Crystal structure of recombinant chicken triosephosphate isomerase-phosphoglycolohydroxamate complex at 1.8-A resolution.

ARTICLE *in* BIOCHEMISTRY · APRIL 1994

Impact Factor: 3.02 · Source: PubMed

7 AUTHORS, INCLUDING:

Gregory A. Petsko
Brandeis University
485 PUBLICATIONS **25,543** CITATIONS

SEE PROFILE

Dagmar Ringe
Brandeis University
270 PUBLICATIONS **13,412** CITATIONS

SEE PROFILE

Crystal Structure of Recombinant Chicken Triosephosphate Isomerase-Phosphoglycolohydroxamate Complex at 1.8-.ANG. Resolution

Zhidong Zhang, Shigetoshi Sugio, Elizabeth A. Komives, Kathleen D. Liu, Jeremy R. Knowles, Gregory A. Petsko, and Dagmar Ringe

Biochemistry, **1994**, 33 (10), 2830-2837• DOI: 10.1021/bi00176a012 • Publication Date (Web): 01 May 2002

Crystal Structure of Recombinant Chicken Triosephosphate Isomerase–Phosphoglycolohydroxamate Complex at 1.8-Å Resolution[†,‡]

Zhidong Zhang,[▽] Shigetoshi Sugio,[▽,§] Elizabeth A. Komives,[∥] Kathleen D. Liu,[⊥] Jeremy R. Knowles,[⊥] Gregory A. Petsko,[▽] and Dagmar Ringe*[,▽]

Departments of Biochemistry and Chemistry, Rosenstiel Basic Medical Sciences Research Center, Brandeis University, Waltham, Massachusetts 02254-9110, Department of Chemistry, University of California at San Diego, La Jolla, California 92092, and Department of Chemistry, Harvard University, Cambridge, Massachusetts 02138

Received October 11, 1993; Revised Manuscript Received December 13, 1993[⊗]

ABSTRACT: The crystal structure of recombinant chicken triosephosphate isomerase (TIM, E.C. 5.3.1.1) complexed with the intermediate analogue phosphoglycolohydroxamate (PGH) has been solved by the method of molecular replacement and refined to an *R*-factor of 18.5% at 1.8-Å resolution. The structure is essentially identical to that of the yeast TIM–PGH complex [Davenport, R. C., et al. (1991) *Biochemistry 30,* 5821–5826] determined earlier and refined at comparable resolution. This identity extends to the high-energy conformations of the active-site residues Lys13 and Ser211, as well as the positions of several bound water molecules that are retained in the active site when PGH is bound. Comparison with the structure of uncomplexed chicken TIM shows that the catalytic base, Glu165, moves several angstroms when PGH binds. This movement may provide a trigger for a larger conformational change, one of 7 Å, in a loop near the active site, which folds down like a lid to shield the bound inhibitor and catalytic residues from contact with bulk solvent. These same conformational changes were seen in crystalline yeast TIM upon binding of PGH; their occurrence here in a different crystal form of TIM eliminates the possibility that they are an artifact of crystal packing.

Triosephosphate isomerase (TIM, EC 5.3.1.1) catalyzes the central reaction in the glycolytic pathway, the interconversion of the three-carbon sugar phosphates dihydroxyacetone phosphate (DHAP) and D-glyceraldehyde 3-phosphate (GAP) (Figure 1). TIM from all organisms is a dimer of identical subunits, each of molecular weight about 26 500. 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). Since only GAP can proceed down the glycolytic pathway, TIM is absolutely essential for efficient energy production. Indeed, people who are defective in one of the alleles for the TIM gene are found to suffer from chronic hemolytic anemia and neuromuscular disorders (Valentine et al., 1983).

TIM has been the subject of extensive biophysical studies. Viscosity-dependence measurements have established that the rate of the enzymatic reaction is limited solely by diffusion of substrate or product, either into or out of the active site, and that all of the purely chemical steps are faster (Albery & Knowles, 1976; Blacklow et al., 1988). A series of isotope-exchange experiments have led to the derivation of a free-energy diagram for the reaction (Albery & Knowles, 1976). The results from these studies indicate that TIM, which catalyzes isomerization 10^{10} times faster than a simple base of pK_a around 7, has reached evolutionary perfection as a catalyst (Albery & Knowles, 1976), in the sense that there is no selective pressure for further improvement in catalytic efficiency. Other studies have demonstrated that the enzyme suppresses the elimination of the phosphate group from the intermediates of catalysis (Pompliano & Knowles, 1990). The elimination reaction, which produces toxic methylglyoxal and inorganic phosphate, is more rapid than isomerization under nonenzymatic conditions (Richard et al., 1984, 1985; Pompliano & Knowles, 1990).

A variety of techniques have been used to probe the structural basis for the catalytic power of TIM. Isotope-exchange studies suggested that a single enzymic base is required (Rieder & Rose, 1959); affinity-labeling experiments (Coulson et al., 1970; Waley et al., 1970; Hartman, 1970) showed that the base is the carboxylate group of Glu165. Spectroscopic studies implied that an electrophilic residue facilitates the enzymatic reaction by polarizing the substrate (Belasco & Knowles, 1980; Komives et al., 1991).

The crystal structures of chicken and yeast triosephosphate isomerase (cTIM, yTIM) (Banner et al., 1975; Lolis et al., 1990), the medium-resolution structure of the cTIM–DHAP Michaelis complex (Alber et al., 1981), and the high-resolution structure of yTIM complexed with the intermediate analogue phosphoglycolohydroxamate (PGH) (Davenport et al., 1991) are all in accord with these deductions from solution studies and have allowed the active-site residues to be identified; the trypanosomal TIM crystal structure also has the same active site and a very similar overall structure (Wierenga et al., 1991). In the active site of cTIM, in addition to the catalytic base Glu165, the side chains of Asn11 and Lys13 (Asn10 and Lys12 in the yeast enzyme sequence), His95, Ser96, and Glu97 are all positioned so as to interact with the sugar portion of the substrate or with each other. Substrate binding induces a large conformational change: a flexible loop of 11 amino acids, residues 166–176, moves by about 7 Å as a rigid "lid" to fold down over the active site and interact with the phosphate group of bound substrates and inhibitors (Alber et al., 1981, 1987; Joseph et al., 1990; Pompliano et al., 1990).

[†] Supported in part by a grant from the Lucille P. Markey Charitable Trust and by a grant from the National Institutes of Health (GM-26788).

[‡] The atomic coordinates are available from the Brookhaven Protein Data Bank under code name 1TPH.

* Author to whom correspondence should be addressed.

[▽] Brandeis University.

[§] Present address: The Central Research Laboratories, The Green Cross Corporation, 2-51-1 Shodai-Otani, Hirakata, Osaka 573, Japan.

[∥] University of California at San Diego.

[⊥] Harvard University.

[⊗] Abstract published in *Advance ACS Abstracts,* February 15, 1994.

0006-2960/94/0433-2830$04.50/0

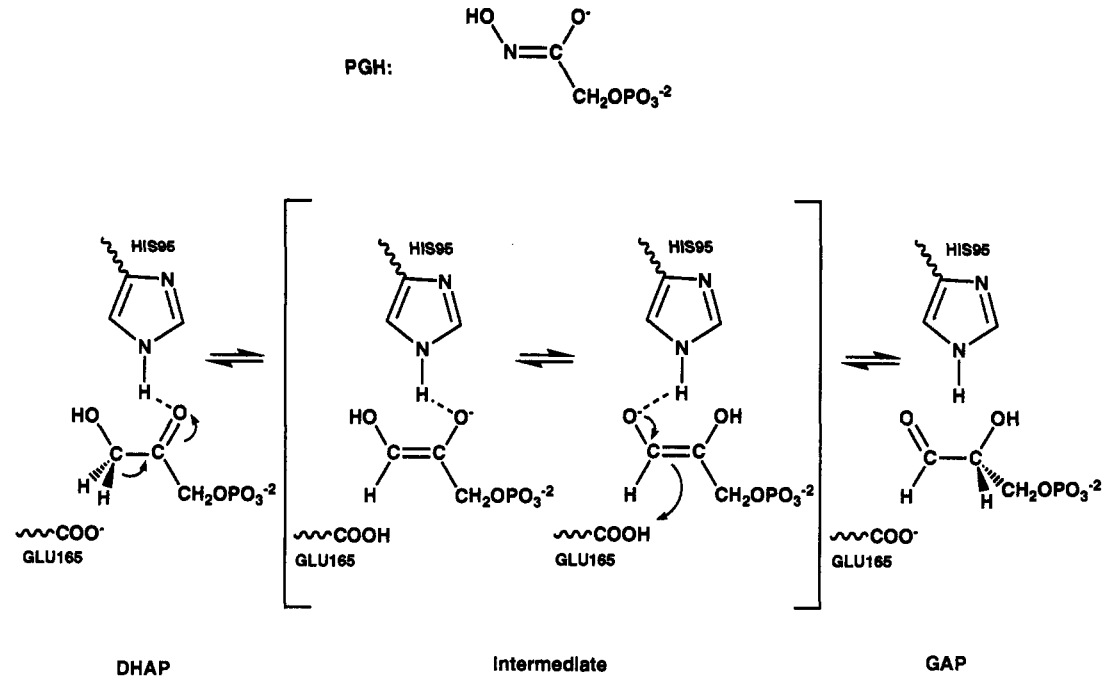

FIGURE 1: Catalytic mechanism of triosephosphate isomerase. Proton abstraction from DHAP by the catalytic base Glu165 leads to an enediol/endiolate intermediate. His95 acts as a general acid in this reaction. PGH is the inhibitor molecule of TIM.

These structural studies have led to a more detailed mechanism for TIM catalysis that is in good agreement with the earlier suggestions: Glu165 acts as a general base to abstract the *pro-R* proton from C1 of DHAP, with electrophilic catalysis by His95 and possibly also Asn11 and Lys13 serving to make the hydrogen more acidic and to stabilize the partial negative charge that forms on the carbonyl oxygen in the transition state (Figure 1). Formation of the *cis*-enediol (or enediolate) intermediate is aided by His95 also acting as a general acid to donate a proton to the substrate carbonyl oxygen. Then Glu165 transfers a proton back to C2 of the intermediate to give GAP. The role of the flexible loop in this mechanism is to stabilize the substrate in a planar conformation that favors enolization over elimination and to stabilize the charged intermediate. Closure of the loop also protects the active site from contact with bulk water (Alber et al., 1987; Pompliano et al., 1990). The latest studies have shown that loop closure is also coupled to proton transfer from the substrate (Sampson & Knowles, 1992).

The evolution of catalytic perfection in TIM has been studied by the techniques of site-directed and random mutagenesis. Knowles and co-workers altered either the active-site base (E165D) or the catalytic electrophile (H95N and H95Q) (Nickbarg et al., 1988; Blacklow et al., 1990). The major effect in each case is relative destabilization of the transition states for the two chemical (enolization) steps that constitute the catalytic reaction. When the genes encoding these sluggish mutant isomerases were subjected to random mutagenesis and a selection for isomerases of increased catalytic potency was then performed, pseudorevertant enzymes with dramatic increases in activity were found with the same second-site substitution, Ser96 to proline. The E165D,S96P double mutant isomerase is 20-fold better in terms of k_{cat}/K_m than the E165D single mutant from which it is derived, and the H95N,S96P pseudorevertant is about 60 times more active than its H96N parent. Surprisingly, such a pseudoreversion was only observed in the recombinant chicken mutant enzymes, whereas neither of the same two double mutants of recombinant yTIM gave any activity increase compared to its corresponding single mutant, even though for the chicken and yeast enzymes the active-site residues are exactly the same, the residues in the flexible loop are highly conserved, and the overall amino acid sequence identity is 53% (Lolis et al., 1990).

In order to understand the basis of pseudoreversion in the cTIM mutants and to explain the difference between the chicken and yeast enzymes at the atomic level, we undertook a serial study of all of the enzymic species involved. We began by solving the three-dimensional structure of a new crystal form of recombinant cTIM complexed with the intermediate analogue phosphoglycolohydroxamate (PGH) at 1.8-Å resolution. The structure has been refined to an *R*-factor of 18.5%. Because all of the other cTIM mutants crystallize isomorphously with this crystal, we can use the cTIM–PGH structure as the starting model for a series of structure determinations and comparisons.

In this paper we present the cTIM–PGH complex structure and its comparison with the uncomplexed cTIM structure (Banner et al., 1975). The results are consistent with the general features of the accepted mechanism and offer more accurate structural information about the high catalytic efficiency of the enzyme. Another comparison, between the chicken and yeast TIM–PGH complex structures (Davenport et al., 1991), offers hints as to the structural basis of the different behaviors of the enzymes from yeast and chicken. Furthermore, the cTIM–PGH structure discussed in this paper provides crucial information for understanding the basis for pseudoreversion when residue serine 96 is replaced by proline in the chicken enzyme.

MATERIALS AND METHODS

Protein Production and Crystallization. The gene for chicken muscle TIM was subcloned into a derivative of pBS+/− that has been described (Blacklow & Knowles, 1990). This phagemid vector allowed the efficient production of single-stranded DNA for mutagenesis. The phagemid contained, on an *Eco*RI to *Pst*I fragment, the trc promoter upstream from the complete TIM gene and allowed the production of 50–80 mg of protein per liter of cells. The expression vectors were used to transform *Escherichia coli* strain DF502, which

is a strep^R, tpi$^-$ strain that was kindly provided by D. Fraenkel and has been previously described (Straus & Gilbert, 1985).

Large amounts of pure protein were prepared by growing the bacterial transformants in a final volume of 10 L of M63 salts (Miller, 1972) containing casamino acids (0.5%, w/v), glucuronolactone (0.4%, w/v), glycerol (0.1%, w/v), $MgSO_4$ (1 mM), thiamin (1 mg/mL), L-histidine (80 mg/L), streptomycin (100 mg/L), and ampicillin (200 mg/L). Cells were harvested after 12–20 h by centrifugation at 3000*g*. The cells were lysed in a continuous-flow French pressure cell (Aminco, Urbana, IL), and the lysate was centrifuged at 8500*g* for 1 h to remove cell debris. The ammonium sulfate fraction from 55% to 90% saturation was collected and dialyzed against TE buffer (10 mM Tris-HCl, pH 7.8, and 1 mM EDTA) overnight. The following day, the crude protein was loaded onto a 300-mL column of QAE Sephadex A-120 equilibrated with TE buffer and eluted with a linear gradient of 0 to 300 mM KCl (1L to 1L). The proteins were finally purified on a MonoQ 10/10 column using the same gradient. Purity of the proteins was assessed by silver staining of overloaded 15% SDS– PAGE gels (Laemmli, 1970). Concentration of the protein was afforded by Centriprep and Centricon concentration (Amicon, Danvers, MA). The purified protein was assayed for conversion of glyceraldehyde 3-phosphate to dihydroxyacetone phosphate according to the method of Blacklow and Knowles (1990), and the k_{cat} and K_m values obtained were identical to those previously reported.

The purified protein was dialyzed into MilliQ H_2O and concentrated to 20 mg/mL (calculated by taking the absorbance at 280 nm and multiplying by the extinction coefficient of cTIM of 1.21 mg/OD unit). A 0.5 M solution of phosphoglycolohydroxamate (PGH) (Collins, 1974) in water was prepared by dissolving 2 mg of PGH in 25 mL of MilliQ H_2O. For a 1.5 mM final concentration of PGH in the crystallization drop, 115.2 μL of protein was mixed with 4.8 μL of PGH solution. The cTIM–PGH was crystallized by vapor diffusion in hanging drops in the presence of saturating amounts of phosphoglycolohydroxamate phosphate (PGH phosphate) by precipitation with a relatively wide concentration range of poly(ethylene glycol). The standard mother liquor of these crystals contains 100 mM Tris buffer (pH 8.5), 12.5%–21.0% (w/v) PEG(8K), and 200 mM lithium sulfate. The crystals grow as elongated prisms at room temperature within 1 week. The cTIM–PGH crystals have the symmetry of space group $P2_12_12_1$, with one dimer per asymmetric unit and cell dimensions of a = 136.38 Å, b = 74.00 Å, and c = 57.16 Å. The crystals were mounted in quartz capillary tubes before data collection.

Data Collection and Data Processing. A 3.2-Å resolution data set was collected from one cTIM–PGH crystal, at room temperature, on a Rigaku AFC5 diffractometer mounted on a Rigaku RU-200 X-ray generator operated at 50 kV and 100 mA. The diffractometer data set contains 10 230 reflections with a completeness of 94% (Table 1), and only those reflections with intensities greater than 3σ were used in the molecular replacement study.

Two high-resolution data sets were collected, under the same experimental conditions, on the San Diego MarkII multiwire area detector system with two multiwire proportional chambers (Cork et al., 1973; Xuong et al., 1985a). One data set with resolution to 1.95 Å was obtained at the University of California, San Diego. A total of 32 281 unique reflections out of 84 781 observed reflections were collected with one cTIM–PGH crystal from eight orientations based on the data collection strategy of Xuong (Xuong et al., 1985b). Another data set was collected on the same area detector system at the Center for Advanced Biotechnology and Medicine at Rutgers University, Piscataway, NJ. The cTIM–PGH crystal diffracted to 1.81-Å resolution, giving a total of 49 099 Bragg reflections from 143 636 recorded spots. The two data sets were processed using the UCSD data reduction package and then were scaled and merged together. The final 1.81-Å data set contains 50 153 unique reflections with 5% merging R-factor on intensity and a completeness of 96% (Table 1). These high-resolution area detector data were used to refine the initial TIM–PGH structure, calculated from the molecular replacement solution obtained from the 3.2-Å resolution diffractometer data.

Table 1: Statistics of Diffraction Data for the Chicken Triosephosphate Isomerase–PGH Complex[a]

data set	resolution range (Å)	no. of reflections observed	no. of reflections theoretical	completeness (%)
diffractometer data	21.8–3.1	10 230	10 893	94
area detector data I (UCSD)	13.0–1.95	33 262	42 990	77
area detector data II (Rutgers)	9.7–1.8	48 606	52 720	92
merged area detector data	13.0–1.8	50 153	52 935	95

[a] The 3.1-Å low-resolution diffractometer data were only used in the molecular replacement. The 13.0–1.8-Å area detector data, merged from two data sets collected at UCSD and Rutgers, respectively, were used in the refinement of the TIM–PGH model generated from the molecular replacement.

Molecular Replacement. The starting model for the molecular replacement investigation was the unrefined 2.5-Å cTIM dimer with the open flexible loop and no substrate or inhibitors bound at the active site (Banner et al., 1975). The coordinates for this model (1TIM) were obtained from the Brookhaven Protein Data Bank. The molecular replacement program package MERLOT (version 2.3; Fitzgerald, 1988), which includes the fast-rotation function of Crowther (1972), the rotation function of Lattman and Love (1972), and the translation function of Crowther and Blow (1967), was used throughout.

The search model was placed in an orthogonal reference cell with dimensions 110 × 110 × 110 Å^3 to generate the harmonic coefficients for the rotation function. The diffractometer data of cTIM–PGH used to calculate the observed Patterson map were in the 8.0–4.0-Å resolution range with all intensities greater than 3σ. Two peaks with rms values of 10.0σ and 9.4σ above the average (all other peaks had rms values less than 4.5σ) were found from the Crowther rotation function in Eulerian angle search with a 5° grid. These two peaks are related to each other by the internal pseudo-twofold symmetry in the TIM dimer. A subsequent Lattman rotation search with a finer grid spacing of 1° gave a peak at the same position in rotation space. Then a translation function with orientation angles determined by the rotation functions gave a unique solution for the molecular location in the cTIM–PGH unit cell. Finally the orientation and location of the molecular model were refined by R-factor minimization, producing a starting R-factor of 32.9% vs all data from 10–3.2-Å resolution.

Refinement and Model Building. The refinement of the crystal structure was done by restrained least-squares with the PROLSQ program (Hendrickson & Konnert, 1980; Hendrickson, 1985) on a Stardent 2000 workstation; interleaved between refinement sessions were model rebuilding sessions on an Evans & Sutherland PS330 graphics system, using the modeling program FRODO (Jones, 1978; Sack, 1988).

Table 2: Refinement Statistics of the cTIM–PGH Structure

parameters of refinement statistics	values
R-factor (%) (6.0–1.8 Å)	18.5 ($F_o/\sigma(F_o) > 1$)
R-factor (%) (10.0–1.8 Å)	18.8 ($F_o/\sigma(F_o) > 1$)
no. of reflections obsd ($F_o/\sigma(F_o) > 0$)	48 735 (6.0–1.8 Å)
no. of reflections obsd ($F_o/\sigma(F_o) > 1$)	47 198 (6.0–1.8 Å)
no. of protein atoms	3712
no. of inhibitor atoms	20
no. of water molecules	249
rms deviations from ideal geometry	
bond distance (Å)	0.017
angle distance (Å)	0.034
planar distance (Å)	0.048

In the first stage of model building, a difference Fourier electron density map with coefficients ($3F_o - 2F_c$) was calculated for the cTIM–PGH structure. At this point, only the 3.2-Å Rigaku diffractometer data with F_o greater than 2σ were used, and the phases were deduced from the molecular replacement model, with no PGH bound and the flexible loop residues (166–176) deleted. The model also contained no side chains for the active-site residues (Asn11, Lys13, His95, Ser96, and Glu165). The map showed clear electron density over the whole of the dimer model and also gave well-defined electron densities for both the inhibitor (PGH) and the closed flexible loop region. The majority of the manual modifications to the model consisted of introducing two PGH molecules into the active sites of the two subunits and building in the two 11-residue lid loops which close over the active site upon PGH binding.

PROLSQ refinement of the modified structure using the MarkII area detector data gradually extended the resolution range every few cycles, finally including all the reflections up to 1.8-Å resolution. With data below 2.5-Å resolution, an overall *B*-factor was refined. After higher resolution data were added, an individual *B*-factor for each atom was used in refinement. Following refinement, we carried out another round of manual rebuilding, based on difference Fourier maps with coefficients ($3F_o - 2F_c$), ($F_o - F_c$), and $-(F_o - F_c)$. The iterative procedure of PROLSQ refinement and model rebuilding was repeated two more times, giving an *R*-factor of 22.7%.

The final stage of refinement involved picking water molecules. A difference Fourier map with coefficient ($F_o - F_c$) was calculated, and the potential water candidates were picked on the basis of peaks at the 3σ level and higher with reasonable hydrogen-bonding partners. A checking procedure eliminated those candidates with *B*-factors higher than 45.0 Å^2. The final model of TIM complexed with PGH, which contains 249 waters, was refined against the merged area detector data between 6.0- and 1.8-Å resolution using data with $F_{obs} > 1\sigma(F_{obs})$, giving an *R*-factor of 18.5% . The *R*-factor using all data from 10.0–1.8 Å is 18.8% (Table 2).

RESULTS AND DISCUSSION

Topology of Structure. The topology of TIM belongs to a class of eight-stranded α/β barrel proteins, the largest and most regular of all domain structures (Farber & Petsko, 1990). TIM was the first enzyme in this class whose structure was elucidated, and remains the prototype of what has been called the TIM barrel folding pattern. In the TIM structure there is a core of eight twisted parallel β strands arranged into a barrel; eight amphipathic α helices interconnect the parallel β strands and are all packed against the barrel surface. Several

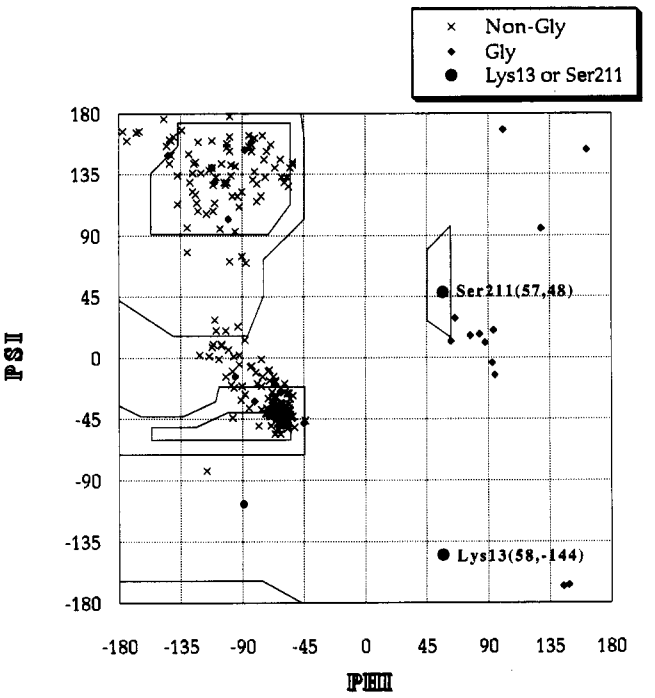

FIGURE 2: Ramachandran plot for the residues of subunit 1 of the cTIM–PGH structure. Lys13 and Ser211 are indicated.

large loops, random coils, and β turns form the conjunctions of the α helices and β strands. 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. Most of the polar amino acids are found at the ends of the barrel, where the barrel is exposed to solvent. The active site is located at the carboxyl-terminal end of the barrel for all members of this family.

Comparison of Subunits and Estimation of Errors. cTIM is a very symmetrical dimer, and this symmetry is maintained when the enzyme is complexed with PGH. The rms difference between the α-carbon positions of the two subunits is 0.37 Å when they are superimposed. Although the two subunits have an identical sequence and therefore must have a similar overall structure, part of the RMS difference between the two subunits probably is due to the different crystal-packing environments that each subunit experiences; also, because some residues around the terminal ends and on the protein surface are solvent exposed, they are quite flexible. On the basis of the assumption that the two chemically identical active sites should have exactly the same conformation, a comparison of the two active sites including atoms of residues Gly10, Asn11, Lys13, His95, Ser96, Glu97, and Glu165 and the active-site flexible loop (residues 166–176) was made, giving an rms deviation of 0.11 Å. We therefore conclude that the rms error in atomic position in this well-refined structure is of the order of 0.2 Å.

The Ramachandran plot for the residues from subunit 1 shows the expected distribution of backbone torsion angles (Figure 2). The second subunit shows little deviation from the first subunit. Only two of the non-glycine residues were found to have high positive ϕ values: the active-site residues Lys13 (58, –144) and Ser211 (57, 48). Apparently these two active-site residues, which are involved in stabilizing the intermediate or transition state, must take unfavorable conformations to do so. Presumably the energy for this is paid for by interactions in the remainder of the protein fold.

Active Site of TIM and Its High Catalytic Efficiency. The TIM catalytic mechanism involves transfers of protons between

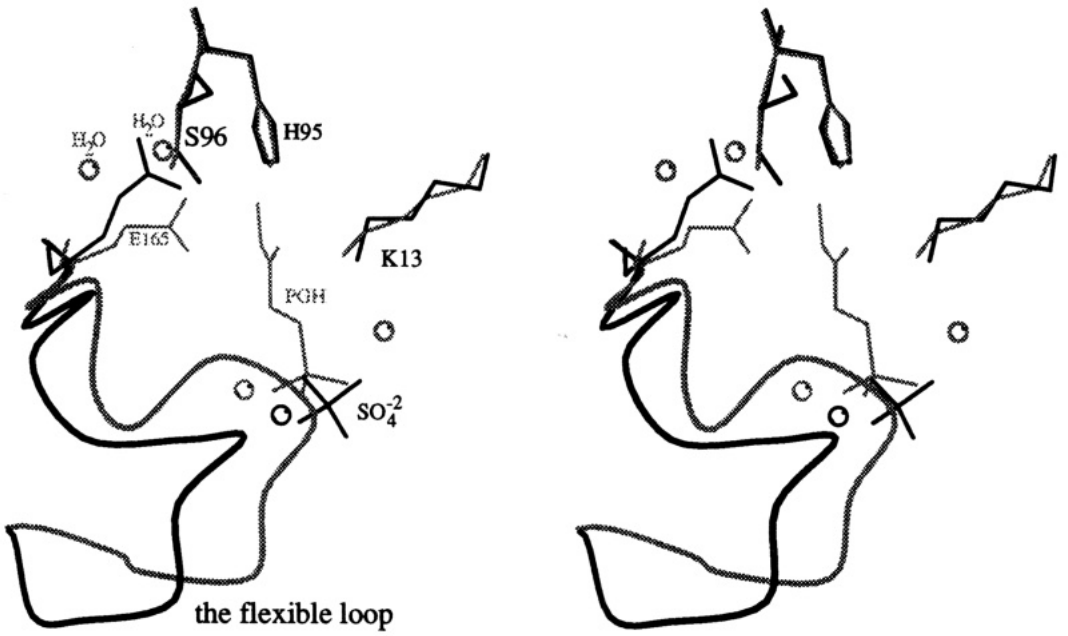

FIGURE 3: Stereo diagram of the comparison of the cTIM structures with PGH bound (gray) and with sulfate ion bound (black). The flexible loop region (residues 166–176) is represented as a smoothed backbone loop for clarity. The important conformational changes, after PGH is bound to the active site, are the movement of Glu165 from the upper position (black) to the lower position (gray) and the movement of the flexible loop from the open form (black) to the closed form (gray). This figure was prepared using the program MOLSCRIPT (Kraulis, 1991).

side chains of active-site residues and the carbon and oxygen atoms of the substrates. In the "forward" direction this involves protonation and deprotonation of DHAP. Under normal aqueous conditions, C–H groups α to a carbonyl carbon are very weak acids (pK_a = 18), while carboxylate groups are weak bases (pK_a = 4); so the nonenzymic proton transfer from DHAP to, say, acetate ion is predictably slow (Richard et al., 1984). In comparison, the enzyme-catalyzed reaction is limited only by diffusion of the substrates into or out of the active site of TIM, which means at least a 10^9-fold enhancement of the proton-transfer rate, suggesting a priori that the enzyme, especially the active-site environment, profoundly affects the chemistry of catalysis. The structure of the active site also suppresses phosphate elimination from the enediolate intermediate during catalysis. In solution the rate of phosphate elimination is comparable to the rate of isomerization, while in the enzyme it is several orders of magnitude slower than isomerization (Pompliano & Knowles, 1990).

PGH is nearly isostructural with the substrate DHAP (Figure 1); consequently PGH is a good inhibitor, and the enzyme-bound form may resemble a catalytic intermediate or transition state (Davenport et al., 1991). Even more important, the model of the active site in the cTIM–PGH complex is similar to that obtained from crystallographic studies of enzyme–substrate complexes of both chicken (6.0-Å resolution) and yeast TIM (3.5-Å resolution) (Alber et al., 1981), which suggests that there are no dramatic differences between the orientations of enzyme-bound PGH and DHAP. So with reasonable caution, the cTIM–PGH structure may be used to define those features of the active site important in the catalytic mechanism.

When a comparison between the structures of the TIM–PGH complex and the uncomplexed TIM (Banner et al., 1975) is made, two significant conformational changes in the active site are observed (Figure 3). The carboxylate group of the catalytic base, Glu165 in TIM–PGH, moves 2–3 Å toward the inhibitor from its position in the unliganded enzyme. In the free enzyme, the carboxylate is oriented deeper in the active-site pocket, and the carboxylate oxygens form two

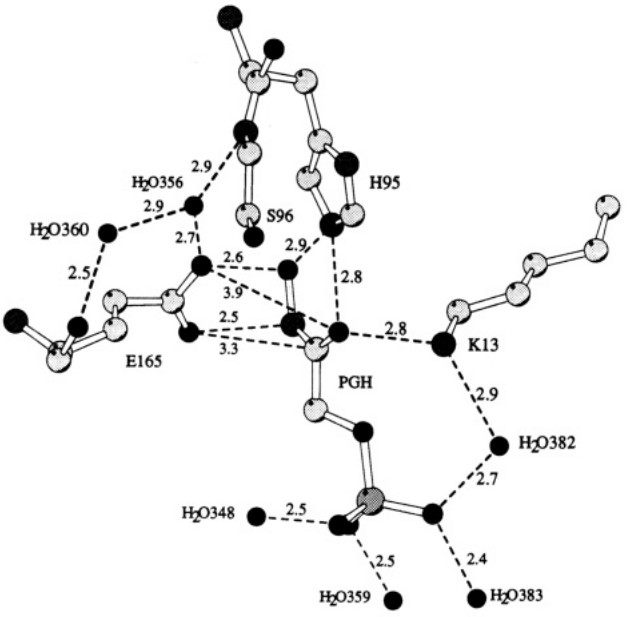

FIGURE 4: Schematic diagram of the active site of the cTIM–PGH structure. Only the residues discussed in the text are shown for clarity. This figure is in the same orientation as Figure 3 and was prepared using the program MOLSCRIPT.

hydrogen bonds with the backbone amide nitrogen and the side-chain hydroxyl group of Ser96, respectively. After inhibitor binding, these hydrogen bonds are broken and the side chain of Glu165 moves into proximity with the reactive triose part of PGH, such that Oϵ2 is 2.5 and 3.3 Å from the expected positions of the C1 and C2 atoms of DHAP, respectively (Figure 4). These distances are consistent with direct action of the carboxylate group in the enolization step, and support the notion that proximity between active-site residues and substrate is very important in efficient catalysis.

The second significant conformational change is the large movement of the flexible loop, residues 166–176 (Figure 3). A comparison of crystal structures shows that in the native enzyme the loop is exposed to solvent, leaving the active site

open for the substrate to enter; in the complexed enzyme, upon the binding of substrate DHAP or competitive inhibitor PGH into the active site, the loop acts like a rigid lid to close down over the active-site mouth, making several interactions with the peripheral phosphate atoms of the ligand. When the open and closed models of cTIM are superimposed by least-squares optimization on all α-carbon atoms of the dimer excluding the loop regions, the root-mean-square (rms) deviation is only 0.39 Å, whereas that for the loop residues is 4.8 Å. The central part of the main chain of the loop moves more than 7 Å in going from the open to the closed form. The amide nitrogen atom of residue Gly171, which is at the center of the loop, is only 2.7 Å from one of the oxygen atoms of the phosphate group in the closed structure. Deletion of the loop residues shows that the loop contributes critically to the stabilization of the enediol phosphate intermediate (Pompliano et al., 1990). The deletion mutant enzyme, which has four residues from 170 to 173 excised, could no longer prevent the loss of the enediol phosphate from the active site and its rapid decomposition to methylglyoxal and inorganic phosphate. Indeed, when glyceraldehyde 3-phosphate (GAP) is the substrate, the enediol phosphate intermediate decomposes 5.5 times faster than it reprotonates to form the product dihydroxyacetone phosphate (DHAP). TIM has evidently evolved its mobile loop to bind the highly reactive intermediate enediol phosphate tightly. However, it is not clear whether the loop movement is triggered or where the trigger of the loop movement is. There are two possibilities: one is that Glu165 also functions as the trigger of the rigid loop, because it is directly involved in the chemical steps of catalysis and so could sense the change from substrate to product; the second possibility is that the central part of the loop initiates the movement by long-range electrostatic interaction with bound DHAP and GAP. Recent experiments have given support to the hypothesis that Glu165 is the trigger because proton transfer to the substrate appears to be concerted with loop closure (Sampson & Knowles, 1992).

The nonaqueous environment of the active site when the loop is closed should increase the basicity of Glu165 over its value in the uncomplexed enzyme. Qualitatively speaking, glutamic acid in simple peptides has a pK_a around 4.3, close to the value of 4.76 for acetic acid in water. Polar but nonaqueous solvents like methanol and DMSO have been shown to increase the pK_a of a carboxylate by 8 units or more; the pK_a of acetic acid is 9.6 in methanol and 12.6 in DMSO (Kemp et al., 1975); in vacuum, the pK_a of acetic acid increases to over 50 (Richard et al., 1984). Although the conformational change in the flexible loop has excluded bulk solvent from contact with the carboxylate in the TIM–PGH complex, there are still two ordered water molecules around the base. One water molecule, H_2O356, is within hydrogen-bonding distance of both the Oϵ1 atom of Glu165 and the N atom of Ser96; the other water, H_2O360, is hydrogen-bonded to the N atom of Glu165 and H_2O356 (Figure 4). The presence of the two waters, especially the hydrogen bonding of H_2O356 to the oxygen of Glu165, should decrease the basicity of the carboxylate group; the extent of this decrease is hard to estimate. Apparently there has been no selective pressure on TIM to eliminate the negative effect of this bound water because the base-catalyzed proton abstraction is still faster than diffusion even with the water present. Therefore the most reasonable model for the enzymic environment would be a low polar but nonaqueous medium, and so it seems likely that the pK_a of Glu165 in the enzyme–substrate complex is close to 7 and possibly higher. On the basis of the pK_a of acetone in water, the proton to be abstracted probably has a pK_a close to 18 in the free substrate (Chiang et al., 1984), so the increased basicity of Glu165 due to solvation by the enzyme, as opposed to bulk water, is important for efficient catalysis.

Structure analysis and site-directed mutagenesis (Lolis et al., 1990; Davenport et al., 1991; Nickbarg et al., 1988; Komives et al., 1991) have implicated His95 in its neutral imidazole form (Lodi & Knowles, 1991) as the general acid in the enzyme-catalyzed reaction. An active-site α helix (95–102) may facilitate the proton transfer from the imidazole Nϵ atom of His95 to the carbonyl oxygen of the substrate DHAP in the enolization step. This two-turn α helix is found at the carboxyl-terminal end of the barrel where the active site is located. The amino-terminal end of helix 95–102 is directed toward the reactive triose part of DHAP. The central axis of the helix dipole goes through the imidazole ring of Glu95 at the N-terminus of the helix, is oriented toward the midpoint between the two carbonyl oxygens of the substrate, and is almost parallel to the proton-transfer direction. Quantum-mechanical calculations (Hol et al., 1978, 1981) have shown that an α helix has a considerable electric field and that the strength of the field increases up to a helix length of about seven residues or two turns, whereafter further elongation has only a marginal effect; the electric potential steeply increases from 0.4 to 1.0 V when the distance to the amino terminus decreases from 6 to 3 Å. Interestingly, these are approximately the distances in this crystal structure from the O2 atom of PGH and the Nϵ atom of imidazole, respectively, to the N-terminus of the dipole. So the electric field, which divergently runs from the N-terminus to the C-terminus of the helix dipole, might accelerate the rate of transfer of the positively charged proton away from the imidazole to the carbonyl oxygen of DHAP in the enolization step, in which the *pro-R* proton of DHAP is abstracted by Glu165 and an intermediate endiol or endiolate arises from proton donation by His95.

It is possible that the two substeps of proton abstraction and proton donation occur concertedly. Gerlt, Gassmann, and co-workers have proposed that protonation of the carbonyl oxygen of a carbon acid to form an enediol is essential to promote proton transfer from carbon (Gerlt et al., 1991; Gerlt & Gassman, 1992). A neutral imidazole should have a pK_a around 16 (Lodi & Knowles, 1991) (the helix dipole may depress this value somewhat), which is close to that expected for the enediol. Matching the pK_a's of the proton donor and acceptor would lead to rapid proton transfer followed by formation of a short, strong hydrogen bond between His95 and the intermediate, stabilizing this species and providing very efficient polarization of the sugar. The crystal structure shows that the hydrogen bond distances between the Nϵ atom of His95 and the O1 and O2 atoms of PGH are 2.9 and 2.8 Å, respectively (Figure 4), consistent with this notion.

From the refined high-resolution structure of the cTIM–PGH complex, it is possible to infer the relative orientations of the carboxylate of Glu165 and the substrate DHAP in the Michaelis complex with some precision. The inhibitor PGH is bound in an extended conformation, with the two catalytically important oxygens cis to one another (Figure 4). This configuration places the σ bond between C1 and the *pro-R* hydrogen of DHAP periplanar with the π orbital of the C2 carbonyl group. Stereoelectronic theory suggests that such an alignment would increase the acidity of the C1 proton (Corey & Sneen, 1955). Protonation states of the intermediate are not shown consistently in the figure to emphasize the point that the exact nature of the intermediate is not known. It

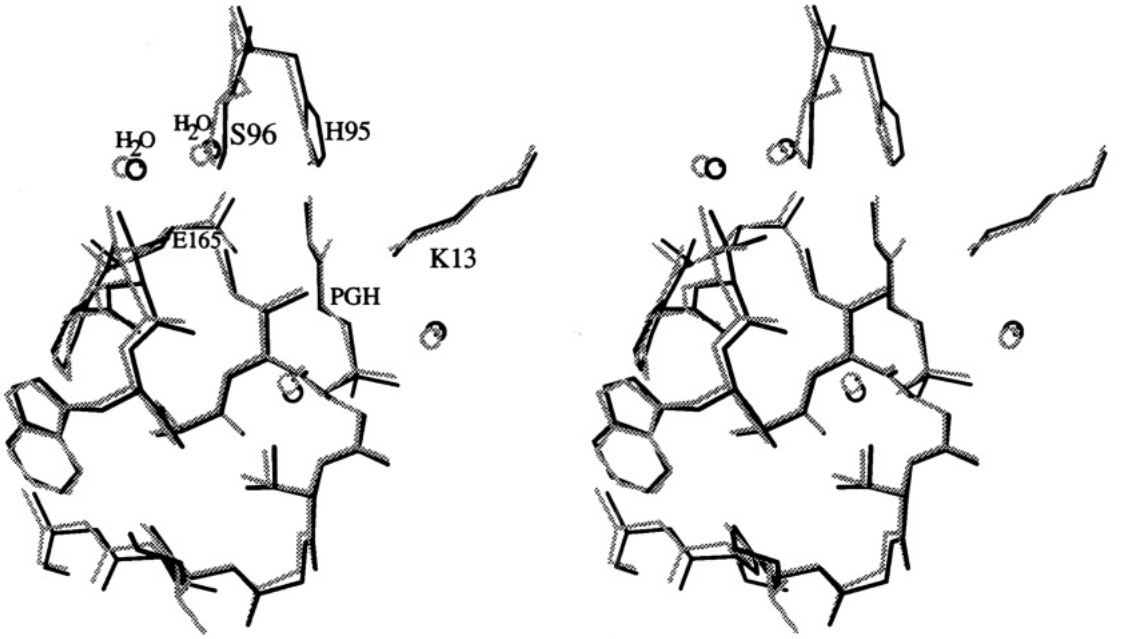

FIGURE 5: Stereo diagram of the comparison of the cTIM–PGH structure (black) with the yTIM–PGH structure (gray). The conformations of the active sites of the chicken and the yeast enzyme are identical, including bound water molecules which are found in identical positions. This figure is in the same orientation as Figure 3 and was prepared using the program MOLSCRIPT.

may be fully protonated as an enediol phosphate, or it may exist as two symmetrical enediolates that differ only in the position of the proton on oxygen. No conclusive evidence for either case exists so far, although Gerlt and Gassmann have suggested that proton transfer to form the enediol will be the preferred pathway (Gerlt & Gassman, 1992). On the other hand, theoretical calculations suggest that formation of the enediolate is also favorable (Bash et al., 1991). Also the observation that His95 is neutral under all conditions favors the hypothesis that the enediolate is the intermediate in the reaction (Lodi & Knowles, 1991).

There is another potential active-site electrophile, Lys13. The terminal N atom of the side chain of Lys13 is 2.8 Å away from the O2 atom of PGH, forming a hydrogen bond (Figure 4). It seems that Lys13 is less likely to be involved in proton transfer than His95, because the distance between the N atom of the Lys13 side chain and the O1 of the substrate is 5.2 Å, whereas the Nϵ atom of His95 is equidistant (2.8 and 2.9 Å) from both substrate oxygens. Lys13 also may form a weak hydrogen bond with the bridging oxygen O3 of the substrate. But it also forms two other hydrogen bonds, with distances of 2.8 Å, with the Glu97 side chain and a bound water molecule.

There are eight hydrogen bond donors around the phosphate group. Four hydrogen bonds are provided by the backbone amides of residues Gly171, Ser211, Gly232, and Gly233; the other four are provided by solvent water molecules H_2O348, H_2O359, H_2O382, and H_2O383 (Figure 3). Those hydrogen bonds stabilize the phosphate group and help to fix the whole substrate molecule in a reactive conformation. Abundant interactions between the phosphate group and the backbone of the enzyme and water molecules are part of how the enzyme anchors the enediol(ate) intermediate and minimizes the possible movement of the phosphate group, leaving no space for the elimination reaction to happen.

In summary, the high-resolution structure of cTIM complexed with the inhibitor PGH and its comparison with the native enzyme illustrate several structural factors which make TIM an efficient catalyst: the proximity effect of the catalytic base Glu165, which is induced by substrate binding; the closure of the rigid loop (residues 166–176) and its stabilization of the enediol phosphate intermediate; the increased basicity of Glu165 by means of the polar but nonaqueous environment of the active site; a possible strong electric field around the catalytic residue His95 and the reactive triose part of the substrate, and its consequent accelerating effect on the proton transfer from the imidazole to the carbonyl oxygen of the substrate; the orientation of the carboxylate of Glu165 which orients its more basic *syn*-orbital toward the proton to be abstracted; the binding of the substrate in a planar conformation that enhances the acidity of the proton to be abstracted and disfavors the elimination side reaction stereoelectronically; stabilization of developing negative charge in the transition state by electrostatic interactions with Lys13 and His95; tight binding of the phosphate group to prevent the elimination reaction; and finally, general acid catalysis by the neutral form of His95 to decrease the pK_a of the carbon acid. Therefore it is believable that in the cases of the pseudorevertant enzymes of E165D/S96P and H95N/S96P a substitution of serine 96 by a nonpolar but bulky proline residue could restore some of the lost enzyme activity by increasing one or more of these positive structural factors. This high-resolution structure provides the basis for interpreting the structural consequences of this pseudoreversion event.

Comparison of Chicken and Yeast TIM–PGH Structures. When the chicken and yeast TIM–PGH structures were superimposed, there was no significant conformational difference in active site between the two enzymes (Figure 5). This result is surprising, because the chicken enzyme could suppress its active-site lesions (e.g., the E165D or H96N single mutants) by the replacement of another active-site residue, serine 96, by proline, whereas for the yeast enzyme the substitution of serine 96 by proline does not improve the activity of either the H95N mutant or the E165D mutant. That result suggested that the active site of cTIM might have some structural differences from the yTIM active site. However, the work reported in this paper shows that this is not the case. Among those amino acid residues which make up the active-site core and enclose the inhibitor PGH, the rigid lid loop takes exactly the same conformation in the two structures except for one unconserved residue, 175 (Thr for the chicken

enzyme; Ala for the yeast enzyme), which is approximately 9 Å away from the sugar part of PGH; the catalytically active residues Glu165, His95, and Lys13 of both TIMs have identical configurations relative to PGH; furthermore, all other core residues, including Asn11, Ser96, Glu97, Gly209, Gly210, Ser211, Leu230, Val231, Gly232, and Gly233, show no differences between the two enzymes. Even the active-site water molecules in the two TIM–PGH complexes are in similar positions. A careful search within a 10-Å radius sphere, which was centered at C1 of PGH, was done, but so far there are no amino acid residues or other structural differences which are suspected to be responsible for the differential behavior between the chicken and yeast enzymes in terms of pseudoreversion. Therefore it will be necessary to obtain crystal structures of those single-site and double-site mutants from both chicken and yeast enzymes to understand why S96P is a pseudorevertant for E165D and H95N in one but not the other.

ACKNOWLEDGMENT

We are grateful to Prof. N.-h. Xuong for making the UCSD Research Resource for Protein Crystallography (NIH-RR01644) available to us for data collection with the expert assistance of Dr. Narendra Narayan. In addition we thank Prof. Ann M. Stock for the use of X-ray equipment at the Center for Advanced Biotechnology and Medicine and the University of Medicine and Dentistry of New Jersey.

REFERENCES

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

Alber, T., 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.

Banner, D. W., Bloomer, A. C., Petsko, G. A., Phillips, D. C., Pogson, C. I., & Wilson, I. A. (1975) *Nature 255*, 609–614.

Bash, P. A., Field, M. J., Davenport, R. C., Petsko, G. A., Ringe, D., & Karplus, M. (1991) *Biochemistry 30*, 5826–5832.

Belasco, J. G., & Knowles, J. R. (1980) *Biochemistry 19*, 472–477.

Blacklow, S. C., & Knowles, J. R. (1990) *Biochemistry 29*, 4099–4108.

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

Blacklow, S. C., Liu, K. D., & Knowles, J. R. (1991) *Biochemistry 30*, 8470–8476.

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

Chiang, Y., Kresge, A. J., & Tang, Y. S. (1984) *J. Am. Chem. Soc. 106*, 460–462.

Collins, K. D. (1974) *J. Biol. Chem. 249*, 136–142.

Corey, E. J., & Sneen, R. A. (1956) *J. Am. Chem. Soc. 78*, 6269–6278.

Cork, C., Fehr, D., Hamlin, R., Vernon, W., Xuong, N.-h., & Perez-Mendez, V. (1973) *J. Appl. Crystallogr. 7*, 319–323.

Coulson, A. F. W., Knowles, J. R., Priddle, J. D., & Offord, R. E. (1970) *Nature 227*,180–181.

Crowther, R. A. (1972) *Molecular Replacement Method* (Rossmann, M.G., Ed.) Gordon and Breach, New York.

Crowther, R. A., & Blow, D. W. (1967) *Acta Crystallogr. 23*, 544–548.

Davenport, R. C., Bash, P. A., Seaton, B. A., Karplus, M., Petsko, G. A., & Ringe, D. (1991) *Biochemistry 30*, 5821–5826.

Farber, G. K., & Petsko, G. A. (1990) *Trends Biol. Sci. 15*, 228–234.

Fitzgerald, P. M. D. (1988) *J. Appl. Crystallogr. 21*, 273–278.

Gandour, R. D. (1981) *Bioorg. Chem. 10*, 169–176.

Gerlt, J. A., & Gassman, P. G. (1992) *J. Am. Chem. Soc. 114*, 5928.

Gerlt, J. A., Kozarich, J. W., Kenyon, G. L., & Gassman, P. G. (1991) *J. Am. Chem. Soc. 113*, 9667–9672.

Hartman, F. C. (1970) *Biochem. Biophys. Res. Commun. 39*, 384–388.

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.

Hendrickson, W. A., & Konnert, J. H. (1980) in *Computing in Crystallography* (Diamond, R., Rameseshan, S., & Venkatesan, K., Eds.) pp 13.01–13.25, Indian Academy of Sciences, Bangalore.

Hol, W. G. J., van Duijnen, P. T., & Berendsen, H. J. C. (1978) *Nature 273*, 443–446.

Hol, W. G. J., Jalie, L. M., & Sander, C. (1981) *Nature 294*, 532–536.

Jancarik, J., & Kim, S.-H. (1991) *J. Appl. Crystallogr. 24*, 409–411.

Jones, T. A. (1978) *J. Appl. Crystallogr. 11*, 268–272.

Joseph, D., Petsko, G. A., & Karplus, M. (1990) *Science 249*, 1425–1428.

Kemp, D. S., Cos, D. D., & Paul, K. G. (1975) *J. Am. Chem. Soc. 97*, 7305–7312.

Komives, E. A., Chang, L. C., Lolis, E., Tilton, R. F., Petsko, G. A., & Knowles, J. R. (1991) *Biochemistry 30*, 3011–3019.

Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946–950.

Laemmli, U. K. (1970) *Nature 227*, 680–685.

Lattman, E. E., & Love, W. E. (1972) *Acta Crystallogr. B26*, 1854–1857.

Lodi, P. J., & Knowles, J. R. (1991) *Biochemistry 30*, 6948–6956.

Lolis, E., Alber, T., Davenport, R. C., Rose, D., Hartman, F. C., & Petsko, G. A. (1990) *Biochemistry 29*, 6609–6618.

Miller, J. H. (1972) *Experiments in Molecular Genetics*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Nickbarg, E. B., Davenport, R. C., Petsko, G. A., & Knowles, J. R. (1988) *Biochemistry 27*, 5948–5960.

Pompliano, D. L., Peyman, A., & Knowles, J. R. (1990) *Biochemistry 29*, 3186–3194.

Richard, J. P. (1984) *J. Am. Chem. Soc. 106*, 4926–4936.

Richard, J. P. (1985) *Biochemistry 24*, 949–953.

Rieder, S. V., & Rose, I. A. (1959) *J. Biol. Chem. 234*, 1007–1010.

Ringe, D., & Petsko, G. A. (1985) *Prog. Biophys. Mol. Biol. 45*, 197–235.

Sack, J. S. (1988) *PS300 FRODO—Molecular graphics program for the PS300*, Version 6.6, Revision A, Howard Hughes Medical Institute, Baylor College of Medicine, Houston, Taxas.

Sampson, N. S., & Knowles, J. R. (1992) *Biochemistry 31*, 8482–8494.

Straus, D., & Gilbert, W. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 2014–2018.

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.

Waley, S. G., Miller, J. C., Rose, I. A., & O'Connell, E. L. (1970) *Nature 227*, 181.

Wierenga, R. K., Noble, M. E. M., Postma, J. P. M., Groendijk, H., Kalk, K. H., Hol, W. G. J., & Opperdoes, F. R. (1991) *Proteins 10*, 33–49.

Xuong, N.-h., Sullivan, D., Nielsen, C., & Hamlin, R. (1985a) *Acta Crystallogr. B41*, 267–269.

Xuong, N.-h., Nielson, C., Hamlin, R., & Anderson, D. H. (1985b) *J. Appl. Crystallogr. 18*, 342–350.