Crystal structure of recombinant triosephosphate isomerase from *Bacillus stearothermophilus*. An analysis of potential thermostability factors in six isomerases with known three-dimensional structures points to the importance of hydrophobic interactions

LUIS F. DELBONI,^{1,5} SHEKHAR C. MANDE,¹ FRANÇOISE RENTIER-DELRUE,³
VÉRONIQUE MAINFROID,³ STEWART TURLEY,¹ FREDERIQUE M.D. VELLIEUX,⁴
JOSEPH A. MARTIAL,³ AND WIM G.J. HOL^{1,2}

¹ Department of Biological Structure, Biomolecular Structure Center and

² Howard Hughes Medical Institute, SL-15, School of Medicine, University of Washington, Seattle, Washington 98195

³ Laboratoire de Biologie Moléculaire et de Génie Génétique, Université de Liège, B6, Sart Tilman, 4000 LIEGE, Belgium

⁴ LCCP, CNRS 38027 Grenoble, Cedex 01, France

(RECEIVED July 31, 1995; ACCEPTED October 5, 1995)

Abstract

The structure of the thermostable triosephosphate isomerase (TIM) from *Bacillus stearothermophilus* complexed with the competitive inhibitor 2-phosphoglycolate was determined by X-ray crystallography to a resolution of 2.8 Å. The structure was solved by molecular replacement using XPLOR. Twofold averaging and solvent flattening was applied to improve the quality of the map. Active sites in both the subunits are occupied by the inhibitor and the flexible loop adopts the “closed” conformation in either subunit. The crystallographic *R*-factor is 17.6% with good geometry. The two subunits have an RMS deviation of 0.29 Å for 248 C α atoms and have average temperature factors of 18.9 and 15.9 Å², respectively. In both subunits, the active site Lys 10 adopts an unusual ϕ, ψ combination.

A comparison between the six known thermophilic and mesophilic TIM structures was conducted in order to understand the higher stability of *B. stearothermophilus* TIM. Although the ratio Arg/(Arg+Lys) is higher in *B. stearothermophilus* TIM, the structure comparisons do not directly correlate this higher ratio to the better stability of the *B. stearothermophilus* enzyme. A higher number of prolines contributes to the higher stability of *B. stearothermophilus* TIM. Analysis of the known TIM sequences points out that the replacement of a structurally crucial asparagine by a histidine at the interface of monomers, thus avoiding the risk of deamidation and thereby introducing a negative charge at the interface, may be one of the factors for adaptability at higher temperatures in the TIM family. Analysis of buried cavities and the areas lining these cavities also contributes to the greater thermal stability of the *B. stearothermophilus* enzyme. However, the most outstanding result of the structure comparisons appears to point to the hydrophobic stabilization of dimer formation by burying the largest amount of hydrophobic surface area in *B. stearothermophilus* TIM compared to all five other known TIM structures.

Keywords: hydrophobicity; thermostability; triosephosphate isomerase

Reprint requests to: Wim G.J. Hol, Biomolecular Structure Center and the Howard Hughes Medical Institute, University of Washington, Box 357742, Seattle, Washington 98195; e-mail: sue@blackhook.bchem.washington.edu or hol@gouda.bchem.washington.edu.

⁵ Present address: Departamento de Física e Ciências dos Materiais, Universidade de São Paulo, Caixa Postal 363, São Carlos, SP Brazil 13560.

Triosephosphate isomerase (TIM) catalyzes the interconversion of D-glyceraldehyde-3-phosphate (GAP) and dihydroxyacetone phosphate (DHAP) in the glycolytic pathway. Only the dimer is fully active (Waley, 1973; Zabori et al., 1980). Catalytic properties of TIM have been studied in detail by numerous investigators, employing kinetic experiments (Knowles & Albery,

1977), structural analysis of transition-state analogues complex (Lolis & Petsko, 1990; Davenport et al., 1991), site-directed mutagenesis experiments (Blacklow et al., 1991), and quantum and molecular mechanical studies (Bash et al., 1991). The fundamental catalytic residues are Lys 10, His 94, and Glu 166, using *Bacillus stearotheophilus* amino acid sequence numbers. It is well known that the catalytic mechanism is accompanied by a conformational change in the active site. The loop formed by residues 168–178, called the “flexible loop,” closes the active site when the substrate binds (Knowles, 1991). Several substrate analogue inhibitor structures have already been reported showing three different conformations of the flexible loop: “open” when there is no ligand at the active site (Noble, 1992), “partially closed” when sulfate is bound (Wierenga et al., 1991a), and “closed” when substrate analogue inhibitors are bound (Lolis & Petsko, 1990; Noble et al., 1991a, 1991b). A unique case has been reported in which *N*-hydroxy-4-phosphono-butanamide, an inhibitor that is one carbon atom longer than the substrate, binds with the open conformation of the loop (Verlinde et al., 1992).

At present, 18 sequences of triosephosphate isomerases are known (Fig. 1) and three-dimensional structures from five different species have been reported: chicken TIM (Banner et al., 1975); yeast TIM (Lolis et al., 1990); trypanosomal TIM (Wierenga et al., 1991b); *Escherichia coli* TIM (Noble et al., 1993); and human TIM (Mande et al., 1994). A detailed comparison of the structures of chicken, yeast, and trypanosomal TIMs has also been carried out (Wierenga et al., 1992). However, none of these enzymes is thermostable. We report in this paper the sixth TIM structure, that of *B. stearotheophilus*, and compare the structure to other TIMs in order to find thermostability factors of *B. stearotheophilus* TIM.

B. stearotheophilus TIM is a dimer formed by two monomers of 252 amino acids each. TIM from *B. stearotheophilus* is active up to 65 °C, significantly higher than *E. coli*, which is active up to around 50 °C (Rentier-Delrue et al., 1993). The former protein therefore is characterized as thermophilic and the latter as mesophilic. The comparison of sequence and structure of *B. stearotheophilus* TIM with their homologues may show the reasons for greater thermostability of the former enzyme. We do not expect to find a single outstanding difference that can account for the greater thermostability of the *B. stearotheophilus* protein, but there may be several small contributions to thermostability which, when added up, can explain the difference between the mesophilic and moderate thermophilic proteins. Factors like hydrogen bonds, salt bridges, and hydrophobicity have been discussed by several authors (Creighton, 1991; Fontana, 1991; Matthews, 1993) as being important in thermostability. The reasons for thermostability vary from protein to protein (Jaenicke, 1991) even though the energy differences involved are small (Rees & Adams, 1995). This means that proteins may have achieved an increased thermostability by different combinations of small modifications that have improved the stability of the monomer as well as the oligomer. The results of our comparisons of six TIM structures are reported in this paper.

Results

Quality of the structure

The crystallographic studies resulted in a well-defined 2.8-Å structure of *B. stearotheophilus* TIM with an *R*-factor of 0.176 and good stereochemistry. Refinement statistics of the

Copja	1	10	20	30	40	50	60	70	80
Maize	NGKFFVVG	WKKCN	GTSE	QVKKIVTLIN	EA	EV	PSE	DVVEV	VSPPPVLPF
Human	APSRKFFVVG	WKKCN	GTTO	QVKKIVTLIN	EA	EV	PSE	DVVEV	VSPPPVLPF
Maize	APSRKFFVVG	WKKCN	GTTO	QVKKIVTLIN	EA	EV	PSE	DVVEV	VSPPPVLPF
Rabbit	APSRKFFVVG	WKKCN	GRKK	NGSLITLTLN	AA	EV	PAD	T	EVV
Mouse	APSRKFFVVG	WKKCN	GRKK	NGSLITLTLN	AA	EV	PAD	T	EVV
Chick	APSRKFFVVG	WKKCN	GRKK	NGSLITLTLN	AA	EV	PAD	T	EVV
Latch	APSRKFFVVG	WKKCN	GRKK	NGSLITLTLN	AA	EV	PAD	T	EVV
Culta	APSRKFFVVG	WKKCN	GRKK	NGSLITLTLN	AA	EV	PAD	T	EVV
Drome	APSRKFFVVG	WKKCN	GRKK	NGSLITLTLN	AA	EV	PAD	T	EVV
Emeni	APSRKFFVVG	WKKCN	GRKK	NGSLITLTLN	AA	EV	PAD	T	EVV
Schpo	APSRKFFVVG	WKKCN	GRKK	NGSLITLTLN	AA	EV	PAD	T	EVV
Yeast	APSRKFFVVG	WKKCN	GRKK	NGSLITLTLN	AA	EV	PAD	T	EVV
Trybb	APSRKFFVVG	WKKCN	GRKK	NGSLITLTLN	AA	EV	PAD	T	EVV
Bacst	APSRKFFVVG	WKKCN	GRKK	NGSLITLTLN	AA	EV	PAD	T	EVV
Ecoli	APSRKFFVVG	WKKCN	GRKK	NGSLITLTLN	AA	EV	PAD	T	EVV
Corgl	APSRKFFVVG	WKKCN	GRKK	NGSLITLTLN	AA	EV	PAD	T	EVV
Morsp	APSRKFFVVG	WKKCN	GRKK	NGSLITLTLN	AA	EV	PAD	T	EVV
Consen	APSRKFFVVG	WKKCN	GRKK	NGSLITLTLN	AA	EV	PAD	T	EVV

Fig. 1. Sequence alignment by using the program GCG (Genetics Computer Group, Inc., Copyright 1982–1992, Wisconsin, USA) of all known TIM sequences. In order: Copja, *Coptis japonica* (Okada et al., 1989); Maize, *Zea mays* (maize) (Marchionni & Gilbert, 1986); Human, *Homo sapiens* (Maquat et al., 1985); Macmu, *Macaca mulatta* (rhesus macaque) (Old & Mohrenweiser, 1988); Rabbit, *Oryctolagus cuniculus* (rabbit) (Corran & Waley, 1975); Mouse, *Mus musculus* (mouse) (Cheng et al., 1990); Chick, *Gallus gallus* (chicken) (Straus & Gilbert, 1985); Latch, *Latimeria chalumnae* (latimeria) (Kolb et al., 1974); Culta, *Culex tarsalis* (mosquito) (Tittiger et al., 1993); Drome, *Drosophila melanogaster* (fruit fly) (Shaw-Lee et al., 1991); Emeni, *Emericella nidulans* (*Aspergillus nidulans*) (McKnight et al., 1986); Schpo, *Schizosaccharomyces pombe* (fission yeast) (Russell, 1985); Yeast, *Saccharomyces cerevisiae* (baker's yeast) (Alber & Kawasaki, 1982); Trybb, *Trypanosoma brucei* (Swinkels et al., 1986); Bacst, *Bacillus stearotheophilus* (Rentier-Delrue et al., 1993); Ecoli, *Escherichia coli* (Pichersky et al., 1984); Corgl, *Corynebacterium glutamicum* (Eikmanns, 1992); Morsp, *Moraxella sp* (Rentier-Delrue et al., 1993). The last line shows the consensus of residues that are conserved in all TIM sequences.

Fig. 2. ϕ, ψ plot of both subunits of *B. stearotherophilus* TIM. For details about unallowed ϕ, ψ combination, see text.

2.8-Å structure of *B. stearotherophilus* TIM are shown in Table 1. There are eight residues lying outside the allowed region in the Ramachandran plot (Fig. 2). Although the residues Thr 64, Met 65, Ala 102, Ser 212, Met 365, and Ser 512 are very close to the allowed regions (residue numbers 1–252 refer to “first” subunit, 301–552 to the “second” subunit), Lys 10 and

Table 1. Refinement statistics of *B. stearotherophilus* TIM structure^a

Parameter	Value
<i>R</i> -factor	0.176
Resolution range	8.0–2.8 Å
Number of reflections	12,119
Model:	
Number of protein atoms	3,769
Number of ligand atoms	18
RMS deviations from ideality:	
Covalent bond lengths	0.010 Å
Bond angles	1.65°
Dihedral angles	24.3°
Improper angles	1.48°

^a Improper angles are used to maintain planarity or chirality of the relevant groups during refinement.

Lys 310 are far from allowed ϕ, ψ combinations. This lysine has the same “unallowed” ϕ, ψ combination in other TIM structures elucidated so far, including the higher resolution structures of trypanosomal as well as yeast TIMs (Lolis et al., 1990; Wierenga et al., 1992; Noble et al., 1993; Mande et al., 1994). In the *B. stearotherophilus* structure, all atoms of this residue have low temperature factors. It is well known that this lysine residue is essential for activity (Straus et al., 1985). Lys 10 (and Lys 310) makes very strong hydrogen bonds with one of the carboxyl oxygens of the inhibitor 2PG (Fig. 3).

Almost all the residues of the dimer are well defined in the structure. However, the C-terminal residues Glu 252 and Glu 552

Fig. 3. Active site of subunit 1 of *B. stearotherophilus* TIM showing the interactions between the inhibitor (2PG) and the enzyme molecules. All hydrogen bonds are shown as dashed lines. For details about hydrogen bonds of 2PG and TIM, see Table 2.

exhibited no density in any electron density map and were not included in the model. Side chains of residues Lys 175, Gln 225, His 251, from subunit 1, and Gln 320 and Glu 538, from subunit 2, have no density and were treated as alanine. The equivalent residues in the other subunit have clear density due to crystal contacts. Gln 20 is the only solvent-exposed residue that does not make any interaction with other atoms but is still visible.

Structural features of *B. stearrowthermophilus* TIM

Overall comparison of the monomers

The overall structure of the *B. stearrowthermophilus* TIM is shown in Figure 4 and Kinemage 1 and is very similar to the other TIM structures reported. Minor differences occur in the loop regions, as for instance in (1) the loop defined by residues 25–36, (2) loop 5 (which comprises residues 127–137, including a small helix), and (3) residues 152–155. These residues are not well conserved in other sequences (Fig. 1). The flexible loop (loop 6), which shows variable conformations in different TIM structures due to differences in crystallization conditions and presence or absence of inhibitors (Wierenga et al., 1992), adopts the “closed” conformation in both independent subunits of the *B. stearrowthermophilus* TIM structure.

The two refined monomers in the asymmetric unit have an RMS difference of 0.44 Å when all 1,003 equivalent main-chain atoms are superposed. The RMS difference drops to 0.29 Å when 31 main-chain atoms deviating more than three times the overall standard deviation are removed during superpositions (Fig. 5). The regions that differ more than 1.5 Å are either involved in crystal contacts or belong to the poorly defined C-terminal region. The largest deviation is shown by residue 54, where the C α of residue Asp 54 is 1.9 Å apart from residue Asp 354. The side chain of Asp 54 makes extensive crystal contacts with a symmetry-related subunit 2 through well-defined hy-

Fig. 4. Subunit 1 of *B. stearrowthermophilus* TIM as viewed down the β -barrel axis. The active site residues Lys 10, His 94, and Glu 166 2PG molecule are shown. Also indicated are two sites of interest, 12 and 190, for analysis of thermostability factors, and Tyr 209 for the variability among TIM structures in regard to interaction between TIM and the inhibitor. The flexible loop is shown in black. Helix, β -strands, and loops have been assigned by DSSP program (Kabsch & Sander, 1983). The figure was prepared by using MOLSCRIPT (Kraulis, 1991).

Fig. 5. RMS deviation between subunit 1 and 2. This plot was calculated for all main-chain atoms. For details, see text.

drogen bonds with residues Arg 495, His 488, and Ser 491, and has relatively low *B*-factors for the oxygens O $^{\delta 1}$ and O $^{\delta 2}$ (27.9 Å 2 and 19.2 Å 2 , respectively). On the other hand, Asp 354(O $^{\delta 2}$) in the other subunit makes only one hydrogen bond, with Lys 326(N $^{\epsilon}$) from the same subunit, whereas the other oxygen points to the solvent. Also, the *B*-values of O $^{\delta 1}$ and O $^{\delta 2}$ are significantly higher (65.9 Å 2 and 66.9 Å 2 , respectively) for Asp 354 compared with those of Asp 54. Hence, the largest difference in structure is the result of differences in crystal contacts.

Average main-chain *B*-factors for the two subunits are 18.9 and 15.9 Å 2 , respectively, and the average side-chain *B*-factors are 23.0 and 19.9 Å 2 , respectively. The *B*-factor plot of both subunits can be seen in Figure 6 and Kinemage 2. Quite intriguingly, there is a significant difference in the *B*-factors of the N-terminal and C-terminal halves of subunit 1. The average *B*-factor for the 120 N-terminal residues is 13.8 Å 2 and the average *B*-factor for the 131 C-terminal residues is 23.6 Å 2 for subunit 1. This difference is not observed in subunit 2. The reason for this contrast in temperature factors of the N-terminal and C-terminal halves of subunit 1 becomes apparent when intersubunit interactions and crystal contacts are evaluated. The N-terminal half of each subunit is involved in subunit–subunit interactions, therefore perhaps has intrinsically lower *B*-factors than the C-terminal halves. The C-terminal part of subunit 1 shows the lowest number of crystal contacts which could explain the higher average *B*-factor for this part of subunit 1. On the other hand, the C-terminal half of subunit 2 is involved in extensive crystal contacts and thus has temperature factors comparable to its N-terminal half. Thus, an unusual combination of intersubunit interactions and crystal contacts seems to manifest itself as a variation in *B*-factors of the two halves of subunit 1.

The two monomers in the asymmetric unit are related by a nearly perfect twofold symmetry axis with a 179.2° rotation and a translation component of 0.004 Å along the local symmetry axis. There are three insertions of one residue each in the *B. stearrowthermophilus* TIM sequence when compared to the *Trypanosoma brucei* model used in molecular replacement. These are all in the loop regions. Two of these insertions were very easily modeled into the density, but the last insertion was only possible to model late in the refinement through analysis based on an omit map.

Fig. 6. Main-chain *B*-factor plot of (A) subunit 1 and (B) subunit 2. The contrast in temperature factors between the N- and C-terminal halves of subunit 1 is evident. Such a difference in *B*-factors is not observed in subunit 2. For details, see text.

Active site and 2-PG binding

Both active sites of the *B. stearotherophilus* TIM contain the inhibitor 2PG and Kinemage 1. The active site Lys 10 and His 94 are well positioned to carry out the enzymatic reaction. The third residue essential for the TIM activity, Glu 166, adopts the "swung out" ($\chi_1 = -55^\circ$ in subunit 1 and $\chi_1 = -58^\circ$ in subunit 2) conformation in both subunits, which is reported to be unsuitable for the catalysis (Wierenga et al., 1992). The flexible loop adopts the "closed" conformation in both subunits, in agreement with previous results (Lolis & Petsko, 1990; Wierenga et al., 1992).

The interactions of 2PG with both subunits of *B. stearotherophilus* TIM are very similar to the previous TIM structures reported (Lolis & Petsko, 1990; Mande et al., 1994) (Table 2). The phosphate oxygens make hydrogen bonds with the main-chain amino group of residues Gly 172, Ser 212, Gly 233, and Gly 234 (Fig. 3). The main-chain amino group of Ser 212 also makes a hydrogen bond with Ala 170. The equivalent serines in other structures, namely, yeast TIM complexed with 2PG (Lolis et al., 1990), subunit 2 of *T. brucei* TIM complexed with 2-phosphoglycerate (Noble et al., 1991a), and glycerol-3-phosphate (Wierenga et al., 1991b), show the same hydrogen bond with the inhibitor. However, among these structures, only *T. brucei* TIM in complex with 2-phosphoglycerate exhibits the hydrogen bond interaction between Ser 212 and Ala 170.

Table 2. Hydrogen bonds between 2-PG and *B. stearotherophilus* TIM

2-PG	TIM	Distance (Å)
Subunit 1		
O11	His 94 N ^{ε2}	3.0
O11	Lys 10 N ^δ	2.8
O12	His 94 N ^{ε2}	2.9
O14	Ser 212 N	2.9
O14	Gly 172 N	2.7
O15	Gly 234 N	3.0
O16	Gly 233 N	3.0
Subunit 2		
O21	His 94 N ^{ε2}	2.8
O21	Lys 10 N ^δ	2.7
O22	His 94 N ^{ε2}	3.0
O24	Gly 172 N	3.3
O25	Gly 234 N	3.3
O26	Gly 312 N	2.8
O26	Gly 172 N	2.9

Among the conformational changes that take place upon substrate binding is a flip by almost 180° of the peptide between Gly 211 and Ser 212. This rotation places the main-chain nitrogen of Ser 212 in a favorable position to interact with the phosphate oxygens of the substrate. The carbonyl oxygen of Gly 211, which now points inward as a result of this change, finds a hydrogen bonding partner in the hydroxyl group of Tyr 209. Also, the hydroxyl group of Tyr 209 hydrogen bonds with the main-chain amino group of Ser 177. Both these interactions of the Tyr 209 hydroxyl group are also observed in yeast TIM complexed with 2PG and in the second subunit of *T. brucei* TIM complexed with glycerol-3-phosphate, sulfate, and 2-phosphoglycerate. On the other hand, this interaction was not observed in human TIM (Mande et al., 1994; despite the fact that subunit 2 was complexed with 2PG), neither in *E. coli* TIM (Noble et al., 1993) nor in chicken TIM (Banner et al., 1975), or in subunit 1 of *T. brucei* TIM complexed with 2-phosphoglycerate (Noble et al., 1991a), sulphate, and glycerol-3-phosphate (Wierenga et al., 1991b). In all the structures in which Ser 211 is not in the same conformation as in that of *B. stearotherophilus*, the hydroxyl oxygen of Tyr 209 has no partner for hydrogen bonds, which is rather unusual. Due to these differences in different structures, the conformation adopted by this residue during the isomerization reaction still remains a mystery.

The N^δ of active site residue Lys 10 makes a hydrogen bond with one of the carboxyl group oxygens of 2PG. The N^{ε2} of His 94 makes two hydrogen bonds with the carboxyl group oxygens of 2PG (Fig. 3). The bridging oxygen of 2PG makes a hydrogen bond with the Gly 233 amino group. Hence, each oxygen of the inhibitor is making specific interactions with the enzyme.

The C^α of Gly 172 moves approximately 8 Å from the molecular replacement model of *T. brucei* TIM, in which the flexible loop was in the "open" conformation, to the final model of *B. stearotherophilus* TIM, in which the flexible loop is in the "closed" conformation. Yet, among the flexible loop residues, Gly 172 is the only residue that makes a hydrogen bond with 2PG. The other contacts between the loop in the closed con-

formation and the protein involve only extensive Van der Waals interactions by Ile 171. Apparently, the large conformational change of the flexible loop upon inhibitor binding is due to rather subtle interactions of the loop with inhibitor and protein.

Comparison of TIM structures

The overall structure of *B. stearothermophilus* TIM is very similar to the five other known TIM structures. Differences primarily exist in loop regions. The alignment of all known TIM sequences can be seen in Figure 1. A list of the reported TIM structures and the RMS difference between structurally equivalent C α positions of *B. stearothermophilus* and respective TIM structures is shown in Table 3. Among the TIMs with known three-dimensional structure, *B. stearothermophilus* TIM is closest to chicken TIM, with 42.3% sequence identity. The three-dimensional superpositions show that it deviates most from *E. coli* TIM, with an RMS difference of 1.7 Å between structurally equivalent C α positions.

Search for stability factors of *B. stearothermophilus* TIM

Factors that affect thermostability of proteins have been a subject of numerous investigations (see the reviews by Ikai, 1980; Merkle et al., 1981; Vihinen, 1987; Menendez-Arias & Argos, 1989; Zhang et al., 1991; Eriksson et al., 1992; Matthews, 1993; Vriend & Eijssink, 1993). A comparison of structures from thermophilic and mesophilic organisms has been reported recently for glyceraldehyde-3-phosphate dehydrogenase (Korn-dörfer et al., 1995), 3-phosphoglycerate kinase (Davies et al., 1993), and malate dehydrogenase (Kelly et al., 1993). TIMs, however, form an excellent example for such analysis because in this family six three-dimensional structures are now available, four mesophilic (those of chicken, human, *T. brucei*, and *E. coli*), one close to thermophilic (yeast), and one thermophilic (*B. stearothermophilus*). Nonetheless, a comparison based on structural features of each protein is difficult because each protein structure was determined by a different group, refined by a different approach, and at a different resolution (Table 3). Fine details of each structure, such as hydrogen bonds, salt bridges, etc., depend to a considerable extent on the program used in the refinement. With these inadequacies in mind, we have analyzed a limited set of characteristics of the six structures

in an attempt to understand higher stability of *B. stearothermophilus* TIM.

Sequences and stability

Amino acid compositions of all the TIMs with known three-dimensional structures are listed in Table 4. The table does not show any striking differences in the amino acid composition of *B. stearothermophilus* TIM compared to its mesophilic homologues. In the following paragraphs, two differences in the amino acid compositions and their significance to thermostability, if any, are analyzed.

Arginines. The number of arginines, or, rather, a larger Arg/(Arg+Lys) ratio, has been reported to be higher in thermostable proteins (Merkler et al., 1981; Mrabet et al., 1992). It is evident from Table 4 that this ratio is indeed higher in *B. stearothermophilus* TIM than in other TIMs. However, this is due to a lower content of lysines rather than a higher content of arginines. In order to explore if arginines indeed play a role in stabilization of *B. stearothermophilus* TIM, we examined their interactions in the TIM structures (Kinemage 3). As anticipated, almost all the arginines in *B. stearothermophilus* TIM reside on the surface of the protein, with the single exception of Arg 190, which is completely buried. The guanidinium moiety of Arg 190 makes extensive hydrogen bonds with the carboxyl groups of Asp 228 and the carbonyl oxygens of Ala 202, Ile 205 and the O $^{1-}$ of Gln 226.

The interactions of arginines with other residues in *B. stearothermophilus* TIM are in good agreement with those observed in other proteins (Nandi et al., 1993; Borders et al., 1994). The most frequent interaction of the guanidinium group is with the charged carboxylate oxygens of Asp and Glu, whereas all other interactions are with main-chain carbonyl oxygens. Stacking interactions of the guanidinium group of arginines with aromatic side chains (Flocco & Mowbray, 1994), as reported for other proteins, are also observed in *B. stearothermophilus* TIM. This is exemplified by an interaction between Arg 97 and Phe 101, in which the guanidinium plane of Arg 97 and the ring of Phe 101 make an angle of around 30°, and a strong van der Waals contact of 3.5 Å. Thus, interactions of arginines with other protein residues in *B. stearothermophilus* TIM follow a pattern similar to that observed for other proteins.

Arg 48 and Arg 219 of *B. stearothermophilus* TIM are of particular interest to analyze with respect to thermostability, because no other sequence has an arginine at this position (Fig. 1).

Table 3. The five reported structures of TIMs: human, chicken, *T. brucei*, yeast, and *E. coli*^a

Source	Resolution (Å)	R-factor	Number of solvent atoms	Number of inhibitor atoms	Refinement protocol	Reference	RMS difference (Å)	% Sequence identity
Human	2.8	0.167	—	9	XPLOR	Mande et al., 1994	1.22	41.1
Chicken	2.5	0.17	—	—		Artymuik, pers. comm.	1.23	42.7
<i>T. brucei</i>	1.83	0.183	278	18	PROLSQ/TNT	Wierenga et al., 1991b	1.43	40.7
Yeast	2.5	0.18	24	18	PROLSQ/XPLOR	Lolis et al., 1990	1.48	38.5
<i>E. coli</i>	2.6	0.119	172	—	XPLOR/TNT	Noble et al., 1993	1.72	39.7
<i>B. stearothermophilus</i>	2.8	0.176	—	18	XPLOR	This work	0.0	100.0

^a The RMS deviations in Å reported are between structurally equivalent C α positions when *B. stearothermophilus* TIM structure is superimposed onto other respective TIM structures.

Table 4. Amino acid composition of all six known TIM structures^a

	Amino acid composition (%)					
	Human	Chicken	<i>T. brucei</i>	Yeast	<i>E. coli</i>	<i>B. stearothermophilus</i>
Gly	10.1	10.9	7.6	8.9	8.6	8.7
Ala	11.3	11.3	13.3	10.1	17.7	11.5
Val	10.1	8.9	10.0	10.5	7.5	9.9
Leu	6.1	6.9	6.8	7.7	6.7	7.5
Ile	6.1	6.9	7.6	6.1	7.5	6.4
Ser	4.8	5.3	7.2	6.5	4.3	5.2
Thr	5.7	4.1	4.8	4.9	3.5	4.8
Phe	3.2	3.2	3.2	4.5	2.4	3.6
Tyr	1.6	1.6	2.0	2.4	2.0	1.6
Trp	2.0	2.0	2.0	1.2	0.8	0.8
Lys	8.1	8.9	6.8	8.5	6.3	4.8
Arg	3.2	2.8	4.0	3.2	3.1	4.4
His	1.6	3.2	2.0	1.2	3.1	2.4
Asp	4.8	5.3	3.2	6.1	3.9	4.8
Glu	6.9	6.9	5.2	6.9	8.2	7.9
Asn	3.2	2.8	4.0	4.9	3.5	2.0
Gln	4.4	3.6	5.2	2.8	4.3	5.6
Cys	2.0	1.6	1.2	0.8	1.2	1.6
Met	0.8	0.8	0.8	0.0	2.8	1.6
Pro	4.0	2.8	2.8	2.8	2.8	5.2

^a Numbers indicated are composition percentage of the entire dimer. It is clear that *B. stearothermophilus* has the highest percentage of Arg and the lowest of Lys. Also, *B. stearothermophilus* has the highest percentage content of Pro.

Of the two arginines, Arg 219 makes a salt bridge with the succeeding residue, Asp 220. The other arginine, Arg 48, is involved in an intersubunit salt bridge with Asp 347. A similar charge-charge interaction equivalent to that between Arg 48 and Asp 347 is also seen in the structures of human, chicken, and yeast TIMs, albeit between Lys 13 and Asp 349 (human and chicken) and Lys 17 and Asp 346 (yeast). The role of Arg 48 of *B. stearothermophilus* TIM in *E. coli* TIM is performed by Arg 18, which is involved in an intersubunit salt bridge with Asp 350. *T. brucei* TIM is the only structure that does not show a similar intersubunit charge-charge interaction in this region. Thus, even at the interface, arginines do not seem to favor better thermostability of *B. stearothermophilus* TIM.

Because the guanidinium group of Arg has an increased hydrogen bonding capability compared to a Lys, we inspected the hydrogen bonding interactions of all arginines of the six TIM structures. The number of hydrogen bonding interactions per an arginine residue for the six TIM structures are 2.27 (*B. stearothermophilus*), 3.64 (chicken), 3.13 (*E. coli*), 2.20 (*T. brucei*), 3.13 (human), and 1.69 (yeast). These interactions obviously do not favor *B. stearothermophilus* TIM over other TIMs. Thus, although all the arginines of *B. stearothermophilus* TIM are involved in extensive interactions, we cannot conclude that any of these interactions contributes specifically to increasing its thermostability when compared to the mesophilic TIMs. Hence, in this case, it is difficult to correlate the higher ratio of Arg/(Arg+Lys) to the better thermostability of *B. stearothermophilus* TIM compared to other species.

Prolines. From Table 4 it is clear that *B. stearothermophilus* TIM has the highest number of prolines (5.2%), being close to the human enzyme (4%) and far exceeding all other TIM se-

quences (2.8%). Several investigators have reported that proline residues are responsible for thermostability by decreasing the entropy of the unfolded state (Watanabe et al., 1991; Nicholson et al., 1992; Hardy et al., 1993). Of the 13 proline residues in *B. stearothermophilus* TIM (Kinemage 3), 5 are located at the N1 position of the helices, 4 in the loops joining α -helices and successive β -strands, and 3 between β -strands and following α -helices. Prolines in proteins are known to have preference for the N1 position of helices (Richardson & Richardson, 1988). Among the TIM structures, chicken and human TIMs have three, *E. coli* and yeast TIMs have one, and *T. brucei* TIM has two prolines at the N1 positions of helices. Thus, the five "N1 prolines," as well as other prolines in *B. stearothermophilus* TIM, appear to be favoring its higher stability.

Role of asparagine 12. Residue 12 is highly conserved as an Asn among the known TIM sequences (Fig. 1 and Kinemage 3). The conservation of this Asn is due to the crucial role of its side-chain amide nitrogen to preserve the unusual conformation of Lys 10 in the active site. The amide nitrogen of its side chain hydrogen bonds to the carbonyl oxygen of residue 9 and the main-chain nitrogen of residue 11. Residue Asn 12 is followed by Gly in most of the sequences. As is well known, an Asn followed by Gly in a polypeptide readily undergoes deamidation at elevated temperatures (Creighton, 1993). If such an Asn is located at an interface of two monomers, then deamidation would introduce a negatively charged Asp at the interface, thereby reducing the stability of the oligomeric protein. Indeed, this Asn in yeast TIM has been shown to undergo deamidation at higher temperatures, with a concomitant loss of stability (Ahern et al., 1987). *B. stearothermophilus* TIM seems to have evolved an elegant way to avoid such a situation. This Asn has been replaced in the *B.*

stearothermophilus TIM sequence by a His, thus placing the N^{δ2} nitrogen of its imidazole at the same position as the N^δ of Asn and thereby preserving the same crucial structural role as played by Asn, while at the same time avoiding the risks of deamidation at higher temperatures by irreversible covalent modification. Interestingly, the TIM sequence of *Thermotoga maritima*, a hyperthermophile, also has a His at this position (F. Rentier-Delrue, unpubl. results). Thus, as we had noted in our earlier analysis, this mutation has a positive effect on *B. stearothermophilus* TIM stability (Rentier-Delrue et al., 1993).

Hydrogen bonds and salt bridges

A calculation of hydrogen bonds in all the six available structures indicates that the number of hydrogen bonds in *B. stearothermophilus* TIM is not significantly different from the other TIM structures (data not shown). The human and *E. coli* structures show more hydrogen bonds than *B. stearothermophilus* TIM. These results are the same when either the program WHATIF (Vriend, 1990) or XPLOR (Brünger, 1990a) is used. The quality and the energy of hydrogen bonds are very difficult to compare because either the resolution of each structure varies so much (from 1.83 to 2.8 Å) or the refinements were done by different programs. Therefore, attempts to establish the contribution of hydrogen bonds to thermostability are not pursued further here.

Salt bridges were calculated by the program WHATIF (Vriend, 1990). Counting only the number of salt bridges, *B. stearothermophilus* TIM ranks fifth, i.e., only *E. coli* enzyme has fewer number of salt bridges. The intersubunit salt bridges are well conserved in all TIM structure (for more detail in this analysis see Wierenga et al., 1992). Therefore, salt bridges can not be considered as a positive factor of thermostability in *B. stearothermophilus* TIM. On the other hand, here also the difference in the structures due to resolution limits and refinement protocol makes it difficult to arrive at absolutely definite conclusions.

Buried surfaces

Hydrophobic effects are known to contribute significantly to the stability of proteins (see, e.g., Dill, 1990) and have been reported to be the main factors in the stability of hyperthermophilic proteins (Kirino et al., 1994). In order to assess the effect of hydrophobic stabilization of *B. stearothermophilus* TIM upon dimer formation, we determined the buried accessible sur-

face areas of the known TIM structures (Kinemage 3). The estimate of the buried accessible surface area was taken to be the difference between the sums of accessible surfaces of the individual monomers and that of the dimer. During these calculations, carbon and sulfur atoms were considered apolar, whereas nitrogen and oxygen were considered to be polar. This indeed is a crude approximation, but should nevertheless give a fair estimate while comparing six different structures with the same procedure. Another complication while comparing buried surface areas arises from the variable conformation of the flexible loops in different structures. We therefore removed the flexible active site loop of all subunits in all the TIM structures during these calculations. Results of the accessible area calculations are summarized in Table 5. Several interesting features stand out in this comparison. These are (1) the highest ratio of the polar to apolar accessible surface occurs *B. stearothermophilus* TIM, (2) burial of the largest percentage of monomer area upon dimer formation in *B. stearothermophilus* TIM, and (3) most important of all, the burial of the largest hydrophobic surface area in *B. stearothermophilus* TIM upon dimer formation: 314–534 Å² more than in other TIMs (Table 5). If we use the reported value of 25 cal mol⁻¹ Å⁻² (Chothia, 1974; Richards, 1977), there is a striking gain of 7.85 to 13.35 kcal mol⁻¹ in stability of the *B. stearothermophilus* TIM compared to other TIMs. Thus, in comparison to the other TIM structures, burial of hydrophobic area stands out as an important factor.

Yet another factor important for protein stability is the number of cavities in the structure and the buried cavity surface area (Erikson et al., 1992). When the cavities in the different TIM structures are analyzed, we find, interestingly, that *B. stearothermophilus* TIM has the smallest number of cavities, the smallest total cavity volume, as well as the smallest surface lining these cavities (Table 6). The buried cavity surface area of *B. stearothermophilus* TIM is 76.1–355.8 Å² less than other TIMs. The reduced cavity volume and cavity surface area of the *B. stearothermophilus* TIM structure thus again favors its thermal stability.

Conclusions

As reported previously, there are several factors that contribute to the thermostability of proteins in general. From our analysis of six TIM structures, we conclude that the main thermophilic factors in *B. stearothermophilus* TIM are: (1) a large number of prolines, decreasing the flexibility of the monomers in the un-

Table 5. Accessible surface areas in Å² of the six dimeric structures^a

	Dimer			Buried upon dimerization			% Buried in dimerization		
	All	Polar	Apolar	All	Polar	Apolar	All	Polar	Apolar
<i>B. stearothermophilus</i>	17,497	7,761	9,736	3,490	1,109	2,381	20	6	14
Chicken	19,056	8,140	10,916	3,175	1,110	2,066	17	6	11
<i>E. coli</i>	18,885	7,934	10,952	3,243	1,341	1,902	17	7	10
<i>T. brucei</i>	18,827	8,043	10,784	3,027	1,042	1,985	16	5	11
Human	18,306	7,782	10,524	3,325	1,258	2,067	18	7	11
Yeast	18,732	7,989	10,743	3,167	1,320	1,847	17	7	10

^a Buried areas are calculated by subtracting the accessible area of the dimer from the sums of the accessible areas of the two individual monomers. Carbons and sulfurs are considered apolar, whereas nitrogens and oxygens are considered polar.

Table 6. List of cavities and the total volume ($\text{\AA}^3$) and the surface areas ($\text{\AA}^2$) lining these cavities in the six TIM structures

	Cavities in TIM structures		
	Number of cavities	Surface area ($\text{\AA}^2$)	Volume ($\text{\AA}^3$)
<i>B. stearothermophilus</i>	4	181	103
Chicken	8	395	234
<i>E. coli</i>	5	275	157
<i>T. brucei</i>	7	380	234
Human	5	257	157
Yeast	10	536	327

^a Cavity calculations were performed using GRASP (Nicholls et al., 1991).

folded state; (2) the smallest number and volume of cavities in *B. stearothermophilus* TIM; (3) replacement of a structurally crucial Asn by a His; and (4) extensive hydrophobic interactions between the monomers upon the dimer formation that are significantly greater than in all other TIMs.

Material and methods

B. stearothermophilus TIM was overexpressed and purified as described by Rentier-Delrue et al. (1993). Crystals were obtained by the hanging drop method: 2 μL of protein solution of 4 mg/mL in 5 mM MES, pH 6.5, 1 mM EDTA, 1 mM sodium azide, and 2 mM 2-phosphoglycolic acid was mixed with 2 μL of the reservoir solution, which consisted of 25% (w/v) PEG 4000 in 100 mM acetate buffer, pH 5.23, 8% isopropanol, 2 mM dithiothreitol, 1 mM EDTA, 1 mM sodium azide. The crystal was mounted in a siliconized glass capillary. Data were collected from a single crystal with dimensions $0.75 \times 0.25 \times 0.10 \text{ mm}^3$ on a Siemens X100 area detector, using CuK α radiation from a Rigaku RU200 rotation-anode generator. Data was processed by the XENGEN (software copyright 1988, Genex Corp., Gaithersburg, Maryland, USA) package. Statistics describing the data are shown in Table 7.

The structure was solved by molecular replacement using the *T. brucei* TIM structure (Wierenga et al., 1991b) as a starting model. The rotation search in XPLOR (Brünger, 1990a) gave

Table 7. Data collection statistics of *B. stearothermophilus* TIM

Parameter	Value
Space group	P2 ₁ 2 ₁ 2
Cell dimensions	$a = 78.61 \text{ \AA}$ $b = 108.42 \text{ \AA}$ $c = 71.21 \text{ \AA}$
Collected reflections	64,784
Unique reflections	16,996
Completeness	89.5% (data to 2.8 $\text{\AA}$)
R_{sym}	9.1% (on intensities; data to 2.8 $\text{\AA}$)

two peaks that were 7.3σ ($\theta_1 = 160.0^\circ$, $\theta_2 = 52.5^\circ$, $\theta_3 = 75.0^\circ$) and 7.1σ ($\theta_1 = 217.5^\circ$, $\theta_2 = 75.0^\circ$, $\theta_3 = 60.0^\circ$) above the mean. The peaks are related by a rotation of nearly exactly 180° . The next highest peak was 4.5σ . PC refinement (Brünger, 1990b) was used to optimize and improve the solutions.

The solution of the rotation function was used in the translation search, which gave only one peak with a 14σ signal above the mean value, the next highest peak being only 4.4σ above the mean. At this stage, the model was replaced by a polyaniline model. Subunit 1 was then used to generate subunit 2 by the non-crystallographic twofold axis and rigid body refinement of the dimer was applied, which gave an R -factor of 47.1% (for 12,119 reflections between 8.0–2.8 $\text{\AA}$, overall B of 20.0 $\text{\AA}^2$).

Twofold averaging and solvent flattening was applied using the DEMON package (Vellieux, 1993). Six cycles of averaging in the range of 15.0–3.2 $\text{\AA}$ resolution was followed by a resolution extension to 2.8 $\text{\AA}$ in 13 steps of 0.03 $\text{\AA}$, finalized by an additional six cycles at 2.8 $\text{\AA}$ resolution. An averaged and solvent-flattened weighted map (all maps calculated were SIGMAA weighted [Read, 1986]) was used to model 60% of the side chains into this density by O (Jones et al., 1991). Subsequently the slow-cooling simulated annealing refinement protocol of XPLOR was applied (Brünger et al., 1990) with noncrystallographic restraints between the monomers. The R -factor dropped to 33.5% (for 12,119 reflections between 8–2.8 $\text{\AA}$, with an overall B -factor of 15.0 $\text{\AA}^2$). A new map was calculated, making it possible to fit 15% more of the side chains. Slow-cooling simulated annealing was then applied without noncrystallographic restraints, followed by an individual B -factor and positional refinement, which dropped the R -factor to 26.2%. The map showed very poor density for several loops, mainly the flexible loops that were in the open conformation in the molecular replacement model. Therefore, omit maps over all poorly defined density were calculated that showed, in almost all cases, good density for the loops, and the main chain was adjusted on the graphics. The most impressive result was in the flexible active site loops, which were in the "open" conformation in the molecular replacement model, appearing in the omit map clearly in the "closed" conformation. Positional and B -factor refinement after some model building dropped the R -factor to 18.7%. At this stage, it was very clear that the density for the inhibitor in both subunits could be contoured at the 3σ level for the entire 2PG and up to 7σ for the phosphate. After the inclusion of the inhibitor molecules, some additional manual adjustments and refinement, the R -factor dropped to 17.9%. Final model building and refinement yielded a R -factor of 17.6%, with very good geometry. The refinement statistics can be seen in Table 1.

All comparisons and analysis were done using either O (Jones et al., 1991) or BIOGRAF (version 3.21, Molecular Simulations Inc., Burlington, Massachusetts, USA). The hydrogen bonds and salt bridges were calculated using either WHATIF (Vriend, 1990) or XPLOR (Brünger, 1990a). Accessible surfaces were calculated using modified RESAREA and SURFACE programs of the CCP4 suite (1994). Sequence alignment was done using GCG (Genetic Computer Group, Inc., copyright 1982–1992, Wisconsin, USA). Ramachandran plots were calculated with the PHIPSI program of the BIOMOL package. The secondary structure was assigned by using DSSP (Kabsch & Sander, 1983). Calculation of cavities was done by using GRASP (Nicholls et al., 1991) suite, first calculating the molecular surface area, and then searching for the enclosed inaccessible cavities.

Acknowledgments

We thank Tjaard Pijning of University of Groningen for help during crystallizations and the members of the Biomolecular Structure Center of the University of Washington, Seattle, for useful discussions. We also thank Dr. Rik Wierenga for stimulating discussions. Financial support for the project was received from EEC-BRIDGE project on triosephosphate isomerase engineering and from the School of Medicine, University of Washington, Seattle.

References

- Ahern TJ, Casal JI, Petsko GA, Klibanov AM. 1987. Control of oligomeric enzyme thermostability by protein engineering. *Proc Natl Acad Sci USA* 84:675-679.
- Alber T, Kawasaki G. 1982. Nucleotide sequence of the triose phosphate 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.
- Borders CL Jr, Brodwater JA, Bekeny PA, Salmon JE, Lee AS, Eldridge AM, Pett VB. 1994. A structural role for arginine in proteins: Multiple hydrogen bonds to backbone carbonyl oxygens. *Protein Sci* 3:541-548.
- Brünger AT. 1990a. *X-PLOR user manual*. Yale University Press, Yale.
- Brünger AT. 1990b. Extension of molecular replacement: A new search strategy based on Patterson correlation refinement. *Acta Crystallogr A* 46:46-57.
- Brünger AT, Krokowski A, Erikson J. 1990. Slow cooling protocols for crystallographic refinement by simulated annealing. *Acta Crystallogr A* 46:585-593.
- CCP4 suite. 1994. The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr D* 50:760-763.
- Cheng J, Mielnicki LM, Pruitt SC, Maquat LE. 1990. Nucleotide sequence of murine triosephosphate isomerase cDNA. *Nucleic Acids Res* 18:4261.
- Chothia C. 1974. Hydrophobic bonding and accessible surface area in proteins. *Nature* 248:338-339.
- Corran PH, Waley SG. 1975. The amino acid sequence of rabbit muscle triose phosphate isomerase. *Biochem J* 145:335-344.
- Creighton TE. 1991. Stability of folded conformation. *Curr Opin Struct Biol* 1:5-16.
- Creighton TE. 1993. *Proteins: Structure and molecular properties*. New York: W.H. Freeman and Co. pp 9-10.
- 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.
- Davies GJ, Gamblin SJ, Littlechild JA, Watson HC. 1993. The structure of a thermally stable 3-phosphoglycerate kinase and a comparison with its mesophilic equivalent. *Proteins Struct Funct Genet* 15:283-289.
- Dill KA. 1990. Dominant forces in protein folding. *Biochemistry* 29:7133-7155.
- 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.
- Eriksson AE, Baase WA, Zhang XJ, Heinz DW, Blaber M, Baldwin EP, Matthews BW. 1992. Response of a protein structure to cavity-creating mutation and its relation to the hydrophobic effect. *Science* 225:178-183.
- Flocco MM, Mowbray S. 1994. Planar interactions of arginine and aromatic side-chains in proteins. *J Mol Biol* 235:709-717.
- Fontana A. 1991. Analysis and modulation of protein stability. *Curr Opin Biotech* 2:551-561.
- Hardy F, Vriend G, Veltman OR, van der Vinne B, Venema G, Eijsink VG. 1993. Stabilization of *Bacillus stearothermophilus* neutral protease by introduction of prolines. *FEBS Lett* 317:89-92.
- Ikai A. 1980. Thermostability and aliphatic index of globular proteins. *J Biochem (Tokyo)* 88:1895-1898.
- Jaenicke R. 1991. Protein stability and molecular adaptation to extreme conditions. *Eur J Biochem* 202:715-728.
- Jones TA, Zou JY, Cowan SW, Kjeldgaard M. 1991. Improved methods for building models in electron density maps and the location of errors in these models. *Acta Crystallogr A* 47:110-119.
- Kabsch W, Sander C. 1983. Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* 22:2577-2637.
- Kelly CA, Nishiyama M, Ohnishi Y, Beppu T, Birktoft JJ. 1993. Determinants of protein thermostability observed in the 1.9 Å crystal structure of malate dehydrogenase from the thermophilic bacterium *Thermus flavus*. *Biochemistry* 32:3913-3922.
- Kirino H, Aoki M, Aoshima M, Hayashi Y, Ohba M, Yamagashi A, Wakagi T, Oshima T. 1994. Hydrophobic interaction at the subunit interface contributes to the thermostability of 3-isopropylmalate dehydrogenase from an extreme thermophile *Thermus thermophilus*. *Eur J Biochem* 220:275-281.
- Knowles JR. 1991. Enzyme catalysis: Not different, just better. *Nature* 350:121-124.
- Knowles JR, Alberly WJ. 1977. Perfection in enzyme catalysis: The energetics of triosephosphate isomerase. *Acc Chem Res* 10:105-111.
- Kolb E, Harris JI, Bridgen J. 1974. Triose phosphate isomerase from the coelacanth. An approach to the rapid determination of an amino acid sequence with small amounts of material. *Biochem J* 137:185-197.
- Korndörfer I, Steipe B, Huber R, Tomschy A, Jaenicke R. 1995. The crystal structure of holo-glyceraldehyde-3-phosphate dehydrogenase from the hyperthermophilic bacterium *Thermotoga maritima* at 2.5 Å resolution. *J Mol Biol* 246:511-521.
- Kraulis PJ. 1991. MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures. *J Appl Crystallogr* 24:946-950.
- 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-phosphoglycerate at 2.5 Å resolution: Implications for catalysis. *Biochemistry* 29:6619-6625.
- Mande SC, Mainfroid V, Kalk KH, Goraj K, Martial JA, Hol WGJ. 1994. Crystal structure of recombinant human triosephosphate isomerase at 2.8 Å resolution. Triosephosphate isomerase-related human genetic disorders and comparison with the trypanosomal enzyme. *Protein Sci* 3:810-821.
- 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. 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 differential loss of introns. *Cell* 46:143-147.
- Menendez-Arias L, Argos P. 1989. Engineering protein thermal stability. Sequence statistics point to residue substitutions in α -helix. *J Mol Biol* 206:397-406.
- Merkler DJ, Farrington GK, Becktel WJ. 1981. Protein thermostability. Correlation between macroscopic parameters and growth temperature for closely related thermophilic and mesophilic bacilli. *Int J Pept Protein Res* 18:430-442.
- Mrabet NT, Van de Broeck A, Van den Brande I, Stanssens P, Laroche Y, Lambeir AM, Matthijssens G, Jenkins J, Chiadmi M, Van Tilbeurgh H, Rey F, Janin J, Quax WJ, Lesters I, Maeyer MD, Wodak SJ. 1992. Arginine residues as stabilizing elements in proteins. *Biochemistry* 31:2239-2253.
- Nandi CL, Singh J, Thornton JM. 1993. Atomic environments of arginine side chains in proteins. *Protein Eng* 6:247-259.
- Nicholls A, Sharp KA, Honig B. 1991. Protein folding and association: Insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins Struct Funct Genet* 11:282-296.
- Nicholson H, Tronrud DE, Becktel WJ, Matthews BW. 1992. Analysis of the effectiveness of proline substitutions and glycine replacements in increasing the stability of phage T4 lysozyme. *Biopolymers* 32:1431-1441.
- Noble ME, Verlinde CL, Groendijk H, Kalk KH, Wierenga RK, Hol WG. 1991a. Crystallographic and molecular modeling studies on trypanosomal triosephosphate isomerase: A critical assessment of the predicted and observed structures of the complex with 2-phosphoglycerate. *J Med Chem* 34:2709-2718.
- Noble MEM. 1992. X-ray crystallographic studies of triosephosphate isomerase [thesis]. Chapter 5. Groningen: State University of Groningen.
- Noble MEM, Wierenga RK, Lambeir AM, Opperdoes FR, Thunnissen AM,

- Kalk KH, Groendijk H, Hol WGJ. 1991b. The adaptability of the active site of trypanosomal triosephosphate isomerase as observed in crystal structures of three different complexes. *Proteins Struct Funct Genet* 10(1):50–69.
- Noble MEM, Zeelen JPh, Wierenga RK, Mainfroid V, Goraj K, Gohimont AC, Martial JA. 1993. Structure of triosephosphate isomerase from *Escherichia coli* determined at 2.6 Å resolution. *Acta Crystallogr D* 49(4):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.
- Pichersky E, Gottlieb LD, Hess JF. 1984. Nucleotide sequence of the triose phosphate isomerase gene of *Escherichia coli*. *Mol Gen Genet* 195:314–320.
- Read RJ. 1986. Improved Fourier coefficients for maps using phases from partial structures with errors. *Acta Crystallogr A* 42:140–149.
- Rees DC, Adams MWW. 1995. Hyperthermophiles: Taking the heat and loving it. *Structure* 15:251–254.
- Rentier-Delrue F, Mande SC, Moyens S, Terpstra P, Mainfroid V, Goraj K, Lion M, Hol WGJ, Martial JA. 1993. Cloning and overexpression of the triosephosphate isomerase genes from psychrophilic and thermophilic bacteria: Structural comparison of the predicted protein sequences. *J Mol Biol* 229:85–93.
- Richards FM. 1977. Areas, volumes, packing and protein structure. *Annu Rev Biophys Bioeng* 6:151–176.
- 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 triose phosphate isomerase (Tpi) 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.
- Straus D, Raines R, Kawashima E, Knowles JR, Gilbert W. 1985. Active site of triosephosphate isomerase: In vitro mutagenesis and characterization of an altered enzyme. *Proc Natl Acad Sci USA* 82:2272–2276.
- Swinkels BW, Gibson WC, Osinga KA, Kramer R, Veeneman GH, van Boom JH, Borst P. 1986. Characterization of the gene for the 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.
- Vellieux FMD. 1993. Density modification (DEMON). Grenoble, France.
- Verlinde CL, Witmans CJ, Pijning T, Kalk KH, Hol WG, Callens M, Opperdoes FR. 1992. 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(12):1578–1584.
- Vihinen M. 1987. Relationship of protein flexibility to thermostability. *Protein Eng* 1:477–480.
- Vriend G. 1990. WHAT IF: A molecular modeling and drug design program. *J Mol Graph* 8:52–56.
- Vriend G, Eijssink V. 1993. Prediction and analysis of structure, stability and unfolding of thermolysin-like proteases. *J Comput Aided Mol Des* 7(4):367–396.
- Waley SG. 1973. Refolding of triosephosphate isomerase. *Biochem J* 135:165–172.
- Watanabe K, Chishiro K, Kitamura K, Suzuki Y. 1991. Proline residues responsible for thermostability occur with high frequency in the loop regions of an extremely thermostable oligo-1,6-glucosidase from *Bacillus thermooglucosidarius* KP1006. *J Biol Chem* 266(36):24287–24294.
- Wierenga RK, Noble MEM, Davenport RC. 1992. Comparison of the refined crystal structure 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.
- Zabori S, Rudolph R, Jaenicke R. 1980. Folding and association of triosephosphate isomerase from rabbit muscle. *Z Naturforsch* 35C:999–1004.
- Zhang XJ, Baase WA, Matthews BW. 1991. Toward a simplification of the protein folding problem: A stabilizing polyalanine α -helix engineered in T4 lysozyme. *Biochemistry* 30:2012–2017.