Structure of D-tagatose 3-epimerase-like protein
from *Methanocaldococcus jannaschii*

Keiko Uechi,^a Goro Takata,^a
Kazunari Yoneda,^b Toshihisa
Ohshima^c and Haruhiko
Sakuraba^{d*}

^aRare Sugar Research Center, Kagawa
University, 2393 Ikenobe, Miki-cho, Kita-gun,
Kagawa 761-0701, Japan, ^bDepartment of
Bioscience, School of Agriculture, Tokai
University, Aso, Kumamoto 869-1404, Japan,
^cDepartment of Biomedical Engineering,
Osaka Institute of Technology, 5-16-1 Ohmiya,
Asahi-ku, Osaka 535-8585, Japan, and ^dFaculty
of Agriculture, Kagawa University, 2393
Ikenobe, Miki-cho, Kita-gun, Kagawa 761-0701,
Japan

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

Received 21 February 2014
Accepted 13 May 2014

PDB reference: D-tagatose 3-epimerase-like
protein, 3wqo

The crystal structure of a D-tagatose 3-epimerase-like protein (MJ1311p) encoded by a hypothetical open reading frame, MJ1311, in the genome of the hyperthermophilic archaeon *Methanocaldococcus jannaschii* was determined at a resolution of 2.64 Å. The asymmetric unit contained two homologous subunits, and the dimer was generated by twofold symmetry. The overall fold of the subunit proved to be similar to those of the D-tagatose 3-epimerase from *Pseudomonas cichorii* and the D-psicose 3-epimerases from *Agrobacterium tumefaciens* and *Clostridium cellulolyticum*. However, the situation at the subunit–subunit interface differed substantially from that in D-tagatose 3-epimerase family enzymes. In MJ1311p, Glu125, Leu126 and Trp127 from one subunit were found to be located over the metal-ion-binding site of the other subunit and appeared to contribute to the active site, narrowing the substrate-binding cleft. Moreover, the nine residues comprising a trinuclear zinc centre in endonuclease IV were found to be strictly conserved in MJ1311p, although a distinct groove involved in DNA binding was not present. These findings indicate that the active-site architecture of MJ1311p is quite unique and is substantially different from those of D-tagatose 3-epimerase family enzymes and endonuclease IV.

1. Introduction

D-Tagatose 3-epimerase (D-TE) family enzymes catalyze C3 epimerization of various keto-sugars, including rare sugars that exhibit unique properties and may even exert beneficial health effects in humans (Levin, 2002; Matsuo *et al.*, 2003). It is therefore hoped that D-TE family enzymes can serve as useful catalysts for the commercially viable production of rare sugars (Granström *et al.*, 2004; Izumori, 2006). 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). This enzyme is thus expected to be key to the mass production of D-psicose from D-fructose (Takeshita *et al.*, 2000). To date, the crystal structures of D-TE from *P. cichorii* (Yoshida *et al.*, 2007), D-psicose 3-epimerases (D-PEs) from *Agrobacterium tumefaciens* (Kim *et al.*, 2006) and *Clostridium cellulolyticum* (Chan *et al.*, 2012), and L-ribulose 3-epimerase (L-RE) from *Mesorhizobium loti* (Uechi *et al.*, 2013) have all been solved. Extensive analysis of these structures has provided detailed information about the substrate-binding sites and has led to elucidation of the substrate-recognition and catalytic mechanisms of the enzymes.

In recent years, much attention has been paid to the isolation and characterization of enzymes from hyperthermophiles, because their greater thermostability represents a significant advantage over their counterparts from mesophiles. If D-TE family enzymes were present in hyperthermophiles, their greater stability would make them more amenable to practical application. Although genome analysis has enabled identification of some D-TE homologues from hyperthermophiles, the true functions of these enzymes remain unclear. For example, TM0416p, which is encoded by a hypothetical open reading frame (ORF ID TM0416) in the genome of the hyperthermophilic bacterium *Thermotoga maritima*, is predicted to be a D-TE homologue (Kim *et al.*, 2006). However, structural analysis indicates that

the hydrophobic pocket around the substrate in TM0416p differs entirely from those of the D-TE family enzymes so far described, although the enzyme shows strict conservation of the key residues involved in catalysis (Sakuraba *et al.*, 2009). The substrate specificity of TM0416p is therefore thought to differ from that of other D-TE family enzymes.

MJ1311p is a functionally unidentified protein encoded by a hypothetical open reading frame (ORF ID MJ1311) in the genome of the hyperthermophilic archaeon *Methanocaldococcus jannaschii*. When the databases were screened for homologues of the D-TE family enzymes using *BlastP*, the amino-acid sequence of MJ1311p exhibited only 24% identity to the sequence of *P. cichorii* D-TE. However, sequence alignment indicates that the key residues involved in the catalytic reaction of D-TE family enzymes are conserved in MJ1311p. In *P. cichorii* D-TE, a Mn^{2+} ion coordinated by Glu152, Asp185, His211 and Glu246 (Yoshida *et al.*, 2007) plays a pivotal role in catalysis by anchoring the bound ketohexose, while Glu152 and Glu246 carry out deprotonation/protonation at the C3 position. An alignment analysis using *ClustalW* (Chenna *et al.*, 2003) suggested these four residues are strictly conserved in MJ1311p as Glu159, Asp192, His220 and Glu260, respectively (Fig. 1). Thus, three-dimensional structural analysis of MJ1311p may help shed light on the diversity of D-TE-related proteins in hyperthermophiles.

In the present study, the gene encoding MJ1311p was expressed in *Escherichia coli* and the recombinant protein produced as an inclusion body was refolded. In addition, the crystal structure of the protein was solved at a resolution of 2.64 Å. Structural comparison revealed that the active-site architecture of MJ1311p is unique and distinctly different from those of D-TE family enzymes.

2. Materials and methods

2.1. DNA manipulation and protein expression

An expression vector encoding MJ1311p was constructed by first amplifying the MJ1311 gene fragment using PCR. The oligonucleotide primers used in the amplification were 5'-TTCATATGAAG-

CGTAAACTAAAT-3', which contains a unique *NdeI* restriction site (bold) overlapping the 5' initiation codon, and 5'-AT-**GGATCCTTATTCCTCTATTTTC**-3', which contains a unique *BamHI* restriction site (bold) proximal to the 3' end of the termination codon. Chromosomal *M. jannaschii* DNA (43067D) was obtained from the American Type Culture Collection (Manassas, Virginia, USA) and used as the template. The amplified 0.9 kbp fragment was digested with *NdeI* and *BamHI* and ligated with the expression vector pET-11a (Novagen, Madison, Wisconsin, USA) previously linearized using *NdeI* and *BamHI*, yielding pET/MJ1311. *E. coli* strain BL21 (DE3) CodonPlus RIPL (Agilent Technologies, Santa Clara, California, USA) was then transformed with pET/MJ1311, after which the transformants were cultivated at 310 K in 1 l Luria-Bertani medium containing 50 µg ml⁻¹ ampicillin until the optical density at 600 nm reached 0.6. Expression was then induced by adding 1.0 mM isopropyl β-D-1-thiogalactopyranoside to the medium and cultivation was continued for an additional 3 h at 310 K.

2.2. Purification from inclusion bodies

To purify MJ1311p from the inclusion bodies, *E. coli* transformants were harvested by centrifugation, suspended in 50 ml 10 mM potassium phosphate buffer pH 7.0 and disrupted by sonication. After sonication, the resultant lysate was centrifuged (20 000g for 10 min at 277 K) and the pellet was suspended in 20 ml 10 mM Tris-HCl buffer pH 7.5 containing 1.0 mM EDTA, 4% Triton X-100, incubated at room temperature for 30 min and centrifuged again. This procedure was repeated twice. The resultant pellet was washed twice with 20 ml Milli-Q water, after which 10 ml of denaturant solution (50 mM Tris-HCl buffer pH 7.5 containing 6 M guanidine-HCl, 0.2 M NaCl, 1 mM EDTA) was slowly added to the centrifuge tube without disturbing the pellet. The tube was then left to stand overnight at 277 K, which allowed the inclusion bodies to dissolve spontaneously. The tube was then centrifuged at 20 000g for 10 min at 277 K to remove any insoluble material. The protein concentration of the resultant supernatant was calculated based on the absorbance at 280 nm (Gill

MJ1311p	1	MKRKTKLDKSKWVLDKFGVSSLVFLPESLTSSMEKIAEHNFDWEIVCEGTHYLSPKN	60
DTE	1	-----MNKVG-----MFYTYWSTEWMDV-FPATAKRIAGLGFDLMEISLGEFHNLSDAK	48
		* * *	** ** *
MJ1311p	61	IKYLMELDRDYEVIVVHAPFS-DLNPASMNERVRKLTVEECIRDAIEGAFELDS----EV	115
DTE	49	KRELKAVADDLGLTVMCCIGLKSEYDFASPKSVRDAGTEYVKRLDDCHLLGAPVFAGL	108
		* *	** ** *
MJ1311p	116	VVVHPGYIPELWSNYVSEILDNNFSTLSEIVEIAEDYGIKIGLENMPNFRGVLGITPESL	175
DTE	109	TFCAWPQSPPLDMKDKRPYVDRAIESVRRVIKVAEDYGIYALVVVNRFEQWLCNDAKEA	168
		* *	***** ** *
MJ1311p	176	LEIVKDIDSKNLGITFDIGHANTAGNPAEFVEKLQNIIGIGIIHVHADNNGYDDEHLKIG	235
DTE	169	IAFADAVIDSPACKVQLDTFHMNIET--SFRDAILACKGKMGHFLGEAN-----RLPPG	221
		** * *	* ** *
MJ1311p	236	EGNINFIEVLEKLKEIGYDGVISIENKNIRDVKSKEILKEYLEIVN----EKVAEKEKI	291
DTE	222	EGRLPWDEIFGALKEIGYDGTIVMEP-FMRKGGSVSRAVGVWRDMSNGATDEEMDERARR	280
		** * ***** *	* * *
MJ1311p	292	EE-----	293
DTE	281	SLQFVRDKLA	290

Figure 1

Amino-acid sequence alignment of MJ1311p and *P. cichorii* D-TE (DTE). Sequences were aligned using *ClustalW*. Asterisks indicate conserved residues. The four residues involved in Mn^{2+} -ion coordination in *P. cichorii* D-TE are shown in red.

Table 1

Data-collection and refinement statistics for MJ1311p.

Values in parentheses are for the highest resolution data shell; r.m.s.d. is the root-mean-square deviation.

	Native	Thimerosal
Data collection		
Wavelength (Å)	1.0	1.0
Space group	$I2_12_12_1$	$I2_12_12_1$
Unit-cell parameters		
<i>a</i> (Å)	97.1	97.3
<i>b</i> (Å)	102.4	102.6
<i>c</i> (Å)	132.3	132.4
Resolution range (Å)	50.0–2.64 (2.69–2.64)	50.0–2.93 (2.98–2.93)
Total No. of reflections	217334	158717
No. of unique reflections	19676	14533
Multiplicity	11.0 (11.2)	10.9 (11.3)
Completeness (%)	99.9 (100)	99.9 (100)
$R_{\text{merge}}^{\dagger}$	0.043 (0.283)	0.069 (0.275)
$\langle I/\sigma(I) \rangle$	20.5 (11.5)	17.0 (13.2)
Phasing (SIRAS)		
No. of sites		4
Figure of merit	0.49	
Refinement		
Resolution range (Å)	33.1–2.64	
$R/R_{\text{free}}^{\ddagger}$ (%)	20.8/24.7 (27.5/32.5)	
No. of protein atoms	4262	
No. of water molecules	47	
No. of ligands	4 Mn^{2+} ions	
Average <i>B</i> factor (Å ²)	53.5	
R.m.s.d.		
Bond lengths (Å)	0.007	
Bond angles (°)	1.17	
Ramachandran statistics (%)		
Favoured	96.1	
Allowed	3.9	
Outliers	0	

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

& von Hippel, 1989) and adjusted to 1 mg ml^{−1} by dilution with the denaturant solution.

The solubilized enzyme (100 mg protein in 100 ml solution) was gently dropped into 1 l of refolding buffer (0.1 M Tris–HCl pH 7.5 containing 1 mM MnCl₂, 0.4 M L-arginine) and incubated for 16 h at 277 K. The resultant enzyme solution containing the refolded MJ1311p was then concentrated to a volume of 50 ml using Vivaflow 50 ultrafiltration modules (30 000 MWCO module ×6; Sartorius Stedim Biotech, Göttingen, Germany) before stepwise dialysis against (i) 50 mM Tris–HCl pH 7.5 containing 0.2 M L-arginine, 1 mM MnCl₂, 0.2 M NaCl, (ii) 50 mM Tris–HCl pH 7.5 containing 0.1 M L-arginine, 1 mM MnCl₂, 0.2 M NaCl and (iii) 50 mM Tris–HCl pH 7.5 containing 1 mM MnCl₂, 0.2 M NaCl. The dialysate containing the refolded enzyme was heated at 353 K for 10 min and then clarified by centrifugation (20 000g for 10 min at 277 K). The supernatant was then subjected to gel filtration on a Superdex 200 column (2.6 × 60 cm, GE Healthcare, Uppsala, Sweden) equilibrated with 10 mM potassium phosphate buffer pH 7.0. The MJ1311p-containing fractions were checked by SDS–PAGE, pooled and used for biochemical and structural experiments. All solutions used for the refolding procedures were filtered through a 0.45 µm membrane filter (Advantec, Dublin, California, USA) to remove dust and any other impurities.

2.3. Crystallization and data collection

For crystallization trials, the purified MJ1311p was concentrated to 10.3 mg ml^{−1}. Crystallization of the enzyme was then accomplished using the sitting-drop vapour-diffusion method. Drops (1 µl) of

protein solution were mixed with an equal volume of reservoir solution composed of 0.2 M ammonium acetate, 30% isopropanol, 0.1 M Tris–HCl buffer pH 7.5 and equilibrated against 0.1 ml reservoir solution. The crystals appeared within 3 d at 293 K and reached maximum dimensions of 0.2 × 0.2 × 0.1 mm within one week. The crystals were found to belong to the orthorhombic space group $I2_12_12_1$, with unit-cell parameters *a* = 97.1, *b* = 102.4, *c* = 132.3 Å. Diffraction data were collected to 2.64 Å resolution using monochromated radiation at a wavelength of 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 30% (v/v) ethylene glycol and cooled to 100 K in a stream of nitrogen gas. The data were processed using the *HKL*-2000 package (HKL Research, Charlottesville, Virginia, USA; Otwinowski & Minor, 1997).

2.4. Phasing, refinement and structural analysis

A heavy-atom derivative was prepared by soaking the crystals for 16 h in reservoir solution containing 2 mM ethylmercurithiosalicylic acid sodium salt (thimerosal). Phase calculation was carried out using the single isomorphous replacement with anomalous scattering (SIRAS) method with *SOLVE* (Terwilliger & Berendzen, 1999). The SIRAS map at 2.64 Å was subjected to maximum-likelihood density modification followed by autotracing using *RESOLVE* (Terwilliger, 2000). The resultant model was built using *Coot* (Emsley *et al.*, 2010), and refinement to a resolution of 2.64 Å was carried out using *REFMAC5* (Murshudov *et al.*, 2011) and *CNS* (Brünger *et al.*, 1998). After several cycles of inspection of the $2F_o - F_c$ and $F_o - F_c$ electron-density maps, the model was rebuilt. 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. The atomic coordinates and structure factors of MJ1311p have been deposited in the Protein Data Bank under accession code 3wqo.

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. Hydrogen bonds were identified using *CCP4mg* (McNicholas *et al.*, 2011). Molecular-graphics figures were created using *PyMOL* (<http://www.pymol.org>).

3. Results and discussion

3.1. Overall structure and structural homologues

The structure of MJ1311p was determined using SIRAS and was refined at a resolution of 2.64 Å. The *R* factor and *R*_{free} values for the final model were 20.8 and 24.7%, respectively (Table 1). The asymmetric unit consisted of two homologous subunits (one dimer) with a solvent content of 50.6%, which corresponded to a Matthews coefficient (Matthews, 1968) of 2.5 Å³ Da^{−1}. By gel-filtration analysis, the molecular weight of MJ1311p was determined to be about 60 000. Since the subunit molecular weight was calculated to be about 30 000 by SDS–PAGE, the protein was confirmed to have a dimeric structure in solution. In the initial electron-density map, two strong densities (5–6σ) owing to metal ions were observed within the active-site cavity. Because the protein was refolded in buffer containing MnCl₂, Mn²⁺ ions were modelled into these densities, and the final structure showed good geometry without Ramachandran outliers. The model of the dimer contained 538 amino-acid residues, four Mn²⁺ ions and

47 water molecules. In this model, the N-terminal region (amino acids 1–15 in subunit *A* and 1–16 in subunit *B*), amino-acid residues 26–28 in subunit *B* and the C-terminal region (amino acids 287–293 in subunit *A* and 287–293 in subunit *B*) were disordered and were not visible in the electron-density map. Consequently, the present description of the subunit structure is based on subunit *A*.

The subunit structure of MJ1311p is a monomer folded into an (α/β)₈ barrel carrying four additional helical segments, $\alpha 1'$, $\alpha 2'$, $\alpha 4'$ and $\alpha 6'$, which are inserted before $\alpha 1$, $\alpha 2$, $\alpha 4$ and $\alpha 6$, respectively (Fig. 2*a*). When this model of the MJ1311p monomer was sent to the DALI server (Holm & Rosenström, 2010) for identification of proteins with similar structures (as of December 24, 2013), the three proteins with the highest structural similarity were a putative sugar isomerase from *Pectobacterium atrosepticum* (PDB entry 3kct, r.m.s.d. 2.3 Å; Joint Center for Structural Genomics, unpublished work), L-xylulose 5-phosphate 3-epimerase from *E. coli* (PDB entries 3cqh, 3cqi, 3cqi and 3cqi, r.m.s.d. 2.5 Å; Shi *et al.*, 2008) and D-TE from *P. cichorii* (PDB entries 2qul, 2qum, 2qun and 2ou4, with r.m.s.d.s between 2.3 and 2.4 Å; Yoshida *et al.*, 2007). The MJ1311p monomer also showed high structural similarity to a D-TE-related protein from *T. maritima* (TM0416p; PDB entry 2zvr, with r.m.s.d. 2.1 Å; Sakuraba *et al.*, 2009), D-PE from *A. tumefaciens* (PDB entries 2hk1 and 2hk0, with r.m.s.d.s between 2.4 and 2.5 Å; Kim *et al.*, 2006) and D-PE from *C. cellulolyticum* (PDB entries 3vni, 3vnj, 3vnk and 3vnl, with r.m.s.d.s between 2.4 and 2.5 Å; Chan *et al.*, 2012). As shown in Fig. 2(*b*), the main-chain coordinates of the MJ1311p monomer were quite similar to those of *P. cichorii* D-TE, *A. tumefaciens* D-PE, *C. cellulolyticum* D-PE and *T. maritima* TM0416p. Taken together, these findings indicate that the structure of the MJ1311p monomer exhibits significant similarity to those of D-TE family enzymes.

On the other hand, the quaternary-structural arrangement of MJ1311p is notably different from those of D-TE family enzymes. It

has been reported that *P. cichorii* D-TE exhibits a dimeric structure (Yoshida *et al.*, 2007), whereas *C. cellulolyticum* and *A. tumefaciens* D-PEs exhibit tetrameric structures (Chan *et al.*, 2012; Kim *et al.*, 2006). The arrangement of the four subunits of the latter two enzymes is essentially the same. In *P. cichorii* D-TE, the amino-acid residues belonging to loop regions between $\beta 4$ and $\alpha 4$, $\beta 5$ and $\alpha 5$, $\alpha 6'$ and $\alpha 6$, $\beta 7$ and $\alpha 7$, and $\alpha 8'$ and $\alpha 8$, and two short α -helices $\alpha 6'$ and $\alpha 8'$ are involved in intersubunit interactions. A similar dimeric structure was also seen with the *A. tumefaciens* D-PE tetramer, and nearly all of the residues that contribute to intersubunit interactions are conserved between *P. cichorii* D-TE and *A. tumefaciens* D-PE (Yoshida *et al.*, 2007). In our model, MJ1311p assembled into a dimer. However, the situation at the subunit–subunit interface differs substantially from those at the interfaces of the *P. cichorii* D-TE and *A. tumefaciens* D-PE subunits. The intersubunit interactions within the MJ1311p dimer are mainly between the residues around two loop regions between $\alpha 1'$ and $\alpha 1$ and between $\beta 3$ and $\alpha 3$, and two α -helices, $\alpha 3$ and $\alpha 4'$, in both subunits. Among these amino-acid residues, nine (Leu24, Leu27, Pro28, Phe81, Leu84, Val93, Leu96, Leu126 and Trp127) form a total of 23 hydrophobic interactions between the two subunits, and six (Asp83, Asn85, Ser88, Arg92, Asp103 and Tyr122) form seven direct hydrogen bonds between the two subunits. In addition, three ion-pair interactions between the subunits are formed by three residues (Arg92, Glu99 and Asp103). As the result of the formation of this unique dimer interface, the possible active site of one subunit appears to be formed by cooperating with the other subunit (see below).

3.2. Active site

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 MJ1311p enabled comparison of the residues involved in Mn^{2+} and substrate

Figure 2 Overall structure of the MJ1311p monomer. (*a*) α -Helices (numbered from $\alpha 1$ to $\alpha 8$), β -strands (numbered from $\beta 1$ to $\beta 8$) and loop regions are indicated in red, yellow and green, respectively. (*b*) Structural comparison of MJ1311p and other D-TE family enzymes. The monomer structure of MJ1311p (green) is superimposed on those of *P. cichorii* D-TE (cyan), *A. tumefaciens* D-PE (magenta), TM0416p (yellow) and *C. cellulolyticum* D-PE (salmon).

binding (Fig. 3*a*). The Mn^{2+} ion is coordinated by Glu152, Asp185, His211 and Glu246 in *P. cichorii* D-TE, and these four residues are strictly conserved in MJ1311p as Glu159, Asp192, His220 and Glu260, respectively, as expected (Fig. 3*a*). However, the location of the Mn^{2+} ions in MJ1311p was found to differ from that in *P. cichorii* D-TE. In our model, two Mn^{2+} ions (Mn1 and Mn2) are present within the possible active-site cavity. Mn1 is coordinated in a distorted tetrahedral-form structure with four coordination bonds formed by His78 (NE2, 2.4 Å), His119 (NE2, 2.2 Å), Glu159 (OE1, 2.3 Å) and one water molecule (HOH1118, 2.6 Å). On the other hand, Mn2 is coordinated in a distorted square-planar arrangement with four coordination bonds formed by His195 (NE2, 2.4 Å), His231 (NE2, 2.4 Å), Asp229 (OD2, 2.1 Å) and a water molecule (W1, 2.8 Å). Of these residues, His78, His119 and His231 are replaced by Cys66, Leu108 and Arg217, respectively, in *P. cichorii* D-TE, although Glu159 and His195 in MJ1311p are conserved as Glu152 and His188, respectively. The residue corresponding to Asp229 is not present in *P. cichorii* D-TE. Moreover, 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, are also less conserved in MJ1311p; Arg217 is replaced by His231 and the residue corresponding to Glu158 is absent in MJ1311p, although His188 in *P. cichorii* D-TE is conserved as His195.

In general, a Mn^{2+} -ion coordination site in proteins has the following characteristics (Harding, 2006; Harding *et al.*, 2010): (i) the usual donor atoms are the side-chain O atom of Asp/Glu and the side-chain N atom of His, (ii) the usual coordination number is five

(tetragonal pyramidal or trigonal bipyramidal) or six (octahedral), but four coordination bonds (tetrahedral or square planar) are also reported when water molecules have not been identified in the metal coordination sphere, and (iii) the typical distances between the donor atoms and metal ion are 2.12–2.25 Å. The Mn^{2+} -ion coordination sites in our model preserve these features, except that some distances between the atoms and Mn^{2+} ion are outside the typical values. Because MJ1311p was denatured in the presence of EDTA and refolded in buffer containing $MnCl_2$, we modelled Mn^{2+} ions at these positions. However, the possibility of the coordination of other metals could not be ruled out.

To identify proteins with similar active-site architectures, the coordinates of the six amino-acid residues interacting with the two Mn^{2+} ions in MJ1311p were sent to the SPASM server (with a maximum superpositioning r.m.s.d. of 2 Å; Madsen & Kleywegt, 2002). The only protein found to have a similar structural motif was a DNA-repair enzyme, endonuclease IV (EndoIV), from *E. coli* (PDB entry 1qtw; Hosfield *et al.*, 1999), which contains a trinuclear zinc centre with nine coordinating residues. When the EndoIV structure was superimposed on MJ1311p, the main-chain coordinates of the protein were found to be closely related to those of the MJ1311p monomer (data not shown). Furthermore, the nine residues that interact with the three Zn^{2+} ions in EndoIV (His69, His109, Glu145, Asp179, His182, His216, Asp229, His231 and Glu261) are completely conserved in MJ1311p as His78, His119, Glu159, Asp192, His195, His220, Asp229, His231 and Glu260, respectively (Fig. 3*b*). Two of the three Zn^{2+} ions (Zn1 and Zn2) are located at positions similar to

Figure 3

Stereographic close-up of the substrate-binding site. (a) The structure of *P. cichorii* D-TE (PDB entry 2qun, cyan with blue labels) is superimposed on that of MJ1311p (green with black labels). Residues *Glu125, *Leu126 and *Trp127 from the adjacent subunit in MJ1311p are shown in yellow. A D-fructose molecule bound to *P. cichorii* D-TE is shown as a stick model in salmon. Mn^{2+} ions in MJ1311p and *P. cichorii* D-TE are shown in purple and wheat, respectively. O and N atoms are shown in red and blue, respectively. (b) Comparison of the metal-binding sites in MJ1311p (green with black labels) and EndoIV (PDB entry 1qtw, cyan with blue labels). Mn^{2+} ions (Mn1 and Mn2) and a water molecule (W1) in MJ1311p are shown in purple and red, respectively, and Zn^{2+} ions (Zn1, Zn2 and Zn3) in EndoIV are shown in orange. Atoms are coloured as in (a).

those of the two Mn^{2+} ions (Mn1 and Mn2) in MJ1311p. Within the possible active-site pocket of MJ1311p, however, no electron density for a metal ion was observed at the position corresponding to that of the third Zn^{2+} ion (Zn3) in the EndoIV structure. Instead, a water molecule (W1) was situated near this position (Fig. 3*b*). Although Mn^{2+} ions were modelled in our structure, the strict conservation of the specific residues comprising the trinuclear zinc centre supports the idea that the metals are bound in a similar fashion in MJ1311p and EndoIV: the aforementioned five His, two Glu and two Asp residues in MJ1311p may be involved in the coordination of three Zn^{2+} ions. On the other hand, Arg37, which is important for apurinic/aprimidinic site recognition in EndoIV (Garcin *et al.*, 2008), was not found in MJ1311p. What is more, the distinct crescent-shaped groove involved in DNA binding in EndoIV (Hosfield *et al.*, 1999) was not present in MJ1311p and, as mentioned above, MJ1311p exhibits a dimeric structure, whereas EndoIV is monomeric. Within MJ1311p, *Glu125, *Leu126 and *Trp127 from one monomer were found to be located over the metal-ion-binding site of the other monomer and appeared to contribute to the active site, narrowing the substrate-binding cleft (asterisks indicate residues in the neighbouring subunit; Fig. 3*a*). These results indicate that the active-site architecture of MJ1311p is quite unique and differs substantially from those of D-TE family enzymes and EndoIV. Although the true function of MJ1311p remains unknown, the present structure may provide critical information for probing the putative substrate of this enzyme.

Data collection was performed at the Photon Factory BL5A. We thank Drs Tanaka and Demura for their kind assistance with the data collection.

References

- 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.* **D54**, 905–921.
- Chan, H.-C., Zhu, 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.
- Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T. J., Higgins, D. G. & Thompson, J. D. (2003). *Nucleic Acids Res.* **31**, 3497–3500.
- Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* **D66**, 486–501.
- Garcin, E. D., Hosfield, D. J., Desai, S. A., Haas, B. J., Björas, M., Cunningham, R. P. & Tainer, J. A. (2008). *Nature Struct. Mol. Biol.* **15**, 515–522.
- Gill, S. C. & von Hippel, P. H. (1989). *Anal. Biochem.* **182**, 319–326.
- Granström, T. B., Takata, G., Tokuda, M. & Izumori, K. (2004). *J. Biosci. Bioeng.* **97**, 89–94.
- Harding, M. M. (2006). *Acta Cryst.* **D62**, 678–682.
- Harding, M. M., Nowicki, M. W. & Walkinshaw, M. D. (2010). *Crystallogr. Rev.* **16**, 247–302.
- Holm, L. & Rosenström, P. (2010). *Nucleic Acids Res.* **38**, W545–W549.
- Hosfield, D. J., Guan, Y., Haas, B. J., Cunningham, R. P. & Tainer, J. A. (1999). *Cell*, **98**, 397–408.
- 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, K., Kim, H. J., Oh, D. K., Cha, S. S. & Rhee, S. (2006). *J. Mol. Biol.* **361**, 920–931.
- Levin, G. V. (2002). *J. Med. Food*, **5**, 23–36.
- Lovell, S. C., Davis, I. W., Arendall III, W. B., de Bakker, P. I., Word, J. M., Prisant, M. G., Richardson, J. S. & Richardson, D. C. (2003). *Proteins*, **50**, 437–450.
- Madsen, D. & Kleweg, G. J. (2002). *J. Appl. Cryst.* **35**, 137–139.
- 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.* **D67**, 386–394.
- 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.* **D67**, 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.
- Sakuraba, H., Yoneda, K., Satomura, T., Kawakami, R. & Ohshima, T. (2009). *Acta Cryst.* **F65**, 199–203.
- Shi, R., Pineda, M., Ajamian, E., Cui, Q., Matte, A. & Cygler, M. (2008). *J. Bacteriol.* **190**, 8137–8144.
- Takeshita, K., Suga, A., Takada, G. & Izumori, K. (2000). *J. Biosci. Bioeng.* **90**, 453–455.
- Terwilliger, T. C. (2000). *Acta Cryst.* **D56**, 965–972.
- Terwilliger, T. C. & Berendzen, J. (1999). *Acta Cryst.* **D55**, 849–861.
- Uechi, K., Sakuraba, H., Yoshihara, A., Morimoto, K. & Takata, G. (2013). *Acta Cryst.* **D69**, 2330–2339.
- Yoshida, H., Yamada, M., Nishitani, T., Takada, G., Izumori, K. & Kamitori, S. (2007). *J. Mol. Biol.* **374**, 443–453.