Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallographic binding studies with an engineered monomeric variant of triosephosphate isomerase

Mikko Salin,[a,b] Evangelia G. Kapetaniou,[a,b] Matti Vaismaa,[c] Marja Lajunen,[c] Marco G. Casteleijn,[d] Peter Neubauer,[b,d]‡ Laurent Salmon[e] and Rik K. Wierenga[a,b]*

[a]Department of Biochemistry, PO Box 3000, 90014 University of Oulu, Finland, [b]Biocenter Oulu, PO Box 5000, 90014 University of Oulu, Finland, [c]Department of Chemistry, PO Box 3000, 90014 University of Oulu, Finland, [d]Bioprocess Engineering Laboratory, Department of Process and Environmental Engineering, PO Box 4300, 90014 University of Oulu, Finland, and [e]Université Paris-Sud, Equipe de Chimie Bioorganique et Bioinorganique, ICMMO, CNRS-UMR 8182, Bâtiment 420, 15 Rue Georges Clemenceau, F-91405 Orsay, France

‡ Current address: Institute of Biotechnology, Technische Universität Berlin, Ackerstrasse 71–76, D-13355 Berlin, Germany.

Correspondence e-mail: rik.wierenga@oulu.fi

Received 5 February 2010
Accepted 30 June 2010

PDB References: monomeric triosephosphate isomerase mutant, 2x16; 2x1s; 2x1r; 2x2g; 2x1t; 2x1u.

Crystallographic binding studies have been carried out to probe the active-site binding properties of a monomeric variant (A-TIM) of triosephosphate isomerase (TIM). These binding studies are part of a structure-based directed-evolution project aimed towards changing the substrate specificity of monomeric TIM and are therefore aimed at finding binders which are substrate-like molecules. A-TIM has a modified more extended binding pocket between loop-7 and loop-8 compared with wild-type TIM. The A-TIM crystals were grown in the presence of citrate, which is bound in the active site of each of the two molecules in the asymmetric unit. In this complex, the active-site loops loop-6 and loop-7 adopt the closed conformation, similar to that observed in liganded wild-type TIM. Extensive crystal-soaking protocols have been developed to flush the bound citrate out of the active-site pocket of both molecules and the crystal structure shows that the unliganded open conformation of the A-TIM active site is the same as in unliganded wild-type TIM. It is also shown that sulfonate compounds corresponding to the transition-state analogue 2-phosphoglycolate bind in the active site, which has a closed conformation. It is also shown that the new binding pocket of A-TIM can bind 3-phosphoglycerate (3PGA; an analogue of a C4-sugar phosphate) and 4-phospho-D-erythronohydroxamic acid (4PEH; an analogue of a C5-sugar phosphate). Therefore, these studies have provided a rationale for starting directed-evolution experiments aimed at generating the catalytic properties of a C5-sugar phosphate isomerase on the A-TIM framework.

1. Introduction

The $(\beta/\alpha)_8$ protein fold, commonly known as the TIM fold, is nature's favorite choice as a scaffold for enzymatic activity. About 20 different classes of enzymes contain a TIM-barrel structural domain (Nagano *et al.*, 2002). The scaffold of this fold was first observed in the crystal structure of triosephosphate isomerase (TIM; EC 5.3.1.1). It is formed by eight parallel β-strands (the central barrel) surrounded by eight α-helices at the exterior of the protein (Wierenga, 2001). The active sites of all TIM-barrel enzymes are at the C-terminal end of the β-strands (the catalytic end), while the loops at the N-terminal end of the β-strands (the stability end) are important for the stability of the fold. The active sites are shaped by residues in the eight loops following the eight β-strands, as visualized in Fig. 1 for TIM. Mutagenesis of these loops can therefore result in novel active sites without altering the overall stability of the fold. For these reasons, $(\beta/\alpha)_8$-barrels provide an ideal scaffold for directed-evolution experiments aimed at producing new non-natural enzymes

that are tailored to catalyse specific reactions (Sterner & Hocker, 2005). This technique, in which large genetic libraries of the target protein are used, mimics natural evolution by combining random mutagenesis with screening for, or selection of, enzyme variants with desirable new properties that are useful for laboratory-scale or industrial-scale applications (Claren *et al.*, 2009; Reetz *et al.*, 2006). We have initiated a protein-engineering project to create a monomeric $(\beta/\alpha)_8$ protein with new catalytic properties. As a reference molecule, we use a monomeric variant of TIM from *Trypanosoma brucei brucei*.

TIM is a dimeric glycolytic enzyme that interconverts dihydroxyacetone phosphate (DHAP) and D-glyceraldehyde 3-phosphate (D-GAP) during glycolysis. It is therefore a sugar phosphate isomerase which exclusively catalyses the interconversion of the C3-sugar phosphates DHAP and D-GAP. TIM is known to be a very efficient enzyme. Its reaction mechanism has been extensively studied (Knowles, 1991). The reaction cycle involves the shuttling of protons; cofactors or metal ions are not required by the enzyme. The residues directly involved in catalysis are Asn11 and Lys13 from loop-1 together with His95 and Glu167 from loop-4 and loop-6, respectively. In addition, four main-chain NH groups of residues from loop-6, loop-7 and loop-8 form a highly specific phosphate dianion-binding site for the phosphate moiety of the substrate (Kursula & Wierenga, 2003). Crystallographic studies have shown that loop-6 and loop-7 have an open conformation in the absence of ligand and a closed conformation in its presence (Joseph *et al.*, 1990; Noble *et al.*, 1993). The tip of loop-6 (Gly173) moves approximately 7 Å upon ligand binding, whereas for loop-7 the conformational switch involves reorientation of two peptide planes (Gly211–Gly212 and Gly212–Ser213; Casteleijn *et al.*, 2006). The concerted movement of loop-6 and loop-7 is important for both binding and catalysis. Loop-8 does not change conformation upon ligand binding and is not involved in catalysis, but through the interactions of two peptide NH groups in a 3_{10}-helical stretch, N(Gly234) and N(Gly235), it binds the phosphate moiety of the ligand. These two hydrogen bonds are complemented by two additional hydrogen bonds to N(Gly173) of loop-6 and N(Ser213) of loop-7. The active-site binding pocket extends from the catalytic residues at its bottom (His95 and Glu167) to the 3_{10}-helical stretch of loop-8, which binds the phosphate group. The tight phosphate-binding pocket limits the substrate specificity of TIM to the C3-sugar phosphates DHAP and D-GAP. The active site is at the dimer interface, but the catalytic residues of the active site, as well as the residues of the binding pocket, are from the same subunit.

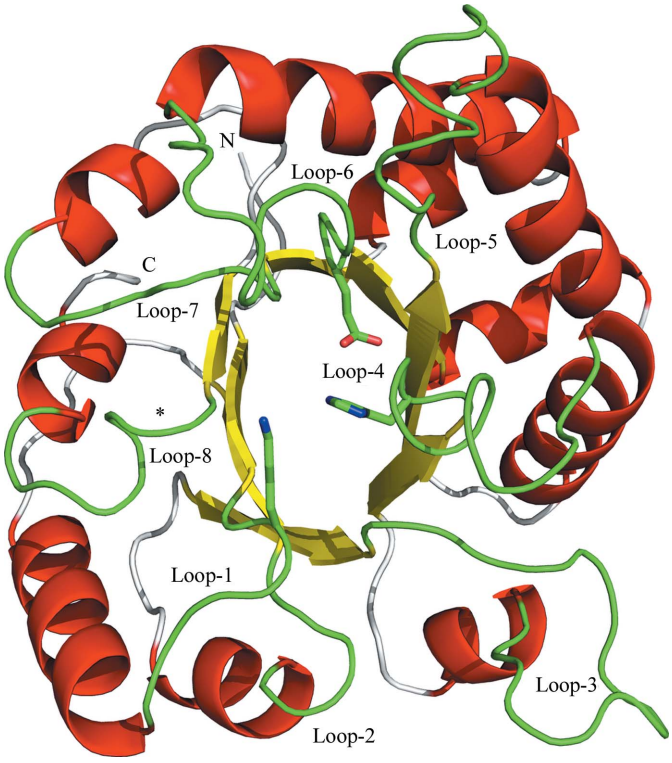

Figure 1
The TIM barrel of triosephosphate isomerase. The side chains Lys13 (loop-1), His95 (loop-4) and Glu167 (loop-6) identify the catalytic site. The loops at the catalytic end are coloured green; the loops at the stability end are coloured grey. In A-TIM loop-1, loop-3 and loop-8 have been shortened. The extended new binding pocket of A-TIM between loop-7 and loop-8, which has been generated because of the loop-8 shortening, is marked with an asterisk. The N-terminus and C-terminus are labelled N and C, respectively.

The dimer-interface region is composed of loop-3, which is inserted into a pocket near the active site of the other subunit shaped by loop-1 and loop-4. This insertion pulls loop-1 and loop-4 into the active-site cavity, which also leads to the formation of a salt bridge between residues Lys13 and Glu97 (Wierenga, Noble, Vriend *et al.*, 1991). Our previous experiments involving deleting/mutating loop-3 converted the dimer into a monomer (Borchert *et al.*, 1993). The new monomer is still a competent TIM, but with an approximately 1000-fold lower k_{cat} and a tenfold higher K_m when compared with the wild-type dimer (Schliebs *et al.*, 1997; Thanki *et al.*, 1997). The catalytic rates of the monomeric TIMs are approximately 1 s^{-1} and the K_m values for substrate are in the millimolar range. Crystallographic characterization of these monomeric TIMs has shown that the mode of binding of the substrate analogues is very similar to that seen in wild-type TIM. In addition, these studies have shown a large structural plasticity for loop-1 and loop-4, in particular in the unliganded form (Borchert *et al.*, 1995), whereas in wild-type TIM these two dimer-interface loops are rather rigid. A common feature of both the wild-type dimeric TIM and monomeric TIM structures is the presence of open (unliganded) and closed (liganded) states of loop-6 and loop-7.

In order to explore the catalytic space of this monomeric TIM framework further, we set out to change its substrate specificity using the following three-step structure-based directed-evolution protocol.

(i) Shortening loop-8 *via* rational modelling approaches. This should allow the binding of larger sugar-like molecules in the active site.

(ii) Characterizing the binding properties of the new variant, in particular to verify whether extended substrate-like molecules bind to the new binding pocket.

(iii) Optimizing new catalytic properties *via* directed-evolution approaches.

Using this protocol, we aim to find new enzymatic activities for this simple monomeric protein. These studies will also address questions about the enzyme mechanistics of the catalytic process. For example, the open/closed mechanism of loop-6/loop-7 is absolutely essential for the TIM-catalysed reaction as it prevents an undesirable phosphate-elimination side reaction (Knowles, 1991; Richard, 1984). This side reaction occurs easily with C3-sugar phosphates once the enediolate intermediate has been formed in the catalytic cycle, because the phosphate moiety is a good leaving group in C3-sugar phosphates. However, the chemical properties of substrate-like molecules with a more extended carbon skeleton are such that this elimination reaction does not occur and therefore the tight phosphate binding and the open/closed mechanism of loop-6 are not relevant to these types of molecules. It will be interesting to see whether suitable mutations can be found *via* directed-evolution approaches that allow good binding of the extended sugar-like molecules with respect to the catalytic residues without the requirement for loop-6/loop-7 closure and without tight phosphate anchoring. Before initiating this directed-evolution approach [step (iii)], it is essential to first complete steps (i) and (ii) of the structure-based protocol. We have previously reported the loop-8 shortening achieved in step (i) (Norledge *et al.*, 2001; Alahuhta, Casteleijn *et al.*, 2008), *via* which A-TIM was generated. In A-TIM loop-8 is shortened by three residues and Val233 has been changed to an alanine (Fig. 2). Initial characterization results [step (ii)] have also been reported (Alahuhta, Casteleijn *et al.*, 2008). Here, we describe the further characterization of A-TIM with more extended substrate-like molecules. The initial characterization showed that A-TIM can bind the transition-state analogue 2PG (Fig. 3). 2PG is a classical TIM transition-state analogue (Wolfenden, 1969); its carboxylic acid moiety is bound to the carboxylic acid moiety of the catalytic base Glu167. In addition, it was found that the A-TIM active site can bind a citrate molecule. This molecule is devoid of a phosphate moiety, but nevertheless its interactions with the catalytic base and with loop-6 and loop-7 are rather similar to those seen for the 2PG mode of binding: loop-6 and loop-7 are in the closed conformation and hydrogen bonded to the ligand and the catalytic Glu167 is in the swung-in competent conformation.

Owing to the shortening of loop-8 the binding pocket is now much more extended and the catalytic site is now connected *via* the truncated phosphate-binding pocket to a hydrophobic pocket shaped by Phe242 and Ile245 (Fig. 4). In this paper, we report two additional A-TIM binding studies. Firstly, the feasibility of binding an extended sulfonate analogue of 2PG [3-(propylsulfonyl)propionic acid; P3SP; Fig. 3] with its tail in this new hydrophobic pocket near Phe242 and Ile245 was explored. However, these experiments showed that the carbon tail does not bind in the new hydrophobic binding pocket. Secondly, further crystallographic binding studies showed that molecules such as 3-phosphoglycerate (3PGA; a C4-sugar phosphate analogue) and 4-phospho-D-erythronohydroxamic acid (4PEH; a C5-sugar phosphate analogue) bind to the A-TIM active site in a similar, but not identical, fashion to that observed in wild-type TIM binding studies. Therefore, these studies have now provided a rationale for completing the structure-based directed-evolution protocol, which is now aimed at generating the catalytic properties of a C5-sugar phosphate isomerase on the A-TIM framework.

Numbering	233	238	242	245
Secondary structure	β β	α α α α α α α		
Wild-type TIM	L V G G A S L	K P E	F V	D I
A-TIM	L A G - - - L	K P E	F V	D I
Secondary structure	β β	α α α α α α α		

Figure 2
The shortening of loop-8. In A-TIM three residues have been deleted and Val233 has been changed to an alanine.

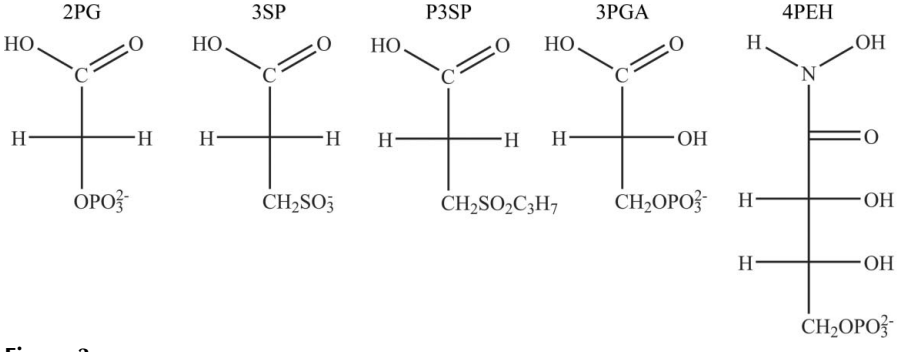

Figure 3
A schematic representation of the compounds used in the A-TIM crystallographic binding experiments: 2PG, 2-phosphoglycolic acid; 3SP, 3-sulfopropionic acid; P3SP, 3-(propylsulfonyl)propionic acid; 3PGA, 3-phosphoglyceric acid; 4PEH, 4-phospho-D-erythronohydroxamic acid.

2. Materials and methods

All commercially available reagents were used as obtained without further purification. The cation-exchange resin CM Sepharose Fast Flow used in the protein-purification protocol was obtained from Amersham Biosciences. D-3-Phosphoglycerate (3PGA) was purchased from Sigma–Aldrich. The purity of this compound was checked by ESI mass spectrometry using a Q-TOF I from Waters/Micromass.

2.1. Synthesis of substrates

2.1.1. Synthesis of 3-sulfopropionic acid (3SP). The synthesis of 3SP was carried out as described previously (Doi *et al.*, 1985). Hydrogen peroxide (VWR; 30%, 2.9 ml) was added to a solution of 3-thiolpropionic acid (Aldrich; 1.0 g, 9.5 mmol) and acetic acid (Riedel de Haen; 10 ml) with stirring at 308 K. After 30 min at 363 K, the product 3SP was isolated by the removal of volatile compounds under vacuum. 3SP was obtained as white crystals (1.6 g): ^{1}H NMR (200 MHz, DMSO-d$_6$, p.p.m.), δ_H

7.4 (br s, OHs), 2.60–2.80 (2H, m, $OHSO_2C\mathbf{H}_2$), 2.40–2.60 (2H, m, $C\mathbf{H}_2CO_2H$); ^{13}C NMR (50 MHz, DMSO-d_6, p.p.m.), δ_C 173.4 (C═O), 47.9, 30.5; HRMS (ESI^-) calculated for $C_3H_5O_5S$ 152.9858, found 152.9866.

2.1.2. Synthesis of 3-(propylsulfonyl)propionic acid (P3SP). The synthesis of 3-(propylthio)propionic acid was carried out as described by Vaismaa *et al.* (2007). The oxidation of 3-(propylthio)propionic acid to P3SP was carried out as follows. A solution of cold 30% hydrogen peroxide (VWR; 2 ml, 19 mmol) was injected into a 308 K solution of 3-(propylthio)propionic acid (950 mg, 6.4 mmol) and 5 ml acetic acid (Riedel de Haen) in a two-necked round-bottom flask equipped with a condenser, a stirring bar and a thermometer. The subsequent reaction was highly exothermic and the temperature of the solution increased rapidly to a maximum of 353 K, after which it decreased to around 343 K. The temperature was maintained at 378 K for an additional 20 min and left to cool. 50 ml water was added and solvents were removed under high vacuum. P3SP was obtained as a white powder (1.15 g, 6.4 mmol) with a melting point of 363.5 K (differential scanning calorimeter): ^{1}H NMR (200 MHz, $CDCl_3$, p.p.m.), δ_H 8.17 (1H, br s, OH), 3.31 (2H, t, J = 7.3 Hz, $SO_2C\mathbf{H}_2CH_2CO$), 3.05–2.9 (4H, m, $C\mathbf{H}_2SO_2CH_2C\mathbf{H}_2CO$), 1.91 (2H, m, $CH_3C\mathbf{H}_2$), 1.10 (3H, t, J = 7.4 Hz, CH_3); ^{13}C NMR (50 MHz, $CDCl_3$, p.p.m.), δ_C 175.6 (C═O), 55.3, 47.7, 26.6, 16.0, 13.3 (CH_3); HRMS (ESI^-) calculated for $C_6H_{11}O_4S$ 179.0378, found 179.0396.

2.1.3. Synthesis of 4-phospho-D-erythronohydroxamic acid (4PEH). The synthesis of 4PEH was performed as described elsewhere (Burgos & Salmon, 2004).

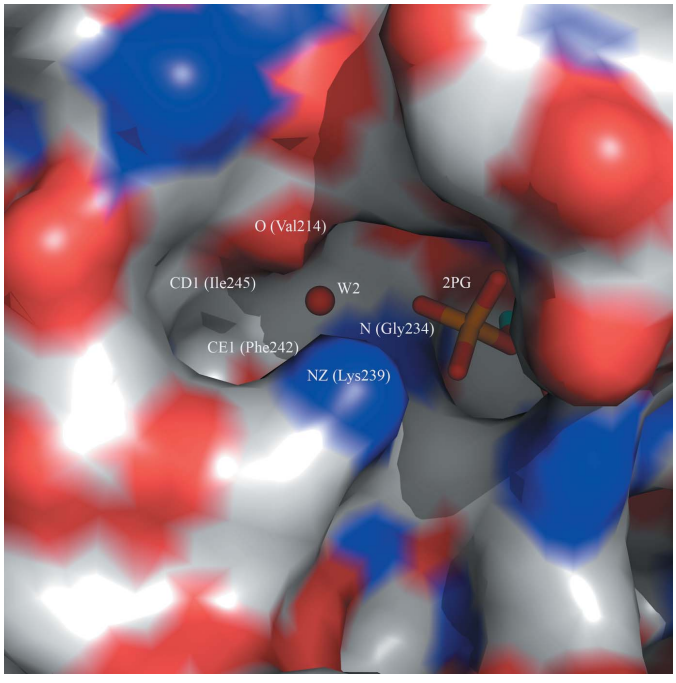

Figure 4
The molecular surface of the A-TIM–2PG complex near the hydrophobic binding pocket, shaped by Phe242 and Ile245 at its bottom (oxygen, red; nitrogen, blue; carbon, grey). The polar entry formed by O(Val214)–Wat2–N(Gly234) is also highlighted.

2.2. Protein expression and purification

The A-TIM gene, which is ultimately a mutant of *T. brucei* TIM, was inserted into the pET-3a plasmid (Novagen, Madison, Wisconsin, USA). The plasmid was used to transform *Escherichia coli* strain BL21 (DE3) pLysS (Invitrogen, Carlsbad, California, USA), which served as the expression host. Protein expression was performed as a bioreactor fed-batch cultivation in a 10 l Biostat bioreactor (Sartorius) equipped with efficient stirring devices. The pH was kept at 7.0 by controlled feeding of 25% ammonium hydroxide. The cultivation medium was a defined Mineral Salt Medium (Neubauer *et al.*, 1995) with an initial glucose concentration of 10 g l^{-1}. After heat sterilization, 2 ml l^{-1} 1 *M* $MgSO_4$ and 2 ml l^{-1} trace-element solution (Holme *et al.*, 1970) were added together with 0.1 g l^{-1} thiamine hydrochloride through a sterile filter (0.2 µm). The culture was inoculated to an initial OD_{600} of 0.4 and cultivated at 310 K. The feed solution contained 650 g l^{-1} glucose but no extra salts. The feeding rate was 169 g glucose per hour. During all fed-batch cultivations an extra 2 ml l^{-1} of 1 *M* $MgSO_4$ solution was added during cultivation approximately per every 10 unit increase in OD_{600}. The cultivation was performed in a glucose-limited fed-batch mode such that anaerobic conditions were avoided. Cells were induced at an OD_{600} of about 40 with 0.4 m*M* isopropyl β-D-1-thiogalactopyranoside (IPTG) and grown for another 4 h (to a final OD of 120).

The protein-purification protocols were similar to those described in Casteleijn *et al.* (2006) with two minor modifications: (i) 150 m*M* ammonium sulfate was added to the lysis buffer and (ii) the ammonium sulfate precipitation steps were omitted. In the final step of purification a gel-filtration step was carried out and only the peak fractions were used in the crystallization experiments. The protein concentration was determined by measuring the absorption at 280 nm with a Nanodrop spectrophotometer. The molar extinction coefficient of 40 575 M^{-1} cm^{-1} for monomeric TIM was calculated using the *ProtParam* tool (Gasteiger *et al.*, 2003). The final protein yield was 88 mg purified A-TIM per litre of culture medium. The protein sample (approximately 10 g l^{-1}) in 20 m*M* triethanolamine buffer pH 8.0, 1 m*M* reduced dithiothreitol, 1 m*M* EDTA, 1 m*M* sodium azide was flash-frozen under liquid nitrogen in 60 µl aliquots in storage buffer.

2.3. Crystallization

The initial crystallization experiments were performed at 295 K using the hanging-drop method as described previously (Alahuhta, Casteleijn *et al.*, 2008). Briefly, drops were made by mixing 2 µl of approximately 10 g l^{-1} protein solution in storage buffer with 2 µl well solution (20% PEG 6000 in 100 m*M* citrate pH 5.5). In order to improve the quality and size of the crystals, microseeding protocols were developed using the small seed crystals that were obtained using the standard crystallization procedure. Microseeding was performed at 277 K. The small seeds were transferred using a hair from a standard crystallization drop containing different-sized crystals to a fresh drop. After 5 min, these drops were trans-

ferred to 288 K. After 3 d or after crystal growth ceased, a temperature gradient from 288 to 273 K was performed in a temperature-controlled crystallization box in order to extend the crystal-growth time. The temperature was then set to 277 K and 2 µl fresh protein solution was added to the drop. The last step was repeated when the crystals did not grow any further. This was repeated until the crystals were considered to be sufficiently large. On average, the crystal diameter was approximately 0.4 mm.

2.4. Ligand-exchange protocols

The well solution for each of the crystallization experiments was 20% PEG 6000 and 100 m*M* citric acid pH 5.5. A citric acid molecule was found to bind tightly to the A-TIM active site (Alahuhta, Salin *et al.*, 2008). Therefore, several extensive crystal-washing protocols (Hassell *et al.*, 2007) were developed to remove or exchange the citrate molecule, as listed below. In these protocols the crystals were transferred to 5 µl drops.

2.4.1. The complex with sulfate. Sulfate-liganded A-TIM crystals were obtained *via* a cross-linking protocol (Lusty, 1999). Firstly, the skin covering the drop was gently removed. Subsequently, 5 µl 25% glutaraldehyde was added to a dialysis button floating on the well solution. The system was sealed, transferred to room temperature and incubated for 100 min. The cover glass was opened and the crystals were transferred directly to the soaking solution (1.75 *M* ammonium sulfate pH 5.7) and equilibrated for 24 h.

2.4.2. The complex with 3SP. The crystals were transferred into a drop of 20% PEG 6000, 100 m*M* 2-(*N*-morpholino)-ethanesulfonic acid (MES), 25 m*M* 3SP pH 5.5 at 277 K. The soak was repeated three times, with the last transfer being just prior to freezing. The time between the soaks was 2–18 h. The total soaking time was 2 d.

2.4.3. The complex with P3SP. The same soaking protocol was used as described for the 3SP experiment, but 50 m*M* P3SP was used instead of 3SP.

2.4.4. The unliganded structure. A crystal was transferred from the original drop to a 5 µl washing-solution drop consisting of 20% PEG 6000, 100 m*M* MES, 100 m*M* D-ribose 5-phosphate (D-R5P) pH 5.5 at 277 K. This washing step was repeated four times, with a time interval of at least 30 min between transfers. The total washing time was 1 d. Subsequently, the crystal was transferred to 20% PEG 6000, 100 m*M* MES pH 5.5 at 277 K for 1 d.

2.4.5. The complex with 3PGA. The same washing protocol was used as described for the unliganded crystal. The final additional soaking steps were carried out by transferring the crystal to the following solution: 20% PEG 6000, 100 m*M* MES, 25% ethylene glycol, 5 m*M* 3PGA pH 5.5. This transfer was repeated two times, with the last transfer being just prior to freezing. The total soaking time was 1 d.

2.4.6. The complex with 4PEH. The same washing protocol was used as described for the unliganded crystal. The final additional soaking steps were performed by transferring the crystal into 20% PEG 6000, 100 m*M* sodium acetate, 20 m*M* 4PEH pH 5.5. The crystal was soaked three times in this solution, with the last soak being just prior to freezing. The total soaking time was 3 d.

2.5. Data collection, data processing and structure refinement

After completing the soaking protocol, the crystals were soaked briefly in cryoprotectant solution and flash-frozen in a liquid-nitrogen stream (100 K). In all cases, the cryoprotectant solutions were identical to the soaking solutions except that 25% ethylene glycol was included in the mother liquor. As an exception, cryoprotection of the P3SP crystal was achieved by covering the drop with 100% mineral oil. As tabulated in Table 1, data were collected on the beamlines of EMBL/DESY, Hamburg, Germany as well as with a Nonius FR591 home source equipped with Montel mirrors and a MAR 345 detector. The data were integrated, merged and scaled using the *XDS* program package (Kabsch, 2010) with the *XDSi* v.1.7 interface (Kursula, 2004) and *SCALA* (Evans, 2005) from the *CCP*4 program suite (Collaborative Computational Project, Number 4, 1994). The low-resolution data in the sulfate data set were of low quality owing to a misaligned beamstop and were therefore not included in the final data set. The *TRUNCATE* program (French & Wilson, 1978) from the *CCP*4 program suite (Collaborative Computational Project, Number 4, 1994) was used to convert intensities to amplitudes. The final statistics for data collection are shown in Table 1. Initial phases were obtained by molecular replacement with *MOLREP* (Vagin & Teplyakov, 1997) or *Phaser* (McCoy *et al.*, 2007). In all cases, one molecule of the A-TIM–citrate structure (PDB code 2vek; Alahuhta, Salin *et al.*, 2008) without citrate and water molecules was used as the search model. A 5% set of randomly selected reflections was excluded from the refinement and was used to calculate the free *R* factor ($R_{\rm free}$; Brünger, 1997). The test set was kept the same for all of the structures. Before starting the refinement, loop-6 and loop-7 were deleted from the molecular-replacement model. The structures were initially refined using *REFMAC* (Murshudov *et al.*, 1997) and finally using *phenix.refine* from the *PHENIX* software suite (Adams *et al.*, 2002). *Coot* (Emsley & Cowtan, 2004) was used for manual model building and fitting of $F_{\rm o} - F_{\rm c}$ and $2F_{\rm o} - F_{\rm c}$ maps. The geometry libraries for the new ligands were generated with *Coot* using the *Smiles* option to define their covalent structure. Water molecules were added to the refined model using both *phenix.refine* and *Coot*. The final refinement statistics of the structures are summarized in Table 2.

2.6. Structure analysis

The structures were inspected and analysed using *Coot* (Emsley & Cowtan, 2004) and *MolProbity* (Chen *et al.*, 2010). The quality of the structures was also checked using *PROCHECK* (Laskowski *et al.*, 1993). The Ramachandran plots were analyzed using the program *MolProbity* (Chen *et al.*, 2010). *BAVERAGE* from the *CCP*4 program suite (Collaborative Computational Project, Number 4, 1994) was used for *B*-factor analysis. Structure-based superpositions were

Table 1
Data-processing statistics.

	A-TIM (unliganded)	A-TIM–3SP	A-TIM–P3SP	A-TIM–3PGA	A-TIM–4PEH	Sulfate complex
Source	DESY X11	Home source	DESY X11	Home source	DESY X12	Home source
Wavelength (Å)	0.8499	1.5418	0.8499	1.5418	1.0000	1.5418
Detector	MAR 555	MAR 345	MAR CCD 165	MAR 345	MAR Mosaic 225	MAR 345
Temperature (K)	100	100	100	100	100	100
Space group	$P2_1$	$P2_1$	$P2_1$	$P2_1$	$P2_1$	$P2_1$
Unit-cell parameters						
a (Å)	45.5	45.8	45.9	45.3	45.6	45.7
b (Å)	85.3	86.2	86.9	85.8	86.9	85.4
c (Å)	55.9	56.3	56.3	55.6	56.3	56.3
$\alpha = \gamma$ (°)	90.0	90.0	90.0	90.0	90.0	90.0
β (°)	98.6	98.3	98.6	97.7	97.3	98.9
Resolution range (Å)	44.98–2.13 (2.19–2.13)	15.00–1.93 (2.00–1.93)	15.00–1.97 (2.50–1.97)	15.00–1.90 (1.95–1.90)	50.00–1.83 (1.88–1.83)	8.22–1.84 (1.94–1.84)
$R_{\rm merge}$ (%)	6.9 (47.2)	9.7 (39.2)	8.8 (28.1)	6.6 (46.5)	6.1 (45.6)	2.8 (33.6)
$I/\sigma(I)$	14.3 (3.0)	15.7 (4.1)	13.1 (5.1)	13.5 (2.9)	18.4 (3.0)	10.2 (3.0)
Completeness (%)	98.4 (97.6)	99.7 (97.6)	90.6 (85.1)	99.0 (98.9)	99.7 (99.6)	98.8 (100)
Redundancy	3.6 (3.5)	5.6 (4.6)	4.2 (4.2)	3.6 (3.4)	3.8 (3.7)	2.8 (2.5)
Unique reflections	23310 (1709)	32470 (3209)	28058 (13391)	32859 (2351)	38001 (2786)	36615 (5428)
Wilson B factor (Å^2)	35	21	25	27	24	20

Table 2
Refinement statistics.

	A-TIM (unliganded)	A-TIM–3SP	A-TIM–P3SP	A-TIM–3PGA	A-TIM–4PEH	Sulfate complex
Resolution range (Å)	44.98–2.13	15.00–1.93	14.92–1.98	14.72–1.90	43.43–1.83	8.22–1.84
$R_{\rm cryst}$ (%)	18.0	16.0	16.1	16.8	17.2	18.3
$R_{\rm free}$ (%)	26.1	20.0	22.6	22.4	22.0	23.7
Protein atoms	3709	3544	3588	3598	3590	3597
Waters	400	590	515	532	605	434
Ligand molecules	0	2	2	1	1	2
Sulfates	0	2	3	0	0	4
R.m.s.d. bond lengths (Å)	0.007	0.006	0.007	0.007	0.007	0.007
R.m.s.d. bond angles (°)	1.1	1.0	1.0	1.1	1.0	1.0
R.m.s.d. B (Å^2)						
Main chain	2.4	1.4	1.9	1.7	1.7	1.7
Side chain	3.6	2.2	3.1	2.8	2.6	2.9
Average B (Å^2)						
Molecule A (main chain)	30	14	21	21	17	19
Molecule B (main chain)	26	12	18	19	15	18
A ligand	—	14	27	—	—	33
B ligand	—	10	22	38	35	21
PDB code	2x16	2x1s	2x1r	2x2g	2x1t	2x1u

performed using the *SSM* protocol (Krissinel & Henrick, 2004) as implemented in *Coot* (Emsley & Cowtan, 2004). Structure comparisons were carried out using known structures of TIM with the following PDB codes: 5tim (subunit *A* of this PDB entry has the open unliganded conformation of trypanosomal TIM; Wierenga, Noble, Postma *et al.*, 1991), 1n55 [the closed liganded (with 2PG) conformation of leishmanial TIM; Kursula & Wierenga, 2003], 2vel (A-TIM liganded with 2PG; Alahuhta, Salin *et al.*, 2008), 2vek (A-TIM complexed with citrate; Alahuhta, Salin *et al.*, 2008) and 1tph (chicken TIM complexed with PGH; Zhang *et al.*, 1994). Figs. 1, 4, 5, 6, 7 and 8 were produced with *PyMOL* (DeLano, 2002).

3. Results

In this crystal form of A-TIM grown in the presence of 100 m*M* citrate pH 5.5 there are two molecules per asymmetric unit. Previous studies have shown that the citrate in both active sites can be replaced by 2PG. This A-TIM–2PG complex (PDB code 2vel; Fig. 4) will be used as the reference structure. In each of the two molecules in the asymmetric unit (as well as in the A-TIM–citrate complex; PDB code 2vek) the loops of the active site have adopted the closed form, as in the liganded closed structure of wild-type TIM (Casteleijn *et al.*, 2006). In each of the two molecules loop-1 exhibits the highest *B* factors, but a complete C^{α} trace could be built with the same structure for both molecules. Molecule *B* is more tightly packed and has a lower average *B* factor (Table 2); therefore, this molecule was taken as the reference molecule. The crystal-soaking experiments indicated that the cocrystallized citrate molecule is also most tightly bound to molecule *B*, whereas the citrate molecule bound to molecule *A* is more easily removed (Alahuhta, Salin *et al.*, 2008).

3.1. The mode of binding of the sulfate ion

Cross-linked crystals were prepared according to a protocol described previously (Lusty, 1999) and discussed elsewhere

(Wine *et al.*, 2007). The cross-linked crystals were used to study the mode of binding of sulfate. This was tested by transferring such a cross-linked crystal directly into 1.75 *M* ammonium sulfate solution. The cross-linked crystal diffracted to 1.84 Å resolution when analysed using the home X-ray source. The data are of high quality; the positions of the S atoms are marked by high peaks in the anomalous difference map despite the low redundancy of the data set (Table 1). The highest peaks are at the S atoms of Cys126 (6.9σ and 5.5σ, respectively, for molecules *A* and *B*). The peaks in this map for the S atoms of the sulfates in the active sites of molecules *A* and *B* are 5.1σ and 4.5σ, respectively. The active-site sulfate position is almost identical to the mode of binding of the phosphate moiety of 2PG (as also shown in Fig. 7) and loop-6 and loop-7 adopt the closed conformation in both molecules.

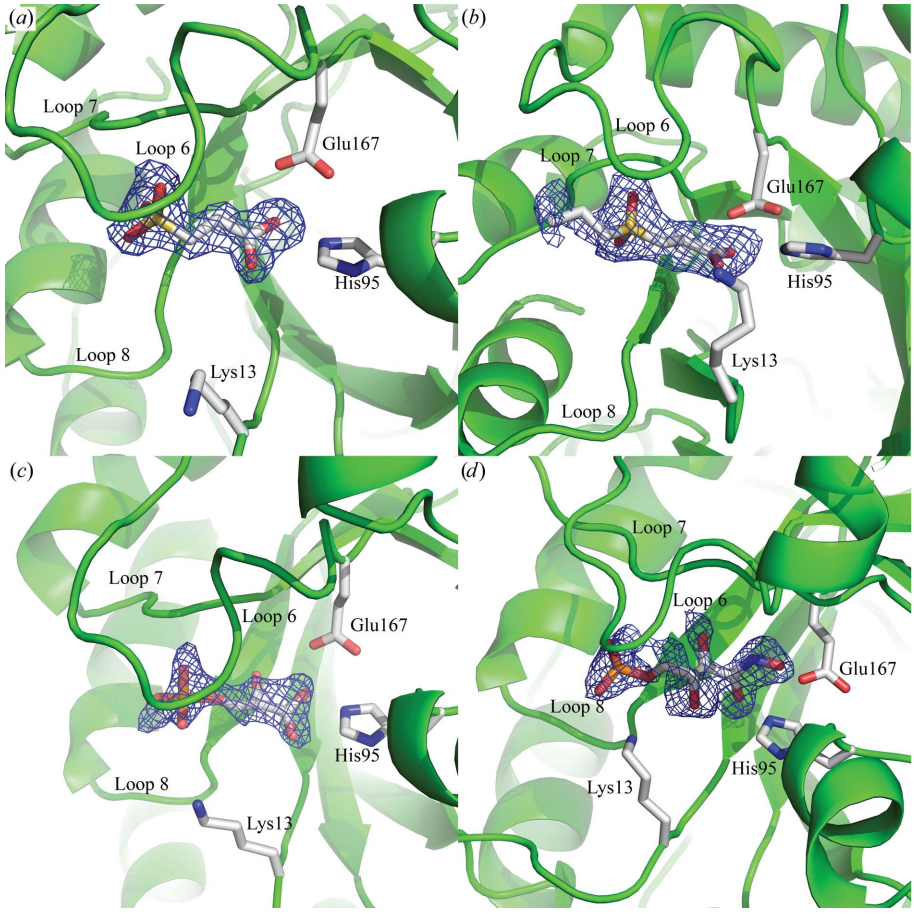

Figure 5
Electron-density ($F_o - F_c$)α_c OMIT maps. The maps were calculated after OMIT refinement, leaving out the highlighted ligand from the model. (*a*) The A-TIM–3SP structure (contoured at 3σ). (*b*) The A-TIM–P3SP structure (contoured at 2σ). (*c*) The A-TIM–3PGA structure (contoured at 3σ). (*d*) The A-TIM–4PEH structure (contoured at 2σ).

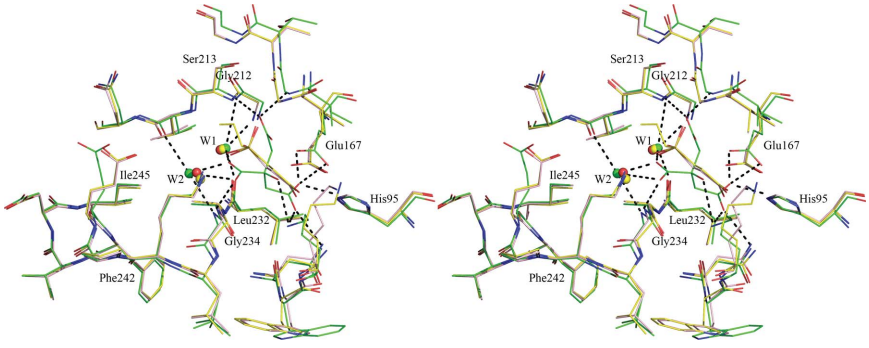

Figure 6
Stereo diagrams of the superimposed C^{α} traces of molecule *B* of A-TIM complexed with citric acid (green; PDB code 2vek), 3SP (salmon pink) and P3SP (yellow). The positions of W1 and W2 are marked; these waters are coloured using the same colouring code as for the protein structures. The side chains of Phe242 and Ile245 are labelled. The hydrogen-bonding interactions that anchor citrate and W1 and W2 to the protein binding site are highlighted by dashed lines.

3.2. The mode of binding of 3SP and P3SP

In both molecules the active-site loops are in the closed conformation, with the 3SP or P3SP molecule bound well in each active site (Fig. 5). The sulfonate moiety binds in the original phosphate-binding site and 3SP binds in an identical manner to that observed for 2PG (PDB code 2vel). The P3SP hydrophobic carbon tail points into the solvent and does not bind into the extended hydrophobic pocket between loop-7 and loop-8 (Fig. 6). The entry into this pocket is rather polar and a water molecule, W2, hydrogen bonded to O(Val214) (loop-7) and N(Gly234) (loop-8) occupies this space (Fig. 4). W2 is also hydrogen bonded to water W1 and both these waters are shown in Fig. 6. W1 is hydrogen bonded to N(Ser213) (loop-7) and O(Leu232) (loop-8). Both water molecules are conserved in the structures of liganded complexes of wild-type TIM and both waters are hydrogen bonded to the same phosphate O atom in these complexes (Kursula *et al.*, 2004) and also to citrate in the A-TIM–citrate complex (Fig. 6). Apparently, the properties of the new binding pocket between loop-7 and loop-8, being rather polar at its entrance, do not favour the binding of the hydrophobic tail of the extended sulfonate molecule into this pocket.

3.3. The unliganded structure

A crystal-washing protocol in which the crystal was washed with mother

liquor in the presence of 100 m*M* MES and 100 m*M* D-R5P followed by a final soak omitting D-R5P from the mother liquor resulted in a crystal structure in which the active sites of both molecules were unliganded. Control experiments in which the final soak was omitted showed that D-R5P is bound in the active site of molecule *B* with low occupancy (data not shown). The crystal structure of the unliganded form at 2.13 Å resolution (Table 2) shows that the active-site loops (loop-6 and loop-7) of both molecules adopt the open conformation, having the same conformation as previously observed in wild-type TIM (Casteleijn *et al.*, 2006). As in wild-type TIM, the *B* factors of loop-6 of both molecules are relatively high in the unliganded structure.

3.4. The mode of binding of 3PGA and 4PEH

Subsequently, crystallographic binding studies were performed with two other substrate sugar phosphate analogues of different length that contained a phosphate moiety. These binding studies were performed with 5 m*M* 3PGA and 20 m*M* 4PEH, respectively (Fig. 3), and data sets were collected at 1.9 and 1.83 Å resolution, respectively (Table 1). The mode of binding of 3PGA and 4PEH is well defined by the electron-density maps (Fig. 5). 3PGA is the carboxylate transition-state analogue of a C4-sugar phosphate molecule, like 2PG is the carboxylate transition-state analogue of the C3-sugar phosphate DHAP, the substrate of TIM (Lolis & Petsko, 1990). 4PEH is the hydroxamate reaction intermediate analogue of the C5-sugar phosphate D-R5P, which is the substrate of D-R5P isomerase (RPI). 4PEH is a hydroxamate and is therefore related to phosphoglycolohydroxamate (PGH), which is the corresponding hydroxamate reaction intermediate analogue of DHAP, the substrate of TIM. 4PEH and PGH are well studied reaction intermediate analogues of RPI (Roos *et al.*, 2005) and TIM (Lolis & Petsko, 1990), respectively.

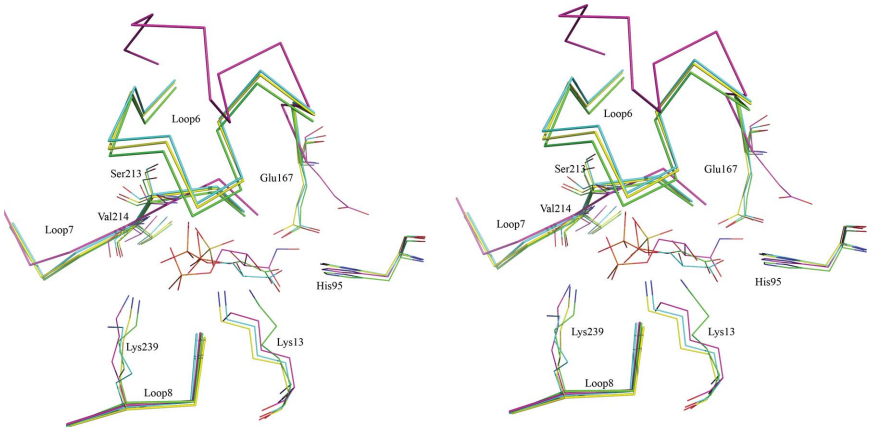

Figure 7
Open and closed states of A-TIM–ligand structures. Stereo diagrams of the superimposed C^{α} traces of the active-site regions (molecule *B*) of A-TIM complexed with sulfate (yellow, closed), 3PGA (blue, closed) and 4PEH (magenta, open). The A-TIM–2PG complex structure (green; PDB code 2vel, closed) has been included as a reference molecule.

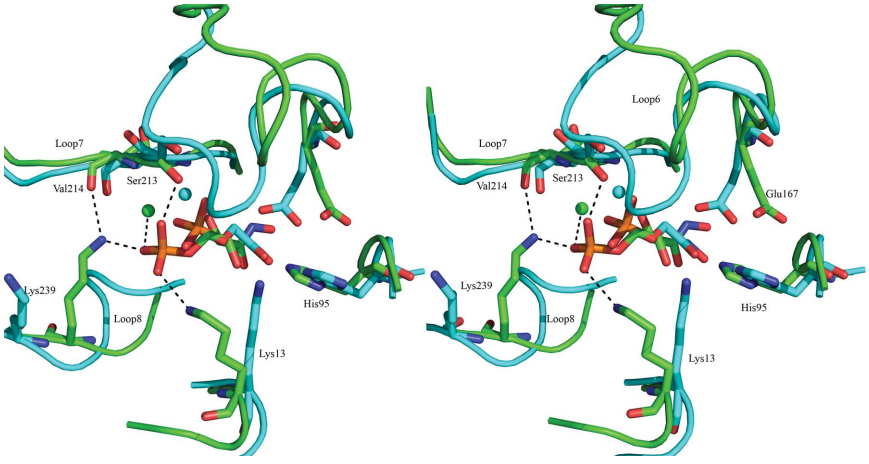

Figure 8
Stereo diagrams of the superimposed C^{α} traces of chicken wild-type TIM complexed with PGH (blue; PDB code 1tph) and A-TIM complexed with 4PEH (green), showing the novel phosphate-binding site near the shortened loop-8. The dashed black lines highlight the hydrogen-bonding interactions of the 4PEH phosphate moiety with the protein part in the A-TIM–4PEH complex.

3PGA and 4PEH bind only in the active-site pocket of molecule *B*. A superposition of the active sites of these complexes is depicted in Fig. 7.

3.4.1. The mode of binding of 3PGA. In molecule *B* both loop-6 and loop-7 are in the closed conformation, with 3PGA bound in the active site (Fig. 5). The phosphate moiety binds 1.2 Å further away from the catalytic centre (near Glu167) compared with the phosphate moiety of 2PG (Fig. 7). The phosphate moiety binds to loop-6 and loop-7, as in wild-type TIM; the interactions of the carboxylic acid O atoms and the Glu167 side chain are also similar to those in the wild-type TIM–2PG complex.

3.4.2. The mode of binding of 4PEH. In molecule *B* both loop-6 and loop-7 are in the open conformation, with 4PEH bound in the active site. The relatively high *B* factors of these two loops indicate high conformational flexibility, as is usual for the open conformation. Compared with the phosphate moiety of the A-TIM–2PG complex, it can be noted that the 4PEH phosphate moiety has moved 1.8 Å away from the catalytic centre (near Glu167) and is bound to loop-8. This mode of binding is not possible in wild-type TIM, but owing to the shortening of loop-8 in A-TIM this new mode of phosphate binding becomes possible. The hydrogen-bonding pattern of the phosphate moiety is different compared with the mode of binding of 2PG. Ser213 has adopted a new conformation that enables hydrogen bonding to the phosphate moiety (Figs. 7 and 8). The phosphate O atoms are also hydrogen bonded to NZ(Lys13), NZ(Lys239), N(Gly234) and a water molecule. Like W2 in the 2PG complex, this

water molecule is hydrogen bonded to O(Val214) and N(Gly234). Comparison of the structure of A-TIM complexed with 4PEH and the structure of chicken TIM complexed with PGH (PDB code 1tph; Zhang *et al.*, 1994; Fig. 8) shows that the hydroxamate moiety of 4PEH is bound in a noncompetent way, being bound to an open active site of A-TIM with the catalytic glutamate side chain in the swung-out conformation. In this mode of binding the 4PEH molecule is nicely hydrogen bonded to the protein, as visualized in Fig. 9; the two hydroxyl moieties of 4PEH also make hydrogen-bonding interactions with main-chain carbonyl O atoms and water molecules.

4. Discussion

In these studies, A-TIM crystals were used to screen the possible modes of binding of a range of different compounds. Firstly, careful crystallization procedures were developed to obtain the high-quality crystals that were used in these efforts. In the course of these studies it was noticed that the citrate molecule is tightly bound to molecule *B*. Therefore, it was important to find a crystal-handling protocol by which the citrate could be flushed out of both active sites. The structure of the crystals obtained using this protocol showed that this is possible as both A-TIM molecules adopt the open unliganded conformation. The experiments with 3SP and P3SP show that the new hydrophobic pocket between loop-7 and loop-8 is unsuitable as a binding pocket for analogues of 2PG with an extended hydrophobic carbon tail. The polar protein environment at the entrance of this pocket, highlighted by the presence of waters W1 and W2, disfavours the binding of such an apolar hydrocarbon tail in this pocket. Interestingly, the mode of binding of 3SP, P3SP (and also 2PG and 3PGA) and citrate reveals that these compounds show common interactions with the catalytic residues, as well as with the closed phosphate-binding loops (loop-6, loop-7 and loop-8; Fig. 7). In particular, it is found that the hydrogen-bonding interactions between the common carboxylate group of the ligand and the side chain of Glu167 are conserved (Fig. 6). The O(ligand)–N(Gly173) (loop-6) and O(ligand)–N(Ser213) (loop-7) hydrogen bonds are also conserved in each of these structures, as in the mode of binding of citrate (Fig. 6). The common hydrogen-bond interactions of sulfate, 3SP, P3SP and 3PGA

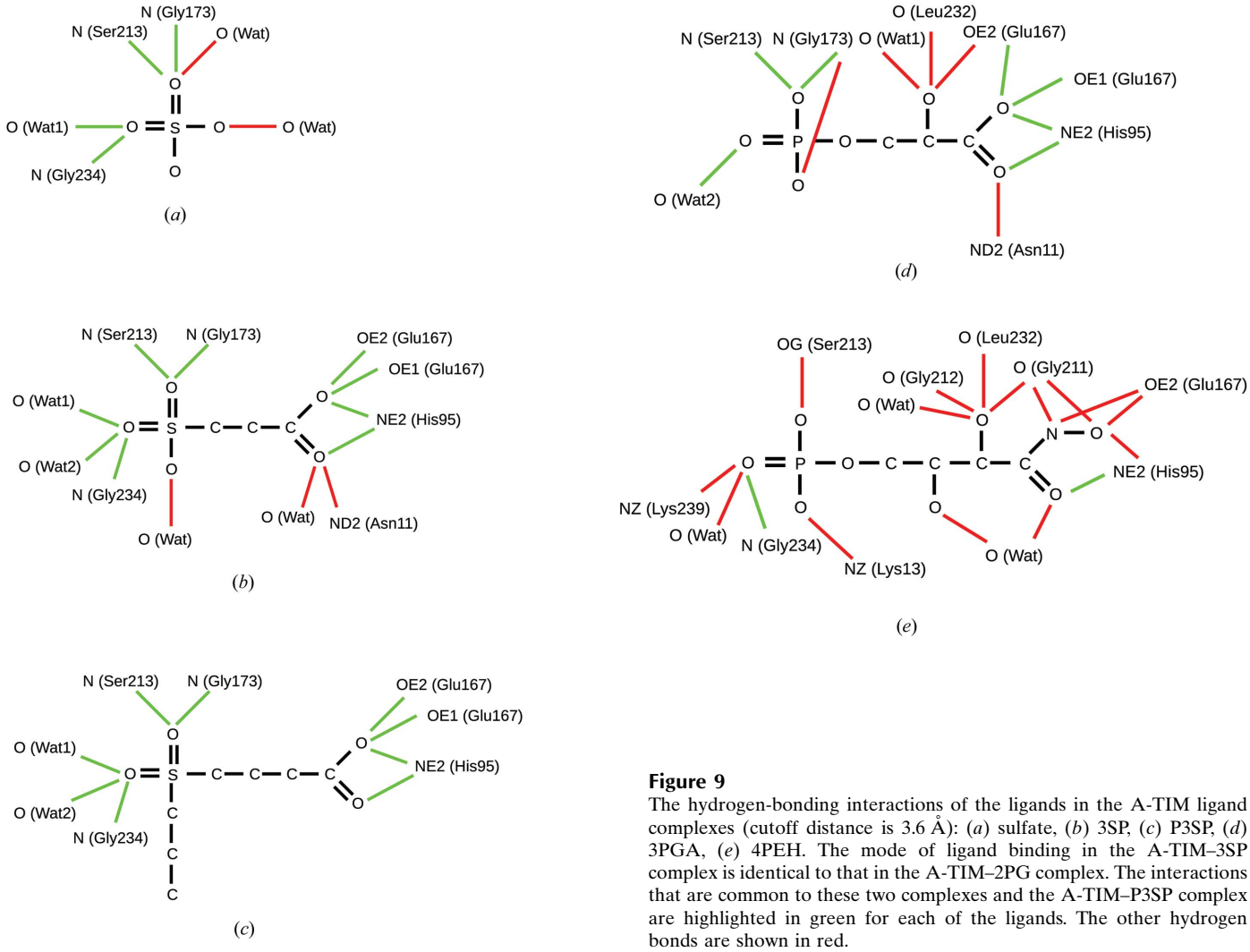

Figure 9
The hydrogen-bonding interactions of the ligands in the A-TIM ligand complexes (cutoff distance is 3.6 Å): (*a*) sulfate, (*b*) 3SP, (*c*) P3SP, (*d*) 3PGA, (*e*) 4PEH. The mode of ligand binding in the A-TIM–3SP complex is identical to that in the A-TIM–2PG complex. The interactions that are common to these two complexes and the A-TIM–P3SP complex are highlighted in green for each of the ligands. The other hydrogen bonds are shown in red.

are also highlighted in Fig. 9. In each of these structures loop-6 and loop-7 are in the closed conformation.

The binding of citrate to A-TIM shows the obvious different binding properties of A-TIM and wild-type TIM (Fig. 6). The affinity/binding properties of A-TIM and wild-type TIM also differ in a more subtle way, which is best appreciated by comparing the covalent structures of 2PG and 3SP (Fig. 3). For 2PG, the phosphate moiety is known to bind to wild-type TIM with the phosphate group in its dianionic state (Campbell *et al.*, 1978, 1979; Kursula & Wierenga, 2003), but the sulfonate moiety of 3SP is monoanionic. Belasco *et al.* (1978) have discussed the very low affinity of the monoanionic sulfate analogue of the substrate, dihydroxyacetone sulfate, for wild-type TIM (at least 100-fold lower affinity compared with DHAP). The crystal structures of the wild-type TIM complexes show that 2PG binds to the phosphate-binding pocket *via* four —NH groups of the three loops loop-6, loop-7 and loop-8 (Kursula & Wierenga, 2003; Kursula *et al.*, 2004). This geometry generates a good phosphodianion-binding pocket, which apparently discriminates against the binding of the monoanion (Belasco *et al.*, 1978). One could speculate that the repulsive electrostatic interactions between the four peptide NH groups of the closed conformation are compensated by the binding of the phosphodianion but not of the monoanion. In A-TIM one of the contributing —NH groups (Gly235) has been deleted because of the shortening of loop-8, thereby apparently facilitating the well defined mode of binding of the monoanionic 3SP.

Interestingly, it is found in these studies that larger sugar phosphate analogues of triosephosphate molecules, such as 3PGA and 4PEH, also bind in the active-site pocket of A-TIM. 4PEH, the hydroxamate analogue of D-R5P, can bind in this pocket because of the shortening of loop-8, but not in the wild-type TIM active site. As shown in Fig. 9, the hydrogen bonds from 4PEH to loop-6 and loop-7 and the catalytic glutamate (conserved in the modes of binding of sulfate, 3SP, P3SP and 3PGA) are different because the protein part has adopted the open conformation. Comparison of the mode of binding of 4PEH to A-TIM with the mode of binding of PGH to wild-type TIM (Fig. 8) indicates that in the A-TIM mode of binding of 4PEH the active-site geometry is not competent for catalysis: loop-6 and loop-7 have adopted the open conformation and Glu167 is in the noncompetent swung-out conformation. Nevertheless, the observed mode of binding of 4PEH shows that a more extended substrate-like molecule can indeed bind to the active site of the A-TIM variant, similarly as observed for wild-type TIM. This is an important result, as it provides a key reference structure for further enzyme-engineering studies as aimed for in the structure-based directed-evolution protocol. Therefore, the directed-evolution experiments of step (iii) in this protocol have now been initiated, using the A-TIM gene as the starting point for making carefully randomized genetic libraries and using as the selection strain an *E. coli* knockout variant in which the RPI genes have been knocked out. If A-TIM variants are obtained with RPI activity through these directed-evolution experiments, then it will be most interesting to see which mutations lead to the new catalytic activity. Subsequent structural enzymological characterization of these variants will then rationalize these sequence changes.

We thank Anni Lahti and Ville Ratas for excellent technical support, Dr Ulrich Bergmann for expert help with the mass-spectrometric measurements and Jaakko Soini for his expertise in fed-batch fermentation. This research was supported by a grant from the Academy of Finland (118167) and by a grant to EGK from the Finnish Cultural Foundation. We thank the staff of beamlines X11 and X12 at EMBL/DESY in Hamburg, Germany for excellent support. The coordinates and structure factors have been deposited with the RCSB for immediate release. The PDB codes are listed in Table 2.

References

Adams, P. D., Grosse-Kunstleve, R. W., Hung, L.-W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W., Read, R. J., Sacchettini, J. C., Sauter, N. K. & Terwilliger, T. C. (2002). *Acta Cryst.* D**58**, 1948–1954.
Alahuhta, M., Casteleijn, M. G., Neubauer, P. & Wierenga, R. K. (2008). *Acta Cryst.* D**64**, 178–188.
Alahuhta, M., Salin, M., Casteleijn, M. G., Kemmer, C., El-Sayed, I., Augustyns, K., Neubauer, P. & Wierenga, R. K. (2008). *Protein Eng. Des. Sel.* **21**, 257–266.
Belasco, J. G., Herlihy, J. M. & Knowles, J. R. (1978). *Biochemistry*, **17**, 2971–2978.
Borchert, T. V., Abagyan, R., Kishan, K. V., Zeelen, J. P. & Wierenga, R. K. (1993). *Structure*, **1**, 205–213.
Borchert, T. V., Kishan, K. V., Zeelen, J. P., Schliebs, W., Thanki, N., Abagyan, R., Jaenicke, R. & Wierenga, R. K. (1995). *Structure*, **3**, 669–679.
Brünger, A. T. (1997). *Methods Enzymol.* **277**, 366–396.
Burgos, E. & Salmon, L. (2004). *Tetrahedron Lett.* **45**, 753.
Campbell, I. D., Jones, R. B., Kiener, P. A., Richards, E., Waley, S. C. & Wolfenden, R. (1978). *Biochem. Biophys. Res. Commun.* **83**, 347–352.
Campbell, I. D., Jones, R. B., Kiener, P. A. & Waley, S. G. (1979). *Biochem. J.* **179**, 607–621.
Casteleijn, M. G., Alahuhta, M., Groebel, K., El-Sayed, I., Augustyns, K., Lambeir, A. M., Neubauer, P. & Wierenga, R. K. (2006). *Biochemistry*, **45**, 15483–15494.
Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* D**66**, 12–21.
Claren, J., Malisi, C., Hocker, B. & Sterner, R. (2009). *Proc. Natl Acad. Sci. USA*, **106**, 3704–3709.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
DeLano, W. L. (2002). *The PyMOL Molecular Viewer.* http://www.pymol.org.
Doi, J. T., Luehr, G. W. & Musker, W. K. (1985). *J. Org. Chem.* **50**, 5716–5719.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Evans, P. (2006). *Acta Cryst.* D**62**, 72–82.
French, S. & Wilson, K. (1978). *Acta Cryst.* A**34**, 517–525.
Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R. D. & Bairoch, A. (2003). *Nucleic Acids Res.* **31**, 3784–3788.
Hassell, A. M. *et al.* (2007). *Acta Cryst.* D**63**, 72–79.
Holme, T., Arvidson, S., Lindholm, B. & Pavlu, B. (1970). *Process Biochem.* **5**, 62–66.
Joseph, D., Petsko, G. A. & Karplus, M. (1990). *Science*, **249**, 1425–1428.
Kabsch, W. (2010). *Acta Cryst.* D**66**, 125–132.
Knowles, J. R. (1991). *Nature (London)*, **350**, 121–124.

Krissinel, E. & Henrick, K. (2004). *Acta Cryst.* D**60**, 2256–2268.
Kursula, P. (2004). *J. Appl. Cryst.* **37**, 347–348.
Kursula, I., Salin, M., Sun, J., Norledge, B. V., Haapalainen, A. M., Sampson, N. S. & Wierenga, R. K. (2004). *Protein Eng. Des. Sel.* **17**, 375–382.
Kursula, I. & Wierenga, R. K. (2003). *J. Biol. Chem.* **278**, 9544–9551.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Lolis, E. & Petsko, G. A. (1990). *Annu. Rev. Biochem.* **59**, 597–630.
Lusty, C. J. (1999). *J. Appl. Cryst.* **32**, 106–112.
McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.
Nagano, N., Orengo, C. A. & Thornton, J. M. (2002). *J. Mol. Biol.* **321**, 741–765.
Neubauer, P., Ahman, M., Tornkvist, M., Larsson, G. & Enfors, S. O. (1995). *J. Biotechnol.* **43**, 195–204.
Noble, M. E., Zeelen, J. P. & Wierenga, R. K. (1993). *Proteins*, **16**, 311–326.
Norledge, B. V., Lambeir, A. M., Abagyan, R. A., Rottmann, A., Fernandez, A. M., Filimonov, V. V., Peter, M. G. & Wierenga, R. K. (2001). *Proteins*, **42**, 383–389.
Reetz, M. T., Wang, L. W. & Bocola, M. (2006). *Angew. Chem. Int. Ed. Engl.* **45**, 1236–1241.
Richard, J. P. (1984). *J. Am. Chem. Soc.* **106**, 4926–4936.
Roos, A. K., Burgos, E., Ericsson, D. J., Salmon, L. & Mowbray, S. L. (2005). *J. Biol. Chem.* **280**, 6416–6422.
Schliebs, W., Thanki, N., Jaenicke, R. & Wierenga, R. K. (1997). *Biochemistry*, **36**, 9655–9662.
Sterner, R. & Hocker, B. (2005). *Chem. Rev.* **105**, 4038–4055.
Thanki, N., Zeelen, J. P., Mathieu, M., Jaenicke, R., Abagyan, R. A., Wierenga, R. K. & Schliebs, W. (1997). *Protein Eng. Des. Sel.* **10**, 159–167.
Vagin, A. & Teplyakov, A. (1997). *J. Appl. Cryst.* **30**, 1022–1025.
Vaismaa, M. J. P., Yliniemelä, S. M. & Lajunen, M. K. (2007). *Z. Naturforsch. B*, **62**, 1317–1323.
Wierenga, R. K. (2001). *FEBS Lett.* **492**, 193–198.
Wierenga, R. K., Noble, M. E., Postma, J. P., Groendijk, H., Kalk, K. H., Hol, W. G. & Opperdoes, F. R. (1991). *Proteins*, **10**, 33–49.
Wierenga, R. K., Noble, M. E., Vriend, G., Nauche, S. & Hol, W. G. (1991). *J. Mol. Biol.* **220**, 995–1015.
Wine, Y., Cohen-Hadar, N., Freeman, A. & Frolow, F. (2007). *Biotechnol. Bioeng.* **98**, 711–718.
Wolfenden, R. (1969). *Nature (London)*, **223**, 704–705.
Zhang, Z., Sugio, S., Komives, E. A., Liu, K. D., Knowles, J. R., Petsko, G. A. & Ringe, D. (1994). *Biochemistry*, **33**, 2830–2837.