Kandavelu Palani,^a Stephen K. Burley^b and Subramanyam Swaminathan^{a*}

^aBiology Department, Brookhaven National Laboratory, Upton, NY 11973, USA, and

^bEli Lilly and Co., San Diego, CA 92121, USA

Correspondence e-mail: swami@bnl.gov

Received 27 June 2012

Accepted 16 November 2012

PDB Reference: alanine racemase, 3co8

Structure of alanine racemase from *Oenococcus oeni* with bound pyridoxal 5'-phosphate

The crystal structure of alanine racemase from *Oenococcus oeni* has been determined at 1.7 Å resolution using the single-wavelength anomalous dispersion (SAD) method and selenium-labelled protein. The protein exists as a symmetric dimer in the crystal, with both protomers contributing to the two active sites. Pyridoxal 5'-phosphate, a cofactor, is bound to each monomer and forms a Schiff base with Lys39. Structural comparison of alanine racemase from *O. oeni* (Alr) with homologous family members revealed similar domain organization and cofactor binding.

1. Introduction

Alanine racemase (Alr; EC 5.1.1.1) belongs to the fold-type III group of pyridoxal 5'-phosphate (PLP; a vitamin B₆ derivative) dependent enzymes and catalyzes the interconversion of D-alanine and L-alanine (Grishin *et al.*, 1995). *Oenococcus oeni* is a Gram-positive bacterium from the *Oenococcus* genus; it has major importance in the field of oenology and is the primary bacterium involved in completing the malolactic fermentation (Kunkee, 1973). The alanine racemases are ubiquitous in bacteria, where they produce D-alanine, an essential component of the peptidoglycan layer that protects cells from osmotic lysis (Walsh, 1989). These enzymes also occur in eukaryotes. In the absence of substrate, the PLP cofactor (a phosphorylated and oxidized form of vitamin B₆; Fenn *et al.*, 2003) forms a Schiff base with a conserved lysine side chain (Shaw *et al.*, 1997; Mustata *et al.*, 2003). Alr utilizes a 'two-base' enzymatic mechanism, which employs two distinct active-site residues (lysine and tyrosine) that function as a catalytic acid and base to generate the enantiomeric product D-alanine from L-alanine (Shaw *et al.*, 1997; Stamper *et al.*, 1998; Morollo *et al.*, 1999; Sun & Toney, 1999) and *vice versa* (Watanabe *et al.*, 2002). Here, we report the X-ray crystal structure of alanine racemase from *O. oeni* with bound PLP cofactor.

2. Materials and methods

2.1. Gene cloning, expression and protein purification

The protocol for cloning, expression and purification is given in detail in the PSI Knowledgebase under Target ID NYSGXRC-11082i (<http://sbkb.org/tt/search?targetid=NYSGXRC-11082i&lab=NYSGXRC>). Briefly, the target gene (Gene ID OE0E_0162) for the full-length target protein from *O. oeni* I (residues 3–371) was PCR-amplified using the following primers: 5'-GAAGCTATTCATCGATCAACGC-3' (forward) and 5'-CATCGACCACCATCCTTTTAAATCG-3' (reverse). The resulting PCR product was cloned with a C-terminal 6×His tag into pSGX3 (BC) and the sequence-verified clone was then transformed into *Escherichia coli* BL21 (DE3) cells. Expression of selenomethionine-derivatized (SeMet) protein was carried out in *E. coli* HY medium by adding 30 ml 1.5× SeMet buffer prior to IPTG induction. The protein was purified on an Ni-NTA column followed by sizing column chromatography.

2.2. Crystallization and data collection

The SeMet protein was concentrated to ~10 mg ml⁻¹ and subjected to sitting-drop vapor diffusion (Greiner 96-well plate) by

mixing 1 μl protein solution and 1 μl reservoir solution from the Index HT screen (Hampton Research). Initial crystals were obtained from a crystallization condition consisting of 0.2 M sodium chloride, 0.1 M bis-tris pH 5.5, 25% (w/v) PEG 3350 (condition F10 of Index HT). Diffraction-quality crystals were obtained by optimizing the above condition using a Linbro 24-well plate and sitting-drop vapor diffusion (2 μl 10 mg ml^{-1} protein solution and 2 μl reservoir solution consisting of 25% PEG 3350, 0.1 M bis-tris pH 5.5, 0.2 M sodium chloride equilibrated against 650 μl reservoir solution). Thick plate-like crystals were obtained within 3 d and were flash-cooled in liquid nitrogen following the addition of 20% (v/v) glycerol to the mother liquor as a cryoprotectant. Diffraction data were collected (360° data set with a 1° rotation angle about φ) using an ADSC Quantum 210 detector under standard cryogenic conditions (100 K) on beamline X12C at the National Synchrotron Light Source (NSLS), Brookhaven National Laboratory and were integrated and scaled using *HKL-2000* (Otwinowski & Minor, 1997). The crystals belonged to the monoclinic space group $P2_1$ (with unit-cell parameters $a = 47.4$, $b = 99.6$, $c = 84.7$ Å, $\beta = 103.3^\circ$). The calculated Matthews coefficient is $2.28 \text{ Å}^3 \text{ Da}^{-1}$, assuming the presence of two molecules per asymmetric unit, which corresponds to 46% solvent content by volume. Data-collection and refinement statistics are provided in Table 1.

2.3. Structure determination

The structure was determined to 1.7 Å resolution *via* single-wavelength anomalous dispersion using SeMet protein. All 16 possible selenium positions (the N-terminal post-translational Met was not included) were located by *SHELXD* (Sheldrick, 2008) and the correct handedness was chosen by *SHELXE* (Sheldrick, 2008). Phase refinement and density modification were performed using *SHARP* (de La Fortelle & Bricogne, 1997) and *SOLOMON* (Abrahams & Leslie, 1996), yielding an experimental electron-density map suitable for automated model building with *ARP/wARP* (Lamzin *et al.*, 2001). 93% of the polypeptide chain was built without manual

Table 1

Data-collection and refinement statistics.

Values in parentheses are for the outermost shell.

Space group	$P2_1$
Unit-cell parameters (Å, °)	$a = 47.4$, $b = 99.6$, $c = 84.7$, $\beta = 103.3$
Method	SAD
Wavelength (Å)	0.9790
No. of unique reflections	80948 (7336)
Resolution range (Å)	50.0–1.7 (1.76–1.70)
Multiplicity	6.6 (4.3)
$R_{\text{merge}}^\dagger$ (%)	6.3 (23.0)
Completeness (%)	98.5 (87.5)
$\langle I/\sigma(I) \rangle$	24.7 (3.0)
Refinement statistics	
Resolution (Å)	38.08–1.70 (1.81–1.70)
No. of unique reflections	80948 (11468)
R factor ‡	0.247 (0.291)
$R_{\text{free}}^\S$	0.259 (0.304)
No. of protein atoms	5548
No. of heteroatoms	30
No. of water molecules	626
R.m.s.d. bond lengths (Å)	0.011
R.m.s.d. bond angles (°)	1.5
Average B factors (Å 2)	
Protein atoms	15.2
Water molecules	24.5
Ligand (PLP)	18.5
Ramachandran plot statistics, residues in (%)	
Core region	91.3
Additionally allowed region	8.5
Generously allowed region	0.2

$^\dagger R_{\text{merge}} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$, where $\langle I(hkl) \rangle$ is the mean intensity of symmetry-related reflections $I_i(hkl)$. $^\ddagger R = \sum_{hkl} ||F_{\text{obs}}| - |F_{\text{calc}}|| / \sum_{hkl} |F_{\text{obs}}|$, where F_{obs} and F_{calc} are the observed and the calculated structure-factor amplitudes, respectively. $^\S R_{\text{free}}$ was calculated using 5% of the data, which were withheld from refinement.

intervention. The remainder was built manually using *O* (Jones *et al.*, 1991). *ARP/wARP* was able to interpret the electron density for both molecules. The final atomic model was refined with *CNS* (Brünger *et al.*, 1998) and contained 705 residues, two PLP cofactors and 626

Figure 1

The structure of *O. oeni* alanine racemase. (a) Ribbon representation of the homodimeric structure showing the protomers in dark blue and red. PLP is shown as a stick model. (b) Ribbon representation of the monomer; helices and β -strands are shown in cyan and magenta, respectively. PLP is shown as a stick model in green and orange.

water molecules, yielding R and R_{free} values of 0.25 and 0.26, respectively. Individual isotropic thermal factors were used in the

Figure 2

The σ_A -weighted $2F_o - F_c$ electron-density map contoured at the 1.0σ level shows the formation of a Schiff base between Lys39 and the PLP cofactor.

refinement and no NCS restraints were used. The following loop segments were not visible in the experimental electron-density map and were not included in structure refinement: 169–177 in molecule *A* and 258–276 in molecule *B*. Refinement statistics are provided in Table 1. *PROCHECK* (Laskowski *et al.*, 1993) was used to determine that 91.3% of the residues lie in the most favored region of the Ramachandran diagram, with 8.5% of the residues in the additionally allowed regions and 0.2% in the generously allowed regions. No residues were found in disallowed regions. All figures were prepared using *PyMOL* (DeLano, 2002). Refined atomic coordinates and experimental structure factors have been deposited with the Protein Data Bank (PDB entry 3co8).

3. Results and discussion

3.1. Overall structure of *O. oeni* Alr

The overall structure of Alr from *O. oeni* reveals that it is a symmetric head-to-tail homodimer (Fig. 1a). Each monomer is

Figure 3

Multiple sequence alignment of six alanine racemase enzymes. Only alanine racemases from Gram-positive bacteria are included. Strictly conserved residues are highlighted in red; highly conserved amino acids, including the catalytic residues Lys39 and Tyr266, are enclosed in boxes. The protein accession numbers are as follows: *O. oeni*, Q07HB4; *E. faecalis*, Q837J0; *B. stearothermophilus*, P10724; *B. anthracis*, Q81VF6; *S. aureus*, P63479; *S. pneumoniae*, P0A2W8.

composed of N- and C-terminal domains. The N-terminal domain (residues 15–235) is an eight-stranded (α/β)₈ TIM-barrel structure which serves as the PLP-binding site. The N-terminal domain is composed of 11 α -helices ($\alpha 1$ – $\alpha 11$) and eight β -strands ($\beta 2$ – $\beta 9$) (Fig. 1b). The first β -strand of the monomer, $\beta 1$, is located at the extreme N-terminus and does not contribute to the TIM barrel. The C-terminal domain, comprising residues 1–14 and 236–370, is composed of five α -helices ($\alpha 12$ – $\alpha 16$) and 11 β -strands ($\beta 10$ – $\beta 20$) (Fig. 1b). The β -strands form two distinct β -sheets (four-stranded and five-stranded).

Conserved catalytic residues are furnished by both protomers (Lys39 and Tyr266', where the prime denotes a residue from the opposing monomer), yielding two active sites per homodimer. In our structure, the PLP cofactor binds covalently *via* an imine bond to the η -amino group of Lys39 from α -helix 2, forming an internal aldimine (Schiff base; Fig. 2). Dimer formation is stabilized by 20 intermolecular hydrogen bonds as determined by *PDBsum* (Laskowski *et al.*, 1997).

3.2. Sequence alignment

Both the UniProt Knowledgebase (<http://www.uniprot.org>) and the Protein Data Bank (<http://www.pdb.org>) contain numerous alanine racemase entries. We restricted our analysis of sequence similarity to alanine racemases from bacteria and then further restricted the analysis to those from Gram-positive bacteria. A *ClustalW2* (Larkin *et al.*, 2007) amino-acid sequence alignment of the six resulting alanine racemases from Gram-positive bacteria for which crystal structures are available [*O. oeni* Q04HB7, *Enterococcus faecalis* Q837J0 (40.0% sequence identity to *O. oeni* Alr), *Bacillus stearothermophilus* P10724 (42.1% identity), *B. anthracis* Q81VF6 (34% identity), *Staphylococcus aureus* P63479 (31% identity) and *Streptococcus pneumoniae* P0A2W8 (35% identity); Watanabe *et al.*, 1999; Couñago *et al.*, 2009; Scaletti *et al.*, 2012; Im *et al.*, 2011] is shown in Fig. 3. The highest sequence similarity was detected within the N-terminal domain and in the vicinity of the active-site residues

Table 2

Conserved catalytic residues of different Alr enzymes and their equivalent positions.

Alr source	Catalytic residues	
<i>O. oeni</i>	Lys39	Tyr266
<i>E. faecalis</i>	Lys40	Tyr267
<i>B. stearothermophilus</i>	Lys39	Tyr265
<i>B. anthracis</i>	Lys41	Tyr274
<i>S. aureus</i>	Lys39	Tyr269
<i>S. pneumoniae</i>	Lys40	Tyr267

(Lys39 and Tyr266 in *O. oeni*; Table 2). The sequence-alignment figure was produced using *ESPrpt* v.2.2 (Gouet *et al.*, 1999). Lys128 (*O. oeni* numbering) is conserved in four of the six Gram-positive bacteria compared here. However, it is alanine in *O. oeni* and asparagine in *B. anthracis*.

3.3. Active site of Alr

All known alanine racemase structures (Noda *et al.*, 2004; Priyadarshi *et al.*, 2009; Watanabe *et al.*, 1999) demonstrate conservation of the active-site residues surrounding PLP (Fig. 3). The highest similarity is observed for residues that interact with the cofactor phosphate group and those involved in the hydrogen-bonding network surrounding the aromatic ring of PLP. The PLP-binding site lies near the center of the barrel of one monomer and close to the C-terminal domain of the homodimeric partner. Covalent anchoring of the cofactor occurs *via* an imine bond to the ϵ -amino group of Lys39, which is thought to play an important role in product release (Toney & Kirsch, 1989; Lu *et al.*, 1993). In *O. oeni* Alr the phosphate group of PLP is involved in forming several hydrogen bonds to the side-chain atoms of residues Tyr43, Ser203, Arg221 and Thr224 and the main-chain atoms of residues Ser203, Gly223 and Thr224 (Fig. 4). Some of these residues are conserved in the alanine racemase from *B. stearothermophilus* (Tyr43, Ser204, Arg219 and Gly221). Residues Tyr44, Ser207, Gly224, Val225 and Tyr356 are involved in hydrogen-bonding interactions with PLP in Alr from *E. faecalis*. The pyridine ring of PLP is held in place by a critical hydrogen bond between its N1 atom and NE of Arg221 (*O. oeni* Alr) or Arg219 (*B. stearothermophilus* Alr). The role of the arginine residue at position 221 can be regarded as a strategy of the enzyme to keep the pyridine N atom of PLP unprotonated throughout the catalytic cycle (Sun & Toney, 1999).

3.4. Comparison of the overall folding of alanine racemases

A *DALI* (Holm & Sander, 1995) search brought out a number of alanine racemase structures from different bacterial species deposited in the Protein Data Bank. To elucidate the important features of this enzyme, alanine racemases from six Gram-positive bacteria as mentioned previously were considered. Their Z-scores and r.m.s. deviations with respect to *O. oeni* alanine racemase (PDB entry 3co8) are as follows: *E. faecalis* (PDB entry 3e5p; Priyadarshi *et al.*, 2009), Z-score = 43.2, 40% sequence identity, r.m.s.d. = 1.8 Å for 371 C α pairs; *B. stearothermophilus* (PDB entry 1sft; Shaw *et al.*, 1997), Z-score = 42.1, 42.1% sequence identity, r.m.s.d. = 1.8 Å for 382 C α pairs; *S. pneumoniae* (PDB entry 3s46; Im *et al.*, 2011), Z-score = 43.5, r.m.s.d. = 2.0 Å for 353 C α pairs; *S. aureus* (PDB entry 4a3q; Scaletti *et al.*, 2012), Z-score = 39.1, r.m.s.d. = 2.0 Å for 333 C α pairs; *B. anthracis* (PDB entry 3ha1; Couñago *et al.*, 2009), Z-score = 41.8, r.m.s.d. = 2.2 Å for 352 C α pairs. Superposition of these structures with *O. oeni* Alr revealed that both the N- and C-termini are structurally well conserved and the cofactor (PLP) binding sites are similar. All of the bacteria mentioned above are involved in peptidoglycon biosynthesis.

Figure 4

Hydrogen-bond network in the *O. oeni* alanine racemase active site. Hydrogen bonds with an interaction distance of less than 3.0 Å are depicted by dotted lines. The active site is formed by both monomers. Residues Lys39, Tyr43, Ser203, Arg221, Gly223 and Thr224 from one monomer (in cyan) and Tyr266 from the opposing monomer (in green) are shown as stick models. PLP is shown as a stick model (salmon).

Figure 5

Superposition of *O. oeni* Alr1 and Alr2 (PDB entries 3co8 and 3hur, respectively). The Alr1 structure is shown in yellow and that of Alr2 in green. A two-stranded β -sheet that is only present in domain 2 of Alr2 is shown in red. The corresponding sequence in Alr1 is highlighted in blue and does not adopt a β -sheet conformation.

Interestingly, *O. oeni* contains two alanine racemases, Alr1 and Alr2, which share a sequence identity of only 32.2% (UniProt IDs Q04HB7 and Q04DI1). The purpose of the presence of two alanine racemase enzymes in this genome is not clear. The crystal structures of both Alrs have been determined (PDB entry 3co8, present study; PDB entry 3hur, New York SGX Research Center for Structural Genomics, unpublished work). They have a similar overall fold apart from a two-stranded β -sheet (residues 345–365) that is present at the C-terminus in Alr2. This seems to be unique to Alr2 since no other alanine racemase presents a similar β -sheet. Remarkably, this two-stranded β -sheet extends from domain 2 (or the C-terminal domain) and covers the TIM barrel as if it were a lid (Fig. 5). While the structure of Alr1 has a PLP bound to the enzyme, it is absent in the Alr2 structure; instead, a sulfate ion occupies the PLP site. Further investigation of the structural differences between these two isozymes is beyond the scope of this study.

4. Conclusions

We have reported the crystal structure of an alanine racemase from *O. oeni* with bound cofactor PLP. The overall structure is very similar to those of the alanine racemases from *E. faecalis* and *B. stearothermophilus*. Alanine racemase is present in many organisms and the structure determination of as many alanine racemases as possible will help in studying the structure–sequence relationship. Sequence analysis reveals that the active-site residues are conserved within this family of proteins. The structural comparison of alanine racemase from *O. oeni* with homologous proteins further demonstrates that the

domain organization and cofactor (PLP) binding site are conserved in this family.

This research was supported by a U54 award from the National Institute of General Medical Sciences to the NYSGXRC (GM074945; PI SKB) under DOE Prime Contract No. DEAC02-98CH10886 with Brookhaven National Laboratory. We gratefully acknowledge data-collection support from beamline X12C at the National Synchrotron Light Source.

References

- Abrahams, J. P. & Leslie, A. G. W. (1996). *Acta Cryst.* **D52**, 30–42.
- 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.
- Couñago, R. M., Davlieva, M., Strych, U., Hill, R. E. & Krause, K. L. (2009). *BMC Struct. Biol.* **9**, 53.
- DeLano, W. L. (2002). *PyMOL*. <http://www.pymol.org>.
- Fenn, T. D., Stamper, G. F., Morollo, A. A. & Ringe, D. (2003). *Biochemistry*, **42**, 5775–5783.
- Gouet, P., Courcelle, E., Stuart, D. I. & Métoz, F. (1999). *Bioinformatics*, **15**, 305–308.
- Grishin, N. V., Phillips, M. A. & Goldsmith, E. J. (1995). *Protein Sci.* **4**, 1291–1304.
- Holm, L. & Sander, C. (1995). *Trends Biochem. Sci.* **20**, 478–480.
- Im, H., Sharpe, M. L., Strych, U., Davlieva, M. & Krause, K. L. (2011). *BMC Microbiol.* **11**, 116.
- Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* **A47**, 110–119.
- Kunkee, R. E. (1973). *The Chemistry of Winemaking*, edited by A. D. Webb, p. 137. Washington: American Chemical Society.
- La Fortelle, E. de & Bricogne, G. (1997). *Methods Enzymol.* **276**, 472–493.
- Lamzin, V. S., Perrakis, A. & Wilson, K. S. (2001). *International Tables for Crystallography*, Vol. F, edited by M. G. Rossmann & E. Arnold, pp. 720–722. Dordrecht: Kluwer Academic Publishers.
- Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J. & Higgins, D. G. (2007). *Bioinformatics*, **23**, 2947–2948.
- Laskowski, R. A., Hutchinson, E. G., Michie, A. D., Wallace, A. C., Jones, M. L. & Thornton, J. M. (1997). *Trends Biochem. Sci.* **22**, 488–490.
- Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
- Lu, Z., Nagata, S., McPhie, P. & Miles, E. W. (1993). *J. Biol. Chem.* **268**, 8727–8734.
- Morollo, A. A., Petsko, G. A. & Ringe, D. (1999). *Biochemistry*, **38**, 3293–3301.
- Mustata, G. I., Soares, T. A. & Briggs, J. M. (2003). *Biopolymers*, **70**, 186–200.
- Noda, M., Matoba, Y., Kumagai, T. & Sugiyama, M. (2004). *J. Biol. Chem.* **279**, 46153–46161.
- Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
- Priyadarshi, A., Lee, E. H., Sung, M. W., Nam, K. H., Lee, W. H., Kim, E. E. & Hwang, K. Y. (2009). *Biochim. Biophys. Acta*, **1794**, 1030–1040.
- Scaletti, E. R., Luckner, S. R. & Krause, K. L. (2012). *Acta Cryst.* **D68**, 82–92.
- Shaw, J. P., Petsko, G. A. & Ringe, D. (1997). *Biochemistry*, **36**, 1329–1342.
- Sheldrick, G. M. (2008). *Acta Cryst.* **A64**, 112–122.
- Stamper, G. F., Morollo, A. A., Ringe, D. & Stamper, C. G. (1998). *Biochemistry*, **37**, 10438–10445.
- Sun, S. & Toney, M. D. (1999). *Biochemistry*, **38**, 4058–4065.
- Toney, M. D. & Kirsch, J. F. (1989). *Science*, **243**, 1485–1488.
- Walsh, C. T. (1989). *J. Biol. Chem.* **264**, 2393–2396.
- Watanabe, A., Kurokawa, Y., Yoshimura, T., Kurihara, T., Soda, K., Esaki, N. & Watababe, A. (1999). *J. Biol. Chem.* **274**, 4189–4194.
- Watanabe, A., Yoshimura, T., Mikami, B., Hayashi, H., Kagamiyama, H. & Esaki, N. (2002). *J. Biol. Chem.* **277**, 19166–19172.