the FEBS Journal

Structures of type B ribose 5-phosphate isomerase from *Trypanosoma cruzi* shed light on the determinants of sugar specificity in the structural family

Ana L. Stern[1,*], Agata Naworyta[1,*], Juan J. Cazzulo[2] and Sherry L. Mowbray[1,3]

1 Department of Molecular Biology, Swedish University of Agricultural Sciences, Uppsala, Sweden
2 Instituto de Investigaciones Biotecnológicas-Instituto Tecnológico de Chascomús (IIB-INTECH), Universidad Nacional de General San Martín-CONICET, Buenos Aires, Argentina
3 Department of Cell and Molecular Biology, Uppsala University, Sweden

Keywords
Chagas disease; enzyme specificity; pentose phosphate pathway; type B ribose 5-phosphate isomerase (RpiB); X-ray crystallography

Correspondence
S. Mowbray, Department of Molecular Biology, Box 590, Biomedical Center, SE-751 24 Uppsala, Sweden
Fax: +46 18 53 69 71
Tel: +46 18 471 4990
E-mail: mowbray@xray.bmc.uu.se

*These authors contributed equally to this work

(Received 13 November 2010, revised 17 December 2010, accepted 23 December 2010)

doi:10.1111/j.1742-4658.2010.07999.x

Ribose-5-phosphate isomerase (Rpi; EC 5.3.1.6) is a key activity of the pentose phosphate pathway. Two unrelated types of sequence/structure possess this activity: type A Rpi (present in most organisms) and type B Rpi (RpiB) (in some bacteria and parasitic protozoa). In the present study, we report enzyme kinetics and crystallographic studies of the RpiB from the human pathogen, *Trypanosoma cruzi*. Structures of the wild-type and a Cys69Ala mutant enzyme, alone or bound to phosphate, D-ribose 5-phosphate, or the inhibitors 4-phospho-D-erythronohydroxamic acid and D-allose 6-phosphate, highlight features of the active site, and show that small conformational changes are linked to binding. Kinetic studies confirm that, similar to the RpiB from *Mycobacterium tuberculosis*, the *T. cruzi* enzyme can isomerize D-ribose 5-phosphate effectively, but not the 6-carbon sugar D-allose 6-phosphate; instead, this sugar acts as an inhibitor of both enzymes. The behaviour is distinct from that of the more closely related (to *T. cruzi* RpiB) *Escherichia coli* enzyme, which can isomerize both types of sugars. The hypothesis that differences in a phosphate-binding loop near the active site were linked to the differences in specificity was tested by construction of a mutant *T. cruzi* enzyme with a sequence in this loop more similar to that of *E. coli* RpiB; this mutant enzyme gained the ability to act on the 6-carbon sugar. The combined information allows us to distinguish the two types of specificity patterns in other available sequences. The results obtained in the present study provide insights into the action of RpiB enzymes generally, and also comprise a firm basis for future work in drug design.

Database
Protein structures and diffraction data have been deposited in the Protein Data Bank (http://www.rcsb.org/pdb) under entry codes 3K7O, 3K7P, 3K7S, 3K8C and 3M1P for the wild-type, mutant/Pi, R5P, 4PEH and mutant/All6P structures, respectively

Structured digital abstract
- MINT-8081804, MINT-8081814: *TcRpiB* (uniprotkb:A1BTJ7) and *TcRpiB* (uniprotkb:A1BTJ7) *bind* (MI:0407) by *x-ray crystallography* (MI:0114)

Abbreviations
Allu6P, D-allulose 6-phosphate; All6P, D-allose 6-phosphate; *Ct*RpiB, *Clostridium thermocellum* RpiB; *Ec*RpiB, *Escherichia coli* RpiB; ESRF, European Synchrotron Radiation Facility; β-ME, β-mercaptoethanol; MESNA, sodium 2-mercapto-ethanesulfonate; *Mt*RpiB, *Mycobacterium tuberculosis* RpiB; *Na*RpiB, *Novosphingobium aromaticivorans* RpiB; PDB, Protein Data Bank; 4PEH, 4-phospho-D-erythronohydroxamic acid; PPP, pentose phosphate pathway; R5P, D-ribose 5-phosphate; Rpi, ribose-5-phosphate isomerase; RpiA, ribose-5-phosphate isomerase A; RpiB, ribose-5-phosphate isomerase B; Ru5P, D-ribulose 5-phosphate; *Sp*RpiB, *Streptococcus pneumoniae* RpiB; TBA, thiobarbituric acid; *Tc*RpiB, *Trypanosoma cruzi* RpiB; *Tc*RpiB-wt, wild-type *Tc*RpiB; *Tm*RpiB, *Thermotoga maritima* RpiB.

Introduction

Trypanosoma cruzi, the parasitic protozoan that causes American trypanosomiasis (known also as Chagas disease), has a functional pentose phosphate pathway (PPP) [1]. This pathway has been proposed to have crucial roles in the protection of trypanosomatids against oxidative stress, as well as in the production of nucleotide precursors [2]. All seven enzymes of the PPP can be detected in all four major stages in the biological cycle of the parasite (i.e. the epimastigote and the metacyclic trypomastigote in the insect vector, and the intracellular amastigote and the bloodstream trypomastigote in the infected mammal) [1].

The PPP consists of two branches. The oxidative branch leads from D-glucose 6-phosphate to D-ribulose 5-phosphate, with the reduction of two molecules of NADP. The non-oxidative, or sugar interconversion, branch ultimately leads back to glycolytic intermediates. Ribose-5-phosphate isomerase (Rpi; EC 5.3.1.6) is a key activity of the non-oxidative branch, catalysing the reversible aldose-ketose isomerization of D-ribose 5-phosphate (R5P) and D-ribulose 5-phosphate (Ru5P) (Fig. 1A). The mechanism is considered to involve two steps: an initial opening of the ring form of the sugar most common in solution, followed by the actual isomerization, which is assumed to proceed via a *cis*-enediolate high energy intermediate.

Known Rpis belong to two completely unrelated protein families, both of which are represented in *Escherichia coli* [3,4]. One of them, type A Rpi (RpiA), is a constitutively expressed 23 kDa protein, whereas the other, type B Rpi (RpiB), is a 16 kDa protein that is under the control of a repressor [5–7]. Expression of either enzyme allows normal growth of the bacterium, although growth of the double mutant *rpiA*$^{-}$/*rpiB*$^{-}$ is severely impaired under all experimental conditions tested, showing that the reaction itself is very

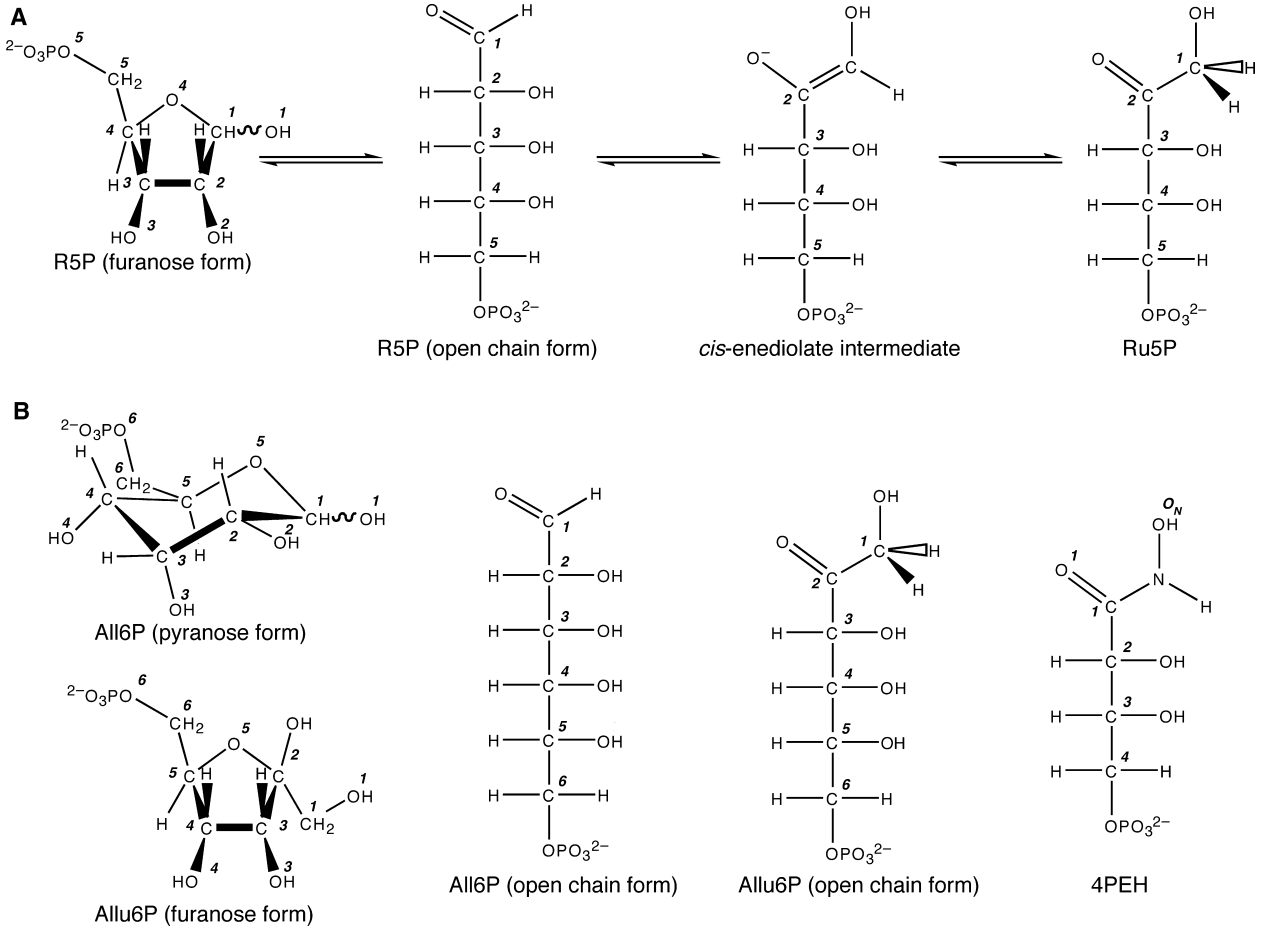

Fig. 1. Reactions and compounds. (A) Isomerization of R5P and Ru5P catalyzed by Rpis. (B) All6P and Allu6P are shown in their open-chain and most common cyclic forms, together with the inhibitor 4PEH. Carbon numbering is given for each sugar.

important for the bacterium [7]. Furthermore, at least one of the known types of Rpi can be identified in every genome sequenced to date. RpiAs are broadly distributed, being found in most eukaryotic organisms, as well as some prokaryotes. Inspection of the protein family database Pfam [8] shows that RpiBs (accession number: PF02502) exist almost exclusively in prokaryotic organisms; there are a few exceptions in the lower eukaryotes, including some trypanosomatids and other parasitic protozoa, as well as some fungi. RpiB-like sequences have also been reported in certain plants, although these are fused to a DNA-damage-repair/toleration protein, and lack some amino acid residues that are linked to binding the substrates.

We recently reported that *T. cruzi* has only a B-type Rpi, which we cloned, expressed and characterized, showing that Cys69 is essential for the isomerization, and that His102 is required for the opening of the furanose ring of R5P [9]. Because RpiBs are absent in all mammalian genomes sequenced so far, this enzyme can be considered as a possible target for the development of new chemotherapeutic agents against the parasite; because the active sites of RpiAs and RpiBs are completely different, the design of highly selective inhibitors should be possible [10].

Among the RpiBs for which biochemical data are available, the sequence of *T. cruzi* RpiB (*Tc*RpiB) was found to be most similar to that of *E. coli* RpiB (*Ec*RpiB) ($\sim$ 40% amino acid identity); it was therefore considered probable that, similar to *Ec*RpiB, *Tc*RpiB would be able to isomerize the 6-carbon sugars D-allose 6-phosphate (All6P) and D-allulose 6-phosphate (Allu6P), in addition to the R5P/Ru5P pair [11,12] (Fig. 1). However, this is not a common property of all RpiBs; the *Mycobacterium tuberculosis* enzyme (*Mt*RpiB) is able to isomerize All6P only with an extremely low catalytic efficiency [13]. Accordingly, we considered it important to perform further studies on *Tc*RpiB specificity. In addition, our previous attempts to identify lead compounds in the development of new drugs against Chagas disease used homology modelling based on *Ec*RpiB; given the moderate sequence identity of the template, it was clearly desirable to obtain the actual 3D structure of *Tc*RpiB.

In the present study, we report that *Tc*RpiB is unable to isomerize All6P, which instead acts as a weak competitive inhibitor of the R5P/Ru5P isomerization. Furthermore, the determination of X-ray structures of wild-type and C69A mutant *Tc*RpiB, with and without bound substrate and inhibitors, allowed us to study in detail the interactions between the enzyme and bound ligands, as well as small conformational changes associated with binding. These studies revealed that the differences in substrate specificity among RpiBs are at least partially the result of changes in the structure of a phosphate-binding loop bordering the active site. Mutation of this loop to make it more similar to that of *Ec*RpiB gave *Tc*RpiB the ability to isomerize All6P. These studies expand our understanding of RpiBs in general and also provide a solid basis for future drug development against *T. cruzi* in particular.

Results

Kinetic studies of wild-type *Tc*RpiB (Tc*RpiB-wt*)

The ability of *Tc*RpiB-wt to isomerize All6P was tested using a discontinuous assay that measures the concentration of Allu6P after derivatization [13]. Isomerization of this 6-carbon sugar could not be detected, even when it was added at a concentration of 30 mM.

The same preparation of *Tc*RpiB had a k_{cat} of 28 s^{-1} and a K_m of 5 mM when R5P was the substrate, measured directly using the A_{290} of Ru5P [14]. The Lineweaver–Burk plot presented in Fig. 2 shows that, when added to the R5P-Ru5P isomerization reaction of *Tc*RpiB, All6P produces the pattern expected for a competitive inhibitor (K_i = 15 mM).

A number of inhibitors that mimic the 6-carbon high-energy intermediate expected for an All6P/Allu6P isomerization [15] were tested [i.e. 5-phospho-D-ribonohydroxamic acid, 5-phospho-D-ribonate, 5-phospho-D-ribonamide, *N*-(5-phospho-D-ribonoyl)-methylamine

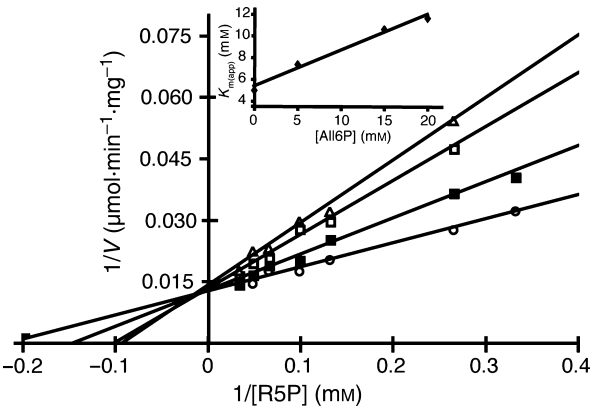

Fig. 2. Inhibition of *Tc*RpiB Rpi activity by All6P. Activity in the isomerization of R5P was tested using a direct spectrophotometric assay, as described within the text. A double-reciprocal (Lineweaver–Burk) plot of initial velocity versus [R5P] is shown, obtained at various concentrations of All6P: 0 mM (open circles), 5 mM (black squares), 15 mM (open squares) and 20 mM (open triangles). The inset graph used for K_i estimations represents the apparent K_m values plotted against [All6P]; the slope of the line is equal to K_m/K_i (R^2 = 0.94).

and *N*-(5-phospho-D-ribonoyl)-glycine]. None of these compounds inhibited *Tc*RpiB significantly, even at concentrations as high as 10 mM. Phosphate did not inhibit at concentrations up to 100 mM.

Structures of *Tc*RpiB and ligand binding

*Tc*RpiB (wild-type or a C69A mutant) was crystallized alone or in the presence of a relevant ligand: phosphate, R5P, 4-phospho-D-erythronohydroxamic acid (4PEH) or All6P (Fig. 1). Data collection and refinement statistics for the five structures solved are summarized in Table 1. All crystals diffracted to high resolution. Most of them exhibited the same space group ($P4_22_12$, with two molecules in the asymmetric unit) with similar cell dimensions; *Tc*RpiB-R5P ($P2_12_12$, with four molecules in the asymmetric unit) was the exception. Each molecule could be traced from residues 1–2 to 152–153 (of a total of 159). The N-terminal 6-His tag (20 residues) was never observed in the electron density. Superimposing the molecules within the various asymmetric units showed that they are very similar, with pairwise rmsd of 0.1–0.2 Å when all C atoms were compared. When aligned using a tighter cut-off (0.5 Å), only residues 39–42 did not always match, showing differences up to 1 Å in some cases. The relatively weak electron density for this segment also suggested some mobility and, in some cases, the conformation could be influenced slightly by crystal packing. However, the stated conclusions apply, regardless of which molecules were used in the comparisons.

For the structures in the $P4_22_12$ space group, the two molecules of the asymmetric unit form a homodimer (Fig. 3A), the major species observed during size exclusion chromatography [9]. Each subunit is based on a Rossmann fold with a five-stranded parallel β-sheet flanked by five-helices, two on one side and three on the other. The sixth (C-terminal) α-helix extends from the main fold and interacts with the second subunit to stabilize the dimer. Dimers interact via crystallographic symmetry to form tetramers. Each subunit of the dimer interacts with both subunits of the second dimer. Hence, residues 113–122 interact with the equivalent regions in one subunit of the second dimer, whereas residues 92–95 make contacts with their equivalents in the other subunit of the second dimer (Fig. 4A). In the case of $P2_12_12$ (*Tc*RpiB-R5P), the four molecules in the asymmetric unit represent the tetramer.

The two active sites of the functional dimer are located in clefts between the subunits, with components drawn from each; residues with numbering < 100 (with the exception of Arg113) from one molecule function together with later residues in the sequence of the other. Strong electron density was seen in both active sites of the wild-type ligand-free structure (Fig. 3B), apparently attached covalently to the active site base, Cys69. In further experiments, reducing agent was included, and protein samples were processed quickly, aiming to avoid potential oxidation of the protein, or reaction between the protein and reducing agent.

The inactive C69A mutant was first crystallized in the presence of high concentrations of phosphate (0.8 M). The observed electron density supported the presence of the ion in each active site (Fig. 3C), although probably at half occupancy. The phosphate, which is largely exposed to solvent, interacts with His11 and Arg113 from one molecule of *Tc*RpiB, together with Arg137′ and Arg141′ (where the prime indicates residues from the other subunit of the functional dimer). We note, however, that multiple conformations of Arg113 are observed in this and all other *Tc*RpiB complex structures. Thus, this side-chain can also interact with Glu112 of the same subunit or Glu118 of a crystallographically-related subunit in the tetramer interface. These multiple conformations do not appear to be related to significant differences elsewhere.

When *Tc*RpiB-wt was crystallized in the presence of R5P, electron density in the active site clearly showed that a linear sugar molecule was bound (Fig. 3D). Again, the phosphate group interacts with His11, Arg113, Arg137′ and Arg141′. The other end of the substrate points into a deep pocket in the enzyme. Moving along the ligand from the phosphate, O4 interacts with His102′ and a water molecule that is in turn within hydrogen-bonding distance of Tyr46, His138′ and Arg141′. O3 hydrogen bonds to Asp10, as well as to the backbone amide nitrogen of Gly70. O2 hydrogen bonds with water, and the backbone nitrogen of Ser71. At the far end, O1 interacts with Asn103′, and the backbone nitrogen of Gly74.

*Tc*RpiB-wt was also crystallized with the linear inhibitor, 4PEH (Fig. 3E; K_i = 1.2 mM) [9]. Hydrogen bonds to the phosphate group are as described above. The O2 and O3 of 4PEH correspond to O3 and O4 of the R5P structure (Fig. 1). Accordingly, O3 of 4PEH interacts with His102′ and a water molecule, whereas O2 hydrogen bonds to Asp10 and to the backbone amide nitrogen of Gly70. O1 of 4PEH interacts with the backbone nitrogen of Ser71 as seen for the O2 interaction in R5P. As for O1 of R5P, the terminal group of the inhibitor has hydrogen bonds to Asn103′ and the backbone nitrogen of Gly74; however, in the

Table 1. Data collection and refinement statistics. Information shown in parentheses refers to the highest resolution shell.

	*Tc*RpiB-wt	*Tc*RpiB-C69A/Pi	*Tc*RpiB-wt/R5P	*Tc*RpiB-wt/4PEH	*Tc*RpiB-C69A/All6P
Data collection statistics					
Data collection beamline/detector	ESRF ID14:1/ ADSC Q210 CCD	ESRF ID14:1/ ADSC Q210 CCD	ESRF ID14:2/ ADSC Q4 CCD	ESRF ID14:2/ ADSC Q4 CCD	ESRF BM30A/ ADSC Q315r CCD
Space group	$P4_22_12$	$P4_22_12$	$P2_12_12$	$P4_22_12$	$P4_22_12$
Cell axial lengths (Å)	93.2, 93.2, 93.7	93.0, 93.07, 93.82	92.3, 92.3, 93.6	92.4, 92.4, 94.0	92.7, 92.7, 93.1
Resolution range (Å)	30.0–2.0 (2.11–2.00)	33.04–1.40 (1.48–1.40)	24.7–1.9 (2.00–1.90)	30.0–2.1 (2.21–2.10)	29.4–2.15 (2.27–2.15)
Number of reflections measured	141 283	455 063	291 291	108 089	131 784
Number of unique reflections	28 418	79 637	63 050	23 175	20 303
Average multiplicity	5.0 (5.0)	5.7 (5.7)	4.6 (4.7)	4.7 (4.0)	6.5 (6.5)
Completeness (%)	99.7 (100.0)	98.2 (99.8)	99.3 (99.1)	96.0 (95.9)	90.3 (92.8)
Rmeas (%)	0.078 (0.641)	0. 097 (0.335)	0.075 (0.332)	0.097 (0.451)	0.108 (0.283)
<<I>/σ<I>>	13.1 (4.6)	11.1 (3.1)	20.3 (5.2)	14.7 (3.6)	15.9 (4.3)
Wilson B-factor ($Å^2$)	31.0	15.5	17.4	22.3	23.5
Refinement statistics					
Resolution range (Å)	30.0–2.0	30.0–1.4	24.7–1.9	30.0–2.1	29.4–2.15
Number of reflections used in working set	26 413	75 593	59 812	21 960	17 931
Number of reflections for R_{free} calculation	1320	3779	2990	1098	896
R-value, R_{free} (%)	19.9, 22.3	22.3, 23.2	17.0, 19.5	21.6, 25.3	18.7, 23.2
Number of nonhydrogen atoms	2498	2603	5358	2481	2583
Number of solvent waters	160	177	480	90	136
Mean B-factor, protein atoms, A and B molecules ($Å^2$)	30.2, 30.5	16.2, 16.1	17.1, 17.7	24.0, 20.0	21.1, 20.5
Mean B-factor, solvent atoms ($Å^2$)	38.5	22.3	28.2	19.5	26.0
Mean B-factor, ligand atoms, ($Å^2$)	–	26.9[a]	21.6	19.4	22.6
Ramachandran plot outliers (nonglycine) (%)[b]	0	0	0.7	1.0	1.4
rmsd from ideal bond length (Å)[c]	0.010	0.006	0.008	0.013	0.010
rmsd from ideal bond angle (°)[c]	1.1	0.9	1.0	1.4	1.1
PDB entry code	3K7O	3K7P	3K7S	3K8C	3M1P

[a] 50% occupancy. [b] Calculated using a strict-boundary Ramachandran plot [16]. The very few (and slight) outliers are in regions of higher mobility. [c] Using the parameters of Engh and Huber [17].

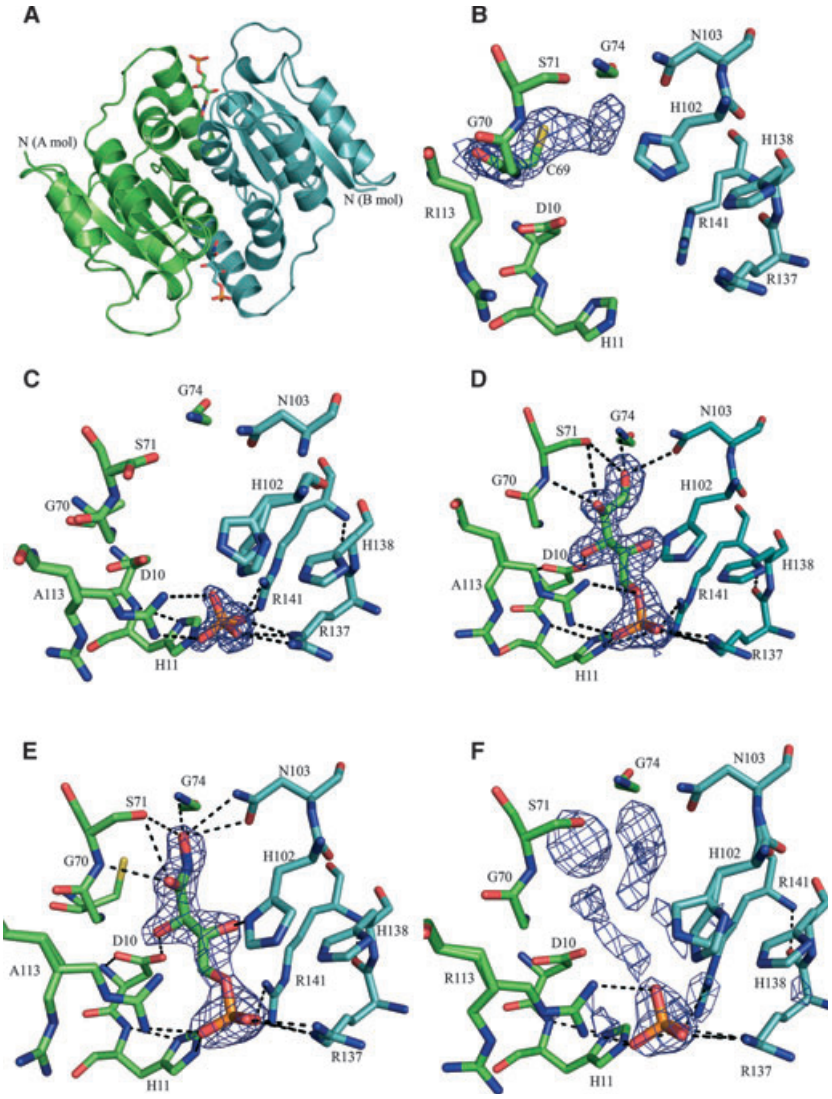

Fig. 3. Structures of *Tc*RpiB. (A) A cartoon drawing shows the overall fold, and the dimer (with subunits coloured cyan and green). The active sites (indicated by linear sugar molecules) are located between the two subunits, with residues contributed by both (as described within the text). (B–F) Showing the active sites in the various structures, solved with similar views and colouring for the carbon atoms. Modelled ligands are shown together with their electron density, using SIGMAA-weighted $|2F_{obs} - F_{calc}|$ maps [18] contoured at 1 σ. (B) *Tc*RpiB-wt, showing possibly oxidized cysteine in the active site (σ = 0.23 e/Å^3). (C) *Tc*RpiB-C69A in complex with phosphate ion (σ = 0.33 e/Å^3). (D) *Tc*RpiB-wt in complex with R5P/Ru5P (σ = 0.33 e/Å^3). (E) *Tc*RpiB-wt in complex with 4PEH (σ = 0.29 e/Å^3). (F) *Tc*RpiB-C69A in complex with All6P (σ = 0.27 e Å^{-3}). Hydrogen bonds as discussed in the text are shown as dashed lines.

4PEH structure, the distance between O_N and Gly74 is shorter (2.7 Å, average of both subunits) than the equivalent distance in R5P (3.0 Å, average of four subunits). The structure of *Tc*RpiB-C69A bound to 4PEH was identical to that of the wild-type complex (not shown).

*Tc*RpiB-C69A was further crystallized with the weaker inhibitor, All6P (K_i = 15 mM). Electron density for this ligand (Fig. 3F) was noticeably poorer in both active sites compared to that seen for other complex structures. The phosphate group lies at the same place, although the rest of the sugar is much less well defined. The electron density suggests that All6P is bound primarily as the linear form, although with mixed binding modes. This density did not improve after cyclic averaging, or when higher concentrations of All6P (upto 50 mM) were included; for these reasons, only the phosphate moiety of the sugar has been modelled in the structure deposited.

Comparison of *Tc*RpiB structures

The various structures of *Tc*RpiB exhibited rmsd in the range of 0.15–0.3 Å when their C atoms were aligned, with most atoms matching within a 0.5 Å cut-off.

When comparing *Tc*RpiB-wt (the ligand-free structure) with the complexes with R5P or 4PEH, the most striking difference is a 1.5–1.8 Å movement of the main chain at residues 10–12. Asp10 and His11 interact with R5P and 4PEH in similar ways, drawing this segment further into the active-site pocket. The movement is coupled to changes in the mobile loop at residues 42–45.

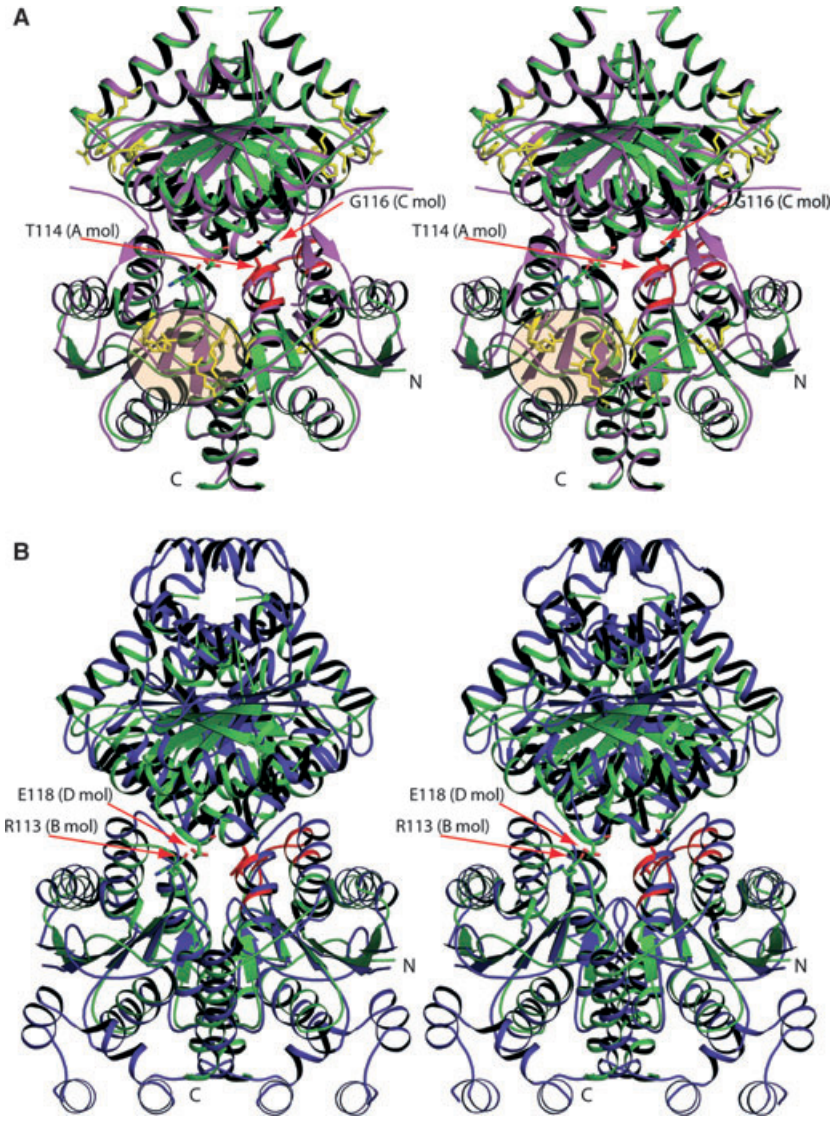

Fig. 4. Comparison of RpiB tetramer structures (stereo views). Tetrameric *Tc*RpiB (green) is superimposed on *Ec*RpiB (magenta) in (A) and *Sp*RpiB (blue) in (B). The N- and C-termini are labelled in molecule A of *Tc*RpiB. In the same molecule, two segments that make contacts in the tetramer interface are coloured red, and two contacting residues (*Tc*RpiB numbering) are labelled. Residues in all four active sites are shown as a yellow stick representation, and the active site of molecule B is circled.

The close similarity between *Tc*RpiB-C69A/Pi and *Tc*RpiB-C69A/All6P indicates that binding phosphate and All6P (of which only the phosphate group is ordered in the electron density) have equivalent effects on the protein. The conformation observed for the mobile loops in these structures is midway between that for the apo/Pi and R5P/4PEH structures, presumably because the phosphate ion interacts with His11 but not Asp10.

Other differences include alternative side-chain conformations that were modelled for His102 and Arg113. The side-chain of His102 in *Tc*RpiB-wt is turned $\sim 90°$ compared to the same residue in the rest of the structures. This residue also has multiple conformations in both structures of mutated protein (i.e. *Tc*RpiB-C69A/Pi and *Tc*RpiB-C69A/All6P). In all the *Tc*RpiB complex structures presented, Arg113 has two different conformations: one pointing towards the phosphate group of the ligand and the other pointing out into solution. In the *Tc*RpiB-wt (i.e. ligand-free) structure, Arg113 is only in the latter conformation (Fig. 3).

Comparison of *Tc*RpiB with other structures

*Tc*RpiB is compared with structures found in the Protein Data Bank (PDB) (including three that are unpublished) in Table 2 and Fig. 4. The majority of Cα atoms match within a 2 Å cut-off when the dimers are compared. As in *Tc*RpiB, a helix at the C-terminus of *Ec*RpiB, *Thermotoga maritima* RpiB (*Tm*RpiB) and *Clostridium thermocellum* RpiB (*Ct*RpiB) (L.W. Kang,

Table 2. Comparison of available RpiB structures with *Tc*RpiB-wt using LSQMAN. Sequences were arranged in order of similarity in a BLAST search. 4PEA, 4-phospho-D-erythronate.

Protein	PDB code	Ligand bound	Number of residues in sequence	Number of Cα atoms within 2 Å cut-off	rmsd to *Tc*RpiB-wt (Å)	Sequence identity of matching residues (%)	Contact in dimer interface, per subunit (Å^2)	Contact in tetramer interface, per dimer (Å^2)	Reference
*Tc*RpiB		–	159	–	0.0	100	1700	1200	Present study
*Ct*RpiB	3HE8	Glycerol	148	258	0.72	48	1939	973	L.W. Kang, J.K. Kim, J.H. Jung and M.K. Hong (unpublished data)
	3HEE	R5P	148	262	0.73	48	1938	986	L.W. Kang, J.K. Kim, J.H. Jung and M.K. Hong (unpublished data)
*Ec*RpiB	1NN4	Pi	150	255	0.9	43	1490	2940, 1320[a]	[10]
	2VVR	–[b]	149	264	0.9	42			[13]
*Tm*RpiB	1O1X	MPD	155	250	0.94	45	1611	1225	[19]
*Mt*RpiB	1USL	Pi	170	262	0.86	40	1990	–	[20]
	2BES	4PEH	172	259	0.84	40		–	[21]
	2BET	4PEA	172	260	0.84	39		–	[21]
	2VVP	R5P/Ru5P	162	259	0.82	40		–	[13]
	2VVQ	4PRH	162	255	0.84	40		–	[13]
	2VVO	All6P	162	257	0.85	42		–	[13]
*Sp*RpiB	2PPW	SO_4	216	168	1.17	28	1711	1890	R. Wu, R. Zhang, J. Abdullah, and A. Joachimiak (unpublished data)
*Na*RpiB	3C5Y	–	231	178	1.11	23	1873	1924	Joint Center for Structural Genomics (unpublished data)

[a] With and without His-tag sequence. [b] The electron density shows a mixture of sugar forms in the active site, although none were included in the PDB file.

J.K. Kim, J.H. Jung and M.K. Hong, unpublished) is an important component of the dimer interface. In *Mt*RpiB, an extension at this end of the protein produces additional interactions that stabilize the dimer. An even longer extension is found in *Streptococcus pneumoniae* RpiB (*Sp*RpiB; R. Wu, R. Zhang, J. Abdullah and A. Joachimiak, unpublished data) and *Novosphingobium aromaticivorans* RpiB (*Na*RpiB; Joint Center for Structural Genomics, unpublished data), which serves primarily to enlarge the structure of the subunit, rather than enhancing dimer interactions. All but *Mt*RpiB form a dimer of dimers (i.e. a tetramer) as a result of crystallographic and/or noncrystallographic symmetry. As for *Tc*RpiB, *Ec*RpiB and *Tm*RpiB tetramers are the consequence of interactions of two segments from each subunit (Fig. 4A). *Ct*RpiB is described as a dimer in the PDB header, although a comparable tetramer is formed by crystallographic symmetry. *Na*RpiB has a four-residue insertion near residue 116 of *Tc*RpiB and, in the resulting tetramer, the second dimer is similarly placed but with a different 'tilt' relative to the first, compared to the above-named structures (Fig. 4B). *Sp*RpiB is described as a dimer in the PDB header, although our analysis suggests that it actually forms a tetramer via a crystallographic symmetry very similar to that found in *Na*RpiB.

In Fig. 5A, the binding of R5P in the active sites of *Tc*RpiB and *Mt*RpiB is compared. Interactions with the substrate are almost completely conserved. The most noteworthy difference is that, in *Tc*RpiB, the catalytic base that transfers a proton between C1 and C2 in the isomerization step is a cysteine (Cys69), whereas, in *Mt*RpiB, the base is a glutamic acid (Glu75) originating later in the sequence but terminating in the same position. The simultaneous transfer of a proton between O1 and O2 is catalysed by the side-chain of Ser71 in both cases. Both enzymes also have the Gly70-Gly74 segment that creates an anion hole stabilizing the *cis*-enediolate intermediate of the reaction. Arg113, a phosphate ligand in the *Mt*RpiB structure, has a different conformation in *Tc*RpiB but is free

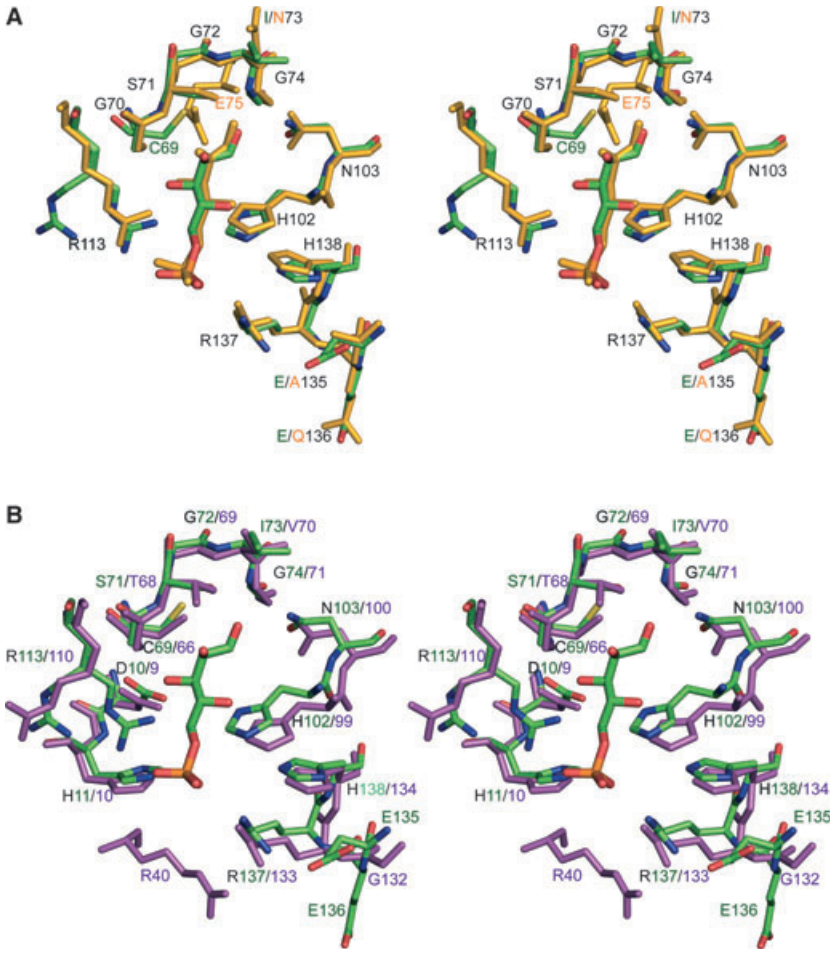

Fig. 5. Structural basis of substrate selectivity. In (A) and (B), the active site of *Tc*RpiB with bound R5P (atomic colours with green carbons) is compared with *Mt*RpiB with bound R5P (orange model) and ligand-free *Ec*RpiB (magenta model), respectively. Residues participating in catalysis, including those forming an anion hole, and interactions with ligand are shown as discussed within the main text. Residues that are the same for both structures under comparison are labelled in black, and the remainder are shown in agreement with the colouring convention for particular structures. The loop altered in the *Tc*RpiB mutant $\triangle^{135}$E136G is also shown in both panels.

to assume a conformation that allows phosphate interactions.

The active site of *Tc*RpiB/R5P is compared with the *Ec*RpiB/apo structure in Fig. 5B. Both enzymes include an active-site cysteine, and the serine (or threonine) and anion hole components are also highly similar. Because these groups are responsible for the catalytic steps, we use them as anchor points in the alignments when considering differences in the rest of the active site that might be linked to substrate specificity. Interactions with Asp10 and His11 (*Tc*RpiB numbering) are likely preserved, although these residues probably move when substrate binds, as noted for the *Tc*RpiB structures above. Arg40 of *Ec*RpiB provides a potential interaction with the phosphate of the substrate that is not present in either *Tc*RpiB or *Mt*RpiB, although it might be more suitable for a substrate longer than R5P. Again, the equivalent of Arg113 is observed in different conformations in the various structures. Residues drawn from the second subunit of the dimer differ more in position relative to the catalytic residues. However, the most striking change is linked to a deletion in the *Ec*RpiB sequence (one residue near 135 in *Tc*RpiB numbering) that moves the equivalents of Arg137 and His138 further away from the catalytic residues; this loop is referred to as the 137-loop in further discussions. This change could additionally affect the relative position of His102.

*Sp*RpiB and *Na*RpiB are less straightforward to compare. The C atoms at the anion hole, including those of the catalytic cysteine and threonine, align very well, and residues equivalent to Tyr46 and Asn103 in *Tc*RpiB are also conserved. However, Asp10 of the *T. cruzi* enzyme is replaced by a glutamate in both *Sp*RpiB and *Na*RpiB, and His11, His102 and His138 are also absent in these two proteins. Furthermore, an insertion in the 137-loop remodels several aspects of the putative phosphate-binding site.

Deletion mutation of *Tc*RpiB (Δ^{135}E136G)

A mutation experiment was undertaken to create a version of *Tc*RpiB that was more similar to *Ec*RpiB

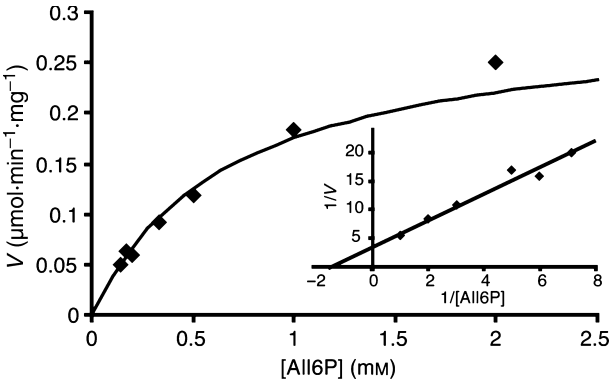

Fig. 6. All6P isomerase activity of *Tc*RpiB-Δ^{135}E136G. A direct plot of All6P isomerase activity of the deletion-mutant enzyme is shown together with the curve calculated from the Michaelis–Menten equation using K_m = 0.7 (mM) and V_{max} = 0.297 (μmol·min^{-1}·mg^{-1}). The inset shows the same data in a Lineweaver–Burk plot.

in the above-mentioned 137-loop (i.e. Δ^{135}E136G). Kinetic analysis indicated that the mutant enzyme had a k_{cat} of 0.15 ± 0.06 s^{-1} and a K_m of 0.8 ± 0.1 mM for the All6P isomerase activity (Fig. 6). When using R5P as a substrate, the k_{cat} of the mutant protein was 16 s^{-1}, and the K_m was 7 mM.

Discussion

We previously experienced problems obtaining complexes of *Ec*RpiB [13], a frustrating contrast to the situation with *Mt*RpiB [13,21]. The difference is attributable to a highly reactive active-site cysteine in *Ec*RpiB. We note further that, in the *Tm*RpiB and *Na*RpiB structures, the active-site cysteine was oxidized (modelled as cysteine sulfonic acid and cysteine-*S*-dioxide, respectively), which may be correlated with the lack of complexes for these enzymes, as well (Table 2). In the present study, we solved a similar problem with *Tc*RpiB (Fig. 3B) by including β-mercaptoethanol (β-ME) in the various protocols, and working quickly. The modified procedure allowed us to obtain clear electron density for a number of complexes (Fig. 3C–F). In the case where R5P was added, the sugar in the active site is expected to be a mixture of R5P and Ru5P. In solution, R5P is present at approximately three-fold higher concentrations than Ru5P [22]; however, it is not possible to make a reliable estimate of the proportions bound to the protein based on the electron density because of the strong similarity of the two sugars.

Our kinetic data show that *Tc*RpiB has values of k_{cat} and K_m similar to those reported previosuly (12 s^{-1} and 4 mM, respectively) [9] and consistent with those normally observed for other RpiBs (Table 3). The 6-carbon sugar, All6P, is not a substrate for *Tc*RpiB, even at a concentration of 30 mM. Considering the sensitivity of the assay, this suggests that k_{cat} in this case is 0.015 s^{-1} or less, if K_m is 20 mM or less. All6P instead acts as an inhibitor of the R5P/Ru5P isomerization of *Tc*RpiB (K_i = 15 mM). However, in the structure with the *Tc*RpiB-C69A mutant, clear electron density was only seen for the phosphate group of All6P. In light of this, it might be appropriate to consider whether the phosphate group accounts for most of the All6P inhibition. Phosphate alone is a very poor inhibitor; no inhibition was observed when it was added at concentrations as high as 100 mM, and the electron density in the complex with phosphate suggests only partial occupancy, indicating that the K_i is in the order of 800 mM. Comparison of the available RpiB structures suggests that allosteric changes do not occur purely as a result of phosphate binding. This type of behaviour has been reported for other enzymes that act on phospho-sugars, even when interactions with the phosphate group account for most of the binding energy; in the case of triose phosphate isomerase, phosphate alone inhibits only weakly, although the phosphate moiety of the substrate is necessary for allosteric changes that make binding much tighter in the transition state [23].

The swap of the catalytic base (i.e. cysteine/glutamate) does not change how the enzymes interact with

Table 3. Comparison of available kinetic data.

	R5P			All6P		
Enzyme	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m (s^{-1}·M^{-1})	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m (s^{-1}·M^{-1})
*Tc*RpiB	5	28[a]	5600	NA[b]	< 0.015	NA[b]
*Tc*RpiB- Δ^{135}E136G	7	16	2300	0.8	0.15	190
*Mt*RpiB[c]	1.0	47	47 000	16	0.22	14
*Ec*RpiB[c]	1.1	52	47 300	0.5	6.1	12 000

[a] The k_{cat} previously reported by Stern *et al.* [9] is 12 s^{-1}, with the difference probably being a result of a change in the purification protocol (see Experimental procedures). [b] No activity observed at a substrate concentration of 30 mM. [c] As reported previously [13].

the 5-carbon substrates; the R5P/Ru5P complexes of *Tc*RpiB and *Mt*RpiB are highly similar (Fig. 5A) Thus, a difference in the base does not explain why *Ec*RpiB can effectively catalyze the All6P/Allu6P conversion but *Tc*RpiB and *Mt*RpiB cannot. Changes in a loop that includes Arg137 (the 137-loop) that make the active site longer in the *E. coli* protein offered a more promising explanation (Fig. 5B); a similar loop in *Ct*RpiB is also associated with an ability to act on both 5- and 6-carbon (nonphosphorylated) sugars [24,25]. This hypothesis was tested by mutation. *Ec*RpiB has a glycine instead of two glutamic acid residues just before the Arg137 of *Tc*RpiB and *Mt*RpiB (Fig. 7). A *Tc*RpiB mutant protein was therefore constructed that has the same sequence as *Ec*RpiB in this loop (i.e. with a glycine replacing Glu135 and Glu136). The isomerization of R5P/Ru5P was not significantly changed in the Δ^{135}E136G mutant (Table 3). However, unlike wild-type *Tc*RpiB, this mutant enzyme was able to isomerize the 6-carbon sugar (Fig. 6), although not with high efficiency. Clearly, other features of the structures must modulate the specificity of *Ec*RpiB, although the results of the present study strongly suggest that the nature of the 137-loop is an important factor. We predict that *Tm*RpiB and *Ct*RpiB, which have a shorter 137-loop similar to that of *Ec*RpiB (and a cysteine base), will also be able to isomerize All6P. *Sp*RpiB and *Na*RpiB are unpublished structures described in the PDB as (putative) RpiBs. However, we are unaware of any biochemical data concerning the activity of these enzymes, and a comparison of site-active features (as described in the Results and shown in Fig. 7) suggests that this is not their functional role. Both of these genomes also have an *Ec*RpiB-like enzyme, and *Streptococcus* has an RpiA as well, meaning that Rpi activity is not a requirement for either *Sp*RpiB or *Na*RpiB.

Two previous studies [23,26] concerning the completely unrelated family of β-barrel enzymes are relevant here. First, for triose phosphate isomerase (which catalyses an aldose/ketose isomerization similar to that of RpiBs), most of the rate acceleration in catalysis is derived from the energy of binding the phosphate group, which is accomplished via conformational changes in the enzyme [23]. Second, for the epimerization at carbon 3 catalyzed by the metal-dependent enzymes D-ribulose 5-phosphate 3-epimerase (which accepts only 5-carbon substrates) and D-allulose 6-phosphate 3-epimerase (which can accept both 5 carbon and 6 carbon substrates), changes in the length and/or structure of the phosphate-binding loop are linked to changes in sugar preferences, primarily expressed as differences in k_{cat} [26]. It appears that the relationships between k_{cat} and K_m in enzymes catalyzing this type of reaction are unusually intimate and complex.

In accordance with the lack of All6P isomerase activity, none of the compounds that mimics the 6-carbon high-energy intermediate expected for an All6P/Allu6P isomerization [15] inhibited *Tc*RpiB. Some of these compounds do inhibit *Mt*RpiB [13]. The difference in the behaviour of the two enzymes appears to arise from the slightly broader active site in *Mt*RpiB as a result of the switch of the catalytic base. Accordingly, the longer linear inhibitors can bind slightly more deeply in the *Mt*RpiB active site than the R5P substrate, in a manner that is blocked by Cys69 of *Tc*RpiB. The broader active site also enables All6P to bind as the ring form in the structure with *Mt*RpiB, in contrast to the disordered (and probably linear) sugar observed in the present study for *Tc*RpiB (Fig. 3F).

Although the dimer comprises a complete functional unit, tetramers are observed in all available RpiB structures except that of *Mt*RpiB (Fig. 4A,B and Table 2). There does not appear to be any aspect of this fold that requires a tetramer for stability. Furthermore, gel filtration experiments at lower protein concentrations typically show dimers as the major form in solution. Although larger aggregates will be favoured at the higher protein concentrations in the crystallizations, the consistency with which highly similar tetramers are observed is intriguing. The reason may lie in the structure of some larger functional assembly as yet unknown.

We aimed to establish the relative frequency of particular features in RpiB sequences. In Fig. 7, the sequences of the structures discussed in the present study are aligned, together with others that are representative. The starting point was the 759 RpiB-family sequences in the Pfam database (http://pfam.sanger.ac.uk/) at the time of writing. Of these, 149 are recognizable galactose-6-phosphate isomerases, namely LacA/Bs (which are considered to function as a heterodimer, and so these sequences are split evenly between LacAs and LacBs). Most of the remainder (n = 541) have a catalytic base like that of *Ec*RpiB (i.e. a cysteine). Among these, 348 have the shorter 137-loop (with the characteristic GGRH, boxed in Fig. 7), whereas 91 have the longer loop (gxxRH); in both groups, the last residue and one other phosphate-binding residue, His11, are sometimes replaced by serines that could be functionally equivalent in phosphate binding. Of the 541 sequences, 102 are distinctly different from *Ec*RpiB in the C-terminal regions and so cannot be analyzed with confidence; *Sp*RpiB and *Na*RpiB fall into this class. Fewer proteins (55) have

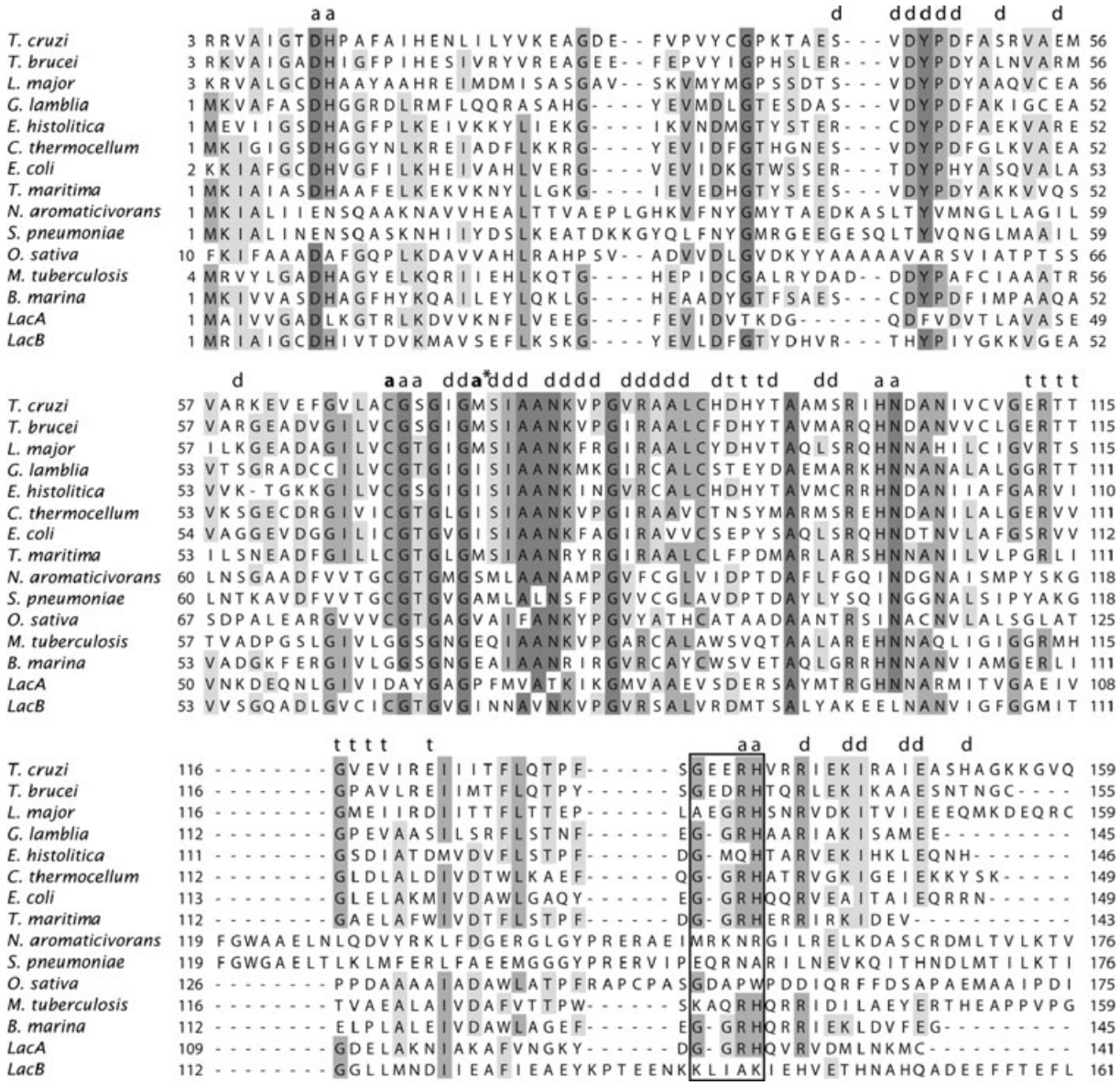

Fig. 7. Sequence comparisons. Alignment of RpiBs from eukaryotic organisms, *Trypanosoma cruzi* (GI: 110984574), *Trypanosoma brucei* (GI: 70834348), *Leishmania major* (GI: 68127548), *Giardia lamblia* (GI: 159119546), *Entamoeba histolytica* (GI: 56468369) and *Oryza sativa* (GI: 22535714), together with bacterial RpiBs from *Clostridium thermocellum* (GI: 266618622), *E. coli* (GI: 85676843), *Thermotoga maritima* (GI: 15643838), *Novosphingobium aromaticivorans* (GI: 146275740), *Streptococcus pneumoniae* (GI: 15900251), *Mycobacterium tuberculosis* (GI: 57116993) and *Blastopirellula marina* (GI: 87310177), as well as subunits A and B of galactose-6-phosphate isomerase from *Lactococcus lactis* (GI: 125907 and GI: 116326614, respectively). Amino acids lining the active site are indicated by 'a'. Bold **a** indicates the catalytic Cys69 from *Tc*RpiB, whereas **a*** indicates the catalytic Glu75 from *Mt*RpiB. Amino acids in the dimer and tetramer interfaces are indicated by 'd' and 't', respectively. The region that contains the loop with a one amino-acid insertion in *Tc*RpiB and *Mt*RpiB with respect to *Ec*RpiB is shown in a box. The alignment was shaded according to the percentage of identical residues.

*Mt*RpiB-like catalytic sites (i.e. a glutamate base). Most of these have the longer 137-loop (gxxRH, where glycine is the most common amino acid in the first position but is not present in all cases); only eight have a shorter loop (GGRH). We conclude that *Ec*RpiB-like sequences are biased towards enzymes expected to process both R5P/Ru5P and 6-carbon sugars, whereas, for *Mt*RpiB-like sequences, the 5-carbon sugars are more likely to be the substrates of interest. It is noteworthy that the LacA subunit from the

galactose-6-phosphate isomerase (i.e. that which contributes its C-terminal part to one active site, with the cysteine base provided by the N-terminal part of LacB) has the GGRH pattern, consistent with its role in processing 6-carbon sugars. The function of the second active site of the LacAB heterodimer is not yet clear; it is lined with residues well conserved within the LacAB family but different from those of RpiB.

Little information exists concerning the ability of various organisms to metabolize D-allose or D-allulose (also known as D-psicose). *E. coli*, which can grow on D-allose as a carbon source, is an enteric organism, and so the ability to use this rare sugar may offer a biological advantage, given that it is not metabolized by the host, and indeed acts as an immunosuppressant [27]. Of the eight organisms with glutamate-base enzymes including a shorter 137-loop, we could only find data for the marine bacterium *Reinekea*, which can metabolize D-allulose [28]. This organism is unusual in that it has an RpiA in combination with the *Mt*-type RpiB enzyme. However, D-allose and D-allulose are sufficiently rare that other 6-carbon sugars should be considered as the 'real' substrates of this class of enzyme. Two recent studies showed that *Ct*RpiB, which has the shorter 137-loop, is able to isomerize a number of 5- and 6-carbon sugars that lack phosphate groups. Best among the sugars tested (and better than D-ribose and D-allose) was L-talose, with a physiologically uninteresting K_m of 37 mM but a k_{cat} of 13 500 s^{-1} at 65 °C [24,25]; higher affinity might well be observed with the phosphorylated sugar. Therefore, tests of L-talose 6-phosphate isomerase activity among the RpiBs that possess the GGRH sequence appear to be worthwhile.

In the present study, we show how small differences in the RpiB sequences are linked to substantial effects on substrate specificity as well as inhibition patterns of the respective enzymes. Such information will be of great importance when designing new and highly specific RpiB inhibitors, especially those that will not interact with the completely different active site of human RpiA.

Experimental procedures

Chemicals and reagents

R5P was purchased from Sigma (St Louis, MO, USA), and sodium 2-mercapto-ethanesulfonate (MESNA) was obtained from Fluka Chemie GmbH (Buchs, Switzerland). Oligonucleotide primers were provided by Gibco, Life Technologies (Gaithersburg, MD, USA). Restriction endonucleases were obtained from New England Biolabs (Beverly, MA, USA). *E. coli* strain BL21 codon Plus (DE3) and the QuickChange™ site-directed mutagenesis kit were purchased from Stratagene (La Jolla, CA, USA). The compounds All6P, 5-phospho-D-ribonohydroxamic acid (monohydroxylammonium salt), *N*-(5-phospho-D-ribonoyl)-glycine (trisodium salt), *N*-(5-phospho-D-ribonoyl)-methylamine (disodium salt), 5-phospho-D-ribonamide (disodium salt), 5-phospho-D-ribonate (trisodium salt) and 4PEH were kind gifts from Laurent Salmon (University of Paris-Sud, Orsay, France), and were prepared as described previously [15,21].

Site-directed mutagenesis

The *Tc*RpiB mutant Δ^{135}E136G was constructed by PCR using methods described previously [9], with the mutagenic primers: 5′-CCG TTT AGC GGC GGG CGC CAT GTA CGA CG-3′ and 5′-CGT CGT ACA TGG CGC CCG CCG CTA AAC GG-3′.

Protein purification

Proteins were over-expressed and purified as described previously [9], with an added step consisting of size-exclusion chromatography (Hiload™ 16/60 Superdex™ 75; Pharmacia Biotech, Uppsala, Sweden) on a column equilibrated with 20 mM Tris/HCl (pH 7.6), 150 mM NaCl, 1 mM EDTA (adjusted to pH 8) and 10 mM β-ME. Most of the other buffers, starting with the cell disruption step, also contained β-ME but at a concentration of 1 mM. The exceptions included the first trials with wild-type protein, and later work with the C69A mutant.

Kinetic assays

For kinetic assays, the buffer was changed to 50 mM Tris/HCl (pH 7.6), 150 mM NaCl and 5 mM MESNA. R5P isomerase activity was tested by the spectrophotometric assay, as previously described by Wood [14], in which the production of Ru5P is monitored directly as the change in A_{290} (ε = 72 $M^{-1}{\cdot}cm^{-1}$). All6P isomerase activity was tested by a modified version (scaled down by 10-fold) of the thiobarbituric acid (TBA) assay described previously [13]. This assay mixture contained 50 mM Tris/HCl (pH 7.6), 5 mM MESNA and All6P, in a final volume of 100 μL. After preincubation at 30 °C for 5 min, the reaction was initiated by the addition of 10 μL of enzyme solution (0.83 $mg{\cdot}mL^{-1}$). After 3 min, the reaction was stopped by the addition of 100 μL of concentrated HCl; after mixing, 100 μL of 20 mM TBA in concentrated HCl (prepared fresh daily) was added. Initial rate conditions were confirmed using a control with half the enzyme; a blank of identical composition (but without enzyme) was used as the reference. After dehydration, dephosphorylation, and reaction with TBA, a characteristic yellow TBA adduct of Allu6P is formed

that strongly absorbs at 438 nm (= 27 800 $M^{-1} \cdot cm^{-1}$). Under the conditions of the assay, it would provide data if the k_{cat} is 0.015 s^{-1} or greater.

Crystallization

Before crystallization, protein solutions [in 20 mM Tris/HCl (pH 7.6), 150 mM NaCl, 1 mM EDTA (pH 8), 10 mM β-ME] were concentrated to 10 $mg \cdot mL^{-1}$ and 16 $mg \cdot mL^{-1}$ for *Tc*RpiB-wt and *Tc*RpiB-C69A, respectively. Screening of crystallization conditions was carried out by the sitting-drop vapour-diffusion method, using a Douglas Instruments Oryx robot together with a XYZV Plate Loader (Douglas Instruments, Hungerford, UK). For co-crystallization with 4PEH and R5P, the protein was mixed with ligand (dissolved in water) to give a final ligand concentration of 20 mM. For the *Tc*RpiB-C69A/All6P structure, the concentration of ligand was 10 mM, and no reducing agent was added. Initial crystallization droplets of 0.9 μL were mixed with an equal volume of reservoir solution, and equilibrated against 80 μL of reservoir solution. Crystals grew within 1–2 days at 20 °C under the conditions: *Tc*RpiB-wt, 0.5% (v/v) Jeffamine ED-2001 (Hampton Research, Aliso Viejo, CA, USA), 0.1 M Hepes, 1.1 M Na-malonate (pH 7.0) (condition F10 from the Easy Xtal JCSG+ crystallisation screen; Qiagen, Valencia, CA, USA); *Tc*RpiB-C69A/Pi, 0.8 M Na/K phosphate (pH 8.2) (condition A6 from the Hampton Quick Screen; Hampton Research); *Tc*RpiB-wt/R5P, 20% (w/v) poly(ethylene glycol) 550 monomethyl ether, 0.1 M NaCl, 0.1 M bicine (pH 9) (condition H4 from the Hampton Crystal Screen II; Hampton Research); *Tc*RpiB-wt/4PEH, 20% (w/v) poly(ethylene glycol) 6000, 0.2 M ammonium chloride, 0.1 M Tris/HCl (pH 8) (condition D8 from the Nextal PACT; Qiagen); *Tc*RpiB-C69A/All6P, 20% (w/v) poly(ethylene glycol) 3350, 0.2 M sodium acetate, 0.1 M BIS-TRIS propane (pH 8.5) (condition H7 from the Nextal PACT; Qiagen). Before flash cooling in liquid nitrogen, crystals were transferred to a ligand-containing cryoprotectant solution composed of 9% sucrose, 8% ethylene glycol, 8% glycerol and 2% glucose.

Structure solution and refinement

X-ray data (Table 1) were collected at the European Synchrotron Radiation Facility (ESRF, Grenoble, France), indexed using MOSFLM [29] and processed with SCALA [30], as implemented in the CCP4 interface [31].

The structure of *Tc*RpiB-C69A/Pi was solved by molecular replacement with PHASER [32] using the functional dimer of the ligand-free *Ec*RpiB structure (PDB entry code: 1NN4 with water molecules and histidine tag removed) as the search model. One dimer was found in the asymmetric unit. The correct sequence for *Tc*RpiB was then introduced using a homology model created using the software SOD [33], which was placed at the position of the molecular replacement solution. From this starting point, the software ARP/WARP [34] was used to auto-build 99% of the structure, which was next refined with alternating cycles of REFMAC5 [35] and rebuilding with the software O [36]. Progressive inclusion of water molecules used the ARP-waters command of REFMAC5. In all other cases except *Tc*RpiB-wt/R5P, the space group dimensions were sufficiently similar for the two protein molecules from the first structure to be placed by simple rigid-body refinement in REFMAC5. The space group was different for the *Tc*RpiB-wt/R5P structure and, for this reason, molecular replacement (using PHASER) was carried out with *Tc*RpiB-wt as a search model. Two dimers were located in the asymmetric unit. All models were subjected to restrained refinement, and the resulting maps inspected for bound ligand density. Water molecules were included in later stages of the refinement, after which R5P, 4PEH or All6P were placed manually in O. Statistics for the final models are presented in Table 1.

Sequence comparison, structural analysis and figure preparation

Similar sequences/structures were identified using BLAST [37], PFAM [8] and DALI [38], obtained from GenBank [39], and aligned using CLUSTALW [40]; alignments were edited with JALVIEW [41] with the structural information at hand. Structures were compared using LSQMAN [42] and O. Contact surfaces were explored using PROTORP [43]. Figures were prepared with Microsoft Excel (Microsoft Corp, Redmond, WA, USA), PYMOL (http://www.pymol.org/) and Adobe Illustrator (Adobe Systems, San Jose, CA, USA).

Acknowledgements

The inhibitors tested were kind gifts from Laurent Salmon, Emmanuel Burgos and Sandrine Mariano (University of Paris-Sud, Orsay, France). The authors would also like to thank the ESRF staff members for their support during data collection. This work was supported from a grant from the Swedish Research Council (contract 2007-6136) to S.L.M. A.L.S. was supported by fellowships from the Repsol-YPF Foundation, the Swedish Institute, and the European Molecular Biology Organisation (EMBO). J.J.C. is a member of the Research Career of the Argentinean National Research Council (CONICET).

References

1 Maugeri DA & Cazzulo JJ (2004) The pentose phosphate pathway in *Trypanosoma cruzi*. *FEMS Microbiol Lett* **234**, 117–123.

2 Maugeri DA, Cazzulo JJ, Burchmore RJ, Barrett MP & Ogbunude PO (2003) Pentose phosphate metabolism in *Leishmania mexicana*. *Mol Biochem Parasitol* **130**, 117–125.
3 David J & Wiesmeyer H (1970) Regulation of ribose metabolism in *Escherichia coli*. II. Evidence for two ribose-5-phosphate isomerase activities. *Biochimica et Biophysica Acta* **208**, 56–67.
4 Skinner AJ & Cooper RA (1971) The regulation of ribose-5-phosphate isomerisation in *Escherichia coli* K12. *FEBS Lett* **12**, 293–296.
5 Skinner AJ & Cooper RA (1974) Genetic studies on ribose-5-phosphate isomerase mutants of *Escherichia coli* K-12. *J Bacteriol* **118**, 1183–1185.
6 Essenberg MK & Cooper RA (1975) Two ribose-5-phosphate isomerases from *Escherichia coli* K12: partial characterisation of the enzymes and consideration of their possible physiological roles. *Eur J Biochem* **55**, 323–332.
7 Sorensen KI & Hove-Jensen B (1996) Ribose catabolism of *Escherichia coli*: characterization of the *rpiB* gene encoding ribose phosphate isomerase B and of the *rpiR* gene, which is involved in regulation of *rpiB* expression. *J Bacteriol* **178**, 1003–1011.
8 Finn RD, Tate J, Mistry J, Coggill PC, Sammut SJ, Hotz HR, Ceric G, Forslund K, Eddy SR, Sonnhammer EL *et al*. (2008) The Pfam protein families database. *Nucleic Acids Res* **36**, D281–D288.
9 Stern AL, Burgos E, Salmon L & Cazzulo JJ (2007) Ribose 5-phosphate isomerase type B from *Trypanosoma cruzi*: kinetic properties and site-directed mutagenesis reveal information about the reaction mechanism. *Biochem J* **401**, 279–285.
10 Zhang R-g, Andersson CE, Skarina T, Evdokimova E, Edwards AM, Joachimiak A, Savchenko A & Mowbray SL (2003) The 2.2 Å resolution structure of RpiB/AlsB from *Escherichia coli* illustrates a new approach to the ribose-5-phosphate isomerase reaction. *J Mol Biol* **332**, 1083–1094.
11 Poulsen TS, Chang Y-Y & Hove-Jensen B (1999) D-Allose catabolism of *Escherichia coli*: involvement of *alsI* and regulation of *als* regulon expression by allose and ribose. *J Bacteriol* **181**, 7126–7130.
12 Kim C, Song S & Park C (1997) The D-allose operon of *Escherichia coli* K-12. *J Bacteriol* **179**, 7631–7637.
13 Roos AK, Mariano S, Kowalinski E, Salmon L & Mowbray SL (2008) D-ribose-5-phosphate isomerase B from *Escherichia coli* is also a functional D-allose-6-phosphate isomerase, while the *Mycobacterium tuberculosis* enzyme is not. *J Mol Biol* **382**, 667–679.
14 Wood T (1970) Spectrophotometric assay for D-ribose-5-phosphateketol-isomerase and for D-ribulose-5-phosphate 3-epimerase. *Anal Biochem* **33**, 297–306.
15 Mariano S, Roos AK, Mowbray SL & Salmon L (2009) Competitive inhibitors of type B ribose 5-phosphate isomerases: design, synthesis and kinetic evaluation of new D-allose and D-allulose 6-phosphate derivatives. *Carbohydr Res* **344**, 869–880.
16 Kleywegt GJ & Jones TA (1996) Phi/Psi-cology: Ramachandran revisited. *Structure* **4**, 1395–1400.
17 Engh RA & Huber R (1991) Accurate bond and angle parameters for X-ray protein structure refinement. *Acta Crystallogr* **47**, 392–400.
18 Read RJ (1986) Improved Fourier coefficients for maps using phases from partial structures with errors. *Acta Crystallogr* **42**, 140–149.
19 Xu Q, Schwarzenbacher R, McMullan D, von Delft F, Brinen LS, Canaves JM, Dai X, Deacon AM, Elsliger MA, Eshagi S *et al*. (2004) Crystal structure of a ribose-5-phosphate isomerase RpiB (TM1080) from *Thermotoga maritima* at 1.90 Å resolution. *Proteins* **56**, 171–175.
20 Roos AK, Andersson CE, Bergfors T, Jacobsson M, Karlén A, Unge T, Jones TA & Mowbray SL (2004) *Mycobacterium tuberculosis* ribose-5-phosphate isomerase has a known fold, but a novel active site. *J Mol Biol* **335**, 799–809.
21 Roos AK, Burgos E, Ericsson DJ, Salmon L & Mowbray SL (2005) Competitive inhibitors of *Mycobacterium tuberculosis* ribose-5-phosphate isomerase B reveal new information about the reaction mechanism. *J Biol Chem* **280**, 6416–6422.
22 Horecker BL, Smyrniotis PZ & Seegmiller JE (1951) The enzymatic conversion of 6-phosphogluconate to ribulose-5-phosphate and ribose-5-phosphate. *J Biol Chem* **193**, 383–396.
23 Amyes TL & Richard JP (2007) Enzymatic catalysis of proton transfer at carbon: activation of triosephosphate isomerase by phosphite dianion. *Biochemistry* **46**, 5841–5854.
24 Park CS, Yeom SJ, Kim HJ, Lee SH, Lee JK, Kim SW & Oh DK (2007) Characterization of ribose-5-phosphate isomerase of *Clostridium thermocellum* producing D-allose from D-psicose. *Biotechnol Lett* **29**, 1387–1391.
25 Yoon RY, Yeom SJ, Kim HJ & Oh DK (2009) Novel substrates of a ribose-5-phosphate isomerase from *Clostridium thermocellum*. *J Biotechnol* **139**, 26–32.
26 Chan KK, Fedorov AA, Fedorov EV, Almo SC & Gerlt JA (2008) Structural basis for substrate specificity in phosphate binding (beta/alpha)8-barrels: D-allulose 6-phosphate 3-epimerase from *Escherichia coli* K-12. *Biochemistry* **47**, 9608–9617.
27 Hossain MA, Wakabayashi H, Goda F, Kobayashi S, Maeba T & Maeta H (2000) Effect of the immunosuppressants FK506 and D-allose on allogenic orthotopic liver transplantation in rats. *Transplant Proc* **32**, 2021–2023.
28 Romanenko LA, Schumann P, Rohde M, Mikhailov VV & Stackebrandt E (2004) *Reinekea marinisedimentorum* gen. nov., sp. nov., a novel gammaproteobacterium

from marine coastal sediments. *Int J Syst Evol Microbiol* **54**, 669–673.
29 Leslie AG (2006) The integration of macromolecular diffraction data. *Acta Crystallogr D Biol Crystallogr* **62**, 48–57.
30 Evans P (2006) Scaling and assessment of data quality. *Acta Crystallogr D Biol Crystallogr* **62**, 72–82.
31 Potterton E, Briggs P, Turkenburg M & Dodson E (2003) A graphical user interface to the CCP4 program suite. *Acta Crystallogr* **59**, 1131–1137.
32 McCoy AJ, Grosse-Kunstleve RW, Storoni LC & Read RJ (2005) Likelihood-enhanced fast translation functions. *Acta Crystallogr D Biol Crystallogr* **61**, 458–464.
33 Kleywegt GJ, Zou JY, Kjeldgaard M & Jones TA (2001) Around O. In *International Tables for Crystallography: Crystallography of Biological Macromolecules* (Rossmann MG & Arnold E eds), pp 353–356. Kluwer Academic Publishers, Dordrecht, the Netherlands.
34 Perrakis A, Sixma TK, Wilson KS & Lamzin VS (1997) wARP: improvement and extension of crystallographic phases by weighted averaging of multiple-refined dummy atomic models. *Acta Crystallogr* **53**, 448–455.
35 Murshudov GN, Vagin AA & Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr* **53**, 240–255.
36 Jones TA, Zou J-Y, Cowan SW & Kjeldgaard M (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr* **47**, 110–119.
37 Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W & Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**, 3389–3402.
38 Holm L & Sander C (1993) Protein structure comparison by alignment of distance matrices. *J Mol Biol* **233**, 123–138.
39 Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J & Wheeler DL (2008) GenBank. *Nucleic Acids Res* **36**, D25–D30.
40 Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R *et al.* (2007) Clustal W and Clustal X version 2.0. *Bioinformatics (Oxford, England)* **23**, 2947–2948.
41 Waterhouse AM, Procter JB, Martin DM, Clamp M & Barton GJ (2009) Jalview Version 2 – a multiple sequence alignment editor and analysis workbench. *Bioinformatics (Oxford, England)* **25**, 1189–1191.
42 Kleywegt GJ (1996) Use of non-crystallographic symmetry in protein structure refinement. *Acta Crystallogr* **52**, 842–857.
43 Reynolds C, Damerell D & Jones S (2009) ProtorP: a protein–protein interaction analysis server. *Bioinformatics (Oxford, England)* **25**, 413–414.