Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

Structure of dihydrodipicolinate synthase from *Methanocaldococcus jannaschii*

Balasundaram Padmanabhan,[a]*‡ Richard W. Strange,[b] Svetlana V. Antonyuk,[b] Mark J. Ellis,[c] S. Samar Hasnain,[b] Hitoshi Iino,[d] Yoshihiro Agari,[d] Yoshitaka Bessho[a,d] and Shigeyuki Yokoyama[a,e]*

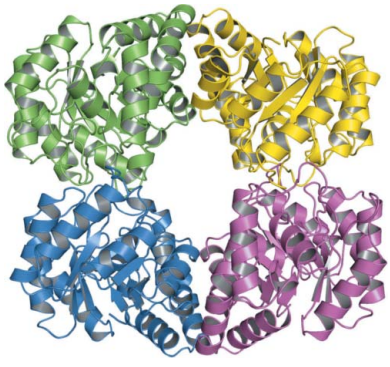
[a]Systems and Structural Biology Center, Yokohama Institute, RIKEN, 1-7-22 Suehiro, Tsurumi, Yokohama 230-0045, Japan, [b]Molecular Biophysics Group, School of Biological Sciences, University of Liverpool, Liverpool L69 7ZB, England, [c]STFC Daresbury Laboratory, Warrington, Cheshire WA4 4AD, England, [d]RIKEN SPring-8 Center, Harima Institute, 1-1-1 Kouto, Sayo, Hyogo 679-5148, Japan, and [e]Department of Biophysics and Biochemistry, Graduate School of Science, The University of Tokyo, Bunkyo-ku, Tokyo 113-0033, Japan

‡ Present address: Aptuit Laurus Private Ltd, ICICI Knowledge Park, Turkapally, Shameerpet, Hyderabad 500 078, India.

Correspondence e-mail: bpadmanabhan@hotmail.com, yokoyama@biochem.s.u-tokyo.ac.jp

Received 25 June 2009
Accepted 4 November 2009

PDB Reference: dihydrodipicolinate synthase, 2yxg, r2yxgsf.



In bacteria and plants, dihydrodipicolinate synthase (DHDPS) plays a key role in the (*S*)-lysine biosynthesis pathway. DHDPS catalyzes the first step of the condensation of (*S*)-aspartate-β-semialdehyde and pyruvate to form an unstable compound, (4*S*)-4-hydroxy-2,3,4,5-tetrahydro-(2*S*)-dipicolinic acid. The activity of DHDPS is allosterically regulated by (*S*)-lysine, a feedback inhibitor. The crystal structure of DHDPS from *Methanocaldococcus jannaschii* (*Mj*DHDPS) was solved by the molecular-replacement method and was refined to 2.2 Å resolution. The structure revealed that *Mj*DHDPS forms a functional homo-tetramer, as also observed in *Escherichia coli* DHDPS, *Thermotoga maritima* DHDPS and *Bacillus anthracis* DHDPS. The binding-site region of *Mj*DHDPS is essentially similar to those found in other known DHDPS structures.

1. Introduction

In prokaryotes and higher plants, DHDPS (DapA; dihydrodipicolinate synthase; EC 4.2.1.52) catalyzes the first unique step leading to (*S*)-lysine biosynthesis *via* the diaminopimelate pathway: the condensation of pyruvate and (*S*)-aspartate β-semialdehyde into (4*S*)-4-hydroxy-2,3,4,5-tetrahydro-(2*S*)-dipicolinic acid (HTPA; see Dobson *et al.*, 2004, and references therein). Lysine, the final product of the lysine-biosynthesis pathway, and the precursor diaminopimelic acid (DAP) are both essential for bacterial survival and growth. Lysine is important for protein biosynthesis and DAP is a structural cross-linking component of the peptidoglycan layer of mycobacterial cell walls (Cummins & Harris, 1956). The unstable product HTPA is obtained in the first catalysis step (Blickling *et al.*, 1997). Dihydrodipicolinic acid can be formed spontaneously from HTPA by the removal of water.

As DapA/DHDPS is expressed in plants and microorganisms but not in animals, this enzyme is a potential target candidate for antibiotics and herbicides (Coulter *et al.*, 1999; Cox *et al.*, 2000; Hutton *et al.*, 2003); however, no potent inhibitor is available to date. We now report the crystal structure of DHDPS from *Methanocaldococcus jannaschii* (*Mj*DHDPS) determined at 2.2 Å resolution.

2. Methods and materials

2.1. Cloning, expression and purification of *Mj*DHDPS

The gene encoding *Mj*DHDPS (gi:15668419) was amplified *via* PCR using *M. jannaschii* DSM 2661 genomic DNA and was cloned into the pET-21a expression vector (Merck Novagen). The expression vector was introduced into *Escherichia coli* Rosetta (DE3) strain (Merck Novagen) and the recombinant strain was cultured in LB medium containing 30 µg ml^{-1} chloramphenicol and 50 µg ml^{-1} ampicillin. The cells (16.4 g) from 4.5 l of medium were collected by centrifugation and were lysed by sonication in 32 ml 20 m*M* Tris–HCl buffer pH 8.0 containing 500 m*M* NaCl, 5 m*M* β-mercaptoethanol and 1 m*M* phenylmethylsulfonyl fluoride on ice. The cell lysate was heat-treated at 363 K for 13 min and centrifuged at 15 000*g* for 30 min at 277 K. The supernatant was desalted by fractionation on a HiPrep 26/10 column (GE Healthcare Biosciences). The sample was applied onto a Toyopearl SuperQ-650M column (Tosoh, Tokyo) equilibrated with 20 m*M* Tris–HCl buffer pH 8.0 and was eluted with a linear (0–0.4 *M*) gradient of NaCl. The target sample, which eluted

in the 0.28 *M* NaCl fraction, was then applied onto a Resource Q column (GE Healthcare Biosciences) equilibrated with 20 m*M* Tris–HCl buffer pH 8.0 and was eluted with a linear gradient of 0–0.4 *M* NaCl. The fractions that eluted in 0.29 *M* NaCl were further purified using a hydroxyapatite CHT5-I column (Bio-Rad Laboratories) with a linear gradient of 0.01–0.5 *M* potassium phosphate buffer pH 7.0. The target sample, which eluted in the 0.11 *M* potassium phosphate fraction, was concentrated and applied onto a HiLoad 16/60 Superdex 200 pg column (GE Healthcare Biosciences) equilibrated with 20 m*M* Tris–HCl buffer pH 8.0 containing 200 m*M* NaCl. The protein sample was analyzed by SDS–PAGE and was confirmed by N-terminal amino-acid sequencing. After concentration to 44.2 mg ml^{-1} by ultrafiltration, the yield was 8.8 mg from 16.4 g of cells.

2.2. Crystallization

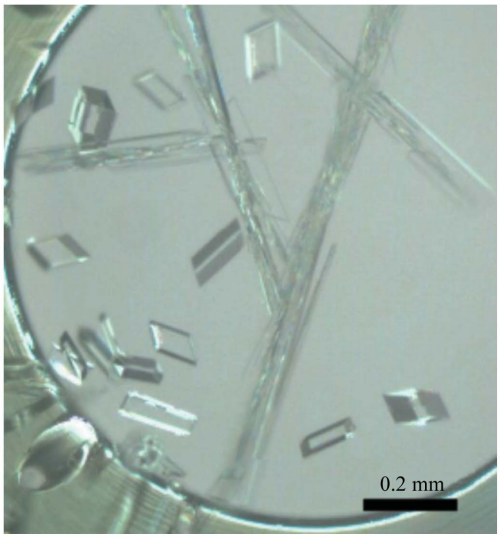
Crystallization was performed by the oil-batch method at 291 K. A 0.5 µl aliquot of crystallization reagent, consisting of 0.1 *M* cacodylate buffer pH 6.5 containing 18%(*w*/*v*) PEG 8000 and 0.2 *M* calcium acetate (Crystal Screen 1 condition No. 44; Hampton Research), was mixed with 0.5 µl 44.2 mg ml^{-1} protein solution. The mixture was then covered with 15 µl of silicone and paraffin oil. Crystals suitable for X-ray data collection appeared within 3 d and reached final dimensions of 0.1 × 0.08 × 0.08 mm (Fig. 1*a*).

2.3. Data collection and processing

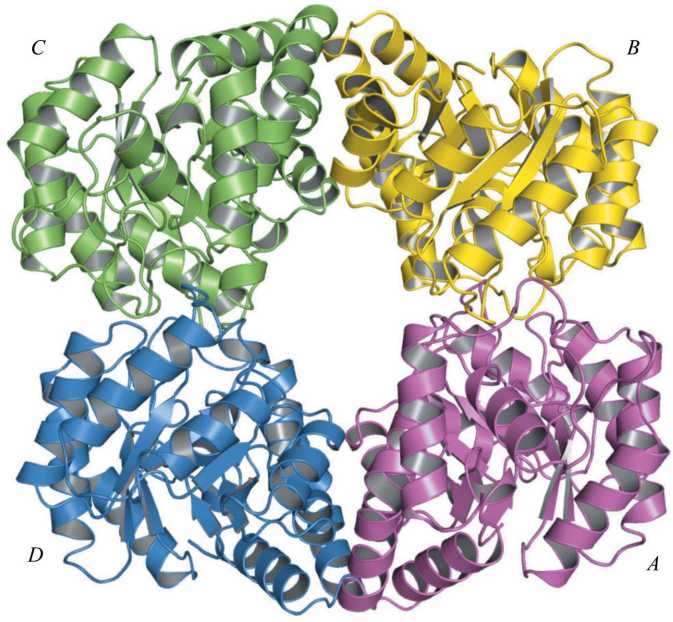
Diffraction data were collected using a MAR Mosaic 225 mm CCD detector on the PX10.1 beamline at the Daresbury Synchrotron Radiation Source (SRS), England. The crystals were directly flash-cooled in a nitrogen-gas stream at 100 K using 10%(*v*/*v*) ethylene glycol as a cryoprotectant in a drop made up of equal volumes of the protein and reservoir solutions. The crystals were maintained at 100 K during data collection. The crystal-to-detector distance was set to 150 mm. The diffraction data were processed with the *HKL*-2000 software suite (*DENZO* and *SCALEPACK*; Otwinowski & Minor, 1997). Crystallographic data and refinement statistics are summarized in Table 1.

2.4. Structure determination and refinement

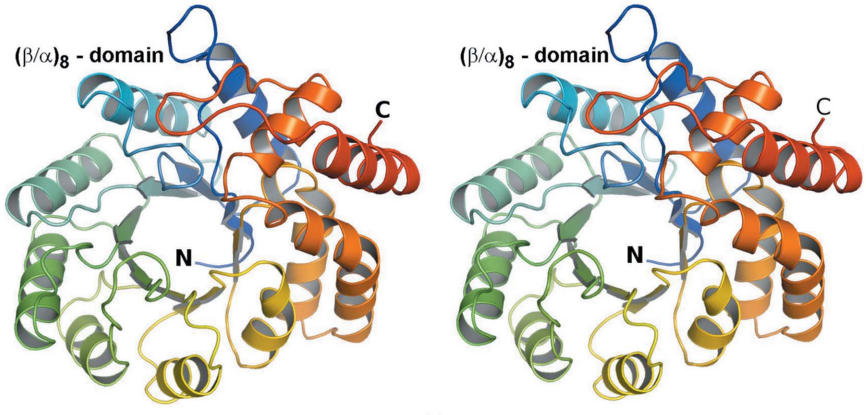
The structure of *Mj*DHDPS was solved by the molecular-replacement method with the program *MOLREP* (Vagin & Teplyakov,



(*a*)



(*b*)



(*c*)

Figure 1
Overall tertiary and quaternary structures of *Mj*DHDPS. (*a*) Crystals of the *Mj*DHDPS protein. (*b*) The quaternary structure of the DHDPS tetramer coloured by subunit. (*c*) Stereoview of a ribbon diagram of the DHDPS subunit, comprising the $(\beta/\alpha)_8$-barrel and the C-terminal domain, and coloured from blue (N-terminus) to red (C-terminus). The view is looking down the axis of the $(\beta/\alpha)_8$-barrel. All figures were produced using *PyMOL* (DeLano, 2002) unless mentioned otherwise.

1997), incorporated in the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994), using the *T. maritima* DHDPS structure (PDB code 1o5k) as a search model. The structure solution confirmed that there were four subunits with approximate 222 symmetry in the asymmetric unit (Fig. 1*b*). The model was refined with *CNS* (Brünger *et al.*, 1998) and *warpNtrace* refinement in *ARP*/*wARP* was subsequently performed (Morris *et al.*, 2003). Surprisingly, tracing using the *warpNtrace* method yielded almost complete chains of the *Mj*DHDPS quaternary structure. Several further rounds of manual fitting and refitting were performed using the program *Coot* (Emsley & Cowtan, 2004), in combination with refinement using *REFMAC*5 (Murshudov *et al.*, 1997), with careful inspection of the $2F_o - F_c$, $F_o - F_c$ and OMIT electron-density maps. The refined model of the homotetrameric form consists of 1853 residues and 700 water molecules, with a final R_{work} of 15.8% and an R_{free} of 22.4% at 2.2 Å resolution. The stereochemistry of the complex structure is excellent, as assessed by *PROCHECK* (Laskowski *et al.*, 1993). Ramachandran plot analysis of this structure revealed that 92.6% of the residues are in the allowed region and 7.0% are in the generously allowed region, with 0.4% in the disallowed region. The functional residue Tyr106 lies in the disallowed region, as was also observed for the equivalent position in other DHDPS structures. The refinement statistics are summarized in Table 1.

3. Results and discussion

3.1. Overall structure

The asymmetric unit contains four *Mj*DHDPS molecules (Fig. 1*c*). The overall tertiary structure of *Mj*DHDPS possesses a $(\beta/\alpha)_8$-barrel fold (or TIM barrel, named after the structure of triosephosphate isomerase) with three additional α-helices (α9–α11) at the

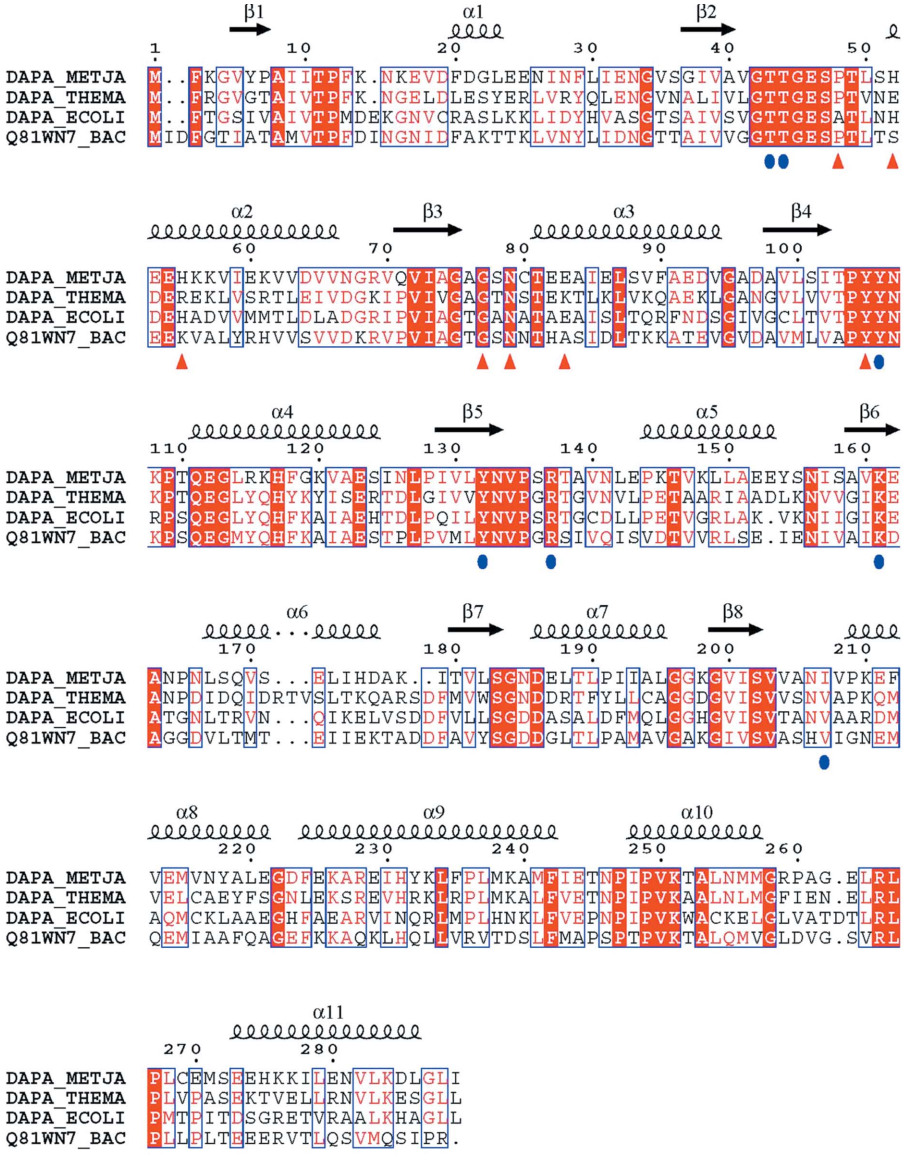


Figure 2
Sequence alignment of DHDPS proteins from *T. maritima* (DAPA_THEMA), *E. coli* (DAPA_ECOLI) and *B. anthracis* (Q81WN7_BAC) with *Mj*DHDPS (DAPA_METJA). The secondary-structural features of *Mj*DHDPS are indicated above the alignment. The potentially important residues in the active-site region and in the allosteric lysine-binding site region are indicated by blue circles and red triangles, respectively. The colours reflect the similarity (white characters on a red background represent conserved residues, red characters represent similarity in a group and blue frames indicate similarity across groups). The sequence was aligned and displayed using *ClustalW* (Thompson *et al.*, 1994) and *ESPript* (Gouet *et al.*, 1999), respectively.

C-terminus of the chain (Figs. 1*c* and 2). As in other TIM-barrel family proteins, the β-strands of the barrel form an intrinsic network of hydrogen-bonding interactions with the neighbouring β-strands and are oriented in the same directions. The overall twist adopted by all of the strands pushes the first and eighth strands close together, which is subsequently responsible for the closure of the barrel by the formation of hydrogen bonds between the main-chain atoms of the β1 and β8 β-strands.

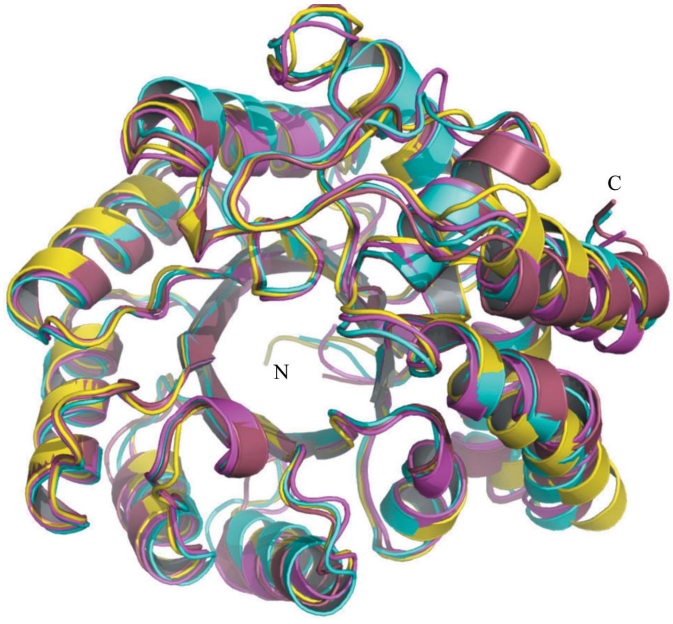
The asymmetric unit comprises four subunits. As observed in *E. coli* DHDPS (Mirwaldt *et al.*, 1995), *T. maritima* DHDPS (Pearce *et al.*, 2006) and *B. anthracis* DHDPS (Blagova *et al.*, 2006), the four subunits of the asymmetric unit clearly form a homotetramer with approximate 222 symmetry (Fig. 1*b*). The homotetramer, which is composed of a dimer of 'tight dimers', is functionally important for its catalytic activity. The *A*–*B* dimer interface is formed by the segments β3–α3, β4–α4, β5–α5, α9–α10 and α10–α11 from each monomer, which contribute a buried surface area of about 2776 Å^2. The *A*–*D* dimer association mainly arises from interhelical interactions. The dimer interface is primarily composed of the α6, α7 and α9 helices and the segments α7–β8 and α8–α9 from each monomer. The α6 and α7 helices of chain *A* interact with α7 of chain *D*, while α6 of chain *D* interacts with α7 and α9 of chain *A*. These interfaces are less extensive, with an average buried surface of 1918 Å^2. There are no contacts between subunits *A* and *C* or between subunits *B* and *D*, and the tetramer contains a large central cavity (Fig. 1*b*). As observed in the other bacterial enzymes, the active site is located in the centre of the $(\beta/\alpha)_8$-barrel in each monomer. The functional residue Lys161, which participates in Schiff-base formation, is located within the β-barrel and the side chain of Tyr132 sits over this residue. The allosteric lysine-binding site is located in the cleft at the tight-dimer interface, with each lysine molecule coordinated by residues from each monomer within the tight dimer.

3.2. Comparison with other DHDPS structures

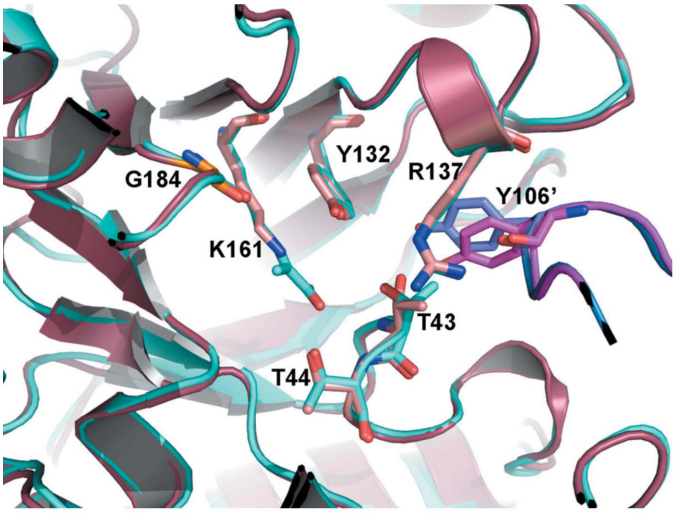
As expected, a *DALI* (Holm & Sander, 1993) search of the *Mj*DHDPS structure against the structures in the Protein Data Bank revealed that the *Mj*DHDPS structure is essentially similar to other known DHDPS structures. The highest *Z*-score value of 45.9 was obtained for *T. maritima* DHDPS (PDB code 1o5k; Pearce *et al.*, 2006), followed by 44.9 and 43.3 for *B. anthracis* DHDPS (PDB code 1xky; Blagova *et al.*, 2006) and *E. coli* DHDPS (PDB code 1dhp; Mirwaldt *et al.*, 1995), respectively. Superposition of the *Mj*DHDPS structure onto the structures of *T. maritima* DHDPS, *B. anthracis* DHDPS and *E. coli* DHDPS yielded r.m.s.d. values of 1.1, 1.2 and 1.4 Å, respectively, for all C$^\alpha$ atoms (Fig. 3*a*).

3.3. The catalytic site region

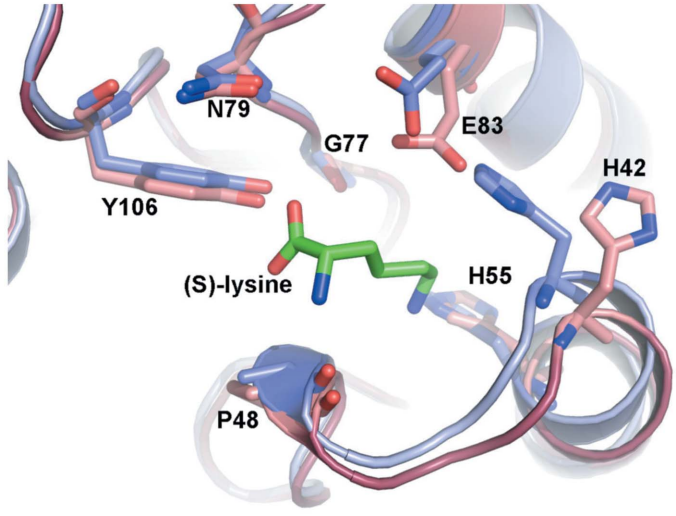
The mechanism of the *E. coli* DHDPS reaction has been studied in detail (Blickling *et al.*, 1997). Immediately after the enzyme binds to the pyruvate substrate, a Schiff base is formed between the substrate and the active-site lysine residue Lys161. The second substrate, aspartate semialdehyde, then binds. This is followed by dehydration and cyclization to form the product, which is subsequently released. It has been proposed that Tyr133 (Tyr132 in *Mj*DHDPS), Thr44 and



(*a*)



(*b*)



(*c*)

Figure 3
Comparison with other DHDPS structures. (*a*) Superposition of *Mj*DHDPS (raspberry) onto the *E. coli* DHDPS (pink), *B. anthracis* DHDPS (yellow) and *T. maritima* DHDPS (cyan) structures for the C$^\alpha$ atoms corresponding to the $(\beta/\alpha)_8$-barrel and the C-terminal domain. (*b*) Superposition of the active sites of the *Mj*DHDPS (wheat) and *T. maritima* DHDPS (cyan) structures. The numbering is shown for the *M. jannaschii* enzyme. (*c*) Superposition of the (*S*)-lysine-binding sites of *E. coli* DHDPS (blue) and *Mj*DHDPS (wheat). The (*S*)-lysine molecule is shown in green. The numbering is shown for the *M. jannaschii* enzyme.

Table 1
Summary of data-collection and refinement statistics.

Values in parentheses are for the highest resolution shell.

Data collection	
Source	SRS, PX10.1
Wavelength (Å)	0.979
Space group	$P2_1$
Unit-cell parameters (Å, °)	a = 80.5, b = 76.5, c = 101.9, γ = 106.9
Resolution (Å)	50–2.20 (2.28–2.20)
Completeness (%)	97.6 (82.6)
Redundancy	3.9 (2.7)
No. of independent observations	58517 (4918)
$R_{\rm merge}$† (%)	6.7 (21.1)
Refinement statistics	
Resolution limit (Å)	20–2.2
R factor/$R_{\rm free}$‡ (%)	15.8/22.4
Mean B factor (Å^2)	26.9
Wilson B factor (Å^2)	27.9
No. of refined atoms	8852
No. of water molecules	700
R.m.s. deviations	
Bond lengths (Å)	0.022
Bond angles (°)	1.8

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$. ‡ $R = \sum_{hkl} ||F_{\rm obs}| - |F_{\rm calc}|| / \sum_{hkl} |F_{\rm obs}|$. $R_{\rm free}$ was calculated with 5% of data that were omitted from refinement.

Thr107 (Tyr106 in *Mj*DHDPS) form a catalytic triad that shuttles protons to and from the active site (Dobson *et al.*, 2004; Fig. 3*b*).

(*S*)-Lysine is an allosteric modulator of DHDPS function and partially inhibits the catalytic activity. Dobson *et al.* (2005) recently determined the high-resolution crystal structure of the *E. coli* DHDPS–(*S*)-lysine complex, which revealed that the (*S*)-lysine-binding site is located in a crevice at the interface of the tight dimer, which is distal from the active site but connected to the active site by a conserved water channel. The allosteric modulator (*S*)-lysine is surrounded by Ala49, His53, His56, Gly78, Asn80, Glu84 and Tyr106 in *E. coli* DHDPS (Figs. 2 and 3*c*).

A comparison of the catalytic site region of *Mj*DHDPS with those of *E. coli* DHDPS, *T. maritima* DHDPS and *B. anthracis* DHDPS revealed that the regions are very similar (Fig. 3*b*). The potential interacting residues and the catalytic site region are highly conserved in the *Mj*DHDPS structure, suggesting that the catalytic function of *Mj*DHDPS may be similar to those of *E. coli* DHDPS, *T. maritima* DHDPS and *B. anthracis* DHDPS. However, further biochemical and structural analyses of *Mj*DHDPS complexes are required in order to understand the enzymatic mechanism of this protein.

We thank Ms Yoko Ukita, Mr Takeshi Ishii and Dr Akeo Shinkai for their assistance with sample preparation. We are grateful to Ms Tomoko Nakayama and Ms Azusa Ishii for clerical assistance. This work was supported in part by the RIKEN Structural Genomics/Proteomics Initiative (RSGI), the National Project on Protein Structural and Functional Analyses, Ministry of Education, Culture, Sports, Science and Technology of Japan. This work was supported by the Science and Technology Facilities Council, Daresbury Laboratory UK and beamline 10.1 at the Synchrotron Radiation Source, which was supported by Biotechnology and Biological Sciences Research Council Grant BB/E001971 (to SSH and RWS).

References

Blagova, E., Levdikov, V., Milioti, N., Fogg, M. J., Kallioman, A. K., Brannigan, J. A., Wilson, K. S. & Wilkinson, A. J. (2006). *Proteins*, **62**, 297–301.
Blickling, S., Renner, C., Laber, B., Pohlenz, H.-D., Holak, T. A. & Huber, R. (1997). *Biochemistry*, **36**, 24–33.
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.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Coulter, C. V., Gerrard, J. A., Kraunsoe, J. A. E. & Pratt, A. J. (1999). *Pestic. Sci.* **55**, 887–895.
Cox, R. J., Sutherland, A. & Vederas, J. C. (2000). *Bioorg. Med. Chem.* **8**, 843–871.
Cummins, C. S. & Harris, H. (1956). *J. Gen. Microbiol.* **14**, 583–600.
DeLano, W. L. (2002). *PyMOL Molecular Viewer.* DeLano Scientific, San Carlos, California, USA. http://www.pymol.org.
Dobson, R. C. J., Gerrard, J. A. & Pearce, F. G. (2004). *Biochem. J.* **377**, 757–762.
Dobson, R. C. J., Griffin, M. D. W., Jameson, G. B. & Gerrard, J. A. (2005). *Acta Cryst.* D**61**, 1116–1124.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Gouet, P., Courcelle, E., Stuart, D. I. & Métoz, F. (1999). *Bioinformatics*, **15**, 305–308.
Holm, L. & Sander, C. (1993). *J. Mol. Biol.* **233**, 123–138.
Hutton, C. A., Southwood, T. J. & Turner, J. J. (2003). *Mini Rev. Med. Chem.* **3**, 115–127.
Laskowski, R. A., Moss, D. S. & Thornton, J. M. (1993). *J. Mol. Biol.* **231**, 1049–1067.
Mirwaldt, C., Korndorfer, I. & Huber, R. (1995). *J. Mol. Biol.* **246**, 227–239.
Morris, R. J., Perrakis, A. & Lamzin, V. S. (2003). *Methods Enzymol.* **374**, 229–244.
Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Pearce, F. G., Perugini, M. A., McKerchar, H. J. & Gerrard, J. A. (2006). *Biochem. J.* **400**, 359–366.
Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). *Nucleic Acids Res.* **22**, 4673–4680.
Vagin, A. & Teplyakov, A. (1997). *J. Appl. Cryst.* **30**, 1022–1025.