Protein Science (1994), *3*:810-821. Cambridge University Press. Printed in the USA.

Crystal structure of recombinant human triosephosphate isomerase at 2.8 Å resolution. Triosephosphate isomerase-related human genetic disorders and comparison with the trypanosomal enzyme

SHEKHAR C. MANDE,[1] VÉRONIQUE MAINFROID,[2] KOR H. KALK,[3] KARINE GORAJ,[2] JOSEPH A. MARTIAL,[2] AND WIM G.J. HOL[1]

[1] Department of Biological Structure, School of Medicine, University of Washington, Seattle, Washington 98195
[2] Laboratoire de Biologie Moléculaire et de Génie Génétique, Université de Liège, Sart Tilman B-4000, Belgium
[3] Department of Biophysical Chemistry, University of Groningen, 9747 AG Groningen, The Netherlands

(RECEIVED December 28, 1993; ACCEPTED March 3, 1994)

Abstract

The crystal structure of recombinant human triosephosphate isomerase (hTIM) has been determined complexed with the transition-state analogue 2-phosphoglycolate at a resolution of 2.8 Å. After refinement, the *R*-factor is 16.7% with good geometry. The asymmetric unit contains 1 complete dimer of 53,000 Da, with only 1 of the subunits binding the inhibitor. The so-called flexible loop, comprising residues 168-174, is in its "closed" conformation in the subunit that binds the inhibitor, and in the "open" conformation in the other subunit. The tips of the loop in these 2 conformations differ up to 7 Å in position. The RMS difference between hTIM and the enzyme of *Trypanosoma brucei*, the causative agent of sleeping sickness, is 1.12 Å for 487 C^{α} positions with 53% sequence identity. Significant sequence differences between the human and parasite enzymes occur at about 13 Å from the phosphate binding site. The chicken and human enzymes have an RMS difference of 0.69 Å for 484 equivalent residues and about 90% sequence identity. Complementary mutations ensure a great similarity in the packing of side chains in the core of the β-barrels of these 2 enzymes. Three point mutations in hTIM have been correlated with severe genetic disorders ranging from hemolytic disorder to neuromuscular impairment. Knowledge of the structure of the human enzyme provides insight into the probable effect of 2 of these mutations, Glu 104 to Asp and Phe 240 to Ile, on the enzyme. The third mutation reported to be responsible for a genetic disorder, Gly 122 to Arg, is however difficult to explain. This residue is far away from both catalytic centers in the dimer, as well as from the dimer interface, and seems unlikely to affect stability or activity. Inspection of the 3-dimensional structure of trypanosomal triosephosphate isomerase, which has a methionine at position 122, only increased the mystery of the effects of the Gly to Arg mutation in the human enzyme.

Keywords: crystal structure; genetic disorder; sleeping sickness; triosephosphate isomerase

Triosephosphate isomerase (TIM) is one of the glycolytic enzymes and catalyzes the interconversion of D-glyceraldehyde 3-phosphate (GAP) and dihydroxyacetone phosphate. TIM has been found to occur in all organisms wherever it has been looked for (Lolis et al., 1990). It is an essential component of the glycolytic pathway. Its deficiency has been shown to enhance levels of dihydroxyacetone phosphate in humans and cause chronic anemia and neuromuscular impairment (Daar et al., 1986; Eber et al., 1991). Due to its important place in glycolysis, TIM also forms an attractive target for drug design against parasites that have the ability to survive in the mammalian bloodstream. Among the various diseases currently ravaging the tropical world, those due to such parasites are some of the most serious ones (e.g., malaria and sleeping sickness). Control of these diseases is presently a major problem and has been a focus of attention of various agencies, including the World Health Organization (WHO,

Reprint requests to: Wim G.J. Hol, Department of Biological Structure, SM-20, School of Medicine, University of Washington, Seattle, Washington 98195; e-mail: hol@xray.bchem.washington.edu.

1992). Blood sugar is the main source of energy for these parasites (Fairlamb, 1989), and thus it is hoped that blocking of glycolysis may lead to elimination of these parasites from the bloodstream.

TIM has been the subject of extensive biophysical and biochemical studies. Its catalytic properties have been studied in detail, and it has been established that the reaction rates are diffusion controlled, thus implying that the evolution of TIM as an enzyme has been nearly perfect (Knowles & Albery, 1977). Structural studies of TIMs in complex with transition-state analogues (Lolis & Petsko, 1990; Davenport et al., 1991), site-directed mutagenesis experiments (Blacklow et al., 1991), and quantum mechanical and molecular mechanical studies (Bash et al., 1991) have established its catalytic mechanism in detail. Historically, TIM also revealed for the first time one of the commonly used protein folds repeatedly observed in a diverse set of enzymes when the crystal structure of chicken TIM was elucidated (Banner et al., 1975). These enzymes follow an 8-fold repeat of a β-loop-α, or more appropriately α-loop-β, motif (Urfer & Kirschner, 1992). The active sites of all these enzymes are invariably located at the C-termini of the barrel (Brändén, 1991). These and other observations have led to an intense debate regarding the evolution of TIM barrel enzymes (Farber, 1993). Due to the unique structural properties, the TIM barrel has been the subject of interesting engineering experiments, such as designing of a synthetic TIM barrel (Goraj et al., 1990) or the exchange of a $\beta\alpha$ unit between TIMs from different species (Mainfroid et al., 1993).

Structurally, TIMs form a well-characterized family. To date, X-ray crystal structures of TIMs from 4 different sources—namely, chicken (Banner et al., 1975), yeast (Alber et al., 1981), the sleeping sickness parasite *Trypanosoma brucei* (Wierenga et al., 1991b), and *Escherichia coli* (Noble et al., 1993b)—have been elucidated. Trypanosomal TIM has been studied to analyze structural changes involved during ligand binding (Noble et al., 1991; Verlinde et al., 1991; Wierenga et al., 1991a). These investigations have established that the flexible loop of about 8 residues closes when binding smaller ligands but not when larger inhibitors are bound (Verlinde et al., 1991). This structural information has been used to design putative inhibitors of the enzyme from *T. brucei* (Verlinde et al., 1992a). The crystal structure of human TIM (hTIM) offers the possibility for a comparative study with trypanosomal TIM structure in order to design specific inhibitors rationally. We also hope that the hTIM structure would aid designing specific inhibitors of TIMs from other parasites, which depend crucially on glycolysis.

TIM is a homodimeric enzyme consisting of around 250 residues. It does not contain any metal ion or cofactor, nor is there any evidence of allostery or cooperativity among subunits important for its function. The numbering scheme used in this paper corresponds to the hTIM sequence (Maquat et al., 1985), unless otherwise specified, from 1 to 248 for monomer 1 and from 301 to 548 for monomer 2. A total of 56 residues have been shown to be conserved through evolution, based on sequence comparison of 13 known sequences (Wierenga et al., 1992). Most of these residues are either part of, or are in the neighborhood of, the active site. The dimer interface is formed mainly by loops 1, 2, and 3 (following the definition of loops in Wierenga et al. [1991a]). Residues important for the activity reside in loops 5, 6, and 7. Loop 6, also called the "flexible loop," undergoes a remarkable conformational change upon ligand binding (Alber et al., 1981; Wierenga et al., 1991a). The tip of the loop moves by as much as 7 Å to form a lid over the bound ligand. We report in this paper the overexpression of hTIM in *E. coli* and the crystal structure of recombinant hTIM in complex with an inhibitor, 2-phosphoglycolate (2-PG), refined at 2.8 Å resolution (Kinemage 1).

Results and discussion

Overexpression and steady-state kinetics

Recombinant hTIM was overproduced in *E. coli* with an optimal production obtained after a 12-h induction. Using the purification protocol described in the Materials and methods, we obtained about 100 mg of hTIM per liter of bacterial culture, with a degree of purity higher than 95%.

The recombinant hTIM was assayed for activity with GAP as substrate. The steady-state kinetic parameters are listed in Table 1. The specific activity of recombinant hTIM is 9,800 units/mg, which is in good agreement with the values previously obtained with natural hTIM (Lu et al., 1984).

Quality of the structure

The refinement statistics are shown in Table 2. The RMS difference between all the main-chain coordinates of the starting and the final model is 1.06 Å. The Ramachandran plot (Ramachandran & Sasisekharan, 1968) of the final model (Fig. 1) shows 5 residues lying outside the allowed region. Three of these have high *B*-factor values: (1) Ser 303, which is close to the N-terminal of subunit 2 and is not very well defined in the electron density; the corresponding residue in subunit 1, Ser 3, however, has allowed ϕ, ψ angles; (2) Ala 546, which is close to the C-terminus of subunit 2; again, corresponding Ala 246 of subunit 1 has allowed main-chain dihedral angles as well as lower temperature factors; (3) Gly 171, which belongs to the flexible loop of subunit 1. Two of the 5 deviating residues, Lys 13 and Lys 313, have low *B*-values. These residues are known to be essential components of the active sites in all known TIMs (Fothergill-Gilmore & Michels, 1993). This lysine residue probably has to adopt an unusual conformation for positioning its side chain at a very precise location in the active site, in agreement with a study by Herzberg and Moult (1991) which reports that residues with deviating stereochemistry tend to cluster near active site regions in proteins.

Structural features of human TIM

Overall comparison of monomers

The overall structure of hTIM shown schematically in Figure 2 and Kinemage 1 resembles, as expected, structures of other

Table 1. *Steady-state kinetic parameters of recombinant hTIM*

K_m (GAP) (mM)	0.49
k_{cat} (min^{-1})	2.7×10^5
k_{cat}/K_m (min^{-1} mM^{-1})	5.5×10^5
K_i (2-PG) (μM)	7.4

Table 2. *Refinement statistics*

Resolution range (Å)	8.0–2.8
Number of reflections	9,625
Number of protein atoms	3,738
Number of inhibitor atoms	9
R-factor	0.167
Geometric parameters	
RMS bond length deviation (Å)	0.019
RMS bond angle deviation (degrees)	3.8
RMS dihedral angle deviation (degrees)	25.5
RMS improper angle deviation (degrees)	1.59

TIMs (Wierenga et al., 1992). Minor structural differences occur in the loop regions, particularly those in the vicinity of the active site. The 2 monomers of the refined model superpose with a 0.49-Å RMS difference between 237 pairs of equivalent C^{α} atoms. Eight residues of the flexible loop (168–175) and 2 N-terminal residues are the only residues that deviate by more than 3 Å, as shown in Figure 3 (see Kinemage 2). The 2 conformations of the loop signify the well-known "open" and "closed" loop conformations (Alber et al., 1981). Average main-chain *B*-factors for the 2 subunits are 13.9 and 18.5 $Å^2$, respectively. Admittedly, the estimation of temperature factors at the limited resolution of 2.8 Å is not very reliable, so this difference will not be discussed further. The first 3 residues in both subunits are poorly defined in the electron density. The flexible loop residues of the second subunit also show high temperature factors with an average of 44.2 $Å^2$. The 2 monomers are related by a near perfect 2-fold symmetry with 179.6° rotation and a translation component of 0.11 Å along the 2-fold axis.

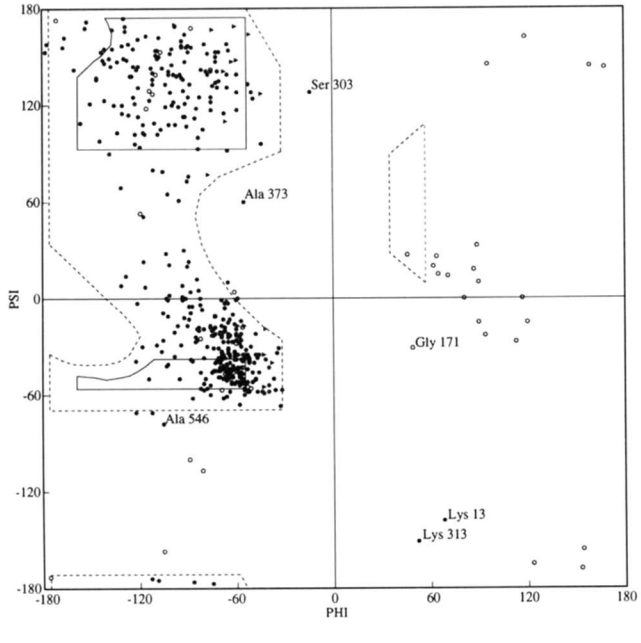

Fig. 1. Ramachandran plot of the refined hTIM structure. Open circles denote glycine residues, and the prolyl residues are indicated in triangles. Residues lying outside the allowed limits are marked (see text for details).

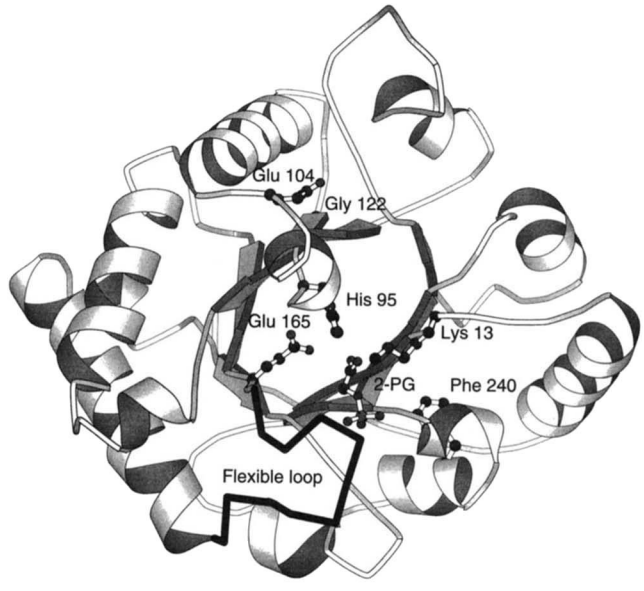

Fig. 2. Schematic ribbon diagram of hTIM drawn using MOLSCRIPT (Kraulis, 1991) showing the inhibitor (2-PG) and the catalytic residues. The sites of thermolabile mutants, 104, 122, and 240, which cause genetic disorders in humans, are also indicated. The highlighted loop undergoes large conformational changes upon ligand binding (see text for details).

Active site and 2-PG binding

Two of the residues implicated for TIM activity, namely Lys 13 and His 95, are well positioned in both subunits to carry out the enzymatic reaction. The third residue essential for TIM activity, Glu 165, has been observed to adopt different conformations in different structures. Its ideal position for carrying out catalysis is "swung in" with $\chi_1 = 60°$ (Knowles, 1991). It is also known to be able to adopt a "swung out" conformation with $\chi_1 = -60°$ (Wierenga et al., 1992). This position is not ideally suited for catalysis. In the hTIM structure, Glu 165 adopts a "swung out" conformation in subunit 1, whereas it is disordered

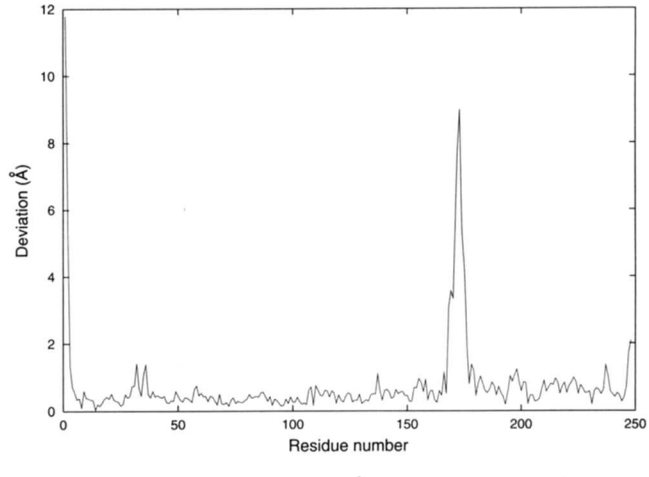

Fig. 3. Plot of positional difference (Å) vs. residue number for the superposition of 2 hTIM monomers. The difference in position of the flexible loop is clearly visible.

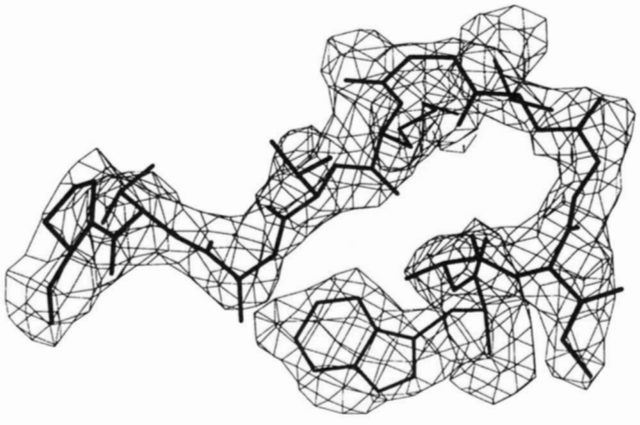

Fig. 4. Electron density map calculated by a $(2F_o - F_c)\alpha_c$ Fourier synthesis contoured at 2σ levels above the mean electron density. The map shows residues Trp 468 through Pro 478, encompassing the flexible loop.

in subunit 2. As mentioned earlier, 2-PG occupies the active site crevice in subunit 2. Also, in this subunit, the flexible loop is "closed" over the inhibitor. In subunit 1 the loop is locked in an "open" conformation due to crystal contacts. The main-chain and the side-chain atoms of the flexible loop of subunit 2 show relatively high thermal parameters. These residues, however, remained well defined in the electron density maps during refinement, as shown in Figure 4.

It was realized during the refinement that, while mounting the crystals in capillaries, the artificial mother liquor was devoid of the inhibitor solution. Higher temperature factors of 2-PG atoms and those of the flexible loop residues of subunit 2, and the lack of density for the side chain of Glu 465, are possibly repercussions of this. 2-PG in hTIM occupies a position similar to that observed in the yeast TIM–2-PG complex (Kinemages 1, 3; Lolis & Petsko, 1990). As mentioned earlier, it does not appear to be fully occupied in the active site, as evidenced by high temperature factors of approximately 45 Å^2. Catalytic residues His 395 ($N^{\epsilon 2}$) and Lys 313 (N^{ζ}) form hydrogen bonds with the carboxylate oxygens of the inhibitor (Fig. 5; Kinemage 3). The terminal phosphate oxygens are hydrogen bonded to the main-chain amide nitrogens of Gly 532, Gly 533, and Gly 471. As may be noted, Gly 471 belongs to the closed flexible loop. Apart from this hydrogen bond, the only other interaction between the flexible loop residues and the inhibitor is a van der Waals contact between 1 of the phosphate oxygens and C^{α} of Ile 470. Additional stability to the phosphate, as shown in Figure 5 and Kinemage 3, is provided by the helix dipole (Hol et al., 1978) of a 3_{10}-helix whose N-terminus points directly to the phosphate. The bridging oxygen of 2-PG does not form any hydrogen bond at all, similar to what was observed in the yeast TIM–2-PG complex. The carbonyl oxygen of residue Gly 510 unusually points toward the phosphate. This carbonyl oxygen in yeast TIM complexed with 2-PG points in the opposite direction. A similar contact has been reported for trypanosomal TIM in complex with sulfate ion (Wierenga et al., 1991b). Whether this conformation in hTIM is genuine, or due to lack of sufficiently high-resolution data, remains undecided.

Human genetic disorders and TIM

It has been reported that mutation of Glu 104 in hTIM to Asp leads to a thermolabile enzyme (Daar et al., 1986). This has also been correlated to a recessive disorder that results in hemolytic anemia and neuromuscular dysfunction in human beings. The importance of Glu 104 in TIM structures and evolution can be understood from its strict conservation over all the known TIM sequences (Fig. 6). Glu 104 in hTIM is almost completely buried at the interface of the 2 monomers (Fig. 7; Kinemage 1). The carboxylate group of Glu is sandwiched between the guanidinium moiety of Arg 98 and N^{ζ} of Lys 112, thereby shielding the 2 positively charged groups from each other. Arg 98 is involved in an intersubunit salt bridge with Glu 377 and is also in contact with the catalytic His 95. All these residues are strictly conserved in all the known TIM sequences. Based on these observations, and through computer modeling studies, Daar et al. (1986) showed that shortening of the Glu 104 side chain by 1 methyl group may lead to an unfavorable interaction between Lys 112 and Arg 98, thereby reducing the stability of the mutant. In the hTIM structure, we also find that the C^{β} and C^{γ} side-chain atoms of Arg 98 are in contact with the carbonyl oxygen of His 95. The interatomic C to O distances range from 3.0–3.4 Å in the 2 hTIM subunits. These distances are of similar magnitude in the structures of trypanosomal and yeast TIMs, determined at a better resolution. Therefore, a perturbation in

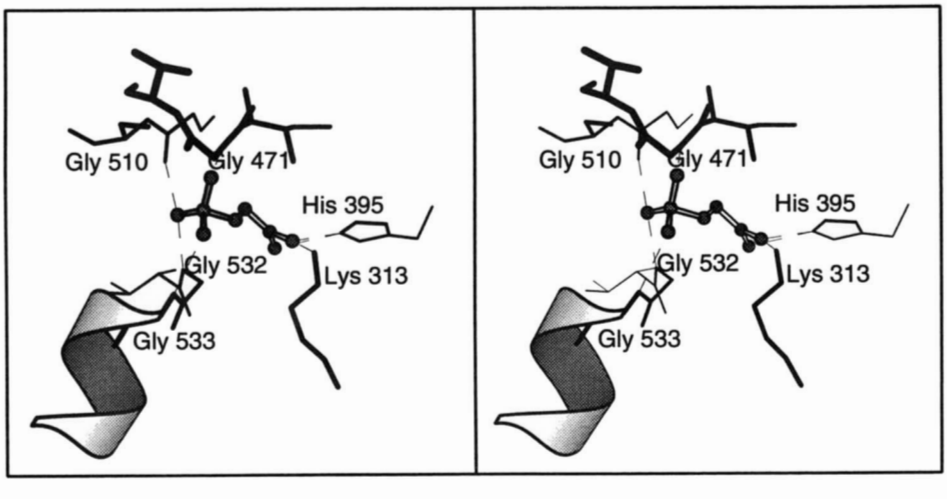

Fig. 5. Interactions of 2-PG with hTIM. The 3_{10}-helix, comprising residues 533–537, points directly on the phosphate moiety. Also shown is the unusual phosphate–carbonyl (Gly 510) interaction. This figure was produced using MOLSCRIPT.

```
                      1          2          3          4          5          6          7          8
 C.japonica .MGRKFFVGG NWKCN.GTSE EVKKIVTLLN EA......EV PSE.DVVEVV VSPPYVFLPF VKNLLR.A.. DFHVAAQNCW VKKGGAFTGE VSAEMLVNLG
      Maize ..GRKFFVGG NWKCN.GTTD QVEKIVKTLN EG......QV PPS.DVVEVV VSPPYVFLPV VKSQLR.Q.. EFHVAAQNCW VKKGGAFTGE VSAEMLVNLG
      Human APSRKFFVGG NWKMN.GRKQ SLGELIGTLN AA......KV PAD.T..EVV CAPPTAYIDF ARQKLD.P.. KIAVAAQNCY KVTNGAFTGE ISPGMIKDCG
 M. mulatta APSRKFFVGG NWKMN.GRKQ NLGELIGTLN AA......KV PAD.T..EVV CAPPTAYIDF ARQKLD.P.. KIAVAAQNCY KVTNGAFTGE ISPGMIKDCG
     Rabbit APSRKFFVGG NWKMN.GRKK NLGELITTLN AA......KV PAD.T..EVV CAPPTAYIDF ARQKLD.P.. KIAVAAQNCY KVTNGAFTGE ISPGMIKDCG
      Mouse APTRKFFVGG NWKMN.GRKK CLGELICTLN AA......NV PAG.T..EVV CAPPTAYIDF ARQKLD.P.. KIAVAAQNCY KVTNGPFTGE ISPGMIKDLG
    Chicken .APRKFFVGG NWKMN.GDKK SLGELIHTLN GA......KL SAD.T..EVV CGAPSIYLDF ARQKLD.A.. KIGVAAQNCY KVPKGAFTGE ISPAMIKDIG
 Coelacanth .APRKFFVGG NWKMN.GDKK SLGELIQTLN AA......KV PFT.G..EIV CAPPEAYLDF ARLKVD.P.. KFGVAAQNCY KVSKGAFTGE ISPAMIKDCG
   Mosquito ..GRKFCVGG NWKMN.GDKA SIADLCKVLT TG......PL NAD.T..EVV VGCPAPYLTL ARSQLP.D.. SVCVAAQNCY KVPKGAFTGE ISPAMLKDLN
 Fruit fly .MSRKFCVGG NWKMN.GDQK SIAEIAKTLS SA......AL DPN.T..EVV IGCPAIYLMY ARNLLP.C.. ELGLAGQNAY KVAKGAFTGE ISPAMLKDIG
 A.nidulans .MPRKFFVGG NFKMN.GNAE STTSIIKNLN SA......NL DKS...VEVV VSPPALYLLQ AREVANK... EIGVAAQNVF DKPNGAFTGE ISVQQLREAN
  S. pombe ..ARKFFVGG NFKMN.GSLE SMKTIIEGLN TT......KL NVG.D.VETV IFPQNMYLIT TRQQVKK... DIGVGAQNVF DKKNGAYTGE NSAQSLIDAG
      Yeast ..ARTFFVGG NFKLN.GSKQ SIKEIVERLN TA......SI .PE.N.VEVV ICPPATYLDY SVSLVKKP.. QVTVGAQNAY LKASGAFTGE NSVDQIKDVG
  T. brucei MSKPQPIAAA NWKCN.GSQQ SLSELIDLFN ST......SI NHD...VQCV VASTFVHLAM TKERLSHP.. KFVIAAQNAI .AKSGAFTGE VSLPILKDFG
 B.stearot. ..MRKPIIAG NWKMH.KTLA EAVQFVEDVK GH......VP PA..DEVISV VCAPFLFLDR LVQAADGT.. DLKIGAQTMH FADQGAYTGE VSPVMLKDLG
   E. coli ..MRHPLVMG NWKLN.GSRH MVHELVSNLR KE......LA GV..AGCAVA IAPPEMYIDM AKREAEGS.. HIMLGAQNVN LNLSGAFTGE TSAAMLKDIG
 C.glutan. .MARKPLIAG NWKMN.LDHQ QAIGTVQKLA FA......LP KEYFEKVDVA VTVPFTDIRS VQTLVEGDKL EVTFGAQDVS QHESGAYTGE VSASMLAKLN
      Morsp ...MQAWVIG NWKQNPATSH DVNALLDDLC AAISTTKQMS HDNSARCQIM VAPSLIHLAA VNGRLKDTSI LCAAQDVSGH SASVGAYTGE CSAQQIVDAG
                       * *                                                                                 ** ***   *

            9          0          1          2          3          4          5                     6          7
 C.japonica IPWVILGHSE RRALLNESNE FVGDKTAYAL SQGLKVIACV GETLEQREAG STISVVAAQT KAIAE..... .KV..SDWTN IVVAYEPVWA IGTGKVASPA
      Maize VPWVILGHSE RRALLGESNE FVGDKVAYAL SQGLKVIACV GETLEQREAG STMDVVAAQT KAIAE..... .KI..KDWSN VVVAYEPVWA IGTGKVATPA
      Human ATWVVLGHSE RRHVFGESDE LIGQKVAHAL AEGLGVIACI GEKLDEREAG ITEKVVFEQT KVIAD..... .NV..KDWSK VVLAYEPVWA IGTGKTATPQ
 M. mulatta ATWVVLGHSE RRHVFGESDE LIGQKVAHAL AEGLGVIACI GEKLDEREAG ITEKVVFEQT KVIAD..... .NV..KDWSK VVLAYEPVWA IGTGKTATPQ
     Rabbit ATWVVLGHSE RRHVFGESDE LIGQKVAHAL SEGLGVIACI GEKLDEREAG ITEKVVFEQT KVIAD..... .NV..KDWSK VVLAYEPVWA IGTGKTATPQ
      Mouse ATWVVLGHSE RRHVFGESDE LIGQKVSHAL AEGLGVIACI GEKLDEREAG ITEKVVFEQT KVIAD..... .NV..KDWSK VVLAYEPVWA IGTGKTATPQ
    Chicken AAWVILGHSE RRHVFGESDE LIGQKVAHAL AEGLGVIACI GEKLDEREAG ITEKVVFEQT KAIAD..... .NV..KDWSK VVLAYEPVWA IGTGKTATPQ
 Coelacanth VTWVILGHSE RRHVFGESDE LIGQKVSHAL SEGLGVVACI GEKLDEREAG ITEGVVFEVT EVIAD..... .DV..KDWSK VVLAYEPVWA IGTGKTASPQ
   Mosquito IGWVILGHSE RRAIFGESDE LIADKVAHAL AEGLKVIACI GETLQEREAG QTEAVCFRQT KAIAD..... .KV..KDWSN VVIAYEPVWA IGTGKTASPE
 Fruit fly ADWVILGHSE RRAIFGESDA LIAEKAEHAL AEGLKVIACI GETLEEREAG KTNEVVARQM CAYAQ..... .KI..KDWKN VVVAYEPVWA IGTGKTATPD
 A.nidulans IDWTILGHSE RRVILKETDE FIARKTKAAI EGGLQVIFCI GETLEEREAN KTIDVVTRQL NAAAK..... .ELSKEQWAK VVIAYEPVWA IGTGKVATTE
  S. pombe ITYTLTGHSE RRTIFKESDE FVADKTKFAL EQGLTVVACI GETLADREAN ETITVVVRQL NAIAD..... .KV..QNWSK IVIAYEPVWA IGTGKTGTPE
      Yeast AKWVILGHSE RRSYFHEDDK FIADKTKFAL GQGVGVILCI GETLEEKKAG KTLDVVERQL NAVLE..... .EV..KDWTN VVVAYEPVWA IGTGLAATPE
  T. brucei VNWIVLGHSE RRAYYGETNE IVADKVAAAV ASGFMVIACI GETLQERESG RTAVVVLTQI AAIAK..... .KLKKADWAK VVIAYEPVWA IGTGKVATPQ
 B.stearot. VTYVILGHSE RRQMFAETDE TVNKKVLAAF TRGLIPIICC GESLEEREAG QTNAVVASQV EKALA..... .GLTPEQVKQ AVIAYEPIWA IGTGKSSTPE
   E. coli AQYIIIGHSE RRTYHKESDE LIAKKFAVLK EQGLTPVLCI GETEAENEAG KTEEVCARQI DAVLK..... .TQGAAAFEG AVIAYEPVWA IGTGKSATPA
   C.glutan CSWVVVGHSE RREYHNESDE LVAAKAKAAL SNGISPIVCV GEPLEIREAG THVEYVVEQT RKSLA..... .GLDAAELAN TVIAYEPVWA IGTGKVASAA
      Morsp ATWTILGHSE RRQYHQESND ALLQKMIHAL SQDLGVVFCI GESQEQYDTK QTLTVIDNQL VVIKEFITQQ PELIDALPTR LIIAYEPVWA IGTGKVPTVA
              *    **** **    *     *  *              *   **                                      ***** *******

            8          9          0          1          2          3          4
 C.japonica QAQEVHFELR KWIKENVGAD VAGSVRIIYG GSVNGANSKE LAGQPDIDGF LVGGASLKPE F.VDIIKSAT VKSS.
      Maize QAQEVHASLR DWLKTNASPE VAESTRIIYG GSVTAANCKE LAAQPDVDGF LVGGASLKPE F.IDIINAAT VKSA.
      Human QAQEVHEKLR GWLKSNVSDA VAQSTRIIYG GSVTGATCKE LASQPDVDGF LVGGASLKPE F.VDIINAKQ .....
 M. mulatta QAQEVHEKLR GWLKSNVSEA VAQSTRIIYG GSVTGATCKE LASQPDVDGF LVGGASLKPE F.VDIINAKQ .....
     Rabbit QAQEVHEKLR GWLKSNVSDA VAQSTRIIYG GSVTGATCKE LASQPDVDGF LVGGASLKPE F.VDIINAKQ .....
      Mouse QAQEVHEKLR GWLKSNVNDG VAQSTRIIYG GSVTGATCKE LATPADVDGF LVGGASLKPE F.VDIINAKQ .....
    Chicken QAQEVHEKLR GWLKSHVSDA VAQSTRIIYG GSVTGGNCKE LASQHDVDGF LVGGASLKPE F.VDIINAKH .....
 Coelacanth QSQELHGKLR KWLKENVSET VADSVRIIYG GSVTGATCKE LASEPDVDGF LVGGASLKPE F.VEYKDVRQ .....
   Mosquito QAQEVHAALR KWFTENVSAD VSAAIRIQYG GSVTAANCRE LAAKPDIDGF LVGGASLKPE F.IQIVNARQ .....
 Fruit fly QAQEVHASLR QWLSDNISKE VSASLRIQYG GSVTAANAKE LAKKPDIDGF LVGGASLKPE F.LDIINARQ .....
 A.nidulans QAQEVHSAIR KWLKDAISAE AAENTRIIYG GSVSEKNCKD LAKEADIDGF LVGGASLKPA F.VDIVNARL .....
  S. pombe EAQEVHAEIR KWATNKLGAS VAEGLRVIYG GSVTGGNCKE FLKFHDIDGF LVGGASLKPE FPTNIVNVHS L....
      Yeast DAQDIHASIR KFLASKLGDK AASELRILYG GSANGSNAVT FKDKADVDGF LVGGASLKPE F.VDIINSRN .....
  T. brucei QAQEAHALIR SWVSSKIGAD VRGELRILYG GSVNGKNART LYQQRDVNGF LVGGASLKPE F.VDIIKATQ .....
 B.stearot. DANSVCGHIR SVVSRLFGPE AAEAIRIQYG GSVKPDNIRD FLAQQQIDGP LVGGASLEPA SFLQLVEAGR HE...
   E. coli QAQAVHKFIR DHIAKV.DAN IAEQVIIQYG GSVNASNAAE LFAQPDIDGA LVGGASLKAD AFAVIVKAAE AAKQA
   C.glutan DAQEVCKAIR GLIVELAGDE VAEGLRILYG GSVKAETVAE IVGQPDVDGG LVGGASLDGE ALPKL..AAN AASVA
      Morsp EVSATHKHIK QTLAGF..AD SLSNMTVLYG GSVNADNANS FAADPMIDGA LVGGASLKAD SFLTIATAFS QASI.
                                        ** **                     *  *******
```

Fig. 6. Sequence alignments of 18 known TIM sequences: *Coptis japonica* (Okada et al., 1989), maize (Marchionni & Gilbert, 1986), human (Maquat et al., 1985), rabbit (Corran & Waley, 1975), mouse (Cheng et al., 1990), chicken (Straus & Gilbert, 1985), coelacanth (Kolb et al., 1974), mosquito (Tittiger et al., 1993), fruit fly (Shaw-Lee et al., 1991), *Aspergillus nidulans* (McKnight et al., 1986), *Schizosaccharomyces pombe* (Russell, 1985), yeast (Alber & Kawasaki, 1982), *Trypanosoma brucei* (Swinkels et al., 1986), *Bacillus stearothermophilus* (Rentier-Delrue et al., 1993), *Escherichia coli* (Pichersky et al., 1984), *Corynebacterium glutamicum* (Eikmanns, 1992), mouse (Cheng et al., 1990), *Macaca mulatta* (Old & Mohrenweiser, 1988), and *Moraxella* sp. (Rentier-Delrue et al., 1993) as calculated by the GCG package (Devereux et al., 1984).

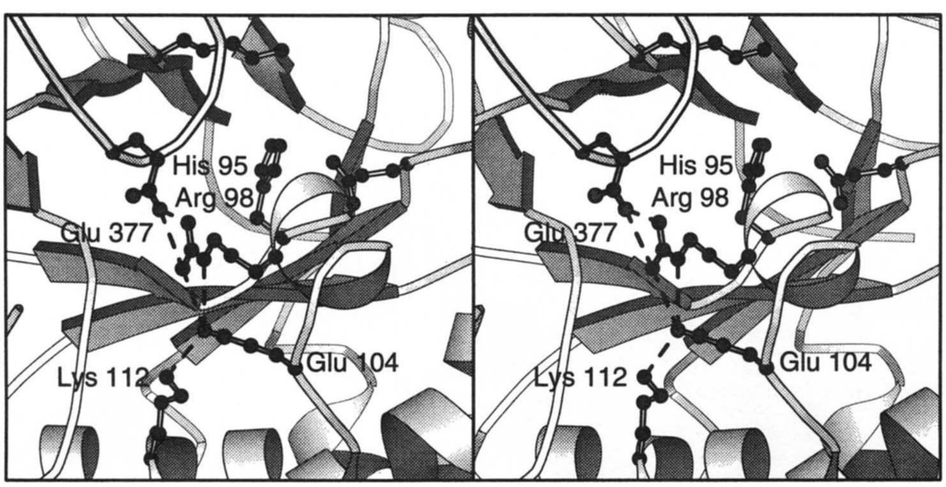

Fig. 7. Site of point mutation for the thermolabile mutant Glu 104 to Asp (Daar et al., 1986) of hTIM. Glu 104 is shown shielding Arg 98 and Lys 112. Arg 98 makes an intersubunit hydrogen bond with Glu 377. Subunit 1 is shown in a ribbon representation, whereas loop 3 extending from subunit 2 is shown as a C^{α} trace. Note also the contact of the C^{β} and C^{γ} atoms of Arg 98 with the carbonyl oxygen of the essential His 95 (see text for further discussion).

the conformation of the Arg 98 side chain may affect the critical positioning of His 95 in the active site. The catalytic role of His 95 is well known (Knowles, 1991; Komives et al., 1991). It has to be placed precisely at a position such that a short helix points on its imidazole, to maintain a neutral charge. It is thus clear that the replacement of Glu 104 by Asp, even though it does not alter the overall charge, is likely to perturb the local structure of the active site through a series of changes. This thus highlights the importance of Glu 104 in TIM sequences.

One more disorder in hTIM resulting in a thermolabile enzyme has recently been reported (Chang et al., 1993). A Hungarian individual has been shown to possess a malfunctional TIM with a mutation of Phe 240 into Ile. At position 240 in the known TIM sequences (Fig. 6), phenylalanine tends to be a highly conserved residue. Undoubtedly the sequence alignment at this position does not appear to be very accurate due to the adjoining insertions and deletions. However, it can be stated that phenylalanine is strictly conserved in all but 1 sequences. *Corynebacterium glutamicum* is the only species in which a replacement at this position is observed. In the crystal structure, Phe 240 is not a direct part of the active site, yet its close proximity to the active site and its critical position appear to be essential factors in maintaining the active site geometry. The phenyl ring of Phe 240 makes several contacts, ranging from 3.4 to 4.0 Å, with residues residing on β-strand 1 as well as a 3_{10}-helix (residues 233–237) (Fig. 8; Kinemage 3). The 3_{10}-helix, as discussed earlier, is important for the phosphate binding. The helix points directly on the phosphate moiety. Any alterations in the direction of the helix or its position may lead to a change in the substrate binding properties of TIM. β-Strand 1 extends up to residue 12. Catalytic Lys 13 is the immediate next residue after the culmination of β-strand 1. Asn 11, which follows residues that are in contact with the side chain of Phe 240, is another evolutionarily conserved residue and makes an intersubunit hydrogen bond. It is therefore understandable that a "conservative" mutation of Phe into Ile leads to different properties of the enzyme.

To understand the mutation in detail, we modeled the Phe 240 to Ile mutation on a computer. The model was not subjected to energy minimization, but the side-chain torsion angles of the Ile were adjusted to avoid short contacts with neighboring residues, thus retaining the rest of the model identical to the wild-type structure. We calculated the cavities in hTIM, both for the wild-type and for the modeled mutant structures using the CAVITY program of the BIOMOL suite of programs. The mutation appears to have created a cavity of volume 15.5 $Å^3$ next to the mutated Ile side chain. The size of this cavity is probably just sufficient to accommodate a water molecule. It is thus obvious that the mutation will lead to minor local structural rearrangements to compensate the cavity creation. These rearrangements are likely to be destabilizing for the association of monomers into a dimer, in view of the fact that many of the interface residues are a part of strand 1. Such cavity creation may also explain the thermolability of the enzyme.

Yet another mutation in hTIM, namely replacement of Gly 122 by an Arg, has been observed to yield a thermolabile electromorph variant of hTIM (Kinemage 1; Perry & Mohrenweiser, 1992). Gly 122 is partly exposed to the solvent. It occurs on the N-terminal side of the TIM barrel and is thus on the directly opposite side of the barrel from the active site. Owing to its remoteness from the active site, the kinetic parameters of the Gly 122 → Arg mutant are also largely unaffected (Perry & Mohrenweiser, 1992). The main-chain dihedral angles, ϕ and ψ, of Gly 122 in the 2 subunits are (−87, 169) and (−76, 143), respectively. These angles are within the allowed range for non-glycine residues, and therefore a mutation by arginine at this position would not be expected to alter the main-chain conformation of the polypeptide. Indeed, Gly 122 is conserved only in 8 of the 18 known sequences (Fig. 6). It is frequently replaced into larger side chains such as Lys, Gln, etc., which are similar in size as Arg. In order to understand the thermolability of the mutated enzyme, we made a mutation on a graphics system by replacing Gly 122 by an arginine residue. In the mutated structure, C^{β} of Arg 122 makes contacts of 2.8 Å with the $C^{\epsilon 3}$ and $C^{\zeta 3}$ atoms of Trp 90 in subunit 1, and contacts of 3.0, 2.8, 2.9, and 3.2 Å, respectively, with the $C^{\delta 2}$, $C^{\epsilon 3}$, $C^{\zeta 3}$, and $C^{\eta 2}$ atoms of Trp 390 in subunit 2. However, by minor adjustments with the χ^2 angle of Trp 90, these short contacts can easily be removed and

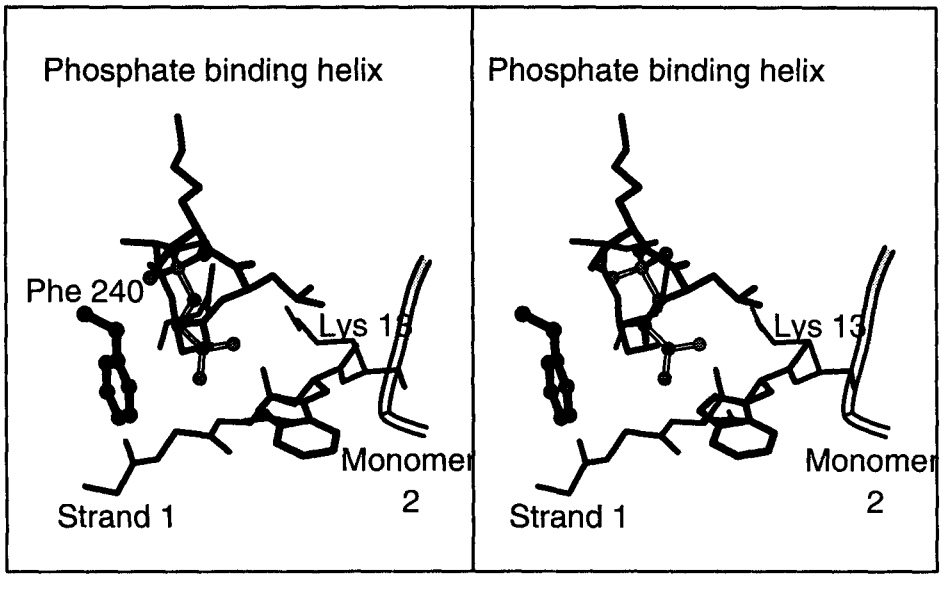

Fig. 8. Mutation of phenylalanine to isoleucine at position 240 resulting in genetic disorder in hTIM (Chang et al., 1993). The phenylalanine and the 2-PG molecule are shown in ball-and-stick representations. The orientation is chosen such that the 3_{10}-helix is seen pointing toward the phosphate. The adjoining β-strand is shown terminating in the catalytic Lys 313.

the mutation can be accommodated without much strain on the local protein structure. The mutated Arg then has sufficient room to expose its charged group to the solvent. Such rearrangements are known to occur in mutant protein structures (Matthews, 1993) and are unlikely to have a dramatic effect on thermostability.

Among the other TIM sequences in which the residue at position 122 is a non-Gly are *T. brucei*, *E. coli*, and *Bacillus stearothermophilus*. The structures of all these enzymes are either known or are being refined (Wierenga et al., 1991b; Noble et al., 1993b; F. Delboni, S.C. Mande, & W.G.J. Hol, pers. comm.). In *E. coli* TIM, Gly 122 is replaced by Thr, but Trp 90 is also replaced by a complementary mutation into a slightly smaller Tyr. In *B. stearothermophilus* TIM, Gly 122 is replaced by Ile and Trp 90 is replaced by Tyr. For a comparison of this thermolabile mutant of hTIM, however, the *T. brucei* structure serves our purposes best because in *T. brucei* Gly 122 is replaced by a Met and the complementary mutation at position 90 has not taken place. This region of the *T. brucei* TIM structure is depicted in Figure 9. Much to our delight, the computer prediction that the side chain of Trp 90 would readjust itself to the Gly 122 → Arg mutation, in order to avoid the short contact between the C^{β} of the residue 122 and the indole ring of Trp 90, holds almost perfectly. Our computer mutation had also shown that no hydrogen bond donors or acceptors of the protein would be buried by the Gly 122 → Arg mutation. The thermolability of the mutant enzyme is therefore very perplexing and cannot be explained in a straightforward manner.

Comparison of human and chicken TIM structures

Human and chicken TIMs are approximately 90% identical in sequence. The 2 structures are consequently very closely related. The RMS difference between 484 C^{α} atoms in the 2 structures is 0.69 Å. It is interesting to compare the sequence variability between the 2 sequences and its correlation to the 2 structures. Most differences, not surprisingly, occur in the loops connecting different secondary structural elements and will not receive further attention. However, 2 mutations in the 2 structures merit discussion.

Asp 17 in chicken TIM corresponds to the N-cap position of α-helix 1 (residues 17–31). The side-chain oxygens of the Asp form hydrogen bonds with the main-chain nitrogens of residues 19 and 20. Residue 17 in hTIM is an arginine and thus has lost

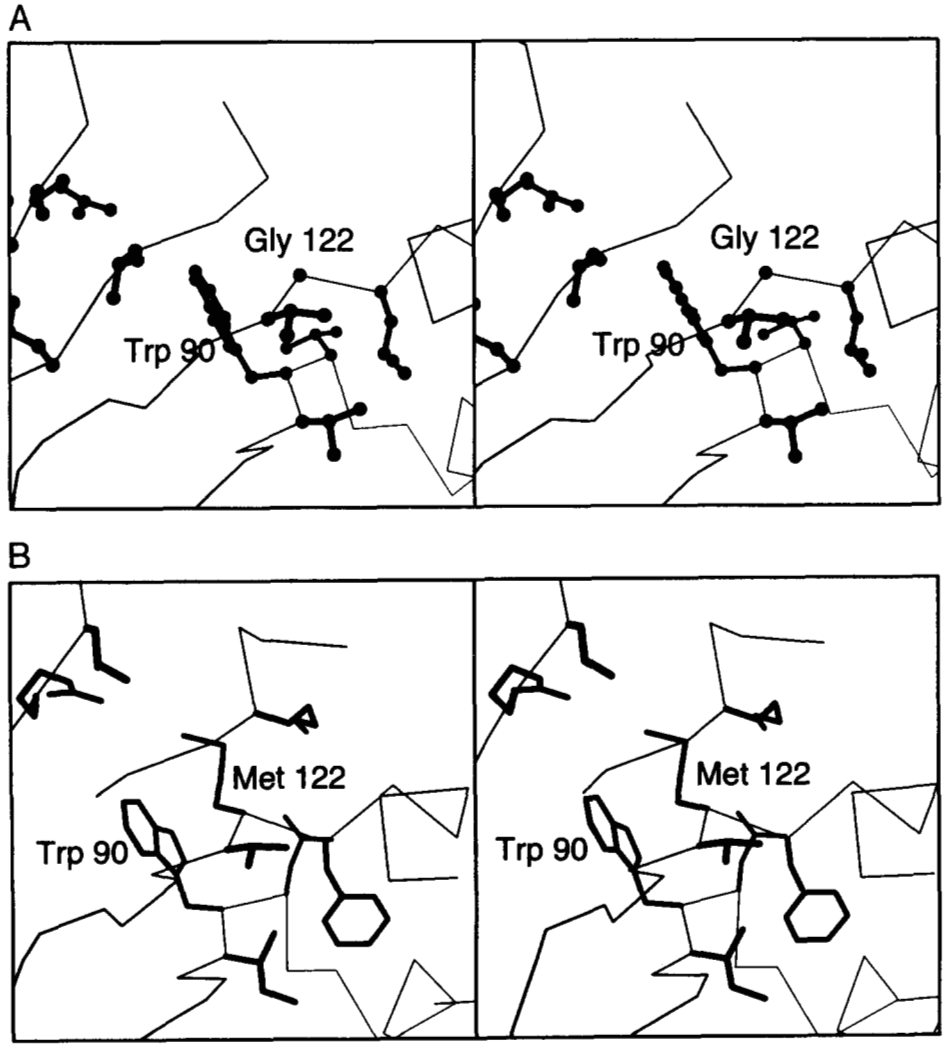

Fig. 9. A: Site of the mysterious point mutation Gly 122 → Arg in hTIM (Perry & Mohrenweiser, 1992). As yet, it is not clear why this substitution should lead to a more thermolabile enzyme compared to the wild-type one. Substitutions at this position are allowed in TIMs from other organisms (see text). **B:** An example of such a mutation, where a methionine replaces glycine at position 122 in trypanosomal TIM compared to hTIM, whereas tryptophan 90 is conserved in both the structures. No major rearrangements occur as a result of this substitution.

its capability of forming a hydrogen bond with the unpaired nitrogens in the first turn of helix 1 (Richardson & Richardson, 1988). However, the longer side chain of the arginine is involved in an intersubunit hydrogen bond with Asn 371. The peptide nitrogens of residues 19 and 20 are thereby exposed to the solvent. Not far from these residues in chicken TIM is Ile 46, which is partly buried in the interface of the monomers and partly exposed. This residue is an Ala in hTIM. The Ile to Ala mutation appears to have made 2 small cavities of approximately 7 $Å^3$ volume at the monomer-monomer interface in hTIM.

Another interesting aspect to compare between the 2 structures is the packing of the β-barrel residues. It has been shown that packing of residues inside the barrel in various TIM barrel enzymes can be best explained as a 3-layer packing (Lesk et al., 1989). Alternate strands of the 8-stranded barrel contribute to packing of each of the layers. Thus, of the 12 residues contributing to the packing of the core of the barrel among human and chicken TIMs, 10 are identical. In the first layer, however, an alanine 42 of hTIM is a glycine in chicken TIM. The reduction of the core volume by 1 methyl carbon is compensated by a substitution of an isoleucine in chicken TIM corresponding to a valine in hTIM at position 92 (Fig. 10). Thus, due to beautifully complementary mutations, the overall packing of the β-barrel residues in the 2 TIMs is very similar.

Comparison of human and trypanosomal TIM structures

As mentioned in the introduction, one of the objectives of undertaking this study was to assess the potential of TIM as a target for rational drug design against African sleeping sickness. This would require designing inhibitors of trypanosomal TIM that do not affect the hTIM function; in other words, the inhibitors should possess a high degree of selectivity. Detailed comparison of the 2 structures is a necessary prerequisite for such a study.

Trypanosomal TIM has been characterized in great detail structurally (Wierenga et al., 1991b; Noble et al., 1993a). Many different crystal forms, as well as several inhibitors' complex structures, have been crystallographically characterized (e.g., Verlinde et al., 1991, 1992b; Noble et al., 1993a). The native crystal structure has been refined to 1.83 Å resolution to an *R*-factor of 0.183. This structure also has a sulfate ion bound in its active site and has both open as well as almost closed conformations of the flexible loop. We therefore compared the cur-

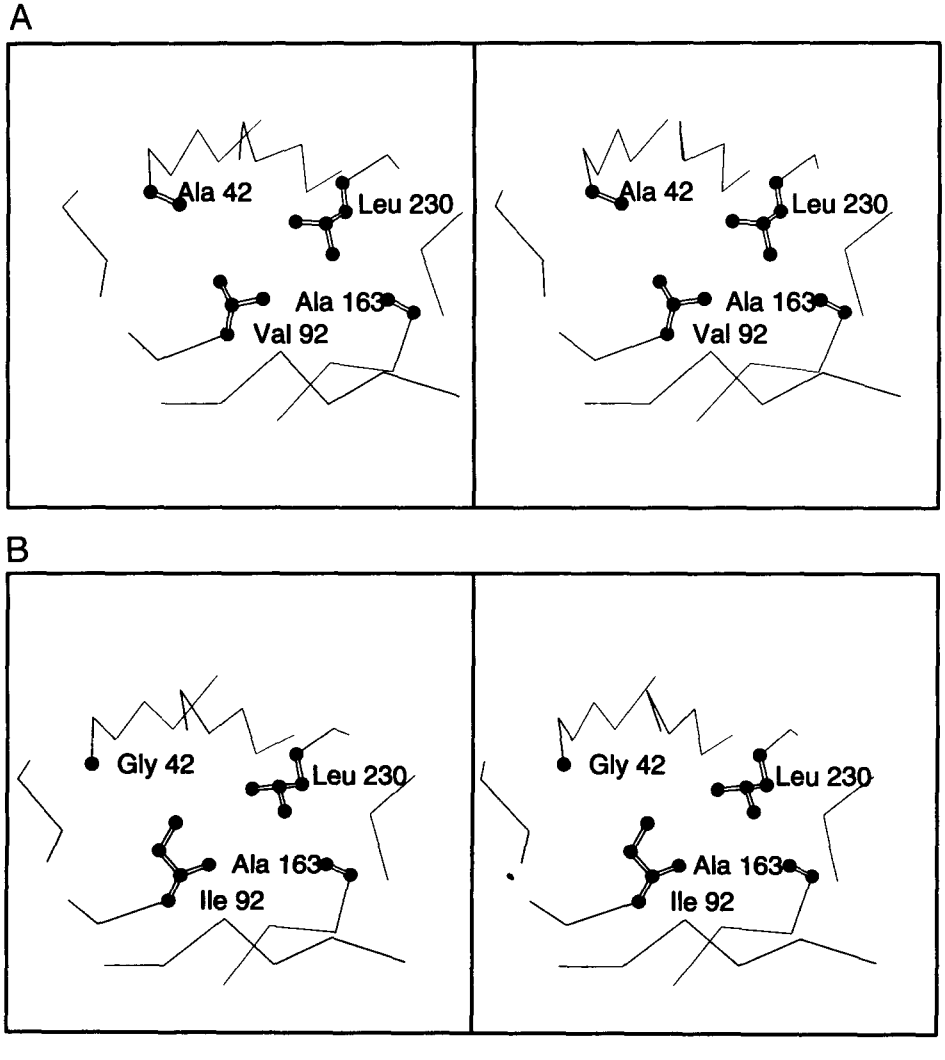

Fig. 10. Variability in the packing of the barrel residues between chicken and hTIMs as viewed down the β-barrel. A mutation of Ala 42 of hTIM (**A**) into a Gly in chicken TIM (**B**), which would have otherwise created a cavity in the core of the barrel, is seen to be elegantly compensated by a complementary mutation of Val into Ile (see text for further discussion).

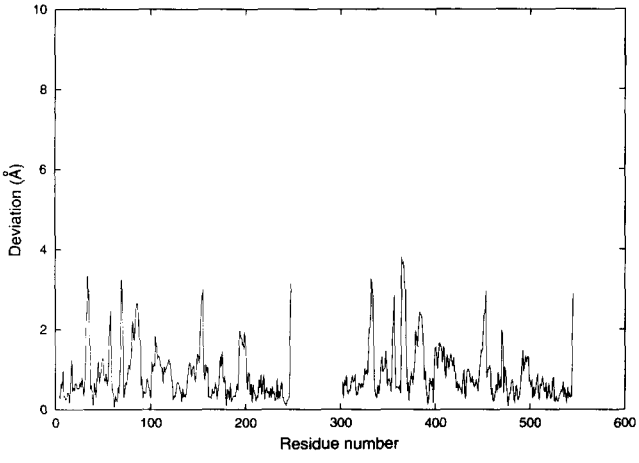

Fig. 11. Positional difference between C^{α} coordinates of trypanosomal and hTIMs. The numbering corresponds to the hTIM sequence. The dimers have been superimposed in such a manner that the subunits with open flexible loops, as well as those with the closed loops, match with each other.

rent 2.8-Å hTIM structure to the 1.83-Å crystal structure of trypanosomal TIM.

The overall percentage identity between the 2 amino acid sequences is about 53% (Maquat et al., 1985; Swinkels et al., 1986). The RMS difference between the coordinates of the 2 structures is 1.12 Å for 487 C^{α} atoms. The RMS difference as a function of residue number is shown in Figure 11, with the numbering of residues corresponding to hTIM. A 1-residue deletion in the hTIM sequence occurs between helix 2 and strand 3, and another 2-residue deletion is in the loop connecting helix 5 and strand 6. A single residue insertion in the hTIM sequence occurs in the loop connecting strand 3 and helix 4. This insertion is close to the interface of monomers. As a consequence of the deletion and the insertion, there are only a few minor local structural rearrangements. These rearrangements do not, however, affect the monomer-monomer interactions nor the packing of secondary structure.

Regions of interest for obtaining selective inhibitors should be close to the active site. Unfortunately, all the residues within about 10 Å of the active sites of the 2 TIMs are very well conserved. However, there are a few interesting substitutions a little further from the active sites. In 1 such difference shown in Figure 12 and Kinemage 1, a pocket in hTIM is lined by residues His 100, Val 101, and Phe 102 on one side, and Tyr 67, Lys 68, and Val 69 on its opposite side. These 6 residues are Ala, Tyr, Tyr, Ile, Ala, and Lys, respectively, in trypanosomal TIM. This pocket has also been a focus of attention for inhibitor design using a "linked-fragment approach" (Verlinde et al., 1991). In yet another difference between the 2 structures, the residues involved in hTIM are Thr 513, Ala 515, Thr 516, and Thr 475. These residues are, respectively, Asn, Lys, Asn, and Val in the trypanosomal TIM structure. Interestingly, the position of the closed flexible loop occurs between the 2 pockets discussed above. These residues thus could provide an anchor for inhibitors interacting with the flexible loop.

The crystal structure of recombinant hTIM has thus provided some new information on the 2 known genetic disorders. The effects of these mutations on the structure of TIM had been predicted prior to this study on the basis of the chicken TIM structure (Daar et al., 1986; Chang et al., 1993), and our conclusions are more or less in agreement with those studies. It is indeed fascinating that a mutation of Glu → Asp, conserving overall charge and reducing volume by only 1 methyl group, can lead to such different enzymatic properties, and most regretfully causes serious suffering to the individuals affected by the genetic disorder. The importance of Phe 240 in providing a floor to the phosphate binding helix and a base to strand 1 is apparent from the crystal structure, as well as from our computer mutation. The third mutation, namely Gly 122 → Arg, is less well understood and remains a mystery that might be resolved once the structure of this intriguing mutant becomes available.

Materials and methods

Cloning, overexpression, and kinetics

The human TIM cDNA sequence was amplified by PCR from a human skin fibroblast cDNA library (Clontech) as described by Friedman et al. (1990). The reaction was carried out at an annealing temperature of 56 °C with 10^7 recombinant phages. We used specific primers whose sequences were deduced from

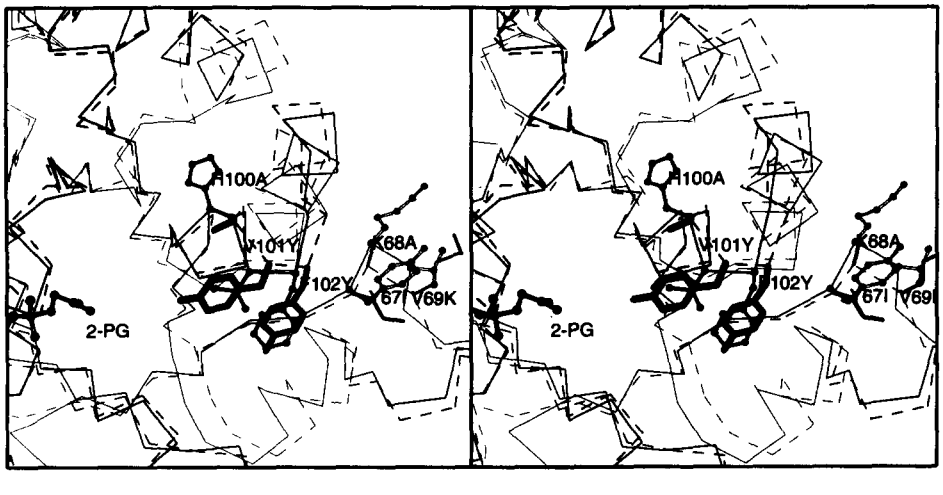

Fig. 12. Closeup of the site showing regions of interest for designing selective inhibitors for trypanosomal TIM. The C^{α} trace of hTIM is shown in thick lines, whereas that for trypanosomal TIM is shown in dashed lines. Also shown at the far left is the 2-PG binding site.

the hTIM nucleotide sequence (Maquat et al., 1985). Both primers contained parts of the extremities of the hTIM cDNA and restriction sites at their 5′-end to allow digestion and insertion of the PCR product at the *Nco*I and *BamH*I sites of the expression vector pARHS-3 (De Moerlooze et al., 1992).

E. coli BL21 (DE3) cells (Studier & Moffatt, 1986) harboring the recombinant pARHS-hTIM vector were grown in LB medium at 37 °C. When the culture reached an OD_{590} of 1, hTIM synthesis was induced with 1 mM IPTG and growth was continued for 12 h. The pelleted cells were resuspended and lysed with a French press (American Instrument Company). The cell debris was removed by centrifugation at 10,000 × *g* for 20 min. The supernatant was applied to a Sephadex G100 column (C26-100, Pharmacia) equilibrated in buffer A (20 mM Tris/HCl, pH 8.0, 100 mM NaCl, containing 0.2 mM EDTA, 0.1 mM phenylmethylsulfonylfluoride, 0.1 mM NaN_3). Fractions containing TIM were dialyzed against buffer A and loaded onto a MonoQ column (HR 10/10, Pharmacia) that had been equilibrated in buffer A. TIM was eluted with a gradient of 0–0.3 M NaCl. The purified protein solution was concentrated by ultrafiltration with Amicon PM-10 filters, and protein concentration was determined by Bradford's (1976) method, using the Bio-Rad protein assay, with bovine serum albumin as standard.

Enzyme assays were based on the decrease of absorbance at 340 nm due to oxidation of NADH in a coupled enzyme assay at 25 °C. The reaction was carried out according Misset and Opperdoes (1984) in 100 mM triethanolamine/HCl, pH 7.6, 20 μg/mL glycerol-3-phosphate dehydrogenase, 0.24 mM NADH, and 0.12–2.28 mM GAP. The inhibition constant (K_i) was obtained with 2-PG, a transition-state analogue (Wolfenden, 1969).

Crystallization and data collection

The protein solution for crystallization was prepared by dialyzing the protein against 5 mM MES buffer at pH 6.5 and 1 mM each of EDTA, NaN_3, and dithiothreitol (DTT). The concentration of the dialyzed protein was adjusted to 4 mg/mL, and then 2 mM 2-PG, an inhibitor of TIM, was added to this solution. The crystals were grown by the hanging-drop vapor-diffusion technique. The drops contained 3 μL each of the well and protein solutions. During the initial crystallization trials, the crystals always grew as thin and long intergrown needles. Failure to obtain good single crystals despite varying the concentration of the precipitant or varying the pH of the well solution prompted us to attempt crystallizations at different temperatures. It was noticed then that the drops which were producing a shower of needles at room temperature remained clear at cold temperatures (4 °C). This observation led to the development of a protocol in which the temperature of the crystallization plates was slowly raised from 4 °C to 18 °C over a period of 18 days. One good crystal of approximately 0.3 × 0.2 × 0.2 mm was obtained with a well solution of 100 mM HEPES buffer, pH 7.5, 20% polyethylene glycol (PEG) 4000, 10% 2-propanol, and 2 mM freshly dissolved DTT.

The X-ray intensities were measured on a FAST area detector. The detector was equipped with a CAD4 κ-goniostat with graphite monochromatized CuKα radiation from an Elliot GX21 rotating anode source. Data were processed using MADNES (Messerschmidt & Pflugrath, 1987) and the BIOMOL suite of programs. Profile fitting of the data was done according to Kabsch (1988). The crystal system was identified as primitive orthorhombic, although the assignments of screw axes could not be made unambiguously. However, later through translation function calculations, the space group was identified as $P2_12_12_1$ with $a = 65.81$, $b = 75.39$, and $c = 92.81$ Å. One dimer of the TIM molecule is present in the asymmetric unit of the crystals. The crystal and data statistics are shown in Table 3. The merging *R*-factor for the data at 2.8 Å resolution is 0.055 with a completeness of about 86%.

Structure solution by molecular replacement

Refined chicken TIM coordinates, kindly provided by Dr. P. Artymiuk, were used as a model for molecular replacement. Human TIM shares approximately 90% sequence identity with chicken TIM (Maquat et al., 1985; Straus & Gilbert, 1985). In the model, side chains of the 22 nonidentical residues were truncated to C^β. There are 2 places in the sequence where a nonglycine residue of chicken TIM is a glycine in hTIM. At these 2 places, the C^β atom of that residue was also removed. The individual atomic temperature factors were retained as in the chicken TIM structure. Structure factors for the model were calculated in an artificial P1 unit cell of dimensions 100 × 110 × 90 Å and angles of 90°. Rotation functions were calculated using Crowther's fast rotation function (Crowther, 1972) adapted to the BIOMOL suite of programs. Crowther and Blow T_2 translation functions (Crowther & Blow, 1967) were calculated using the TRAFUN program in the BIOMOL suite of programs. Packing calculations were carried out using the PACKNG program (Mande & Suguna, 1989). Cross-rotation functions calculated using data between 8 and 4 Å showed 2 clear peaks related by a 180° rotation corresponding to the 2 different superposi-

Table 3. *Crystal and data statistics*

Space group	$P2_12_12_1$
Cell dimensions (Å)	$a = 65.81$
	$b = 75.39$
	$c = 92.81$

V_m [a] = 2.17 Å³/Da (for 1 dimer per asymmetric unit)

	R-factors (%)		Completeness (%)	
Resolution shell	In shell	Cumulative	In shell	Cumulative
1,000–10.00	3.3	3.3	90.4	90.4
10.0–7.07	4.1	3.6	94.5	93.0
7.07–5.77	4.3	3.7	92.2	92.7
5.77–5.00	4.4	3.8	95.1	93.5
5.00–4.47	3.9	3.8	94.0	93.6
4.47–4.08	4.6	3.9	94.9	93.9
4.08–3.78	6.0	4.2	96.4	94.4
3.78–3.54	7.5	4.5	94.8	94.5
3.54–3.33	8.5	4.8	93.8	94.4
3.33–3.16	10.2	5.0	92.1	94.0
3.16–3.02	12.3	5.3	87.0	93.1
3.02–2.89	17.1	5.4	60.7	89.2
2.89–2.80	18.4	5.5	45.9	85.5

[a] V_m is the ratio of unit cell volume in Å³ and the total molecular weight of the protein in the unit cell in Da as defined by Matthews (1968).

tions of the dimers. Translation functions were calculated in the resolution ranges of 10-4 Å, 10-6 Å, 6-4 Å, and 8-4 Å. During each run of the translation functions, 6 sections with u, v, and w equal to 0 and 0.5 were calculated. In all the translation functions, the true peak appeared as the highest peak in the $v = 0.5$ and $w = 0.5$ sections. However, the $u = 0$ as well as $u = 0.5$ sections were much noisier, and made it difficult to identify the correct peak unambiguously. Packing calculations were carried out for primitive orthorhombic space groups involving different combinations of 2-fold and 2_1 screw axes. Only 1 solution compatible with packing as well as the translation function was then obtained. This thus enabled us to recognize the space group as $P2_12_12_1$.

Structure refinement

The chicken TIM model was placed in the unit cell of hTIM according to the molecular replacement solution. To initiate the refinement, 50 cycles of X-PLOR (Brünger, 1990) rigid-body refinement were carried out, first treating the entire dimer as a rigid body and later allowing the 2 monomers to move independently. At the end of rigid-body refinement, the R-factor in the resolution range 100-3.2 Å had dropped from an initial value of 0.397 to 0.366. The crystallographic refinement was carried out by means of X-PLOR's energy minimization option. After 100 cycles of alternate coordinate and restrained B-factor refinement in the resolution range 8-3.2 Å, the hTIM model was completed by mutating the appropriate residues. This was followed by 400 cycles of positional and restrained B-factor refinement in the resolution range 8-2.8 Å, interspersed with model building sessions with O (Jones et al., 1991). The models were checked against σ_A weighted $(2F_o - F_c)$ and $(F_o - F_c)$ maps (Read, 1986).

During the course of refinement, the flexible loop of subunit 2 (residues 468-474) was considerably rebuilt in the difference electron density maps. To avoid model bias, all the loop residues, as well as all residues lying within 5 Å from any loop residues, were deleted from the model. This resulted in removal of a total of 345 atoms, 154 belonging to subunit 1 and 191 to subunit 2. An omit map calculated after refining the partial model indicated that the flexible loop of subunit 2 had undergone a major conformational change compared to the open conformation in the chicken TIM model. The tip of the loop (residue Gly 473) had moved by as much as 7 Å from the initial model. This was not entirely unanticipated because the crystallization buffer contained 2-PG, an inhibitor of TIM. During the final stages of refinement, 2-PG was modeled in the residual density of the $(F_o - F_c)$ maps in subunit 2. The corresponding region, i.e., residues 168-174, in subunit 1 did not show any significant density. The final model contains 3,738 protein atoms and 9 2-PG atoms. No attempt has been made to model the water structure around the protein in view of the limited resolution of the data. At the end of the refinement, the R-factor was 0.167.

Acknowledgments

We thank Tjaard Pijning for help during crystallization and Drs. Andrea Mattevi, Christophe Verlinde, and other members of the Groningen and Seattle protein crystallography groups for many stimulating discussions. We thank Dr. Peter Artymiuk for refined chicken TIM coordinates. We are obliged to Dr. Rik Wierenga for his interest in this work. The project was partly funded through the EEC-BRIDGE project on TIM engineering and by the Belgian Programme on Interuniversity Poles of attraction PAI P3-044 (to J.A.M.).

References

Alber T, Banner DW, Bloomer AC, Petsko GA, Phillips DC, Rivers PS, Wilson IA. 1981. On the three-dimensional structure and catalytic mechanism of triosephosphate isomerase. *Philos Trans R Soc (Lond) B 293*: 159-171.

Alber T, Kawasaki G. 1982. Nucleotide sequence of the triosephosphate isomerase gene of *Saccharomyces cerevisiae*. *J Mol Appl Genet 1*:419-434.

Banner DW, Bloomer AC, Petsko GA, Phillips DC, Pogson CI, Wilson IA, Corran PH, Furth AJ, Milman JD, Offord RE, Priddle JD, Waley SG. 1975. Structure of chicken muscle triosephosphate isomerase determined crystallographically at 2.5 Å resolution using amino acid sequence data. *Nature 255*:609-614.

Bash PA, Field MJ, Davenport RC, Petsko GA, Ringe D, Karplus M. 1991. Computer simulation and analysis of the reaction pathway of triosephosphate isomerase. *Biochemistry 30*:5826-5832.

Blacklow SC, Liu KD, Knowles JR. 1991. Stepwise improvements in catalytic effectiveness: Independence and interdependence in combinations of point mutations of a sluggish triosephosphate isomerase. *Biochemistry 30*:8470-8476.

Bradford MM. 1976. A rapid and sensitive method for the quantitation of protein-dye binding. *Anal Biochem 72*:248-254.

Brändén CI. 1991. The most frequently occurring folding motif in proteins – The TIM barrel. *Curr Opin Struct Biol 1*:978-983.

Brünger AT. 1990. *X-PLOR user manual*. New Haven, Connecticut: Yale University Press.

Chang ML, Artymiuk PJ, Wu X, Hollàn S, Lammi A, Maquat LE. 1993. Human triosephosphate isomerase deficiency resulting from mutation of Phe 240. *Am J Hum Genet 52*:1260-1269.

Cheng J, Mielnicke LM, Pruitt SC, Maquat LE. 1990. Nucleotide sequence of murine triosephosphate isomerase cDNA. *Nucleic Acids Res 18*:4261.

Corran PH, Waley SG. 1975. The amino acid sequence of the rabbit muscle triosephosphate isomerase. *Biochem J 145*:335-344.

Crowther RA. 1972. The fast rotation function. In: Rossmann MG, ed. *The molecular replacement method. Int Sci Rev Ser 13*. New York: Gordon & Breach. pp 173-178.

Crowther RA, Blow DM. 1967. A method of positioning a known molecule in an unknown crystal structure. *Acta Crystallogr 23*:544-548.

Daar IO, Artymiuk PJ, Phillips DC, Maquat LE. 1986. Human triosephosphate isomerase deficiency: A single amino acid substitution results in a thermolabile enzyme. *Proc Natl Acad Sci USA 83*:7903-7907.

Davenport RC, Bash PA, Seaton BA, Karplus M, Petsko GA, Ringe D. 1991. Structure of the triosephosphate isomerase-phosphoglycolohydroxamate complex: An analogue of the intermediate on the reaction pathway. *Biochemistry 30*:5821-5826.

De Moerlooze L, Struman L, Renard A, Martial JA. 1992. Stabilization of T7-promoter-based pARHS expression vectors using the parB locus. *Gene 119*:91-93.

Devereux J, Haeberli P, Smithies O. 1984. A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res 12*:387-395.

Eber SW, Pekrun A, Bardosi A, Gahr M, Krietsch WKG, Kruger J, Matthei R, Schroter W. 1991. Triosephosphate isomerase deficiency: Haemolytic anaemia, myopathy with altered mitochondria and mental retardation due to a new variant with accelerated enzyme catabolism and diminished specific activity. *Eur J Pediatr 150*:761-766.

Eikmanns BJ. 1992. Identification, sequence analysis and expression of a *Corynebacterium glutamicum* gene cluster encoding the three glycolytic enzymes, glyceraldehyde-3-phosphate dehydrogenase, 3-phosphoglycerate kinase and triosephosphate isomerase. *J Bacteriol 174*:6076-6086.

Fairlamb AH. 1989. Novel biochemical pathways in parasitic protozoa. *Parasitology 99*:s93-s112.

Farber GK. 1993. An α/β-barrel full of evolutionary trouble. *Curr Opin Struct Biol 3*:409-412.

Fothergill-Gilmore L, Michels P. 1993. Evolution of glycolysis. *Progr Biophys Mol Biol 59*:105-235.

Friedman KD, Rosen NL, Newman PJ, Montgomery RR. 1990. Screening of λgt11 libraries. In: *PCR protocols: A guide to methods and applications*. San Diego: Academic Press. pp 253-258.

Goraj K, Renard A, Martial JA. 1990. Synthesis, purification and initial structural characterization of octarellin, a de novo polypeptide modelled on the α/β barrel proteins. *Protein Eng 3*:259-266.

Herzberg O, Moult J. 1991. Analysis of the steric strain in the polypeptide backbone of protein molecules. *Proteins Struct Funct Genet 11*:223-229.

Hol WGJ, van Duijnen PT, Berendsen HJC. 1978. The α-helix dipole and the properties of proteins. *Nature 273*:443-446.

Jones TA, Zou JY, Cowan SW. 1991. Improved methods for building models in electron density maps and location of errors in these models. *Acta Crystallogr A 47*:110-119.

Kabsch W. 1988. Evaluation of single-crystal diffraction data from a position-sensitive detector. *J Appl Crystallogr 21*:916–924.

Knowles JR. 1991. Enzyme catalysis: Not different, just better. *Nature 350*:121–124.

Knowles JR, Albery WJ. 1977. Perfection in enzyme catalysis: The energetics of triosephosphate isomerase. *Acc Chem Res 10*:105–111.

Kolb E, Harris JI, Bridgen J. 1974. Triosephosphate isomerase from the coelcanth. *Biochem J 137*:185–197.

Komives EA, Chang LC, Lolis E, Tilton RF, Petsko GA, Knowles JR. 1991. Electrophilic catalysis in triosephosphate isomerase. The role of histidine-95. *Biochemistry 30*:3011–3019.

Kraulis PJ. 1991. MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures. *J Appl Crystallogr 24*:946–950.

Lesk AM, Brändén CI, Chothia C. 1989. Structural principles of α/β proteins: The packing of the interior of the sheet. *Proteins Struct Funct Genet 5*:139–148.

Lolis E, Alber T, Davenport RC, Rose D, Hartman FC, Petsko GA. 1990. Structure of yeast triosephosphate isomerase at 1.9 Å resolution. *Biochemistry 29*:6609–6618.

Lolis E, Petsko GA. 1990. Crystallographic analysis of the complex between triosephosphate isomerase and 2-phosphoglycolate at 2.5 Å resolution: Implications for catalysis. *Biochemistry 29*:6619–6625.

Lu HS, Yuan PM, Gracy RW. 1984. Primary structure of human triosephosphate isomerase. *J Biol Chem 259*:11958–11968.

Mainfroid V, Goraj K, Rentier-Delrue F, Houbrechts A, Loiseau A, Gohimont AC, Noble MEM, Borchert TV, Wierenga RK, Martial JA. 1993. Replacing the $(\beta\alpha)$-unit 8 of *E. coli* TIM with its chicken homologue leads to a stable and active hybrid enzyme. *Protein Eng 6*:893–900.

Mande SC, Suguna K. 1989. Fast algorithm for macromolecular packing calculation. *J Appl Crystallogr 22*:627–629.

Maquat LE, Chilcote R, Ryan PM. 1985. Human triosephosphate isomerase cDNA and protein structure: Studies of triosephosphate isomerase deficiency in man. *J Biol Chem 260*:3748–3753.

Marchionni M, Gilbert W. 1986. The triosephosphate isomerase gene from maize: Introns antedate the plant animal divergence. *Cell 46*:133–141.

Matthews BW. 1968. Solvent content of protein crystals. *J Mol Biol 33*: 491–497.

Matthews BW. 1993. Structural and genetic analysis of protein stability. *Annu Rev Biochem 62*:139–160.

McKnight GL, O'Hara PJ, Parker ML. 1986. Nucleotide sequence of the triosephosphate isomerase gene from *Aspergillus nidulans*: Implications for a difference loss of introns. *Cell 46*:143–147.

Messerschmidt A, Pflugrath J. 1987. Crystal orientation and X-ray pattern prediction routines for area detector diffractometer systems in macromolecular crystallography. *J Appl Crystallogr 20*:436–439.

Misset O, Opperdoes FR. 1984. Simultaneous purification of hexokinase, class-I fructose-bisphosphate aldolase, triosephosphate isomerase and phosphoglycerate kinase from *Trypanosoma brucei*. *Eur J Biochem 144*:475–483.

Noble MEM, Wierenga RK, Lambeir AM, Opperdoes FR, Thunnissen AM, Kalk KH, Groendijk H, Hol WGJ. 1991. The adaptability of the active site of trypanosomal triosephosphate isomerase as observed in the crystal structures of three different complexes. *Proteins Struct Funct Genet 10*:50–69.

Noble MEM, Zeelen JP, Wierenga RK. 1993a. Structures of the "open" and "closed" states of trypanosomal triosephosphate isomerase, as observed in a new crystal form: Implications for the reaction mechanism. *Proteins Struct Funct Genet 16*:311–326.

Noble MEM, Zeelen JP, Wierenga RK, Mainfroid V, Goraj K, Gohimont AC, Martial JA. 1993b. Structure of triosephosphate isomerase from *Escherichia coli* determined at 2.6 Å resolution. *Acta Crystallogr D 49*:403–417.

Okada N, Koizumi N, Tanaka T, Ohkubo H, Nakanishi S, Yamada Y. 1989. Isolation, sequence and bacterial expression of a cDNA for (s)-tetrahydroberberine oxidase from cultured berberine-producing *Coptis japonica* cells. *Proc Natl Acad Sci USA 86*:534–538.

Old SE, Mohrenweiser HW. 1988. Nucleotide sequence of the triosephosphate isomerase gene from *Macaca mulatta*. *Nucleic Acids Res 16*:9055.

Perry BA, Mohrenweiser HW. 1992. Human triosephosphate isomerase: Substitution of Arg for Gly at position 122 in a thermolabile electromorph variant, TPI-Manchester. *Hum Genet 88*:634–638.

Pichersky E, Gottlieb LD, Hess JF. 1984. Nucleotide sequence of the triosephosphate isomerase gene of *Escherichia coli*. *Mol Gen Genet 195*:314–320.

Ramachandran GN, Sasisekharan V. 1968. Conformation of polypeptides and proteins. *Adv Protein Chem 23*:283–438.

Read RJ. 1986. Improved Fourier coefficients for maps using phases from partial structures with errors. *Acta Crystallogr A 42*:140–149.

Rentier-Delrue F, Mande SC, Moyens S, Mainfroid V, Goraj K, Lion M, Hol WGJ, Martial JA. 1993. Cloning and overexpression of the triosephosphate isomerase genes from psychrophilic and thermophilic bacteria. *J Mol Biol 229*:85–93.

Richardson JS, Richardson DC. 1988. Amino acid preferences for specific locations at the ends of α-helices. *Science 240*:1648–1652.

Russell PR. 1985. Transcription of the triosephosphate isomerase gene of *Schizosaccharomyces pombe* initiates from a start point different from that in *Saccharomyces cerevisiae*. *Gene 40*:125–130.

Shaw-Lee RL, Lissemore JL, Sullivan DT. 1991. Structure and expression of the triosephosphate isomerase gene of *Drosophila melanogaster*. *Mol Gen Genet 230*:225–229.

Straus D, Gilbert W. 1985. Chicken triosephosphate isomerase complements an *Escherichia coli* deficiency. *Proc Natl Acad Sci USA 82*:2014–2018.

Studier FW, Moffatt BA. 1986. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J Mol Biol 189*: 113–130.

Swinkels BW, Gibson WC, Osinga KA, Kramer R, Veeneman GH, van Boom JH, Borst P. 1986. Characterization of the gene for microbody (glycosomal) triosephosphate isomerase of *Trypanosoma brucei*. *EMBO J 5*: 1291–1298.

Tittiger C, Whyard S, Walker VK. 1993. A novel intron site in the triosephosphate isomerase gene from the mosquito *Culex tarsalis*. *Nature 361*:470–472.

Urfer R, Kirschner K. 1992. The importance of surface loops for stabilizing an eightfold $\beta\alpha$ protein. *Protein Sci 1*:31–45.

Verlinde CLMJ, Noble MEM, Kalk KH, Groendijk H, Wierenga RK, Hol WGJ. 1991. Anion binding at the active site of trypanosomal triosephosphate isomerase. Monohydrogen phosphate does not mimic sulphate. *Eur J Biochem 198*:53–57.

Verlinde CLMJ, Rudenko G, Hol WGJ. 1992a. In search of new lead compounds for trypanosomiasis drug design: A protein structure-based linked-fragment approach. *J Comput Aided Mol Design 6*:131–147.

Verlinde CLMJ, Witmans CJ, Pijning T, Kalk K, Hol WGJ, Callens M, Opperdoes FR. 1992b. Structure of the complex between trypanosomal triosephosphate isomerase and *N*-hydroxy-4-phosphono-butanamide: Binding at the active site despite an "open" flexible loop conformation. *Protein Sci 1*:1578–1584.

WHO. 1992. *Eleventh program report, tropical disease research*. Geneva: World Health Organization.

Wierenga RK, Noble MEM, Davenport RC. 1992. Comparison of the refined crystal structures of liganded and unliganded chicken, yeast and trypanosomal triosephosphate isomerase. *J Mol Biol 224*:1115–1126.

Wierenga RK, Noble MEM, Postma JPM, Groendijk H, Kalk KH, Hol WGJ, Opperdoes FR. 1991a. The crystal structure of the "open" and the "closed" conformation of the flexible loop of trypanosomal triosephosphate isomerase. *Proteins Struct Funct Genet 10*:33–49.

Wierenga RK, Noble MEM, Vriend G, Nauche S, Hol WGJ. 1991b. Refined 1.83 Å structure of trypanosomal triosephosphate isomerase crystallized in the presence of 2.4 M-ammonium sulphate. A comparison with the structure of the trypanosomal triosephosphate isomerase-glycerol-3-phosphate complex. *J Mol Biol 220*:995–1015.

Wolfenden R. 1969. Transition state analogues for enzyme catalysis. *Nature 233*:704–705.