Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structural insight into L-ribulose 3-epimerase from *Mesorhizobium loti*

Keiko Uechi,[a] Haruhiko Sakuraba,[b] Akihide Yoshihara,[a] Kenji Morimoto[a] and Goro Takata[a]*

[a]Rare Sugar Research Center, Kagawa University, 2393 Ikenobe, Miki-cho, Kita-gun, Kagawa 761-0795, Japan, and [b]Department of Applied Biological Science, Faculty of Agriculture, Kagawa University, 2393 Ikenobe, Miki-cho, Kita-gun, Kagawa 761-0795, Japan

Correspondence e-mail: goro@ag.kagawa-u.ac.jp

Received 18 April 2013
Accepted 2 August 2013

PDB Reference: L-ribulose 3-epimerase, 3vyl

L-Ribulose 3-epimerase (L-RE) from *Mesorhizobium loti* has been identified as the first ketose 3-epimerase that shows the highest observed activity towards ketopentoses. In the present study, the crystal structure of the enzyme was determined to 2.7 Å resolution. The asymmetric unit contained two homotetramers with the monomer folded into an $(\alpha/\beta)_8$-barrel carrying four additional short α-helices. The overall structure of *M. loti* L-RE showed significant similarity to the structures of ketose 3-epimerases from *Pseudomonas cichorii*, *Agrobacterium tumefaciens* and *Clostridium cellulolyticum*, which use ketohexoses as preferred substrates. However, the size of the C-terminal helix (α8) was much larger in *M. loti* L-RE than the corresponding helices in the other enzymes. In *M. loti* L-RE the α8 helix and the following C-terminal tail possessed a unique subunit–subunit interface which promoted the formation of additional intermolecular interactions and strengthened the enzyme stability. Structural comparisons revealed that the relatively small hydrophobic pocket of the enzyme around the substrate was likely to be the main factor responsible for the marked specificity for ketopentoses shown by *M. loti* L-RE.

1. Introduction

Ketose 3-epimerase catalyzes the epimerization of hydroxyl groups at the C-3 position of various keto sugars, including rare sugars that exist naturally only in small amounts. Recent studies have revealed that these rare sugars exhibit unique properties and might even exert beneficial health effects in humans (Levin, 2002; Matsuo *et al.*, 2003, 2004). Therefore, it is hoped that ketose 3-epimerase family enzymes can serve as useful catalysts for the commercially viable production of rare sugars (Granström *et al.*, 2004; Izumori, 2006). D-Tagatose 3-epimerase (D-TE) from *Pseudomonas cichorii* efficiently converts not only D-tagatose to D-sorbose but also D-fructose to D-psicose (Itoh *et al.*, 1994). Thus, this enzyme is expected to be a key enzyme for the mass production of D-psicose from D-fructose (Takeshita *et al.*, 2000). Some ketose 3-epimerases from other sources reportedly show substrate specificities that differ from that of *P. cichorii* D-TE. For example, enzymes from *Agrobacterium tumefaciens* and *Clostridium cellulolyticum* exhibit high specificity for D-psicose and were thus named D-psicose 3-epimerase (D-PE; Kim, Hyun *et al.*, 2006; Mu *et al.*, 2011), and a *Rhodobacter sphaeroides* enzyme shows the highest observed activity towards D-fructose (Zhang *et al.*, 2009). These enzymes use ketohexoses as preferred substrates, although they also show some activity towards ketopentoses.

To date, crystal structures of D-TE from *P. cichorii* (Yoshida *et al.*, 2007) and D-PEs from *A. tumefaciens* (Kim, Kim *et al.*, 2006) and *C. cellulolyticum* (Chan *et al.*, 2012) have been solved. Extensive analysis of these structures has resolved the substrate-binding sites and led to elucidation of the substrate-recognition and catalytic mechanisms of these enzymes.

Recently, a novel ketose 3-epimerase (L-RE) that shows the highest observed activity towards L-ribulose has been identified in *Mesorhizobium loti* Tono (Fig. 1; Uechi *et al.*, 2013). The amino-acid sequence of *M. loti* L-RE showed 31, 29 and 28% identity to those of *P. cichorii* D-TE, *C. cellulolyticum* D-PE and *A. tumefaciens* D-PE, respectively (Fig. 2). Characterization revealed that the enzyme possesses a maximum activity of 416 U mg^{-1} for L-ribulose at 333 K, rendering this enzyme potentially useful for the bioproduction of L-xylulose (Uechi *et al.*, 2013). This is the first example of a ketose 3-epimerase that shows high specificity for a ketopentose. The detailed structural analysis of L-RE could provide insights into the substrate-recognition mechanism of this enzyme, information that is important for facilitating rare-sugar production. In the present study, therefore, the crystal structure of *M. loti* L-RE was determined and its active-site architecture was compared with those of ketose 3-epimerases from other sources. An evaluation was performed of the amino-acid residues that might be responsible for the substrate-specificity differences between *M. loti* L-RE and the other enzymes. In addition, the structural features that contribute to *M. loti* L-RE stability were analyzed by means of site-directed mutagenesis.

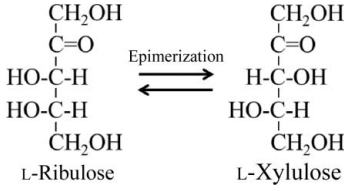

Figure 1
Epimerization reaction of L-ribulose catalyzed by *M. loti* L-RE.

2. Materials and methods

2.1. DNA manipulation

The expression vector (pQE-lre) for *M. loti* L-RE without a His tag was constructed as described previously (Uechi *et al.*, 2013). An expression vector for N-terminally His-tagged *M. loti* L-RE was constructed by first amplifying the *M. loti* L-RE gene fragment by PCR. The oligonucleotide primers used in the amplification were 5′-AGA**CATATG**GCGCGC-ATCGGAATCCATTCT-3′, which contained a unique *Nde*I restriction site (bold) overlapping the 5′ initiation codon, and 5′-AGA**GGATCC**CTACGGCTTGTGACGGGCCGAGAC-3′, which contained a unique *Bam*HI restriction site (bold) proximal to the 3′ end of the termination codon. The expression vector pQE-lre served as the template. The amplified fragment was digested with *Nde*I and *Bam*HI and ligated with the expression vector pColdII (Takara Bio Inc., Shiga, Japan) previously linearized using *Nde*I and *Bam*HI, yielding pC-lre.

Site-directed mutagenesis was accomplished using a PrimeSTAR Mutagenesis Kit (Takara Bio Inc.) according to the manufacturer's instructions. To constract a deletion mutant lacking the C-terminal tail (Ser292–Pro297), the expression vector (pC-lre) for His-tagged *M. loti* L-RE served as the template and the following pair of oligonucleotide primers were used as the mutagenic primers: 5′-ACTCGTC-TAGGGATCCGAATTCAAGCTT-3′ and 5′-GATCCCTAG-ACGACTTCCGCCTTGCGCG-3′. To construct the Y39A mutant, the expression vector (pQE-lre) for nontagged *M. loti* L-RE served as the template and the following pair of oligo-nucleotide primers were used as the mutagenic primers (the

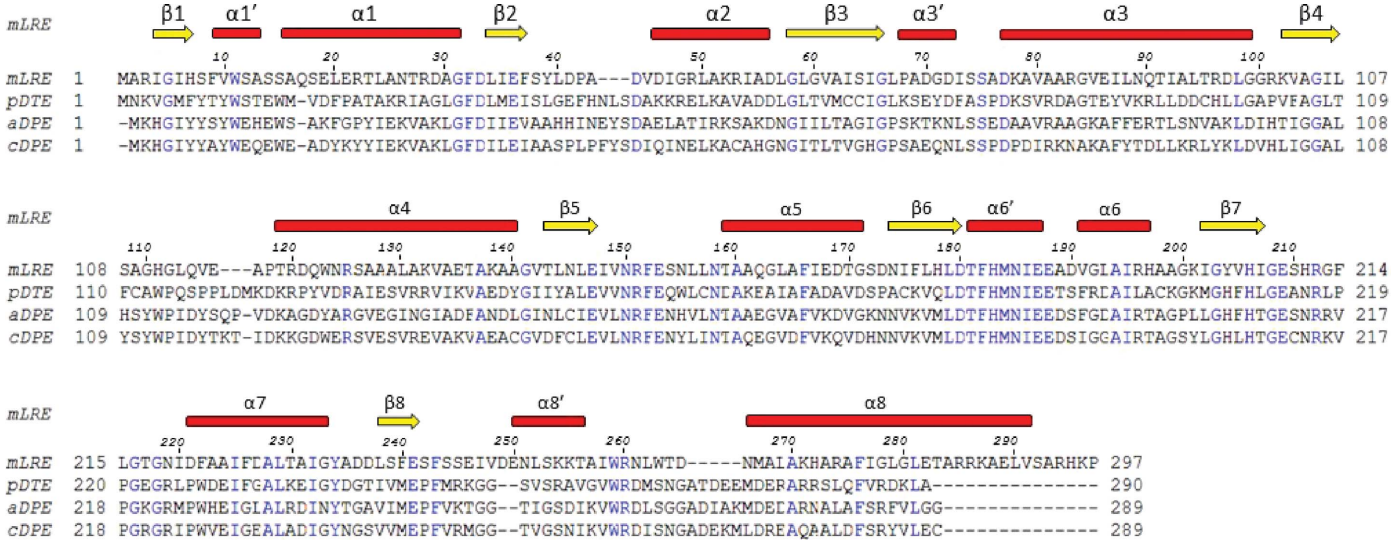

Figure 2
Structure-based amino-acid sequence alignment of *M. loti* L-RE (*mLRE*), *P. cichorii* D-TE (*pDTE*), *A. tumefaciens* D-PE (*aDPE*) and *C. cellulolyticum* D-PE (*cDPE*). The secondary-structural assignments for *mLRE* are shown above the alignment. Conserved residues are shown in blue.

mutations are shown in bold): 5′-TTCTCC**GCG**CTCGACC-CGGCCGATGTC-3′ and 5′-GTCGAG**CGC**GGAGAATTC-GATCAGGTC-3′.

2.2. Protein expression and purification

Expression and purification of nontagged *M. loti* L-RE were carried out as described previously (Uechi *et al.*, 2013), except that gel-filtration column chromatography was employed as the final purification step. *Escherichia coli* Origami B cells (Novagen Inc., Madison, Wisconsin, USA) harbouring the recombinant plasmid pQE-lre were cultivated at 310 K in Luria–Bertani medium containing 100 µg ml^{-1} ampicillin until the OD_{600} reached 0.6–0.7. Enzyme expression was induced by adding 1 m*M* isopropyl *β*-D-1-thiogalactopyranoside (IPTG) to the medium and cultivation was continued at 293 K for an additional 12 h. The cells were then harvested by centrifugation and disrupted by ultrasonication; cell debris was removed by centrifugation and the resulting supernatant was loaded onto a Q-Sepharose HP column (GE Healthcare Bio-Sciences AB, Uppsala, Sweden) equilibrated with 20 m*M* Tris–HCl buffer pH 8.0 (Tris-8 buffer). After washing the column with the same buffer, the enzyme was eluted using a linear gradient of 0.0–1.0 *M* NaCl in the same buffer. The active fractions were then pooled and ammonium sulfate was added to the solution to 2.0 *M*. The enzyme solution was loaded onto a Resource PHE column (GE Healthcare Bio-Sciences AB) previously equilibrated with Tris-8 buffer supplemented with 2.0 *M* ammonium sulfate. After the column had been washed with the augmented buffer, the enzyme was eluted with a linear 2.0–0.0 *M* ammonium sulfate gradient in Tris-8 buffer. The active fractions were pooled and dialyzed against Tris-8 buffer. The dialysate was applied onto a Resource Q column (GE Healthcare Bio-Sciences AB) equilibrated with the same buffer, the column was washed with this buffer and the enzyme was eluted with a linear gradient of 0–250 m*M* NaCl in the buffer. The active fractions were pooled, concentrated and loaded onto a Superdex 200 pg gel-filtration column (GE Healthcare Bio-Sciences AB) equilibrated with 10 m*M* 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid buffer (HEPES) pH 7.5. Expression and purification of the Y39A mutant were performed using the same method as described for nontagged *M. loti* L-RE.

To prepare His-tagged *M. loti* L-RE, *E. coli* Origami B cells (Novagen Inc.) were transformed with pC-lre, after which the transformants were cultivated at 310 K in Luria–Bertani medium containing 100 µg ml^{-1} ampicillin until the OD_{600} reached 0.7. Expression was then induced by adding 1 m*M* IPTG to the medium and cultivation was continued at 288 K for an additional 24 h. The cells were harvested by centrifugation, suspended in 20 m*M* sodium phosphate buffer pH 7.4 containing 300 m*M* NaCl and 10 m*M* imidazole and then disrupted by ultrasonication. The cell debris was removed by centrifugation and the resulting supernatant was loaded onto a HisTrap HP column (GE Healthcare Bio-Sciences AB) equilibrated with the same buffer. After washing the column with this buffer, the enzyme was eluted with a linear gradient to 500 m*M* imidazole in the buffer. The active fractions were collected and dialyzed against 20 m*M* Tris–HCl buffer pH 7.5 (Tris-7.5). The dialysate was then applied onto a Resource Q column (GE Healthcare Bio-Sciences AB) equilibrated with the same buffer and, after washing the column with further buffer, the enzyme was eluted with a linear gradient of 0–250 m*M* NaCl in the same buffer. The active fractions were pooled and the enzyme solution was dialyzed against Tris-7.5. Expression and purification of the C-terminal tail deletion mutant were performed using the same method as described for His-tagged *M. loti* L-RE.

The protein concentration was determined using the Bradford method (Bradford, 1976) and the enzyme purity was checked by SDS–PAGE (Laemmli, 1970).

Table 1
Data-collection and refinement statistics for *M. loti* L-RE.

Values in parentheses are for the highest resolution shell. R.m.s.d., root-mean-square deviation.

Data collection	
Wavelength (Å)	1.0
Space group	$C222_1$
Unit-cell parameters (Å)	
a	202.5
b	206.2
c	122.0
Resolution range (Å)	50–2.7 (2.75–2.70)
Total No. of reflections	1045640
No. of unique reflections	70374
Multiplicity	14.9 (14.6)
Completeness (%)	99.9 (100)
R_{merge}†	0.060 (0.290)
$\langle I/\sigma(I)\rangle$	16.8 (13.5)
Refinement	
Resolution range (Å)	50.0–2.7
R/R_{free}‡ (%)	20.9/26.4 (28.3/39.2)
No. of protein atoms	17936
No. of water molecules	326
No. of ligands	8 Mn^{2+}, 4 MPD
B factors (Å^2)	
Protein	46.8
Manganese ion	33.1
Water	35.1
MPD	61.6
R.m.s.d.	
Bond lengths (Å)	0.007
Bond angles (°)	1.2
Ramachandran statistics (%)	
Favoured	97.3
Allowed	2.7
Outliers	0

† $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the scaled intensity of the *i*th observation of reflection *hkl*, $\langle I(hkl)\rangle$ is the mean value and the summation is over all measurements. ‡ R_{free} was calculated with randomly selected reflections (5%).

2.3. Crystallization and data collection

The purified *M. loti* L-RE without His tag was concentrated to 10.0 mg ml^{-1} by ultrafiltration (Amicon Ultra 30K NMWL) for crystallization trials. The final protein level was checked by absorbance at 280 nm with a NanoDrop ND-2000 (Thermo Fisher Scientific, Dreieich, Germany). Initial screening for crystallization was carried out with Crystal Screen and Crystal

Screen 2 (Hampton Research, Aliso Viejo, California, USA) at 293 K using the sitting-drop vapour-diffusion method, in which a 1 µl drop of protein solution was mixed with an equal volume of reservoir solution and equilibrated against 0.1 ml reservoir solution using CompactClover Crystallization Plates (Emerald BioSystems Inc., Bedford, Massachusetts, USA). Initially, small crystals were grown using reagent No. 35 of Crystal Screen 2: 70% (4*S*)-2-methyl-2,4-pentanediol (MPD) in 0.1 *M* HEPES buffer pH 7.5. This precipitant solution was taken as a starting point and was optimized by variation of the MPD concentration; diffraction-quality crystals (maximum dimensions of 0.4 × 0.2 × 0.2 mm) were obtained using a reservoir solution composed of 50% MPD in 0.1 *M* HEPES buffer pH 7.5. The crystals were grown at 293 K for 1 d and belonged to the orthorhombic space group $C222_1$, with unit-cell parameters a = 202.5, b = 206.2, c = 122.0 Å. Diffraction data were collected at 2.7 Å resolution using monochromated radiation of wavelength 1.0 Å and an ADSC CCD detector system on the BL5A beamline at the Photon Factory, Tsukuba, Japan. All measurements were carried out on crystals cryoprotected with 50% MPD and cooled to 100 K in a stream of nitrogen gas. The oscillation angle per image was set to 1° and the crystal-to-detector distance was 257.6 mm. The data were processed using the *HKL*-2000 package (HKL Research Inc., Charlottesville, Virginia, USA; Otwinowski & Minor, 1997).

2.4. Phasing, refinement and structural analysis

The initial phases for the *M. loti* L-RE structure were determined by molecular replacement using the *MOLREP* program (Vagin & Teplyakov, 2010) in the *CCP*4 program suite (Winn *et al.*, 2011); the structure of chain *A* from *P. cichorii* D-TE (PDB entry 2qul) served as the search model. Further model building was performed using the program *Coot* (Emsley *et al.*, 2010) and refinement to a resolution of 2.7 Å was carried out using *REFMAC*5 (Murshudov *et al.*, 2011) and *CNS* (Brünger *et al.*, 1998). MPD molecules were clearly visible in both the σ_A-weighted $2F_o - F_c$ and $F_o - F_c$ density maps and were included in the latter part of the refinement. In addition, a strong electron-density peak surrounded by four residues at the end of the β-barrels was interpreted to be an Mn^{2+} ion bound in the same position as in *P. cichorii* D-TE. Water molecules were incorporated using *Coot* (Emsley *et al.*, 2010) and model geometry was analyzed using *RAMPAGE* (Lovell *et al.*, 2003). The data-collection and refinement statistics are listed in Table 1.

Ion pairs, with a cutoff distance of 4.0 Å, and hydrophobic interactions were identified using the *WHAT IF* web server (Rodriguez *et al.*, 1998). The number of hydrophobic interactions was determined by calculation of the interatomic contacts between atoms from hydrophobic side chains; a contact was defined as two atoms with van der Waals surfaces separated by <1.0 Å. The number of hydrogen bonds was identified using the *CCP*4*mg* program (McNicholas *et al.*, 2011). Accessible surface area (ASA; radius of probe solvent molecule = 1.4 Å) was calculated using *AREAIMOL* in the *CCP*4 program suite (Winn *et al.*, 2011). Cavity volumes were

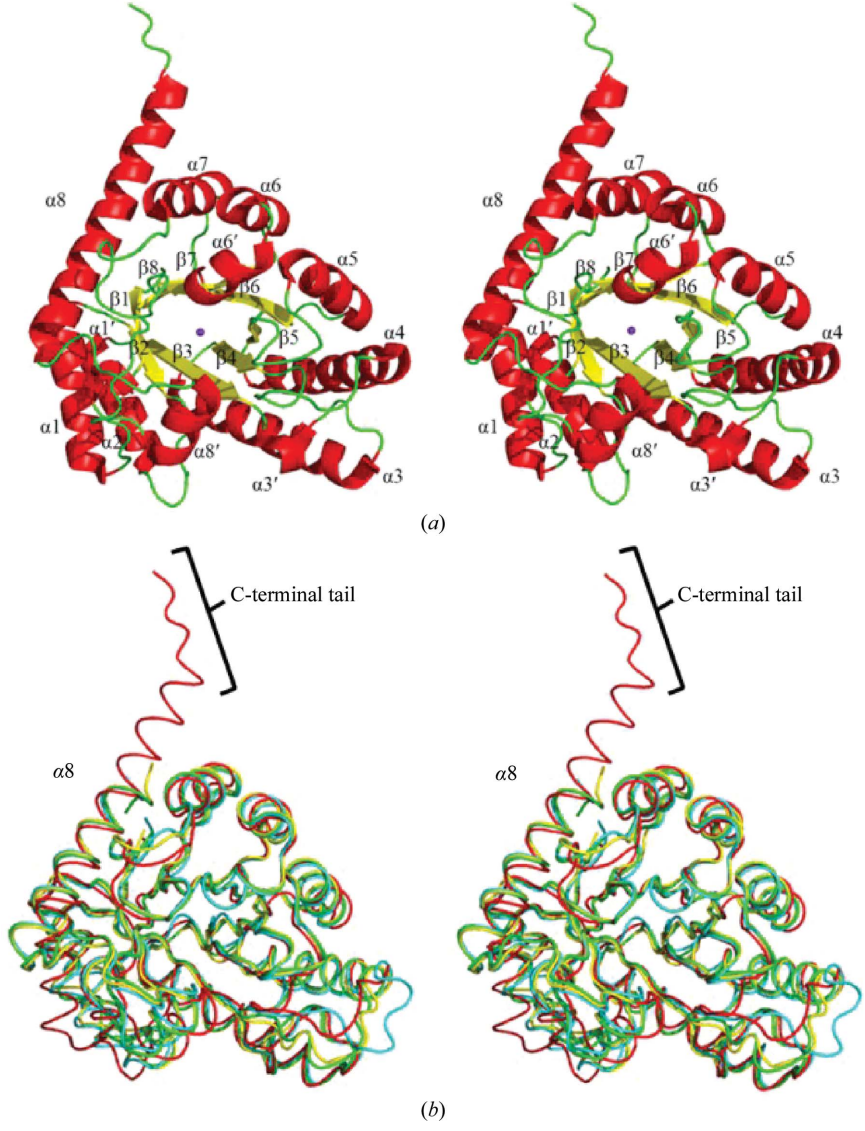

Figure 3
(*a*) Overall structure of the *M. loti* L-RE monomer (stereoview). α-Helices (numbered from α1 to α8) and β-strands (numbered from β1 to β8) are indicated in red and yellow, respectively; Mn^{2+}, purple. (*b*) Structural comparison of *M. loti* L-RE and other ketose 3-epimerases (stereoview). *M. loti* L-RE structure (red) superimposed on those of *P. cichorii* D-TE (cyan), *A. tumefaciens* D-PE (green) and *C. cellulolyticum* D-PE (yellow).

calculated using *Q-SiteFinder* (http://www.modelling.leeds.ac.uk/qsitefinder/; Laurie & Jackson, 2005). Molecular-graphics figures were created using *PyMOL* (http://www.pymol.org/).

2.5. L-RE activity measurement

L-RE activity was assayed by measuring the rate of L-xylulose production from L-ribulose using HPLC as described previously (Uechi *et al.*, 2013). The kinetic parameters of nontagged *M. loti* L-RE and nontagged Y39A mutant were determined as described previously (Uechi *et al.*, 2013). To estimate the thermostability, the effect of temperature on His-tagged *M. loti* L-RE and His-tagged ΔSer292–Pro297 mutant was examined. After incubation of the enzyme (0.3 mg ml^{-1}) over the temperature range 333–343 K in 50 m*M* Tris-7.5, the residual activity was determined at appropriate intervals using the standard assay method.

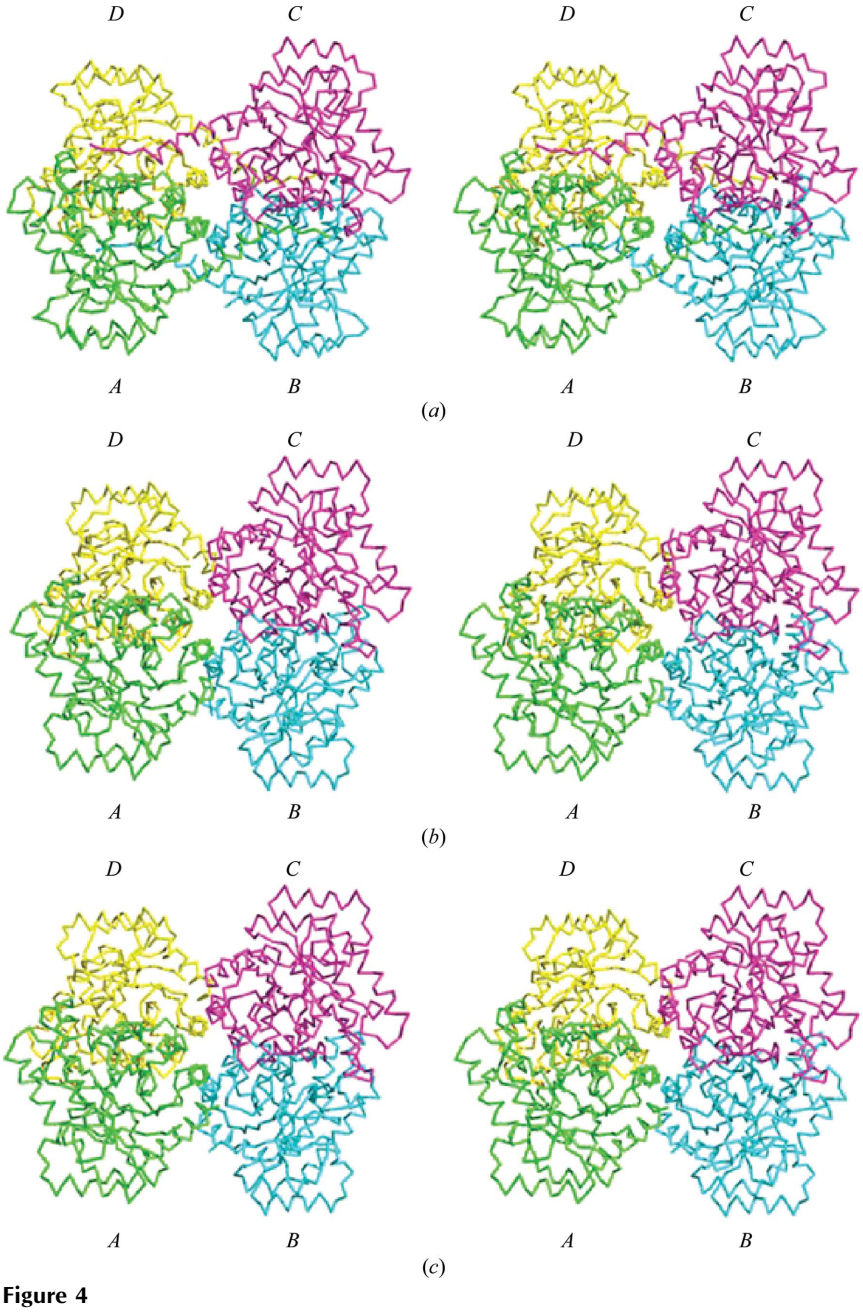

Figure 4
Tetramer structures of *M. loti* L-RE (*a*), *A. tumefaciens* D-PE (*b*) and *C. cellulolyticum* D-PE (*c*) (stereoview). Subunits *A*, *B*, *C* and *D* are shown in green, cyan, magenta and yellow, respectively.

3. Results and discussion

3.1. Overall structure and structural homologues

The structure of *M. loti* L-RE was determined using molecular replacement and refined to 2.7 Å resolution (Table 1). The asymmetric unit consisted of two homotetramers (*A*, *B*, *C*, *D* and *E*, *F*, *G*, *H*) with a solvent content of 50.6%, which corresponds to a Matthews coefficient (Matthews, 1968) of 2.5 Å^3 Da^{-1}. The final structure showed good geometry without Ramachandran outliers. The model of two homotetramers contained the complete ordered residues 1–297 of each subunit, four MPD molecules, eight Mn atoms and 326 water molecules.

The subunit structure of *M. loti* L-RE comprised a monomer folded into an $(\alpha/\beta)_8$-barrel carrying four additional helical segments, $\alpha1'$, $\alpha3'$, $\alpha6'$ and $\alpha8'$, which were inserted before $\alpha1$, $\alpha3$, $\alpha6$ and $\alpha8$, respectively (Fig. 3*a*). When this model of the *M. loti* L-RE monomer was sent to the *DALI* server (Holm & Rosenström, 2010) seeking identification of proteins with similar structures (as of 8 April 2013), the three proteins with the highest structural similarity were D-TE from *P. cichorii* (PDB entries 2qul, 2qum, 2qun and 2ou4, with r.m.s.d.s between 1.6 and 1.7 Å), D-PE from *C. cellulolyticum* (PDB entries 3vni, 3vnj, 3vnk and 3vnl, with r.m.s.d.s between 1.7 and 1.8 Å) and D-PE from *A. tumefaciens* (PDB entries 2hkl and 2hk0, with r.m.s.d.s between 1.8 and 1.9 Å), as expected. The main-chain coordinates of the *M. loti* L-RE monomer were similar to those of *P. cichorii* D-TE and the *C. cellulolyticum* and *A. tumefaciens* D-PEs, although a clear topological difference was found between *M. loti* L-RE and the other three enzymes. Here, the C-terminal $\alpha8$ helix in *M. loti* L-RE was extremely long, and as a result the C-terminal half (282–291) of $\alpha8$ and the following C-terminal tail (292–297) protruded away from the barrel (Fig. 3*b*). In addition, the short α-helix ($\alpha2'$) inserted between $\beta2$ and $\alpha2$ in the other three enzymes was not observed in *M. loti* L-RE.

It has been reported that *C. cellulolyticum* and *A. tumefaciens* D-PE exhibit tetrameric structures (Mu *et al.*, 2011; Kim, Kim *et al.*,

2006), whereas *P. cichorii* D-TE is dimeric (Itoh *et al.*, 1994). In our model, *M. loti* L-RE assembled into a tetramer (Fig. 4*a*) and the arrangement of the four subunits of the enzyme was essentially the same as those of *A. tumefaciens* and *C. cellulolyticum* D-PEs (Fig. 4). The tetramer (*A*, *B*, *C* and *D*) of *M. loti* L-RE had approximate dimensions of 90 × 79 × 79 Å. The *A*–*B*, *A*–*C* and *A*–*D* associations represented three distinct subunit–subunit interfaces (Fig. 5). In the largest interface, *A*–*D*, the amino-acid residues belonging to loop regions between β4 and α4, β5 and α5, α6′ and α6, β7 and α7, α8′ and α8, and two short α-helices α6′ and α8′ were all involved in intersubunit interactions (Fig. 5*a*). Six amino-acid residues [Leu113, Trp118, Phe152, Phe182 (α6′), Phe186 (α6′) and Trp259] formed a total of 47 hydrophobic interactions between subunits, and 12 amino-acid residues [Glu116, Arg151, Ser154, Thr159, Asn185 (α6′), Glu187 (α6′), Glu188, Val191, His211, Asn219, Trp259 and Arg260] formed a total of 17 direct hydrogen bonds between subunits. In addition, four ion-pair interactions formed by four amino-acid residues [Glu116, Glu187, Lys254 (α8′) and Arg260] were observed between subunits. The *A*–*B* interface was formed by the antiparallel alignment of α8 helices together with interactions with residues in the loop between α8′ and α8 (Fig. 5*b*). Between subunits *A* and *B*, four amino-acid residues [Leu262, Phe276 (α8), Leu279 (α8) and Leu290 (α8)] formed 12 hydrophobic interactions and two amino-acid residues [His272 (α8) and Thr283 (α8)] formed two direct hydrogen bonds. In the *A*–*C* interface (Fig. 5*c*), the C-terminal Pro297 formed 18 hydrophobic interactions with Pro118* and Trp123* (α4) and Val291 (α8) formed eight hydrophobic interactions with Leu193* (α6) (asterisks indicate the neighbouring subunit in the *A*–*C* dimer). His295, Lys296 and Pro297, which belonged to the C-terminal tail, formed five hydrogen bonds to amino-acid residues (Arg120*, Gln162* and Asp169*) in the neighbouring subunit. Moreover, Lys296 and Lys287 (α8) formed ion-pair interactions (five in total) with Asp169* and Asp190*, respectively. It is noteworthy that interactions between subunits *A* and *C* could only be observed around the C-terminal part of α8 and the following C-terminal tail (Fig. 5*c*).

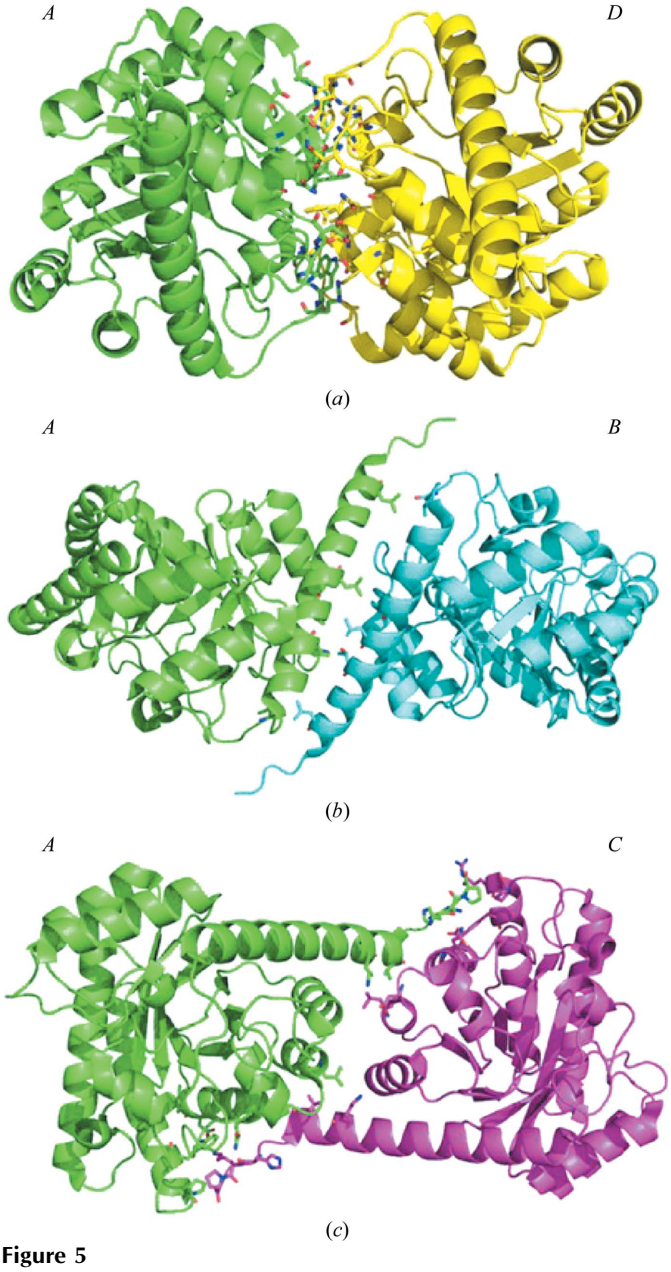

Figure 5
A–*D* (*a*), *A*–*B* (*b*) and *A*–*C* (*c*) interfaces in the *M. loti* L-RE tetramer. Residues involved in intermolecular interactions are shown as stick models and O and N atoms are shown in red and blue, respectively.

Table 2
Comparison of the structural features of *M. loti* L-RE with those of *A. tumefaciens* D-PE, *C. cellulolyticum* D-PE and *P. cichorii* D-TE.

Oligomer* represents the summation of all inter-subunit interactions (*A*–*B*, *A*–*C*, *A*–*D*, *B*–*C*, *B*–*D* and *C*–*D*) in the biological unit.

Enzyme (PDB code)	*M. loti* L-RE (3vyl)	*A. tumefaciens* D-PE (2hk1)	*C. cellulolyticum* D-PE (3vni)	*P. cichorii* D-TE (2qul)
R.m.s.d. (Å)	—	1.8	1.8	1.7
Assembly	Tetramer	Tetramer	Tetramer	Dimer
ASA (Å^2)				
Oligomer	40100	40700	40800	20300 (*A*–*D*)
Monomer	13100	12400	12700	12400
Interface area (Å^2)				
A–*B*	900	500	600	200
A–*C*	1200	300	200	0
A–*D*	1500	1400	1400	1500
Oligomer	7200	4400	4400	3200
No. of ion pairs				
A–*B*	0	0	1	0
A–*C*	5	0	0	0
A–*D*	4	8	8	8
Oligomer*	18	16	18	16
No. of hydrophobic interactions				
A–*B*	12	23	9	0
A–*C*	26	0	0	0
A–*D*	47	45	37	74
Oligomer*	170	137	91	149
No. of hydrogen bonds				
A–*B*	2	0	7	28
A–*C*	5	2	2	0
A–*D*	17	26	34	0
Oligomer*	49	56	85	66

Comparisons of the structural features of *M. loti* L-RE with those of *P. cichorii* D-TE, *A. tumefaciens* D-PE and *C. cellulolyticum* D-PE are shown in Table 2. The *M. loti* L-RE tetramer

showed a similar total ASA (40 100 Å^2) compared with the *A. tumefaciens* and *C. cellulolyticum* D-PE tetramers (40 700–40 800 Å^2). In contrast, the intermolecular ASA of *M. loti* L-RE was totally different from those of the other enzymes. The ASA (1200 Å^2) of the *A*–*C* interface of *M. loti* L-RE was much larger (by 4–6 times) than those of *A. tumefaciens* and *C. cellulolyticum* D-PEs. The *A*–*B* interface of *M. loti* L-RE also showed a larger ASA (900 Å^2) than the other enzymes, while the ASAs of the *A*–*D* interfaces were not very different among these enzymes (1400–1500 Å^2). As a result, the total intermolecular ASA of *M. loti* L-RE was calculated to be a huge 7200 Å^2, suggesting that *M. loti* L-RE might be more stable than *A. tumefaciens* and *C. cellulolyticum* D-PEs. Interestingly, no ion-pair and hydrophobic interactions were detected between subunits *A* and *C* of *A. tumefaciens* and *C. cellulolyticum* D-PEs (Table 2). This suggested that the presence of the long α8 helix and the following C-terminal tail

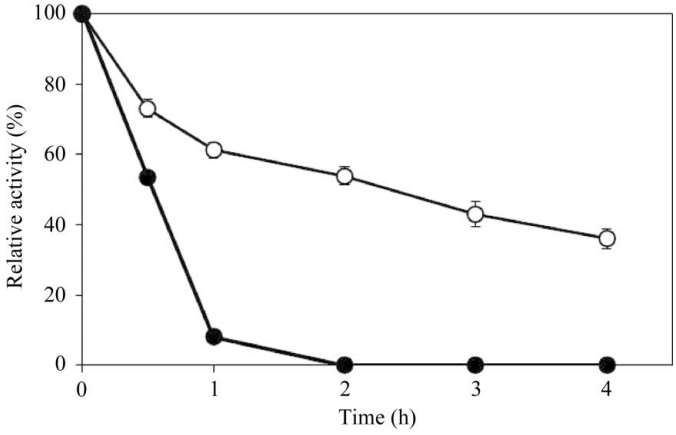

Figure 6
Effect of temperature on the stability of His-tagged *M. loti* L-RE (open circles) and His-tagged C-terminal tail (Ser292–Pro297) deletion mutant (filled circles). Enzymes were incubated at 343 K and residual activity was measured at the indicated times.

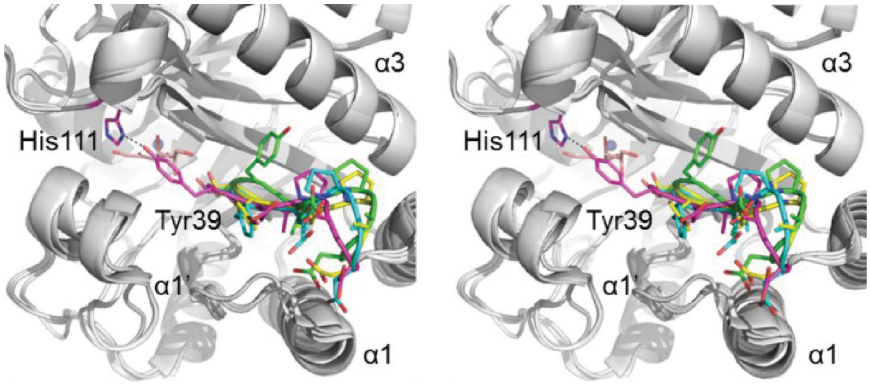

Figure 7
Stereographic close-up of the active-site cavities in subunits *A*, *B*, *C* and *D* of *M. loti* L-RE. The loop regions between β2 and α2 in subunits *A*, *B*, *C* and *D* are shown in green, cyan, magenta and yellow, respectively. The D-fructose molecule of *P. cichorii* D-TE was built in for better understanding (salmon). The Mn^{2+} ion is shown as a light blue sphere. The hydrogen bond between Tyr39 and His111 in subunit *C* is shown as a dashed line. The conformations of Tyr39 in subunits *E* and *F* were equivalent to that in subunit *C* and those in subunits *G* and *H* were comparable to those in subunits *B* and *D*, respectively.

led to the formation of additional intermolecular interactions and contributed substantially to the *M. loti* L-RE tetramer associations.

3.2. Insight into enzyme stability

The effect of temperature on His-tagged *M. loti* L-RE was determined by incubation over the temperature range 333–343 K as described in §2. Nearly full enzyme activity was retained even after incubation at 333 K for 24 h, but 60% of the activity was lost at 343 K after 4 h incubation (Fig. 6). The half-life of enzyme activity at 333 K was more than 48 h and that at 343 K was 2.3 h. As the thermostability of *P. cichorii* D-TE has not been determined, the recombinant enzyme was prepared according to the previously described procedure (Itoh *et al.*, 1994) and subsequent analysis of the effect of temperature on stability showed the half-life of the activity of this enzyme at 333 K to be 30 min. Conversely, it has been reported that the half-life of *A. tumefaciens* D-PE activity at 323 K is 63.5 min (in the presence of Mn^{2+}; Kim, Hyun *et al.*, 2006), while that of *C. cellulolyticum* D-PE at 328 K is 9.5 h (in the presence of Co^{2+}; Chan *et al.*, 2012). Therefore, *M. loti* L-RE appeared to be the most thermostable ketose 3-epimerase among the enzymes described to date. As mentioned above, the amino-acid residues belonging to the C-terminal tail of *M. loti* L-RE formed unique intermolecular interactions. The role of the C-terminal tail was assessed by construction of a *M. loti* L-RE mutant lacking some C-terminal residues (Ser292–Pro297). The molecular weight of the mutant enzyme was determined to be about 130 000 by gel-filtration analysis, indicating that the enzyme was a tetramer. A subsequent thermostability analysis showed that the enzyme was less thermostable, with an enzyme-activity half-life of 39 min at 343 K (Fig. 6). This suggested that the presence of various intersubunit interactions around the C-terminal tail, such as ion pairs, hydrophobic interactions and hydrogen bonds, was essential for the marked thermostability of *M. loti* L-RE. Because the optimum growth temperatures of *M. loti*, *A. tumefaciens*, *C. cellulolyticum* and *P. cichorii* are similar to one another (301–307 K), there is no relationship between the length of the C-terminal region and the growth temperature of the species. Although the significance of the enzyme stability is currently unclear, the thermostable L-RE is expected to have high potential as a catalyst for the bioproduction of L-xylulose.

3.3. Structure of the active site and insight into substrate specificity

When the structures of the active-site cavity in each subunit of *M. loti* L-RE were compared, Tyr39, which was observed to lie at the cavity entrance, adopted at least three different conformations. As shown in Fig. 7, only the Tyr39 side chain in subunit *C* was

situated towards His111. The Tyr39 side chain in subunit *A* was located towards the α3 helix, whereas those in subunits *B* and *D* were located towards the short loop between $\alpha1'$ and α1. The conformations of Tyr39 in subunits *E* and *F* were equivalent to that in subunit *C*, and those in subunits *G* and *H* were comparable to those in subunits *B* and *D*, respectively. As the most clear electron density for Tyr39 and surrounding residues was observed in subunit *C*, the main focus here was on the structure of the active site of subunit *C*.

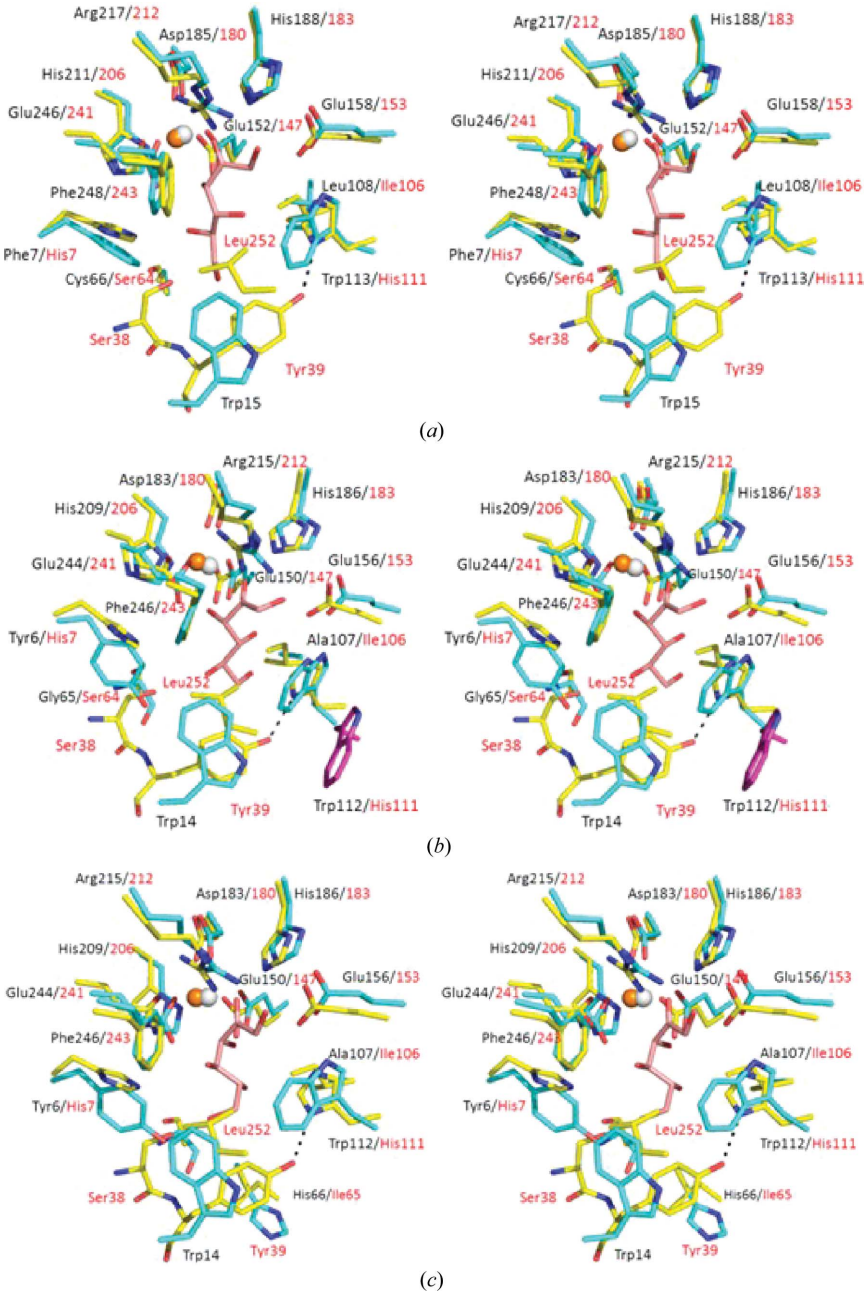

Figure 8
Stereographic close-up of the *M. loti* L-RE substrate-binding site. (*a*) *P. cichorii* D-TE structure (cyan and black labels) superimposed on that of *M. loti* L-RE (yellow and red labels); D-fructose bound to *P. cichorii* D-TE is shown as a stick model in salmon, Mn^{2+} ions are shown in orange and grey for *M. loti* L-RE and *P. cichorii* D-TE, respectively, and the hydrogen bond between His111 and Tyr39 in *M. loti* L-RE is shown as a dashed line. (*b*) Comparison of active-site structures in *M. loti* L-RE (yellow and red labels) and *A. tumefaciens* D-PE (holo structure; PDB entry 2hk1; cyan and black labels); the Trp112 side chain in the apo structure of *A. tumefaciens* D-PE (PDB entry 2hk0) is shown in magenta, the D-fructose bound to *A. tumefaciens* D-PE is shown as a stick model in salmon and Mn^{2+} ions are shown in orange and grey for *M. loti* L-RE and *A. tumefaciens* D-PE, respectively. (*c*) Comparison of active-site structures in *M. loti* L-RE (yellow and red labels) and *C. cellulolyticum* D-PE (cyan and black labels); D-fructose bound to *C. cellulolyticum* D-PE is shown as a stick model in salmon and Mn^{2+} ions are shown in orange and grey for *M. loti* L-RE and *C. cellulolyticum* D-PE, respectively.

The crystal structure of *P. cichorii* D-TE in complex with Mn^{2+} and D-fructose (PDB entry 2qun) has been determined (Yoshida *et al.*, 2007). Superposition of this structure onto the structure of *M. loti* L-RE (1.5 Å r.m.s.d. for the C^{α} atoms of 270 amino-acid residues) enabled comparison of the residues involved in Mn^{2+} and substrate binding (Fig. 8*a*). The Mn^{2+} ion is coordinated by Glu152, Asp185, His211 and Glu246 in *P. cichorii* D-TE and these four residues were strictly conserved in *M. loti* L-RE as Glu147, Asp180, His206 and Glu241, respectively. In addition, Glu158, His188 and Arg217 in *P. cichorii* D-TE, which are responsible for interactions between the enzyme and O1, O2 and O3 of D-fructose, were also conserved in *M. loti* L-RE as Glu153, His183 and Arg212, respectively. The strict conservation of key amino-acid residues involved in catalysis suggested that *M. loti* L-RE catalyzes the epimerization reaction in a manner similar to *P. cichorii* D-TE. In contrast, the amino-acid residues providing a hydrophobic environment around the substrate were significantly different in the two enzymes. In *P. cichorii* D-TE, Phe7, Trp15, Cys66, Leu108, Trp113 and Phe248 create a hydrophobic pocket around the 4, 5 and 6 positions of D-fructose. Of these residues, Phe7, Cys66, Leu108 and Trp113 were replaced by His7, Ser64, Ile106 and His111, respectively, in *M. loti* L-RE, although Phe248 was conserved as Phe243. The residue corresponding to Trp15 was not present in *M. loti* L-RE. Alternatively, Leu252, which belonged to the short $\alpha8'$ helix, was located in a similar position. The clearest difference between the two enzymes was that Ser38 and Tyr39 in subunit *C* of *M. loti* L-RE, which belonged to the loop between $\beta2$ and $\alpha2$, were situated close to the substrate (Fig. 8*a*). *P. cichorii* D-TE does not contain the residues corresponding to Ser38 and Tyr39 in three-dimensional structure because this loop region was replaced by a short $\alpha2'$ helix. These two residues are also lacking in *A. tumefaciens*

and *C. cellulolyticum* D-PEs (Figs. 8*b* and 8*c*). In our model, the Tyr39 side chain in subunit *C* formed a hydrogen bond to the His111 side chain and, *via* this interaction, Tyr39 was held at the cavity entrance, where it completely shielded the substrate-binding pocket (Fig. 8*a*). On the other hand, the Tyr39 side chains in subunits *A*, *B* and *D* made no interactions with other residues and the active-site cavities of subunits *A*, *B* and *D* appeared to be solvent-accessible (Fig. 7).

Comparison of the substrate-complex and apo structures of *A. tumefaciens* D-PE (PDB entries 2hk1 and 2hk0, respectively) revealed that substrate incorporation induced a significant conformational change (Kim, Kim *et al.*, 2006). In contrast, a ligand-induced conformational change was not observed in *P. cichorii* D-TE (Yoshida *et al.*, 2007) and *C. cellulolyticum* D-PE (Chan *et al.*, 2012). In *A. tumefaciens* D-PE, the Trp112 side chain moved towards the bound substrate upon substrate binding, acting as a lid on the active site. In *M. loti* L-RE, this Trp112 was replaced by His111 that formed a hydrogen bond to Tyr39 (Fig. 8*b*). In the case of *M. loti* L-RE, the substrate-induced movement of Tyr39 is therefore likely to take place during the enzyme reaction and Tyr39 may act as a lid on the active site. To assess the role of Tyr39, aY39A mutant was constructed and the V_{max} value of the mutant was observed to be about one-half of the V_{max} value of the wild-type enzyme [200 U mg^{-1} (mutant) *versus* 416 U mg^{-1} (wild type)]. By contrast, the K_m values for L-ribulose of the mutant and wild-type enzymes did not differ greatly [280 m*M* (mutant) *versus* 290 m*M* (wild type)]. This suggests that Tyr39 is not essential for maintaining an affinity for the substrate, but is important for proper enzyme reaction. At the present stage, it remains unclear whether or not Tyr39 completely shields the active-site pocket even in the substrate-complex structure. The structure of the L-RE–substrate binary complex should be a useful focus for further investigation.

M. loti L-RE shows high specificity for ketopentoses, while *P. cichorii* D-TE and *A. tumefaciens* and *C. cellulolyticum* D-PEs use ketohexoses as preferred substrates. The structural features that might be responsible for differences in substrate specificities were assessed by calculating the active-site cavity volumes of these enzymes using *Q-SiteFinder*. As a result, the cavity volume of *M. loti* L-RE subunit *C* (164 Å^3) was found to be markedly smaller than those of *P. cichorii* D-TE (PDB entry 2qun, chain *A*, 456 Å^3), *A. tumefaciens* D-PE (PDB entry 2hk1, chain *A*, 589 Å^3) and *C. cellulolyticum* D-PE (PDB entry 3vnk, chain *A*, 346 Å^3). When the active-site volumes of all eight subunits in the asymmetric unit were examined, the cavity volumes of subunits *C*, *E* and *F* were found to be similar to one another and their average volume was about 192 Å^3. However, the average cavity volume of subunits *B*, *D*, *G* and *H* was estimated to be about 355 Å^3. Subunit *A* exhibited the largest active-site volume of 466 Å^3. The latter two values are equivalent to those observed in *C. cellulolyticum* D-PE and *P. cichorii* D-TE, respectively. Therefore, the difference in the active-site volumes among these subunits is likely to depend on the conformation of Tyr39 in each subunit. On the other hand, the side chains of Ile106 and Leu252, which formed a hydrophobic pocket in the *M. loti* L-RE active site, were located much closer (0.94 Å and 2.6 Å, respectively) to the substrate than those of the corresponding residues (Leu108 and Trp15, respectively) in *P. cichorii* D-TE (Fig. 8*a*). When the *M. loti* L-RE active site was compared with those of *A. tumefaciens* and *C. cellulolyticum* D-PEs, similar observations were obtained (Figs. 8*b* and 8*c*). In the *M. loti* L-RE structure, moreover, the conformations of Ile106 and Leu252 were well conserved in all eight subunits in the asymmetric unit. Taken together, these observations suggest that the relatively small hydrophobic pocket around the substrate is likely to be the main factor responsible for the marked specificity for ketopentoses shown by *M. loti* L-RE.

When the amino-acid sequence of *M. loti* L-RE was submitted to the *BLAST* server (Altschul *et al.*, 1990), the five proteins from other sources with the highest identity were xylose isomerase domain-containing protein from *M. opportunistum* (97% identity; WP_013892747.1), a putative sugar phosphate isomerase/epimerase from *M. australicum* (93%; WP_015315482.1), a xylose isomerase domain-containing protein from *M. ciceri* (93%; WP_013529342.1), a xylose isomerase domain protein from *Mesorhizobium* sp. STM 4661 (90%; WP_006328597.1) and a hypothetical protein from *Rhizobium* sp. 2MFCol3.1 (70%; WP_018899452.1). In all cases, Ser38, Tyr39, Ile106 and Leu252 in *M. loti* L-RE were completely conserved. It would therefore be of interest to test whether these homologues possess L-RE activity.

Collectively, the present findings provide structural insight into the substrate binding and stability of a novel type of ketose 3-epimerase that shows high specificity for ketopentoses. We anticipate that this will provide important information for the effective application of the enzyme in rare-sugar production.

Data collection was performed on BL5A at the Photon Factory. We thank Drs Matsugaki, Igarashi and Wakatsuki for their kind assistance with the data collection. This study was supported in part by KAKENHI (Grant-in-Aid for Scientific Research C) and by the fund for Kagawa University Young Scientists 2006 and 2007.

References

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990). *J. Mol. Biol.* **215**, 403–410.
Bradford, M. M. (1976). *Anal. Biochem.* **72**, 248–254.
Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.
Chan, H.-C., Zyu, Y., Hu, Y., Ko, T.-P., Huang, C.-H., Ren, F., Chen, C.-C., Ma, Y., Guo, R.-T. & Sun, Y. (2012). *Protein Cell*, **3**, 123–131.
Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.
Granström, T. B., Takata, G., Tokuda, M. & Izumori, K. (2004). *J. Biosci. Bioeng.* **97**, 89–94.
Holm, L. & Rosenström, P. (2010). *Nucleic Acids Res*, **38**, W545–W549.
Itoh, H., Okaya, H., Khan, A. R., Tajima, S., Hayakawa, S. & Izumori, K. (1994). *Biosci. Biotechnol. Biochem.* **58**, 2168–2171.
Izumori, K. (2006). *J. Biotechnol.* **124**, 717–722.

Kim, H.-J., Hyun, E.-K., Kim, Y.-S., Lee, Y.-J. & Oh, D.-K. (2006). *Appl. Environ. Microbiol.* **72**, 981–985.
Kim, K., Kim, H.-J., Oh, D.-K., Cha, S.-S. & Rhee, S. (2006). *J. Mol. Biol.* **361**, 920–931.
Laemmli, U. K. (1970). *Nature (London)*, **227**, 680–685.
Laurie, A. T. & Jackson, R. M. (2005). *Bioinformatics*, **21**, 1908–1916.
Levin, G. V. (2002). *J. Med. Food*, **5**, 23–36.
Lovell, S. C., Davis, I. W., Arendall, W. B., de Bakker, P. I., Word, J. M., Prisant, M. G., Richardson, J. S. & Richardson, D. C. (2003). *Proteins*, **50**, 437–450.
Matsuo, T. & Izumori, K. (2004). *Asia Pac. J. Clin. Nutr.* **13**, S127.
Matsuo, T., Tanaka, T., Hashiguchi, M., Izumori, K. & Suzuki, H. (2003). *Asia Pac. J. Clin. Nutr.* **12**, 225–231.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
McNicholas, S., Potterton, E., Wilson, K. S. & Noble, M. E. M. (2011). *Acta Cryst.* D**67**, 386–394.
Mu, W., Chu, F., Xing, Q., Yu, S., Zhou, L. & Jiang, B. (2011). *Agric. Food Chem.* **59**, 7785–7792.
Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Rodriguez, R., Chinea, G., Lopez, N., Pons, T. & Vriend, G. (1998). *Bioinformatics*, **14**, 523–528.
Takeshita, K., Suga, A., Takada, G. & Izumori, K. (2000). *J. Biosci. Bioeng.* **90**, 453–455.
Uechi, K., Takata, G., Fukai, Y., Yoshihara, A. & Morimoto, K. (2013). *Biosci. Biotechnol. Biochem.* **77**, 511–515.
Vagin, A. & Teplyakov, A. (2010). *Acta Cryst.* D**66**, 22–25.
Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.
Yoshida, H., Yamada, M., Nishitani, T., Takada, G., Izumori, K. & Kamitori, S. (2007). *J. Mol. Biol.* **374**, 443–453.
Zhang, L., Mu, W., Jiang, B. & Zhang, T. (2009). *Biotechnol. Lett.* **31**, 857–862.