Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structure determination and refinement at 2.44 Å resolution of argininosuccinate lyase from *Escherichia coli*

Prasenjit Bhaumik, M. Kristian Koski, Ulrich Bergmann and Rik K. Wierenga*

Biocenter Oulu and Department of Biochemistry, University of Oulu, Linnanmaa, Box 3000, FIN-90014 Oulu, Finland

Correspondence e-mail: rik.wierenga@oulu.fi

Received 5 July 2004
Accepted 6 September 2004

PDB Reference: argininosuccinate lyase, 1tj7, r1tj7sf.

Escherichia coli argininosuccinate lyase has been crystallized from a highly concentrated sample of purified recombinant α-methylacyl-CoA racemase, in which it occurred as a minor impurity. The structure has been solved using molecular replacement at 2.44 Å resolution. The enzyme is tetrameric, but in this crystal form there is a dimer in the asymmetric unit. The tetramer has four active sites; each active site is constructed from loops of three different subunits. One of these catalytic loops, near residues Ser277 and Ser278, was disordered in previous structures of active lyases, but is very well ordered in this structure in one of the subunits owing to the presence of two phosphate ions in the active-site cavity. The positions of these phosphate ions indicate a plausible mode of binding of the succinate moiety of the substrate in the competent catalytic complex.

1. Introduction

Argininosuccinate lyase (ASL; EC 4.3.2.1) participates in arginine synthesis in all organisms catalyzing the reversible breakdown of argininosuccinate to arginine and fumarate (Fig. 1). The reaction is also part of the urea cycle, the major pathway for the detoxification of ammonia in ureotelic species. ASL is a member of a superfamily of enzymes that are active as tetramers and catalyze similar β-elimination reactions in which a C—N or C—O bond is cleaved, with subsequent

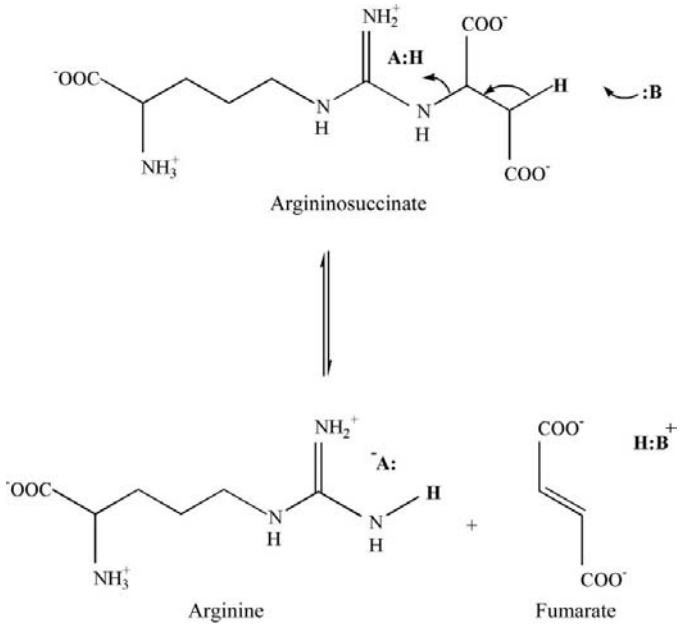

Figure 1
The reaction catalyzed by the enzyme argininosuccinate lyase. A and B refer to protic protein side chains that are presumed to be important for the catalysis. Only functional H atoms are explicitly shown.

release of fumarate as one of the products. Other members of the superfamily with known structures include fumarase (Weaver & Banaszak, 1996), aspartase (Fujii *et al.*, 2003) and adenylosuccinate lyase (Toth & Yeates, 2000). The enzymes in this superfamily catalyze similar reactions but have very little sequence similarity. However, there are three regions in which amino-acid sequences are highly conserved (Turner *et al.*, 1997; Sampaleanu *et al.*, 2002). The crystal structures of ASLs from several species have been studied, in particular the duck δ1 and δ2 eye lens crystallins (Sampaleanu, Vallee, Slingsby *et al.*, 2001; Sampaleanu *et al.*, 2002; Vallee *et al.*, 1999; Abu-Abed *et al.*, 1997), which are inactive and active homologues of ASL, respectively. The structures of human ASL at 4 Å resolution (Turner *et al.*, 1997), inactive human Q286R ASL (Sampaleanu, Vallee, Thompson *et al.*, 2001) and inactive turkey δ1 crystallin (Simpson *et al.*, 1994) have also been studied. These studies have shown that the three highly conserved sequence regions, referred to as C1, C2 and C3, are in loops that surround the catalytic site and provide most of the catalytic residues. Of these loops, the C1 and C2 loops are well ordered in all structures. The C3 loop is in an open disordered conformation in the structures of active lyases; a closed conformation has only been observed in one structure of the inactive duck δ1 crystallin, with one sulfate ion bound in the active site (Sampaleanu, Vallee, Slingsby *et al.*, 2001). Here, we report the 2.44 Å resolution structure of ASL from *Escherichia coli* (eASL). This enzyme (SWISS-PROT entry ARLY_ECOLI) is the gene product of the *argH* gene (Blattner *et al.*, 1993); its sequence identity to human ASL is 45%. The structure analysis shows that in this structure the flexible C3 loop is in a well ordered closed conformation in one of the active sites, with two phosphate ions bound.

2. Materials and methods

2.1. Crystallization

eASL was crystallized unintentionally while optimizing the crystallization conditions of recombinant α-methylacyl-CoA racemase, resulting predominantly in bipyramidal crystals (Bhaumik *et al.*, 2003). These crystals were grown at room temperature in mother liquor consisting of 1.26 *M* ammonium phosphate pH 7.0. The space group was $P6_122$ or $P6_522$ and the unit-cell parameters were $a = b = 121.5$, $c = 255.3$ Å. These crystals were grown routinely from several batches of the purified racemase; the purity at the various stages of purification is shown in Fig. 2. These bipyramidal crystals were well ordered but small. Consequently, studies aimed at finding suitable growth conditions for larger crystals were carried out. The mother liquor and crystals in the crys-

Table 1
Data-collection and refinement statistics.

Values in parentheses are for the highest resolution shell.

Data-collection statistics	
Space group	$P6_122$
Unit-cell parameters	
$a = b$ (Å)	121.47
c (Å)	255.25
Temperature (K)	100
Wavelength (Å)	1.134
Resolution (Å)	35–2.44 (2.53–2.44)
R_{merge}† (%)	9.9 (40.1)
Completeness (%)	99.8 (98.7)
$\langle I\rangle/\langle\sigma(I)\rangle$	20.7 (6.06)
Unique reflections	42164 (4091)
Redundancy	14.4 (6.1)
Mosaicity (°)	0.25
B factor from Wilson plot (Å^2)	35.4
Refinement statistics	
Resolution (Å)	20–2.44
Total No. reflections	39912
No. reflections in working set	37891
R factor (%)	16.6
No. reflections in test set	2021
R_{free} (%)	21.7
Protein atoms	7011
Glycerol atoms	18
Phosphate atoms	10
Water atoms	247
Geometry statistics	
R.m.s.d. bond distance (Å)	0.013
R.m.s.d. bond angle (°)	1.5
R.m.s.d. B (Å^2)	
Main-chain bonded atoms	1.1
Side-chain bonded atoms	2.1
Average B (Å^2)	
Main-chain atoms	16.5
Side-chain atoms	20.1
Solvent molecules	32.7
Ramachandran plot‡	
Most favoured region (%)	94.6
Additionally allowed regions (%)	5.1
Generously allowed regions (%)	0.2
Disallowed regions (%)	0.0

† $R_{merge} = \sum_h \sum_i |\langle I\rangle_h - I_{h,i}| / \sum_h \sum_i I_{h,i}$, where $\langle I\rangle_h$ is the average intensity for reflection I_h and $I_{h,i}$ is the intensity of an individual measurement. ‡ As defined with *PROCHECK* (Laskowski *et al.*, 1993).

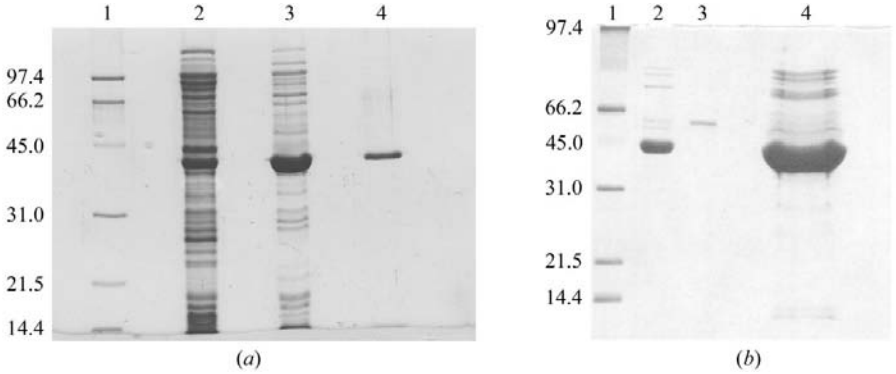

Figure 2
(*a*) Purity of α-methylacyl-CoA racemase at different stages of its purification. Lane 1, molecular-weight standards (kDa); lane 2, 60% ammonium sulfate-precipitated sample dissolved; lane 3, sample after DEAE column; lane 4, sample after octyl Sepharose column. (*b*) Analysis of the crystallization drop. Lane 1, molecular-weight standards (kDa); lane 2, protein solution used for crystallization; lane 3, dissolved bipyramidal crystals; lane 4, crystallization drop after removal of bipyramidal crystals.

tallization drops were analyzed by SDS–gel electrophoresis. For this, the crystals were first transferred to fresh mother liquor and washed several times; the washed crystals were dissolved in water. From this SDS–gel electrophoretic analysis, it was found that the bipyramidal crystals were not of α-methylacyl-CoA racemase but of a very minor contaminant (Fig. 2), while the intact racemase was still present in the mother liquor. A subsequent analysis of the dissolved crystalline protein by mass spectrometry showed it to be *E. coli* ASL. Under the employed crystallization conditions the protein concentration was approximately 24 mg ml^{-1}. Assuming that the impurity was present at approximately 2%, it seems that the eASL concentration was less than 0.5 mg ml^{-1}. Apparently, the high solubility of the racemase and the low solubility of eASL under the employed conditions favoured the crystallization of eASL. In fact, a minor crystal form present as small needles under these conditions (Bhaumik *et al.*, 2003) was shown to be the crystalline racemase and was used for further structural studies of the racemase (Savolainen, Bhaumick *et al.*, manuscript in preparation).

2.2. Data collection, data processing, structure solution and refinement

Diffraction data were collected from a cryocooled crystal at 100 K using a solution of 1.26 *M* ammonium phosphate pH 7.0 containing 30% glycerol as cryoprotectant. The crystal was briefly soaked in the cryoprotectant solution and subsequently flash-cooled by rapidly moving it from the cryo-solution into a cold nitrogen stream. A data set was collected by the rotation method with 0.5° rotations per frame at a wavelength of 1.134 Å at beamline I711, MAX-lab, Lund, Sweden. The image frames of this data set were indexed and integrated using the programs *DENZO*, *SCALEPACK* (Otwinowski & Minor, 1997) and the *CCP*4 package (Collaborative Computational Project, Number 4, 1994). The data-processing statistics are shown in Table 1.

The V_M coefficient (Matthews, 1968) calculation (V_M = 3.1 Å^3 Da^{-1}) indicates the presence of two subunits in the asymmetric unit. The structure was solved by molecular replacement using the program *CNS* (Brünger *et al.*, 1998) with the coordinates of the mutant of duck δ2 crystallin (PDB code 1k7w; Sampaleanu *et al.*, 2002) as a search model as it is the highest resolution (1.96 Å) structure currently available. The molecular-replacement calculations were performed with the *A* subunit of the 1k7w model using space group $P6_122$. The orientations and positions of two subunits were found. After rigid-body refinement, simulated-annealing refinement was performed with *CNS*, followed by a *REFMAC*5 refinement (Murshudov *et al.*, 1997) and a density-modification cycle with the program *DM* (Cowtan, 1994). Finally, an initial model was built using the program *RESOLVE* (Terwilliger, 2000). After another round of *REFMAC*5 refinement, a second automated model-building step was performed using the program *MAID* (Levitt, 2001). From this model (70% complete), the complete model was built by visual inspection of the electron-density map using the graphics package *O* (Jones *et al.*, 1991). In the structure, only two N-terminal residues from the *A* subunit and six N-terminal residues from the *B* subunit are missing.

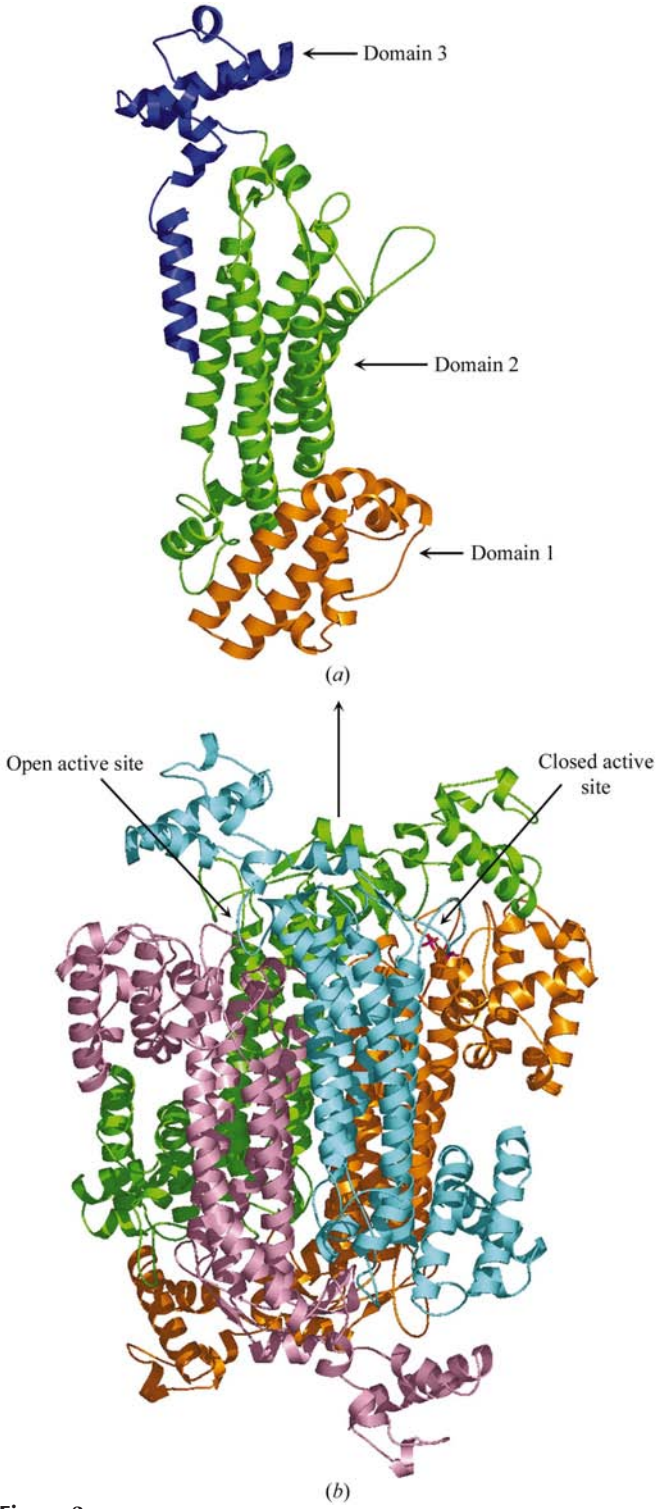

Figure 3
(*a*) The eASL subunit with its three domains: domain 1 (the N-terminal domain), domain 2, and domain 3 (the C-terminal domain). (*b*) The eASL tetramer. The NCS twofold axis (marked by an arrow) runs vertically. Monomers *A*, *B*, *C* and *D* are shown in orange, cyan, green and pink, respectively. The crosses in the closed active site mark the phosphate positions.

These residues could not be built because the electron-density maps of these regions were featureless. The completed model was further refined with *REFMAC*5, using the TLS mode to describe the overall anisotropy of the molecule. After each refinement step, σ_A-weighted $2F_o - F_c$ and $F_o - F_c$ electron-density maps (Read, 1986) were computed and corrections to the model were made with the program *O*. Water molecules with proper hydrogen-bonding coordination were progressively introduced while monitoring the decrease in R_{free}. At the end of the refinement, three glycerol molecules and two phosphate ions were included in the model. *PROCHECK* (Laskowski *et al.*, 1993) was used to analyze the stereochemistry of the model. The final refinement statistics are presented in Table 1.

3. Results and discussion

3.1. Overall structure

The structure of eASL has been refined at 2.44 Å resolution with good refinement statistics (Table 1). In this crystal form there are two subunits per asymmetric unit (Fig. 3). For subunit *A* residues 3–457 could be built, while for subunit *B* the structures of residues 7–457 are defined in the electron-density map. The N-terminal region of subunit *A* is better ordered than that of subunit *B* and has a different structure. Like all other enzymes in the β-elimination superfamily, the eASL subunit consists of three domains. Domain 1 includes residues 3–105, domain 2 residues 106–363 and domain 3 residues 364–457. Domain 2 is elongated and is essentially a bundle of five long helices. Domains 1 and 3 are more compact spherical domains.

The NCS twofold axis for the two subunits of the asymmetric unit intersects with a crystallographic twofold axis (Fig. 3). The intersection of these twofold axes generates the tetrameric molecule. The core of the tetramer is formed by the interactions between the helical bundles of domain 2 of each of the four subunits (Fig. 3). The N-terminal and C-terminal domains protrude out of the core (at opposing ends), giving the tetramer its peculiar shape (Fig. 3). The active site is built from regions of three different subunits (Figs. 3 and 4) such that there are four active sites per tetramer. At each end of the tetrameric helical bundle there are two active sites. In the structure of eASL these two active sites are related by the NCS twofold axis (Figs. 3 and 4). As can be seen in Figs. 3 and 4, the two active sites are dissimilar. One active site is in an open conformation and unliganded, while the other active site is in a closed conformation with two phosphate ions bound. The highly conserved peptide regions C1, C2 and C3 are in loops near the active site as shown in Fig. 4. These regions protrude into the active site from different subunits. The active site complexed with the phosphate ions is formed by C1 residues 106–116 (just after domain 1 of subunit *A*), C2 residues 153–163 (protruding out of domain 2 of subunit *C*) and C3 residues 273–291 (protruding out of domain 2 of subunit *B*). In addition, this active site is shaped by the helical domain of subunit *B*, the N-terminal domain (domain 1) of subunit *A* and the C-terminal domain (domain 3) of subunit *C*, as shown in Fig. 4.

From comparison of the open and closed conformations of the active site, it can be seen that the N-terminal region (residues 3–15), the C-terminal domain and the C3 loop change their conformations. In the closed conformation, the C3 loop and domain 3 move towards the N-terminal domain and the N-terminal residues 3–15 adopt a new conformation in which they interact with the closed C3 loop and the closed C-terminal domain. The C3 loop of the open active site could be built completely (Fig. 4) but has relatively high *B* factors, while the C3 loop of the closed active site is very well

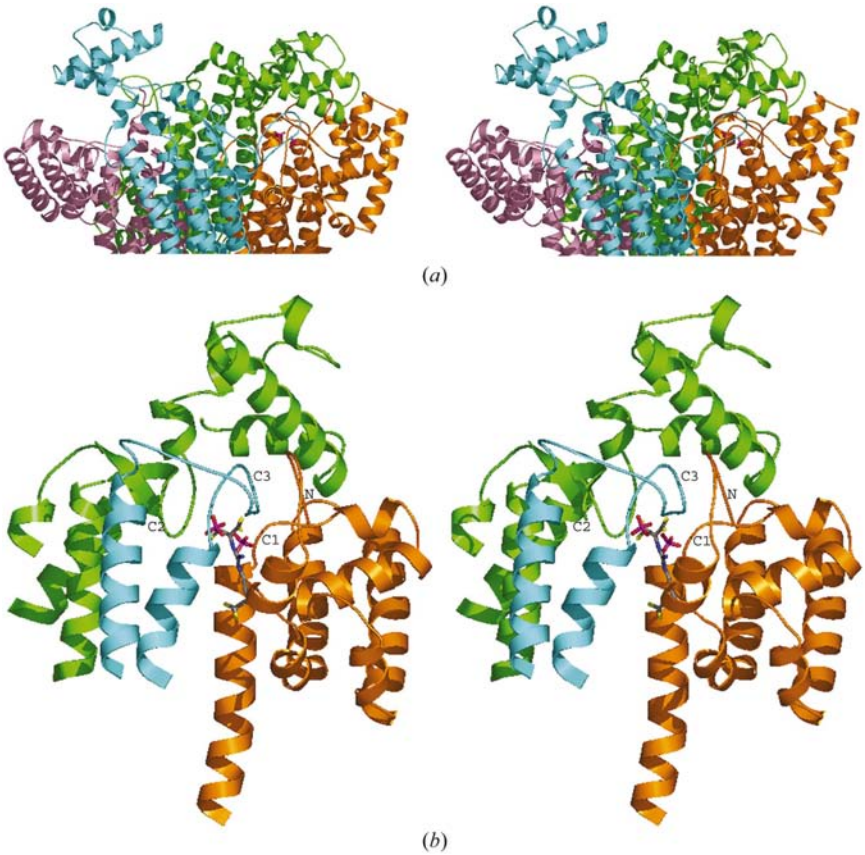

Figure 4
(*a*) A stereoview of the top half of the eASL tetramer showing the two active sites related by the non-crystallographic twofold axis. The colouring code is the same as that in Fig. 3(*b*). The crosses mark the phosphate sites in the closed active site. (*b*) Zoomed-in picture of the closed active site showing phosphate ions and argininosuccinate (grey) as observed in the structure of S283A duck δ2 crystallin (Sampaleanu *et al.*, 2002). The three active-site regions, C1 of subunit *A* (orange), C2 of subunit *C* (green) and C3 of subunit *B* (cyan) are labelled. The N-terminus of subunit *A* (orange), seen near the C1 region of the same subunit, is labelled N.

Table 2
Distances (within 3.5 Å) between the two phosphate ions and neighbouring active-site residues.

Phosphate	eASL residue	Distance (Å)
O_1 (phos-1)	His156*C* $N^{\varepsilon 2}$	2.9
O_2 (phos-1)	Lys283*B* N^{ζ}	2.6
O_2 (phos-1)	Asn285*B* $N^{\delta 2}$	3.1
O_3 (phos-1)	Asn110*A* $N^{\delta 2}$	2.8
O_3 (phos-1)	Thr155*C* $O^{\gamma 1}$	2.6
O_3 (phos-1)	His156*C* $N^{\varepsilon 2}$	3.3
O_4 (phos-1)	Ser277*B* O^{γ}	2.5
O_1 (phos-2)	Asn110*A* $O^{\delta 1}$	2.7
O_2 (phos-2)	Arg109*A* $N^{\eta 2}$	2.8
O_2 (phos-2)	Ser278*B* N	2.9
O_2 (phos-2)	Arg109*A* N^{ε}	3.4
O_3 (phos-2)	Ser108*A* O^{γ}	3.1
O_3 (phos-2)	Ser277*B* O^{γ}	3.0
O_4 (phos-2)	Arg109*A* N	2.6
O_4 (phos-2)	Arg109*A* N^{ε}	3.1
O_4 (phos-2)	Ser278*B* O^{γ}	2.3

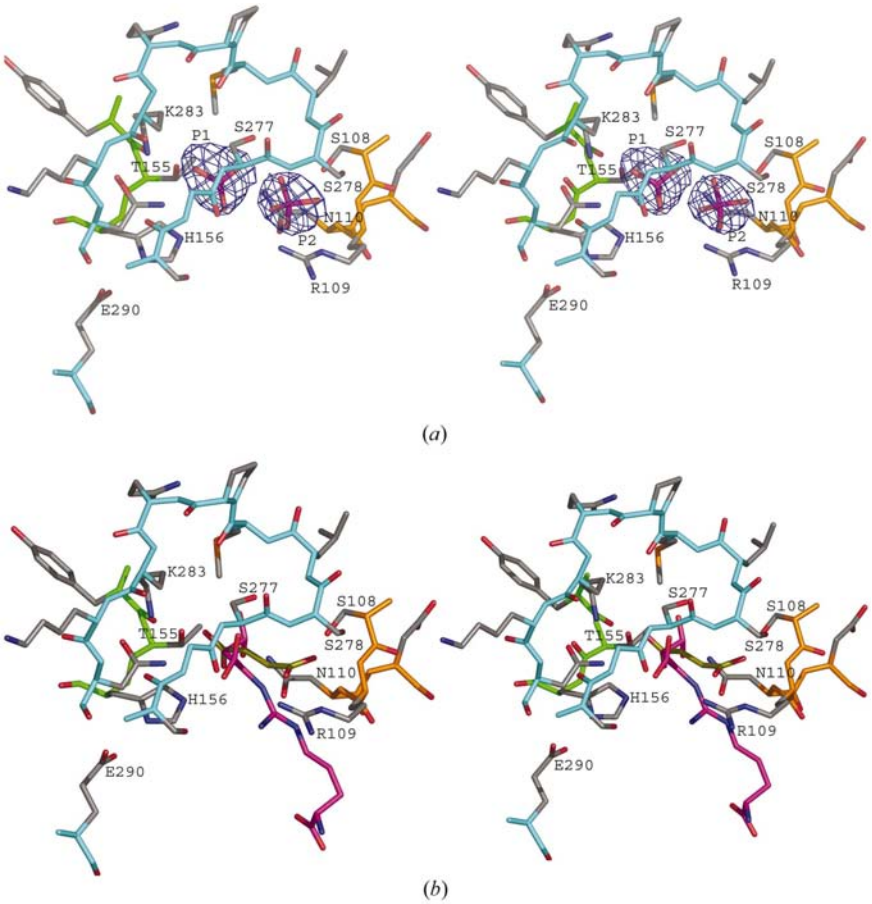

Figure 5
(*a*) The active-site region of eASL where the two phosphate ions are bound. The three different catalytic active-site regions are shown in different colours (C1 in orange, C2 in green and C3 in cyan). The electron density is from the σ_A-weighted $F_o - F_c$ map (Read, 1986), calculated from the model before any ligand had been included in the model. P1 and P2 refer to phos-1 and phos-2, respectively. The protein part includes residues 108–111 (C1 region), 154–156 (C2 region) and 275–285 (C3 region), as well as Glu290. (*b*) The active-site region of eASL with superimposed argininosuccinate, as observed in the structure of S283A duck δ2 crystallin (Sampaleanu *et al.*, 2002). Also shown is the proposed mode of binding of the succinate moiety in a competent active site (see text). The three different catalytic active-site regions are shown in different colours (C1 in orange, C2 in green and C3 in cyan).

defined and has low *B* factors. The N-terminal residues 3–15 are also much better defined in the closed active-site structure. This closed active-site arrangement has been seen in only one other ASL structure: the complex of the inactive duck δ1 crystallin with sulfate. Only one sulfate was present in this active site and in this sulfate-complexed structure the N-terminal region was disordered.

3.2. The closed liganded active site

The two phosphates bound in the closed active site will be referred to as phos-1 and phos-2. The phos-1 site is occupied by the sulfate ion in the structure of the inactive duck δ1 crystallin (data not shown). The phos-1 site is near the C2 region. The phos-2 site is near the C1 region. The C3 loop is near both phosphates (Fig. 5). Initially, it was suspected that a copurified succinate was possibly present in this closed active site; however, the refinement of a succinate did not properly describe the shape and size of the electron-density map, while two phosphates provided a good description of the maps and resulted in clean $F_o - F_c$ difference maps. Fig. 5 shows the difference map of the region of the catalytic site before the model included any active-site ligand; superimposed on the difference density is shown the final model of the two phosphate ions. The presence of the two phosphate ions correlates with the high phosphate concentration in the mother liquor. The presence of negatively charged ions at these positions is also in good agreement with the contacting side chains, as summarized in Table 2. For example, phos-1 and phos-2 make salt-bridge interactions with Lys283 and Arg109, respectively. Moreover, each phosphate O atom is hydrogen bonded to at least one side-chain hydrogen-bonding partner. The phosphates are not hydrogen bonded to any ordered water molecules.

3.3. Implications for the reaction mechanism

There is much structural and enzymological data available on duck δ2 crystallin and its mutated variants (Sampaleanu *et al.*, 2002; Wu *et al.*, 1998). Nevertheless, a precise description of its reaction mechanism has been elusive owing to the absence of a crystal structure of an active metabolic ASL having a liganded competent active site. From mutagenesis data of duck δ2 crystallin (and from sequence conservation) it is clear that the residues crucial for full catalytic activity are His162 and Ser283 (duck δ2 crystallin sequence numbering; Sampaleanu *et al.*, 2002). From

the available data it has been hypothesized that His162 is possibly the catalytic base and Ser283 is possibly the catalytic acid. However, His162 is not fully conserved in the superfamily and the possible role of Ser283 (which is fully conserved) is uncertain owing to the absence of a structure of the closed competent active site. Consequently, this structure of the complexed closed active site of eASL provides new information on the possible catalytic role of this histidine and this serine. For eASL the equivalent residues are His156 and Ser277. A possible mode of binding of the succinate moiety of the substrate argininosuccinate is shown in Fig. 5. This proposed mode of binding is obtained by superposing the carboxylate moieties of *trans*-succinate on the two phosphate ions. Such a superposition is achieved in a straightforward manner as the P—P distance between the two phosphate ions (3.9 Å) is very similar to the C1—C4 distance of the C atoms of the carboxylic acid moieties of succinate in its all-*trans* geometry (3.7 Å). The mode of binding of argininosuccinate in the open form has been observed in the structures of the inactive variants of duck $\delta2$ crystallin S283A (Sampaleanu *et al.*, 2002) and H162N (Vallee *et al.*, 1999). From the comparison of the observed mode of binding of argininosuccinate to the open active site and the predicted mode of binding of succinate to the closed active site in eASL it can be seen that the succinate moiety of argininosuccinate rotates into a new position when the active site closes. This movement of the succinate is required in order to avoid clashes with the closed C3 loop. This rearrangement of the succinate moiety is achieved by enforcing a *trans* conformation for the succinate moiety and a major rotation about the $C^{\zeta}-N^{\eta1}$ bond of the arginine moiety, followed by further rearrangements of the whole molecule. As a result the arginine moiety essentially does not move, except for the guanidino moiety which moves by approximately 1 Å to the His156 side chain. With respect to the succinate moiety, the phos-1 carboxylic acid moiety remains bound in the same pocket, but the phos-2 carboxylic acid moiety moves so that it becomes hydrogen bonded to the side chains of the C1 region. In the proposed orientation of the succinate (Fig. 5) the arginine moiety of the substrate points downward into its corresponding binding pocket and the pro-*R* hydrogen of the C^{β} atom points to the side-chain O^{γ} atom of Ser277 (the distance between succinate C^{β} and Ser277 O^{γ} is 3.0 Å). Kinetic data using a labelled substrate has shown that the pro-*R* hydrogen is abstracted by the enzyme (Wu *et al.*, 1998). This suggests that the Ser277 side chain could initiate the catalysis by abstracting a proton from the C^{β} atom. The Ser277 side chain thus becomes the base; this role will be facilitated by the hydrogen bonding between Ser277 O^{γ} and the carboxylate group of the succinate moiety (Fig. 5). In addition, the nucleophilicity of the Ser277 O^{γ} atom is further enhanced by the oxyanion-hole-like environment made by the peptide NH groups of residues 279 and 280 (Fig. 5). A similar role has been suggested recently for the corresponding serine in aspartase (Fujii *et al.*, 2003). In the proposed mode of binding the His156 side chain points to the guanidino group of the arginine moiety, such that it would be possible to protonate this moiety when it becomes the leaving group in the second half of the reaction. This proposal requires that His156 is doubly protonated, both at $N^{\delta1}$ and at $N^{\varepsilon2}$. In the phosphate complex this indeed is the case, as $N^{\delta1}$ is hydrogen bonded to Glu290 (Fig. 5) and $N^{\varepsilon2}$ to phos-1 (Table 1). Also, in the modelled complex with the succinate moiety His156 is doubly protonated: His156 $N^{\delta1}$ is hydrogen bonded to Glu290 (Fig. 5) and His156 $N^{\varepsilon2}$ is hydrogen bonded to the phos-1 carboxylic acid moiety. The sequence conservation of Glu290 is highly correlated with the sequence conservation of the histidine (Sampaleanu *et al.*, 2002). His156 is conserved as a histidine in the sequences of ASL, adenylosuccinate lyase and fumarase, but not in some other members of this superfamily, such as aspartase (Sampaleanu *et al.*, 2002). At least two possible explanations might rationalize the non-conservation of this histidine in this superfamily. (i) 'Hopping' of essential catalytic residues between catalytic loops of homologous enzymes of a superfamily has been observed (Hasson *et al.*, 1998; Todd *et al.*, 2002); in this case the functionality of His156 is replaced by another histidine in another loop. (ii) The leaving group is protonated by another residue or by a water molecule, for example after a rearrangement of the active site following the first step of the reaction. For the aspartase reaction, it is currently not clear whether the leaving group is ammonia or an ammonium ion (Fujii *et al.*, 2003). Summarizing, the proposed mechanism, in which Ser277 acts as a catalytic base, is consistent with the available enzymological data of the wild-type and mutated $\delta2$ crystallins as well as with the known sequence conservations of the C1, C2 and C3 regions of the superfamily.

We would like to thank Ville Ratas for his excellent technical support for this work. We would also like to thank the staff at MAX-lab, Lund for their help during data collection at beamline I711.

References

Abu-Abed, M., Turner, M. A., Vallee, F., Simpson, A., Slingsby, C. & Howell, P. L. (1997). *Biochemistry*, **36**, 14012–14022.
Bhaumik, P., Kursula, P., Ratas, V., Conzelmann, E., Hiltunen, K., Schmitz, W. & Wierenga, R. K. (2003). *Acta Cryst.* D**59**, 353–355.
Blattner, F. R., Burland, V., Plunkett, G. III, Sofia, H. J. & Daniels, D. L. (1993). *Nucleic Acids Res.* **21**, 5408–5417.
Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, N., 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.
Cowtan, K. (1994). *Jnt CCP4/ESF–EAMCB Newsl. Protein Crystallogr.* **31**, 34–38.
Fujii, T., Sakai, H., Kawata, Y. & Hata, Y. (2003). *J. Mol. Biol.* **328**, 635–654.
Hasson, M. S., Schlichting, I., Moulai, J., Taylor, K., Barrett, W., Kenyon, G. L., Babbitt, P. C., Gerlt, J. A., Petsko, G. A. & Ringe, D. (1998). *Proc. Natl Acad. Sci. USA*, **95**, 10396–10401.
Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.
Laskowski, R. A., MacArthour, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Levitt, D. G. (2001). *Acta Cryst.* D**57**, 1013–1019.

Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
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.
Read, R. J. (1986). *Acta Cryst.* A**42**, 140–149.
Sampaleanu, L. M., Vallee, F., Slingsby, C. & Howell, P. L. (2001). *Biochemistry*, **40**, 2732–2742.
Sampaleanu, L. M., Vallee, F., Thompson, G. D. & Howell, P. L. (2001). *Biochemistry*, **40**, 15570–15580.
Sampaleanu, L. M., Yu, B. & Howell, P. L. (2002). *J. Biol. Chem.* **277**, 4166–4175.
Simpson, A., Bateman, O., Driessen, H., Lindley, P., Moss, D., Mylvaganam, S., Narebor, E. & Slingsby, C. (1994). *Nature Sruct. Biol.* **1**, 724–734.
Terwilliger, T. C. (2000). *Acta Cryst.* D**56**, 965–972.
Todd, A. E., Orengo, C. A. & Thornton, J. M. (2002). *Trends Biochem. Sci.* **27**, 419–426.
Toth, E. A. & Yeates, T. O. (2000). *Structure*, **8**, 163–174.
Turner, M. A., Simpson, A., McInnes, R. R. & Howell, P. L. (1997). *Proc. Natl Acad. Sci. USA*, **94**, 9063–9068.
Vallee, F., Turner, M., Lindley, P. & Howell, P. (1999). *Biochemistry*, **38**, 2425–2434.
Weaver, T. & Banaszak, L. (1996). *Biochemistry*, **35**, 13955–13965.
Wu, C., Lee, H., Wu, S., Chen, S., Chiou, S. & Chang, G. (1998). *Biochem. J.* **333**, 327–334.