser 96. In addition, the interactions of the other active site residues, Lys 12 and His 95, are also discussed. For the first time in any TIM structure, the "flexible loop" has well-defined density; the conformation of the loop in this structure is stabilized by a crystal contact. Analysis of the subunit interface of this dimeric enzyme hints at the source of the specificity of one subunit for another and allows us to estimate an association constant of 10^{14}-10^{16} M^{-1} for the two monomers. The analysis also suggests that the interface may be a particularly good target for drug design. The conserved positions (20%) among sequences from 13 sources ranging on the evolutionary scale from *Escherichia coli* to humans reveal the intense pressure to maintain the active site structure.

Yeast triosephosphate isomerase (TIM;[1] EC 5.3.1.1) is a dimer of identical subunits, each of molecular weight about 26 000. There is no evidence for catalytic cooperativity between the two subunits. It has been shown, however, that the enzyme is only active as a dimer (Waley, 1973; Casal et al., 1987). TIM is the glycolytic enzyme that catalyzes the isomerization of the two products of aldolase, D-glyceraldehyde 3-phosphate and dihydroxyacetone phosphate. Since only GAP continues to proceed along the glycolytic pathway, TIM is absolutely essential for efficient energy production. There are no reported cases of organisms devoid of triosephosphate isomerase activity. Indeed, people who are defective in one of the alleles for the TPI gene are found to suffer from chronic hemolytic anemia and neuromuscular disorders (Valentine et al., 1983).

Triosephosphate isomerase has been the subject of extensive biophysical studies. The free energy profile of the chicken

[†] The research of G.A.P. was supported by grants from NIH and that of F.C.H. by the Office of Health and Environmental Research, U.S. Department of Energy, under Contract DE-AC05-84OR21400 with the Martin Marietta Energy Systems, Inc. We also acknowledge Berlex Laboratories for providing a fellowship for one of us (E.L.).
[‡] Crystallographic coordinates have been submitted to the Brookhaven Protein Data Bank.
[§] Present address: Department of Biochemistry, University of Utah School of Medicine, Salt Lake City, UT 84132.
[∥] Present address: Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139.
[⊥] Present address: Division of Biological Sciences, National Research Council, Ottawa, Canada K1A 0R6.
[#] Permanent address: Biology Division, Oak Ridge National Laboratory, Oak Ridge, TN 37830.

[1] Abbreviations: DHAP, dihydroxyacetone phosphate; GAP, D-glyceraldehyde 3-phosphate; rms, root mean square; TIM, triosephosphate isomerase; TPI, gene coding for triosephosphate isomerase.

0006-2960/90/0429-6609$02.50/0

Table I: TIM Structures

enzyme source	resolution (Å)	*R* factor (%)	active site ligand
chicken[a]	2.5	17[b]	sulfate
chicken[c]	6.0		PGA[d]
yeast[e]	3.5		DHAP[d]
Trypanosoma brucei[f]	2.4	23	
yeast[g]	1.9	21	H_2O
yeast[h]	2.5	18	PGA

[a] Banner et al., 1975. [b] Peter Artemiuk, personal communication. [c] Phillips et al., 1977. [d] Flow cell experiment. [e] Alber et al., 1981a. [f] Wierenga et al., 1987. [g] This study. [h] Lolis and Petsko, 1990.

muscle (Albery & Knowles, 1976) and yeast (Nickbarg & Knowles, 1988) enzymes has been determined and shows the highest barrier in the reaction to be substrate (D-GAP) binding to enzyme. The dependence of the rate of the reaction on the viscosity of the medium confirms that TIM operates at the diffusion-controlled limit (Blacklow et al., 1988). A set of equations based on individual rate constants has been derived that defines the efficiency of any enzyme (Albery & Knowles, 1977). The kinetic parameters of TIM indicate that the enzyme has reached evolutionary perfection (Knowles & Albery, 1977); no mutation can ever increase the rate of the reaction in vivo.

In the classic paper on enzyme inhibition by transition-state analogues, Wolfenden used TIM as the model enzyme and phosphoglycolate as the transition-state analogue (Wolfenden, 1969). TIM has also been used more recently to study subunit interactions (Casal et al., 1987) and enzyme thermostability (Ahern et al., 1987). The nonenzymatic isomerization reaction has been studied with simple organic bases as catalysts (Richard, 1984). In these solution studies it has been determined that abstraction of a proton from the substrate (either DHAP or GAP) forms an enediolate intermediate. This intermediate is more prone to undergo an ω-elimination reaction to form methyl glyoxal and inorganic phosphate than it is to undergo a reprotonation on the adjacent carbon to complete the isomerization reaction.

For the enzymatic reaction the story is very different. TIM catalyzes the isomerization reaction via an enediol(ate)[2] intermediate by almost 10 orders of magnitude relative to an organic base, while limiting production of toxic methyl glyoxal to a negligible amount (Richard, 1984). In order to understand the structural basis of catalysis, the three-dimensional structures of triosephosphate isomerases from chicken (Banner et al., 1975), yeast (Alber et al., 1981a), and *Trypanosoma brucei brucei* (Wierenga et al., 1987) have been solved (Table I). Mutants of the enzyme have also been made in order to decipher the mechanism of the reaction (Straus et al., 1985; Nickbarg et al., 1988). The structure of the yeast enzyme has now been refined at 1.9-Å resolution. We report here details of the structure determination and refinement and discuss the active site and intersubunit interactions of TIM in light of these new, more accurate atomic coordinates.

Materials and Methods

Crystallization and Heavy Atom Search. The crystallization of yeast TIM has been reported previously (Alber et al., 1981b). A 20 mg/mL sample of yeast triosephosphate isomerase in 12% saturated ammonium sulfate, 1 mM EDTA, 1 mM mercaptoethanol, and 50 mM Tris buffer, pH 7.5, was crystallized by the addition of poly(ethylene glycol) of average molecular weight 4000 to a final concentration of about 16% by either vapor diffusion or batch methods. The crystals have the symmetry of space group $P2_1$ with one dimeric molecule in the asymmetric unit. The unit cell of the crystal has dimensions a = 61.15 Å, b = 98.55 Å, c = 49.26 Å, and β = 91.2°.

To diffuse heavy atoms into the crystal lattice, crystals were transferred to sulfate-free mother liquor containing 40% PEG 4000 at 4 °C and then to identical mother liquor containing the heavy atom compound. The soaks were usually carried out in the dark at 4 °C to inhibit any photochemical reactions of the heavy atom compounds. To assess the results of the soaks, precession photographs of various derivative-soaked crystals were compared with those of the native. Two mercury compounds, mercury *meso*-2,3-dimethoxytetramethylenebis-[acetate] (Baker's dimercurial) at a concentration of 1.8 mM and soaking time of 14–22 days and ethyl mercuric phosphate (EMP) at a concentration of 3 mM and soaking time of 44 and 89 h, were found to bind to two sites per asymmetric unit. Dipotassium platinum nitrite at a concentration of 5 mM and a soaking time of 31 days was found to bind to four sites per asymmetric unit.

Data Collection and Reduction. Data were collected at −5 to 5 °C to reduce the rate of radiation damage. By use of two crystals, a native data set was collected to 1.9-Å resolution on a single-counter Nicolet P3 diffractometer equipped with a low-temperature device. Diffractometer measurements were made on the heavy atom derivatives to 3.0-Å resolution, with equivalent reflections measured as h,k,l and $h,-k,l$. Nickel-filtered copper Kα radiation was used for all intensity measurements. Data were collected by the ω-step scan method (Wyckoff et al., 1967). Peak intensities were estimated by integrating over 40% of the width of the reflections. By use of a takeoff angle of 5°, the counts of seven of eleven steps of 0.03° in ω were summed. On average, the full peak width was 0.5° in ω. Negative intensities and those intensities below twice the standard deviation were not included in the final reduced data sets. Backgrounds were estimated by using a radial background curve which was dependent on ϕ and 2θ. Background curves were measured for each crystal at the ϕ value of minimum transmission (ϕ_{min}) and at $\phi_{min+90°}$. Absorption corrections were made by the semiempirical method (North et al., 1968). Radiation damage was monitored by periodically measuring five reflections in different regions of reciprocal space. The two native crystals suffered radiation damage of 17% and 33%. A linear correction function for radiation damage was calculated from the five standard reflections and applied along with Lorentz and polarization corrections. The data from the two crystals were merged with an R_{merge} ($R_{merge} = \sum_{hkl}|I - \bar{I}|/\sum_{hkl}I$) of 11.3% for 2028 common reflections. To correct systematic errors in the measurement of Friedel equivalent reflections, scale factors were applied so that the average value of $F(PH)(h,k,l)/F(PH)(h,-k,l)$ in blocks of reciprocal space approached unity (North et al., 1968). Data sets of the metal-derivatized crystals were analyzed and scaled to the native data set by iterative application of the local scaling procedure (Matthews & Czerwinski, 1975). The method of Patterson origin comparison was used to improve the scaling of the derivative to the native data (Arnone et al., 1971).

Heavy Atom Parameters and Phasing. Patterson maps were calculated by using the isomorphous, anomalous, and "combined" differences as coefficients. The F(HLE) method (Matthews, 1966b; Dodson & Vijayan, 1971) was used to integrate the information from the isomorphous and anomalous

[2] It has not yet been determined whether the intermediate formed on the enzyme is an enediol or an enediolate.

Table II: Multiple Isomorphous Replacement Statistics

derivative	Heavy Atom Sites x	y	z	occupancy	B factor
Baker's dimercurial					
Hg 1	0.22655	0.00000	0.26639	0.64	11.2
Hg 2	0.22579	0.41522	0.00552	0.79	16.9
$K_2Pt(NO_2)_4$					
Pt 1	0.02262	0.03575	0.03141	0.64	93.1
Pt 2	0.03125	0.99716	0.99421	0.40	66.2
Pt 3	0.96994	0.02765	0.03091	0.19	−13.5
Pt 4	0.96373	0.04362	0.11079	0.39	101.0

derivative	Phasing Power as a Function of Resolution (Å) total	10.75	6.80	5.32	4.50	3.99	3.61	3.32	3.00
Baker's dimercurial									
no. of centric reflections	466	63	64	58	65	53	57	56	50
$rms(F_H)/rms(E)$	1.53	2.15	1.76	2.12	1.56	1.56	0.99	1.34	1.01
no. of acentric reflections	11 269	536	948	1190	1427	1563	1772	1911	1922
$rms(F_H)/rms(E)$	2.54	2.64	3.23	3.01	2.36	2.18	2.50	2.42	2.47
$K_2Pt(NO_2)_4$									
no. of centric reflections	477	64	62	58	64	58	61	58	52
$rms(F_H)/rms(E)$	1.42	2.36	1.24	0.93	0.80	1.08	0.93	1.00	0.86
no. of acentric reflections	10 882	488	894	1138	1374	1531	1713	1851	1893
$rms(F_H)/rms(E)$	1.53	2.56	1.90	1.55	1.33	1.41	1.22	1.24	1.27

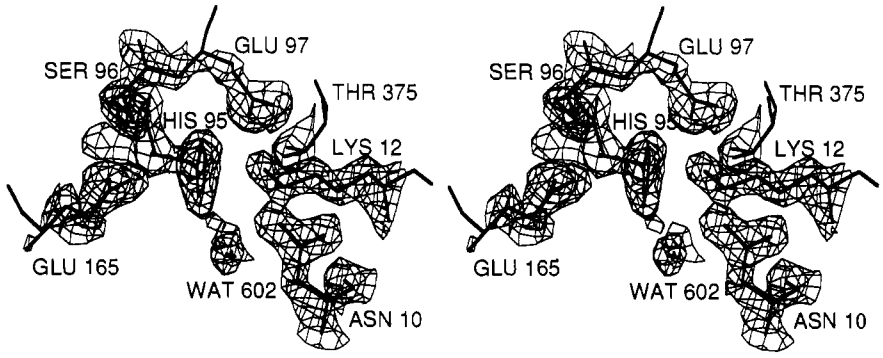

FIGURE 1: (a) Stereoscopic view of the electron density at the active site in subunit 1. The electron density is from a $2F_o - F_c$ omit map (the residues in this figure were deleted) contoured at 1.5 times the standard deviation of electron density in the asymmetric unit. Thr 375 belongs to subunit 2 and is included in this figure to show the proximity of this residue to the active site of subunit 1.

differences. $F(HLE)^2$ Patterson maps were solved by inspection, and the putative heavy atom positions were refined by minimizing the overall difference between $F(HLE)$ and the calculated heavy atom structure amplitude $F_H(hkl)$. The EMP derivative was found to be identical with the Baker's mercurial derivative and was not used further. Phases using isomorphous plus anomalous contributions from the Baker's and platinum derivatives were calculated by standard methods (Blow & Crick, 1959; Matthews, 1966a). Centric isomorphous E values divided by $2^{1/2}$ were used, and the E's for the anomalous differences were set at approximately $^1/_3$ of those for the isomorphous differences. The average figure of merit was 0.76 for all reflections to 3.0-Å resolution. The statistics for the heavy atom derivatives used for phasing are summarized in Table II.

Electron Density and Interpretation. Electron density maps were calculated with the fast Fourier program of R. Fisher. Both the molecular and subunit boundaries were clearly defined. The chain was traced in a minimap (Hill et al., 1972). After 210 of 247 Cα's in one subunit had been located, the fold of the chicken TIM monomer (Banner et al., 1975) could be discerned in the yeast TIM electron density map. The coordinates of 10–20 Cα's were estimated, and rotation and translation matrices were calculated which transformed the unrefined chicken TIM coordinates into the yeast TIM unit cell. This procedure was repeated until the root mean square difference between the Cα coordinate sets was not improved by adding more atoms. The transformed coordinates were used to complete the tracing of the first monomer and to follow the polypeptide chain of the second subunit from beginning to end. After all the Cα's were fit, the remaining backbone and side-chain atoms were fit by using the program FRODO (Jones & Thirup, 1986) on an Evans and Sutherland Picture System 300. Water molecules were assigned (after a number of rounds of refinement; see below) by inspecting all positions where significant peaks were found on difference maps. An indication of the quality of the electron density map after refinement is given in Figure 1, where a delete map was calculated for some of the active site residues.

Refinement. The atomic coordinates were refined by using the restrained-parameter least-squares method of Hendrickson and Konnert (Hendrickson, 1985). The initial model containing only protein atoms had an R value of 43.1% for all observed reflections between 10 and 3.3 Å. This model was subjected to crystallographic refinement in a generally repetitive manner, as follows: (1) least-squares refinement proceeded until the R value stopped decreasing; (2) at this point, the new coordinates were output and $2F_o - F_c$ "delete maps", in which 10-residue segments of the structure were omitted from the refinement and phase calculations, were used to manually rebuild the structure; and (3) after rebuilding, the coordinates were resubmitted for more refinement. The resolution range most influenced by the bulk solvent (10–5 Å) was eliminated from the calculation as soon as the R factor in that shell stopped decreasing. Reflections from higher resolution shells, first to 2.5 Å and then to 1.9 Å, were

Table III

Refinement Statistics		restraints applied, σ (Å)	deviations obsd, Δ (Å)[a]
no. of protein atoms	3766		
no. of water molecules	119		
bond distance		0.030	0.025
angle distance		0.040	0.048
planar distance		0.052	0.055
no. of bad dihedral angles[b]	8		

Resolution Breakdown				
resolution (Å)	no. of reflections	% of data	shell *R* factor	sphere *R* factor
40.00–4.00	5160	98.5	0.193	0.193
4.00–3.15	5158	97.2	0.180	0.187
3.15–2.75	4741	91.5	0.208	0.192
2.75–2.45	5185	81.0	0.223	0.197
2.45–2.23	4987	70.0	0.230	0.201
2.23–2.06	5052	65.2	0.243	0.206
2.06–1.90	4880	48.9	0.253	0.210
40.00–1.90	35167	74.9		0.210

[a] These Δ values are rms deviations from the corresponding values for ideal groups derived from small molecular structural studies (Sielecki et al., 1979). [b] Residues with dihedral angles significantly outside Ramachandran limits: Lys 12, Asn 28, Asn 35, His 103, Arg 303, His 403, Leu 431, Asp 522.

eventually added to the refinement calculations. Waters were added at appropriate positions by analyzing $2F_o - F_c$ maps. Prior to addition of waters and refinement of individual isotropic Debye–Waller factors, the *R* value had dropped to 28.3% for all data between 5- and 1.9-Å resolution. After water molecules and individual Debye–Waller (or *B*) factors were included in the refinement, a weighting scheme designed to even out the *R* value in each resolution range was used. A final structure using this method had an *R* value of 23.1%. The structure (without the water molecules) was then used as the initial model for refinement by simulated annealing according to the program XPLOR (Brunger et al., 1987). The output from this XPLOR refinement was then checked for errors and resubmitted for additional least-squares refinement, this time without the weighting scheme to even out the *R* value in each resolution range. In addition, the scattering contributions from the bulk solvent were also modeled (Fraser et al., 1978) in this later version of the refinement program. The *R* factor of the final model is 21.0% for 35 167 reflections from 40- to 1.9-Å resolution. Refinement statistics are given in Table III.

Results and Discussion

Analysis of the Structure. The topology of TIM consists of an eight-stranded α/β barrel. This topology is one of the most common in metabolic biochemistry. To date, 16 other functionally different enzymes with this fold have been found: pyruvate kinase (Stuart et al., 1979), enolase (Lebioda & Stec, 1988), KDGP aldolase (Mavridis et al., 1982), Taka-amylase (Matsuua et al., 1984), mandelate racemase (D. Neidhart, P. L. Howell, V. M. Powers, G. L. Kenyon, J. A. Gerlt, and G. A. Petsko, unpublished results), α-amylase (Buisson et al., 1987), muscle aldolase (Sygusch et al., 1987), xylose isomerase (Farber et al., 1987; Henrick et al., 1987), glycolate oxidase (Lindqvist & Branden, 1985), RuBisCo (Schneider et al., 1986; Chapman et al., 1988), muconate lactonizing enzyme (Goldman et al., 1987), tryptophan synthase (Hyde et al., 1987), trimethylamine dehydrogenase (Lim et al., 1986), flavocytochrome b_2 (Xia et al., 1987), and the bifunctional enzyme *N*-(5′-phosphoribosyl)anthranilate isomerase–indole-3-glycerol phosphate synthase, where each domain is an α/β barrel (Priestle et al., 1987). TIM was the first enzyme in this class and remains the prototype of what has been called the TIM barrel folding pattern. In the TIM structure, the β barrel is composed of eight parallel β-strands involving residues 5–11, 37–41, 58–63, 89–93, 122–127, 159–165, 205–209, and 227–231.[3] The eight β-strands have been referred to in the past by lower case letters of the alphabet, starting with "a" for the first β-strand and ending with "h" for the last β-strand. Looking down the axis of the barrel, the barrel resembles an ellipse with the lengths of the semiaxes being 12 and 18 Å. The length of the barrel itself is only about 13 Å. The barrel principally arises from the right-hand twist of the β-sheet (Lasters et al., 1988). The angle made by the β-strand axes with the barrel axis is about 35°. Figure 2 contains an atomic listing of the intrachain hydrogen-bonding pattern of the staggered β-strands. The barrel is situated in the protein interior; the strands are interconnected by amphipathic helices that pack against the barrel surface. As can be seen from Figure 2, the amino acids comprising the barrel are predominantly hydrophobic; many of these hydrophobic side chains fill the interior of the barrel and form part of the hydrophobic core of the enzyme. The only polar amino acids are found at the ends of the barrel, especially the amino-terminal end, where the barrel is exposed to solvent.

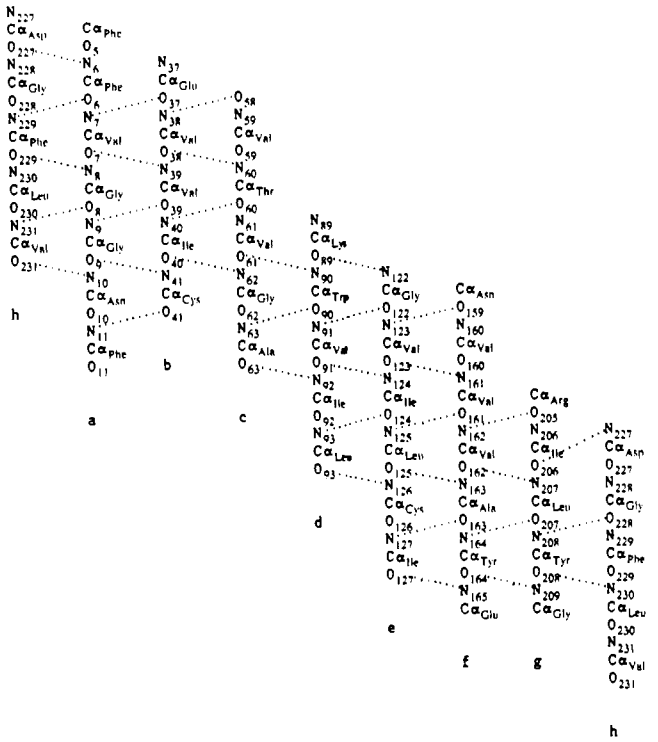

FIGURE 2: Hydrogen-bonding pattern of β barrel that has been unrolled to form a single sheet. The hydrophobic nature of the sheet is obvious from its amino acid sequence. Below each β-sheet is the letter that it is represented by.

All of the β-strands are followed by at least one helix. By use of backbone hydrogen bonds as a criterion, the helical segments consist of residues 16–30, 46–54, 79–87, 95–102, 105–120, 130–137, 138–154, 177–196, 197–204, 213–221, 232–237, and 239–246. Since there are four more helices than there are β-strands in the overall structure, four of the β-strands are followed by a pair of helices. For example, β-strand

[3] In order to conform to the numbering of triosephosphate isomerase from chicken muscle, the first residue in the yeast enzyme is designated Ala 2.

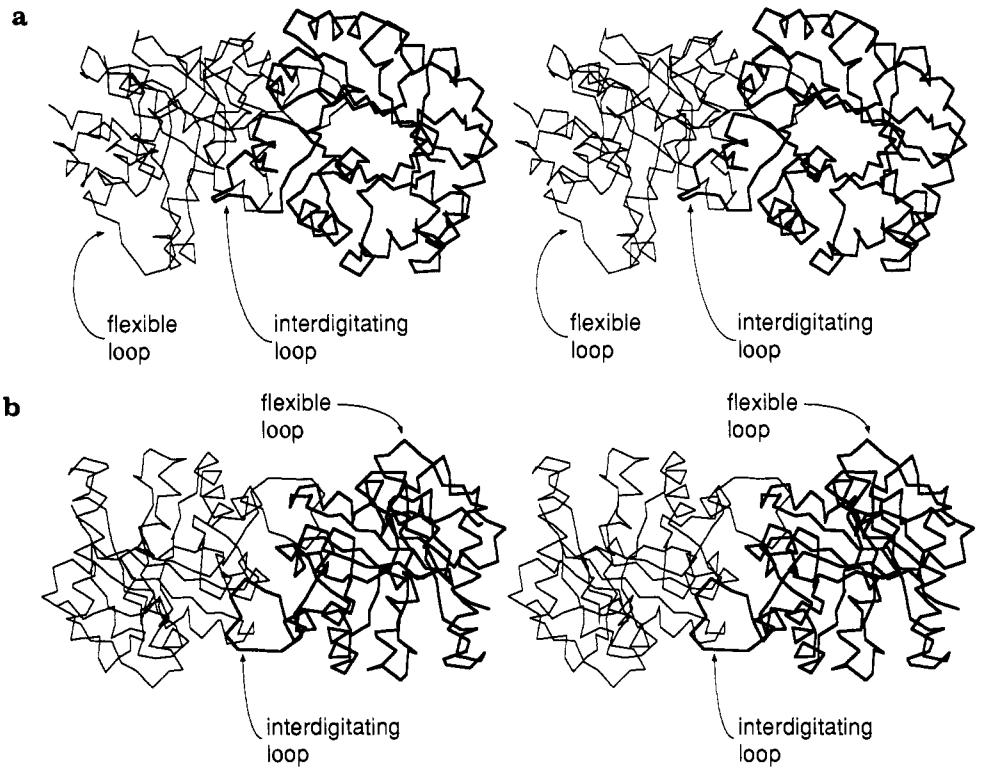

FIGURE 3: (a) Stereoscopic view of the Cα backbone of the dimer of the prototypical α/β barrel protein, triosephosphate isomerase. The subunit on the right is viewed down the barrel axis. (b) The same dimer is viewed along the 2-fold axis in the middle of the molecule, revealing the symmetry between the two subunits.

d is followed by helices D_1 and D_2, representing helices 95–102 and 105–120, respectively. All the paired helices appear very close together in sequence (95–102/105–120, 130–137/138–154, 177–196/197–204, 232–237/239–246). The first helix of these paired helices usually ends as a 3_{10} helix that leads into a very tight turn to begin the next helix. Helices D_1 and D_2 have three residues between them, and for this reason a 3_{10} helix at the end of helix D_1 is not necessary to form a tight turn. The hydrogen-bonding pattern of helices G (213–221) and H_1 (232–237) allow them to be classified entirely as 3_{10} helices.

Nine of the helices, A (16–30), B (46–54), C (79–87), D_2 (105–120), E_2 (138–154), F_1 (177–196), G (213–221), H_1 (232–237), and H_2 (239–246), pack around the cylindrical surface of the barrel, with their main-chain boundaries generally not exceeding the two ends of the barrel. Two helices, D_1 (95–102) and E_1 (130–137), are found at the carboxy-terminal end of the barrel (where the active site is located). The amino-terminal end of helix 95–102 is directed toward the active site of TIM, and it has been speculated that the helix macrodipole may provide electrostatic binding energy for the enzyme–substrate complex (Hol et al., 1978) or electrostatic stablization of the developing enediol(ate) transition state. The presence of helices abutting the carboxy-terminal end of the barrel also explains why polar residues are not found at this end of the β-sheet: the carboxy end of the barrel forms a cavity containing the active site, with α-helix (and loop) residues actually being at the protein–water interface. The remaining helix, F_2 (197–204), is present beyond the amino-terminal end of the barrel.

An analysis of the remaining amino acids revealed the various β-turns that occur in this structure. On the basis of the dihedral angle parameters defined previously (Chou & Fasman, 1977), ideal type 1 β-turns are found at Thr 45 and Ala 224; an ideal type 2 turn is found only at Thr 75; ideal type 3 turns are found at Ala 44, Pro 57, Val 167, Trp 168, and Pro 238. The remaining amino acids are either large loops or random coils that defy further categorization. As one might expect, there are a number of interactions between the side chains of the barrel and the helices that surround it, and in fact, most of these interactions are hydrophobic. There are also numerous side-chain interactions between the barrel and loop and turn residues. This observation should be of importance to those who have speculated that an archetypal α/β barrel protein might be used as a scaffold to design new enzymes by mutating the residues on the loops.

Table IV: Orientation Matrix between the Two Subunits in Yeast TIM

rotation matrix	−1.00000	−0.00168	−0.00198
	−0.00251	0.43452	0.90066
	−0.00066	0.90066	−0.43453
translation vector	120.45	12.84	20.83

Comparison of Subunits and Estimation of Errors. TIM is a very symmetrical dimer (Figure 3). The Cα's from the two subunits superimpose with an rms difference of 0.39 Å from the orientation matrix shown in Table IV. The most significant difference between the two subunits when they are superimposed occurs at the amino end of the chain. The Cα positions for residue 2 differ by 3.2 Å in the superimposed structure. This is not very surprising since the terminal ends of proteins are usually solvent exposed and quite flexible. If residue 2 is eliminated for the purpose of the comparison, the rms difference for the two subunits decreases to 0.34 Å. Although the two subunits have an identical sequence and therefore have a similar overall structure, part of the rms difference between the two subunits must be due to crystal field effects that each subunit experiences. Indeed, a Luzzati plot (not shown) indicates a coordinate error of about 0.25 Å (Luzzati, 1952). A comparison of the two active sites (all atoms in residues Asn 10, Lys 12, His 95, Ser 96, Glu 97, Cys 126, and Glu 165) gives an rms deviation of 0.28 Å. We

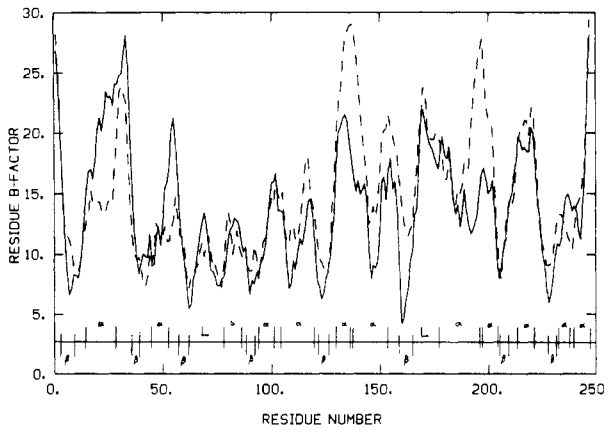

FIGURE 4: Plot of temperature factors versus residue number. The temperature factors are averaged over backbone atoms for each residue. The solid line indicates the first subunit, and the dashed line indicates the second subunit. The secondary structure of TIM is also represented above the sequence. An α between two lines indicates an α-helix, and a β represents a β-strand. Where neither an α or β appears, the structure is either a turn or loop. L represents the interdigitating loop (residues 71–77) and the flexible loop that closes over the active site (residues 167–176).

conclude from these comparisons that the rms error in atomic positions in this structure is on the order of 0.3 Å.

An examination of the *B* factors of the two subunits indicates a similar pattern. The differences consist of higher *B* values for subunit 2 in the amino acid regions of 125–148 and 186–206 (Figure 4). High *B* values have been correlated with incorrectly modeled structures (Chambers & Stroud, 1979), and these *B* values might be an indication that this part of the structure is suspect. However, the quality of the electron density map argues against such an interpretation. An alternative explanation is that the differences in the *B* factors are due to crystal field effects. Indeed, the region of 125–148 in the first subunit makes crystal contacts that are absent in the second. The differences in the remaining region are a little more difficult to assess. This region in both subunits makes significant contacts with symmetry-related molecules, and the reason for the different *B* factors is not well understood.

The general features of the *B*-factor plot are as would be expected, with the regions of high *B* factor being found on the surface of the protein. The loop that undergoes a conformational change upon substrate binding [see following paper in this issue (Lolis & Petsko, 1990)], 167–176, is also a region of high *B* factor, but these values are on average not any greater than those for other loops in the protein.

The Active Site. The active site of TIM, which is shown in Figure 1, is composed of the catalytic base Glu 165, which has been identified by the binding of covalent active-site-directed inhibitors (Miller & Waley, 1971; Hartman, 1971; De la Mare et al., 1972), and the electrophilic residues Lys 12 and His 95, which are thought to polarize the carbonyl oxygens of the substrate (Alber et al., 1987).

The carboxylic acid oxygens of Glu 165 form hydrogen bonds with the amide nitrogen and side-chain hydroxyl group of Ser 96, as well as the Nϵ2 atom of His 95. Comparison of the conformation of Glu 165 to that of yeast TIM with inhibitors bound to the active site (R. C. Davenport, B. Seaton, G. A. Petsko, and D. Ringe, unpublished results; Lolis & Petsko, 1990) reveals that the side chain moves over 2 Å upon binding. The obvious implication of this observation is that serious errors could arise when model building is used to dock substrates or inhibitors in the active site of proteins. It will be difficult to predict what minor rearrangements are necessary so that the substrate is accommodated by the enzyme in its catalytically active conformation. It is becoming increasingly clear that the structures of unbound enzymes do not necessarily exist in the active conformations and that subtle (and in some cases, major) conformational changes are required for catalysis to occur.

The charged state of the two electrophiles, Lys 12 and His 95, has been the subject of great interest. Lys 12, which has been implicated as the amino acid responsible for the specificity of the enzyme for the natural substrates over their phosphonate analogues (Belasco et al., 1978), is involved in a hydrogen bond with the carboxylate of Glu 97. Due to the proximity of Lys 12 to Glu 97 (Figure 1), and the usual pK_a of a Lys amino group, it has generally been agreed that Lys 12 is charged. Glu 97 is also involved in a hydrogen bond with the backbone amide nitrogen of Thr 75 of the *other* subunit. This network of hydrogen bonds involving Lys 12 may explain why only the dimer of TIM is catalytically competent: each subunit provides a "wall" upon which the other active site rests.

A consensus on the charged state of His 95 has been more elusive. The pK_a value of a His side chain free in solution is about 6.3. However, the imidazole of His 95 lies at the amino-terminal end of the helical segment 95–102. Fersht has obtained evidence that the α-helix dipole moment increases the pK_a of an imidazole group at the carboxy-terminal end of the helix (Sali et al., 1988). It should therefore follow that the effective positive charge at the amino-terminal end due to the macrodipole moment of helix 95–102 should decrease the pK_a of His 95. Although hydrogen positions cannot be seen in electron density maps, in many cases it is rather straightforward to infer hydrogen positions based on stereochemical principles. For example, hydrogens on backbone amide nitrogens can be assigned with a high degree of confidence due to the planar requirements of a peptide bond and to knowledge of the typical nitrogen–hydrogen bond distance. We have deduced the charged state of the imidazole ring using these stereochemical considerations along with an analysis of potential hydrogen-bond interactions of the imidazole ring. The imidazole Nϵ2 of His 95 is within hydrogen-bond distance of a solvent oxygen (3.0 Å)[4] and a carboxylate oxygen of Glu 165 (3.3 Å). Since the carboxylate oxygen is unprotonated and both the water and glutamate side chain can function as hydrogen-bond acceptors, we conclude that the Nϵ2 of His 95 is the hydrogen-bond donor and is therefore protonated. The Nδ1 atom is within hydrogen-bond distance (3.0 Å) of the backbone amide nitrogen of Glu 97. However, the requirements for a hydrogen bond between these two atoms are more stringent than for other pairs of polar atoms. All the requirements are satisfied. The bond angle (C_{96}–N_{97}–Nδ1) is 130°, close to the necessary 120°. In addition, the Nδ1 atom is close to the plane of the peptide bond formed by residues 96 and 97. In a hydrogen-bond interaction between His 95 Nδ1 and the backbone amide nitrogen of Glu 97, stereochemical considerations dictate that the amide nitrogen must be the proton donor and the Nδ1 the acceptor. The Nδ1 atom is therefore unprotonated, and the imidazole is in an uncharged state.

Unbound active sites usually refer to enzymes that are not complexed with substrate, inhibitors, or products. But even in the unbound state, it is generally believed that water occupies the active site. Most of the solvent molecules in the TIM active site must be rapidly exchanging with bulk solvent, since not many are observed crystallographically. However,

[4] All distances given are the average of values in both independent subunits in the crystallographic asymmetric unit.

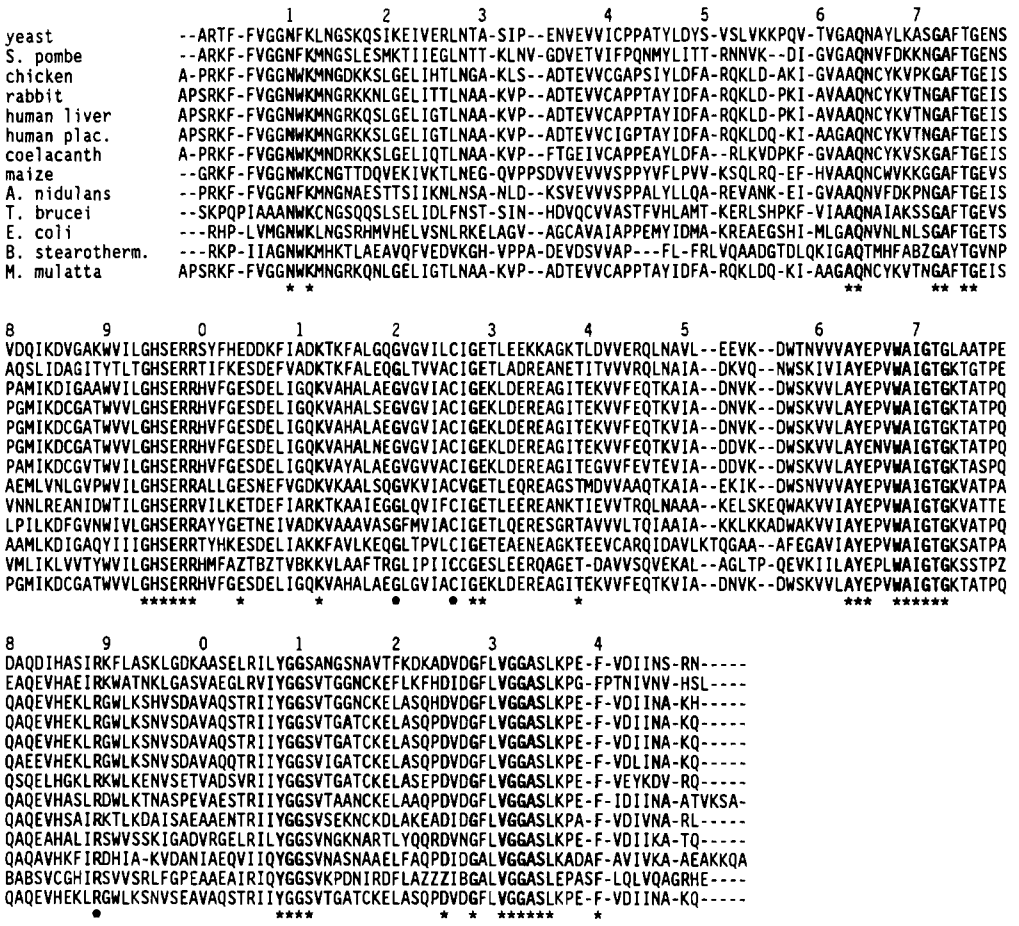

FIGURE 5: Amino acid alignment of sequences from 13 sources: yeast (Alber & Kawasaki, 1982), *Schizosaccharomyces pombe* (Russell, 1985), chicken (Straus & Gilbert, 1985), rabbit muscle (Corran & Waley, 1975), human liver (Maquat et al., 1985), human placenta (Lu et al., 1984), coelacanth (Kolb et al., 1974), maize (Marchionni & Gilbert, 1986), *Aspergillus nidulans* (McKnight et al., 1986), *Trypanosoma brucei* (Swinkels et al., 1986), *Escherichia coli* (Pichersky et al., 1984), *Bacillus stearothermophilus* (Artavanis-Tsakonis & Harris, 1980), and *Macaca mulatta* (Old & Mohrenweiser, 1988). The numbering scheme used is that for yeast TIM, with its first residue beginning at Ala 2. The strictly conserved positions are boldfaced and designated by (*).

there is one ordered water molecule that is present in both subunits. This water molecule (residue 602 in subunit 1 and 604 in subunit 2) is held in place by hydrogen bonds with Nϵ2 (3.0 Å) of His 95 and Nδ2 (3.3 Å) of Asn 10. This water molecule is displaced when substrate or inhibitor binds to the active site (Lolis & Petsko, 1990).

Much has been written in the past of the possible catalytic role of a loop near the active site of TIM (Alber et al., 1981a) that has no electron density and has thus been termed flexible. In a 6.0-Å crystal structure analysis, it has been shown that this loop (residues 167–176) undergoes a conformational change when DHAP is introduced into chicken TIM crystals (Phillips et al., 1977). This movement was confirmed at 3.5-Å resolution in yeast TIM (Alber et al., 1981a). In this refined native crystal structure, the loop seems to be stabilized by two hydrogen bonds with the same loop in a symmetry-related molecule. Therefore, a flexible loop in solution cannot be ruled out. A plot of *B* factor vs residue number (Figure 4) does indicate that this region has a high average *B* factor. Structural comparison of the bound and unbound form of the enzyme (see following paper in this issue) confirms that the loop moves significantly upon binding of substrate (Davenport et al., unpublished results; Lolis & Petsko, 1990).

The Subunit Interface. The major contacts between the two subunits of the TIM dimer are provided by an interdigitating loop, residues 71–77, that extends from one subunit into a pocket near the active site of the other subunit. Examination of the subunit interface indicates that this part of the structure is, in general, more hydrophobic than other surface areas of the monomer structure but that it is not as hydrophobic as the globular interior of the structure (Campbell, 1988). It further seems that a large part of the specificity of one subunit for another is derived from the polar interactions at the interface. This conclusion is underscored by the fact that there are about 15 ordered water molecules at the subunit interface and that 10 of these mediate interactions between the two subunits. It has been hypothesized that these buried water molecules are remnants of a stage of the folding process where the two subunits combine to make an active enzyme (R. C. Davenport, unpublished results).

With the exception of the interdigitating loop, there is very little sequence conservation at the interface, suggesting that monomers from different organisms would not form active heterodimers. This observation can be used to design effective and specific inhibitors of TIM. Since the active site residues of TIM are conserved among different organisms (vide infra), it is unlikely that a species-specific active site inhibitor will be found. However, since TIM is active only as a dimer, and there is a greater diversity at the subunit interface among the various species, it may be possible to design a peptide or other organic molecule to bind to the subunit interface of a TIM from a pathogenic organism and prevent dimerization. The molecule would have to bind to the interface without providing the necessary "wall" to generate a proper active site.

The equilibrium constant between the two monomers and dimer, to our knowledge, has never been determined but is

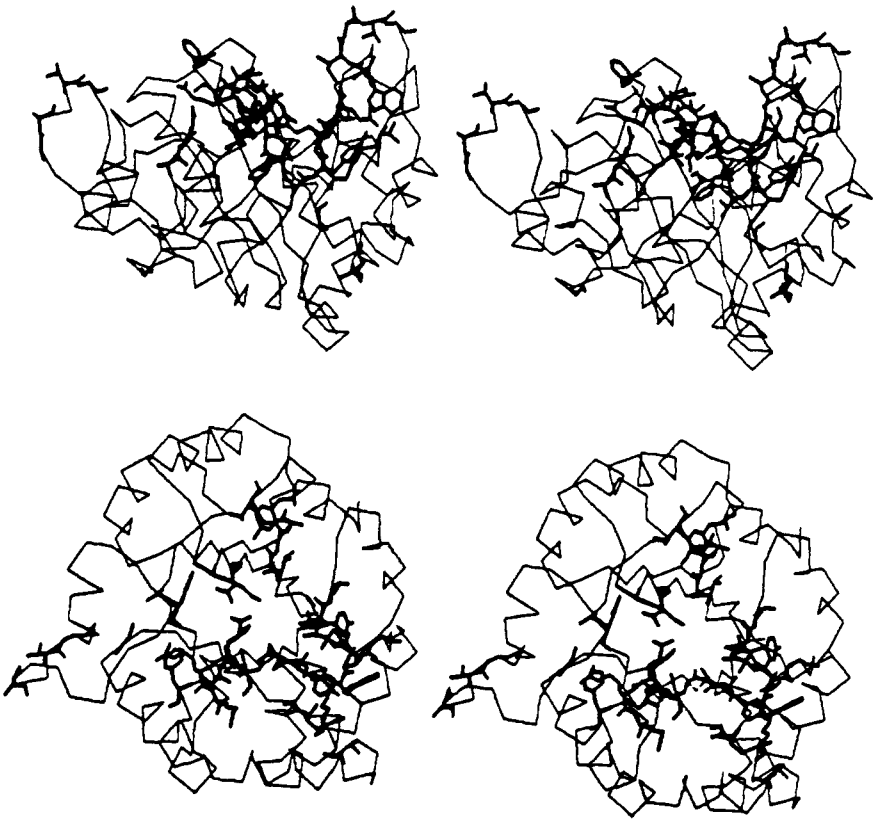

FIGURE 6: Stereoscopic view of two orientations of the Cα's of TIM with the highlighted side chains of strictly conserved residues. The bottom view was generated by rotating the top view 90° along the horizontal axis in the plane of the paper.

believed to be quite high. The crystal structure allows this number to be estimated on the basis of buried surface area. Molecular surface calculations indicate that 575 Å^2/monomer is buried on going from monomer to dimer. Campbell estimated a binding constant of 10^{14} M^{-1} based on the fact that the interface area is 2.5 times greater than that of the antibody Fc fragment and fragment B of protein A, whose binding constant is 10^6 M^{-1} (Campbell, 1988). By use of Richards's value of 20 cal/(Å^2·mol) for the free energy of transferring an amino acid from an aqueous solvent to a less polar environment (Richards, 1977), the binding constant can be estimated to be 10^{16} M^{-1}. If such a magnitude for the binding constant is confirmed experimentally, it would complicate the drug design process suggested above.

Correlation of Sequence and Structure. It is of interest to compare this fungal TIM with the same enzyme from a higher eukaryote. We used the refined coordinates of chicken TIM furnished to us by Peter Artemiuk (personal communication). This structure has been solved to 2.5-Å resolution (Banner et al., 1975) and refined to an *R* value of 17%. The enzymes from both chicken and yeast contain 247 amino acids, with a sequence identity of 53%. For the purpose of comparison, we aligned residues 4–56 and 57–248 from the chicken TIM with residues 3–55 and 57–248 from the yeast enzyme. The staggering of residues at the amino-terminal end is necessary because the two sequences are out of register for the first 56 amino acids due to an insertion in the chicken sequence. An insertion at position 56 of the yeast sequence places the sequences back in register. The rms difference in coordinate position after superposition for 490 of the possible 494 Cα's is 1.0 Å, indicating very close agreement. We also compared all atoms for all the sequence positions that are strictly conserved between chicken and yeast TIM. The rms difference for the 130 residues (969 atoms) is 1.1 Å.

The sequences of triosephosphate isomerase from 13 sources are presently known (Figure 5). Identities among these sequences range from 98% for human and rabbit to 36% for yeast and *Bacillus stearothermophilus*. We have highlighted on the three-dimensional structure of the enzyme all the amino acids that are strictly conserved among the 13 sequences (Figure 6). As is apparent from Figure 6, the conserved residues are clustered around the carboxy terminus of the barrel, where the active site is located. Of the 44 conserved residues, 28 have atoms within 12.0 Å of WAT 602, which occupies the substrate position in the active site of subunit 1. Eleven conserved residues have atoms that are within 12.0–19.0 Å of the active site. Five of these eleven are part of the flexible loop that closes over the active site when substrate binds, and another two (Glu 129 and Thr 139) form part of a hydrogen-bonding network that may be responsible for keeping the loop open when the enzyme is unbound. The four remaining residues that are within 12.0 and 19.0 Å of the active site form ionic interactions in hydrophobic environments of the enzyme. The side chains of Glu 104 and Lys 112 form a salt bridge that if disrupted by replacement of Glu 104 by an aspartic acid leads to a thermolabile enzyme with much reduced activity (Daar et al., 1986). The side chains of Arg 189 and Glu 225 are also involved in a network that stabilizes ionic interactions in a hydrophobic environment. Of the 44 conserved residues, only five do *not* have any atoms within 19 Å of WAT 602. Four of these (residue 72, 73, 75, and 76) are part of the loop that interdigitates with the other subunit and as a result are within 12.0 Å of the water molecule (WAT 604) occupying the active site in the *other* subunit. The remaining residue, Gly 120, is found at the surface of the protein where it terminates a helix and begins a tight turn that leads into a β-strand. Strict conservation of this residue may be important for structural reasons.

From the structural analysis of the sequence variability, the pressure to preserve the active site is clear. In order to

quantitate the evolutionary pressure to maintain the active site, we determined the rms difference between the chicken and yeast TIM atomic coordinates for only the residues that are strictly conserved among all 13 sources. The rms difference for the 44 residues (319 atoms) is 0.9 Å. If the "flexible loop" residues (168–173) are not included in the calculation, the rms difference decreases to 0.6 Å. These numbers can be considered a composite of real differences based on residues that are not conserved but influence the conformation of conserved residues, differences in crystal contacts, and errors in the actual positions of the atoms in the structure. A better indication of how well the structure of the active site is conserved is the rms difference between what are generally acknowledged as active site residues (all atoms in residues Ans 10, Lys 12, His 95, Ser 96, Glu 97, Cys 126, and Glu 165). It is less than 0.4 Å for each subunit and 0.5 Å for the dimer; this number is half the overall difference of the superimposed Cα's in the chicken and yeast TIM structures.

Most of the conserved residues do not make contact with substrate when it is bound in the active site. The structure of the enzyme complexed with phosphoglycolate (Lolis & Petsko, 1990) indicates only 11 of the 44 are within 3.7 Å of the inhibitor. Most of the conserved positions form a layer around the primary active site residues (those that do make contact with substrate) and are presumably important in maintaining a precise active site structure. The analysis also indicates that there are very few positions where a specific amino acid is required to maintain the structural integrity of the α/β barrel. Only five of the conserved residues, Glu 104, Lys 112, Gly 120, Arg 189, and Asp 225, are implicated in maintaining the α/β barrel structure. All of this is consistent with what has been observed for other α/β barrel proteins. Although there are now structures for 17 different proteins with this topology, there is very little sequence homology among any of them. It is clear that the analysis of strictly conserved residues is not sufficient to decipher the sequence requirements for α/β barrels.

Conclusions

We have reported the structure of triosephosphate isomerase at 1.9-Å resolution and have described its active site and subunit interface. The unbound enzyme is the starting point in the reaction coordinate profile, and its structure is indispensable in determining what changes occur during the course of catalysis and how those changes facilitate the large rate accelerations that are observed. In the following paper (Lolis & Petsko, 1990), some of the interactions (between substrate and transition-state analogue) responsible for catalysis are described.

In addition, since triosephosphate isomerase is essential to all organisms, it is an attractive target for drug design (Wierenga et al., 1987). In light of the strong active site sequence identity among various organisms, it seems unlikely that an active-site-directed inhibitor which can discriminate among the enzymes from various species will be found. However, the variability of sequences at the subunit interface suggests it might be an alternative target for drug design on this enzyme. It remains to be determined whether these differences can be exploited to develop an effective drug.

Acknowledgments

We gratefully acknowledge fruitful discussions with Steven C. Almo, Robert L. Campbell, Wim Hol, Diane Joseph, Jeremy Knowles, Elliot Nickbarg, Sir David C. Phillips, and Dagmar Ringe. We are also grateful to Dan Leahy, Steve Hubbard, John Dewan, and Brian Guenther for assistance in producing some of the figures.

Registry No. TIM, 9023-78-3.

References

Ahern, T. J., Casal, J. I., Petsko, G. A., & Klibanov, A. M. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 675–679.

Alber, T., & Kawasaki, G. (1982) *J. Mol. Appl. Genet. 1*, 419–434.

Alber, T., Banner, D. W., Bloomer, A. C., Petsko, G. A., Phillips, D. C., Rivers, P. S., & Wilson, I. A. (1981a) *Philos. Trans. R. Soc. London, B 293*, 159–171.

Alber, T., Hartman, F. C., Johnson, R. M., Petsko, G. A., & Tsernoglou, D. (1981b) *J. Biol. Chem. 256*, 1356–1361.

Alber, T. C., Davenport, R. C., Giammona, D. A., Lolis, E., Petsko, G. A., & Ringe, D. (1987) *Cold Spring Harbor Symp. Quant. Biol. 52*, 603–613.

Albery, W. J., & Knowles, J. R. (1976) *Biochemistry 15*, 5627–5640.

Albery, W. J., & Knowles, J. R. (1977) *Angew. Chem., Int. Ed. Engl. 16*, 285–293.

Arnone, A., Bier, C. J., Cotton, F. A., Day, V. W., Hazen, E. E., Richardson, D. C., Richardson, J. S., & Yonath, A. (1971) *J. Biol. Chem. 246*, 2302.

Artavanis-Tsakonis, S., & Harris, J. J. (1980) *Eur. J. Biochem. 108*, 599–611.

Banner, D. W., Bloomer, A. C., Petsko, G. A., Phillips, D. C., Pogson, C. I., & Wilson, I. A. (1975) *Nature 255*, 609.

Belasco, J. G., Herlihy, J. M., & Knowles, J. R. (1978) *Biochemistry 17*, 2971–2978.

Blacklow, S. C., Raines, R. T., Lim, W. A., Zamore, P. D., & Knowles, J. R. (1988) *Biochemistry 27*, 1158–1167.

Blow, D. M., & Crick, F. H. C. (1959) *Acta Crystallogr. 12*, 794–803.

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

Buisson, G., Duce, E., Haser, R., & Payan, F. (1987) *EMBO J. 6*, 3909–3916.

Campbell, R. L. (1988) Ph.D. Thesis, Massachusetts Institute of Technology.

Casal, J. I., Ahern, T. J., Davenport, R. C., Petsko, G. A., & Klibanov, A. M. (1987) *Biochemistry 26*, 1258–1264.

Chambers, J. L., & Stroud, R. M. (1979) *Acta Crystallogr. B35*, 1861–1874.

Chapman, M. S., Suh, S. W., Curmi, P. M. G., Cascio, D., Smith, W. W., & Eisenberg, D. S. (1988) *Science 241*, 71–74.

Chou, P. Y., & Fasman, G. D. (1977) *J. Mol. Biol. 115*, 135–175.

Corran, P. H., & Waley, S. G. (1975) *Biochem. J. 137*, 335–344.

Daar, I. O., Artymiuk, P. J., Phillips, D. C., & Maqaut, L. E. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 7903–7907.

De la Mare, S., Coulson, A. F. W., Knowles, J. R., Priddel, J. D., & Offord, R. E. (1972) *Biochem. J. 129*, 321–331.

Dodson, E. J., & Vijayan, M. (1971) *Acta Crystallogr. B27*, 2402–2411.

Farber, G. K., Petsko, G. A., & Ringe, D. (1987) *Protein Eng. 1*, 459–466.

Fraser, R. D. B., MacRae, T. P., & Suzuki, E. (1978) *J. Appl. Crystallogr. 11*, 693–694.

Goldman, A., Ollis, D. L., & Steitz, T. A. (1987) *J. Mol. Biol. 194*, 143–153.

Hartman, F. C. (1971) *Biochemistry 10*, 146–154.

Hendrickson, W. A. (1985) in *Methods in Enzymology* Wyckoff, H. W., Hirs, C. H. W., & Timasheff, S. N., Eds.) Vol. 115, pp 252–270, Academic Press, New York.
Henrick, K., Blow, D. M., Carrell, H. L., & Glusker, J. P. (1987) *Protein Eng. 1*, 467–469.
Hill, E., Tsernoglou, D., Webb, L., & Banazak, L. J. (1972) *J. Mol. Biol. 72*, 577–591.
Hol, W. J. G., Van Duijnen, P. T., & Berendsen, H. J. (1978) *Nature 273*, 443.
Hyde, C. C., Padlan, E. A., Ahmed, S. A., Miles, E. W., & Davie, D. R. (1987) *Fed. Proc., Fed. Am. Soc. Exp. Biol. 46*, 2215.
Jones, T. A., & Thirup, S. (1986) *EMBO J. 5*, 819–822.
Knowles, J. R., & Albery, W. J. (1977) *Acc. Chem. Res. 10*, 105–111.
Kolb, E., Harris, J. J., & Bridger, J. (1974) *Biochem. J. 137*, 185–197.
Lasters, I., Wodak, S. J., Alard, P., & Van Cutsem, E. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 3338–3342.
Lebioda, L., & Stec, B. (1988) *Nature 333*, 683–686.
Lim, L. W., Shamala, N., Mathews, F. S., Steenkamp, D. J., Hamlin, R., & Xuong, N. H. (1986) *J. Biol. Chem. 261*, 15140–15146.
Lindqvist, Y., & Branden, C.-I. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 6855–6859.
Lolis, E., & Petsko, G. A. (1990) *Biochemistry* (following paper in this issue).
Lu, H. S., Yuan, P. M., & Gracy, R. W. (1984) *J. Biol. Chem. 259*, 11958–11968.
Luzzati, V. (1952) *Acta Crystallogr. 5*, 802–810.
Maquat, L. E., Chilote, R., & Ryan, P. M. (1985) *J. Biol. Chem. 260*, 3748–3753.
Marchionni, M., & Gilbert, W. (1986) *Cell 46*, 133–141.
Matsuua, Y., Kusunoki, M., Harada, W., & Kakudo, M. (1984) *J. Biochem. 95*, 697–702.
Matthews, B. W. (1966a) *Acta Crystallogr. A20*, 82–86.
Matthews, B. W. (1966b) *Acta Crystallogr. A20*, 230–239.
Matthews, B. W., & Czerwinski, G. W. (1975) *Acta Crystallogr. A31*, 480–487.
Mavridis, J. M., Hatada, M. H., Tulinskky, A., & Lebioda, L. (1982) *J. Mol. Biol. 162*, 419–444.
McKnight, G. L., O'Hara, P. J., & Parker, M. L. (1986) *Cell 46*, 143–147.
Miller, J. C., & Waley, S. G. (1971) *Biochem. J. 123*, 163–170.
Nickbarg, E. B., & Knowles, J. R. (1988) *Biochemistry 27*, 5939–5947.
Nickbarg, E. B., Davenport, R. C., Petsko, G. A., & Knowles, J. R. (1988) *Biochemistry 27*, 5948–5960.
North, A. C. T., Phillips, D. C., & Mathew, F. S. (1968) *Acta Crystallogr. A24*, 351–359.
Old, S. E., & Mohrenweiser, H. W. (1988) *Nucleic Acids Res. 16*, 9055.
Phillips, D. C., Sternberg, M. J. E., Thornton, J. M., & Wilson, I. A. (1977) *Biochem. Soc. Trans. 12*,642–647.
Pichersky, E., Gottlieb, L. D., & Hess, J. P. (1984) *Mol. Gen. Genet. 195*, 314–320.
Priestle, J. P., Grütter, M. G., White, J. L., Vincent, M. G., Kania, M., Wilson, E., Jardetzky, T. S., Kirschner, K., & Jansonius, J. N. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 5690–5694.
Rao, S. T., Hogle, J., & Sundaralingam, M. (1983) *Acta Crystallogr. C39*, 237–240.
Richard, J. P. (1984) *J. Am. Chem. Soc. 106*, 4926.
Richards, F. M. (1977) *Annu. Rev. Biophys. Bioeng. 6*, 151–176.
Russell, P. R. (1985) *Gene 40*, 125–130.
Sali, D., Bycroft, M., & Fersht, A. R. (1988) *Nature 335*, 740–743.
Schneider, G., Lindqvist, Y., Branden, C.-I., & Lorimer, G. (1986) *EMBO J. 5*, 3409–3415.
Straus, D., & Gilbert, W. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 2014–2018.
Straus, D., Raines, R., Kawashima, E., Knowles, J. R., & Gilbert, W. (1985) *Proc. Natl. Acad. Soc. U.S.A. 82*, 2272–2276.
Stuart, D. I., Levine, M., Muirhead, H., & Stammers, D. K. (1979) *J. Mol. Biol. 134*, 109–142.
Swinkles, B. W., Gibson, W. C., Osingas, K. A., Kramer, R., Veeneman, G. H., Van Boon, J. H., & Berst, P. (1986) *EMBO J. 5*, 1291–1298.
Sygusch, J., Beaudry, D., & Allaire, M. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 7846–7850.
Valentine, W. N., Tanaka, K. R., & Paglia, D. E. (1983) in *The Metabolic Basis of Inherited Disease* (Stanbury, J. B., Wyngaarden, J. B., Frederickson, D. S., Goldstein, J. L., & Brown, M. S., Eds.) pp 1606–1682, McGraw-Hill Book Co., New York.
Wierenga, R. K., Kalk, K. H., & Hol, W. J. G. (1987) *J. Mol. Biol. 198*, 109–121.
Wolfenden, R. (1969) *Nature 223*, 704.
Wyckoff, H. W., Doscher, M., Tsernoglou, D., Inagami, Y., Johnson, L. N., Hardman, K. D., Allewell, N. M., Kelly, D. M., & Richards, F. M., (1967) *J. Mol. Biol. 27*, 563–578.
Xia, Z.-X., Shamala, N., Bethge, P. H., Lim, L. W., Bellamy, H. D., Xuong, N. H., Lederer, F., & Mathews, F. S. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 2629–2633.