STRUCTURAL BIOLOGY
COMMUNICATIONS

ISSN: 2053-230X
journals.iucr.org/f

Serendipitous crystallization and structure determination of cyanase (CynS) from *Serratia proteamaculans*

Agata Butryn, Gabriele Stoehr, Christian Linke-Winnebeck and Karl-Peter Hopfner

Acta Cryst. (2015). F**71**, 471–476

STRUCTURAL BIOLOGY COMMUNICATIONS

ISSN 2053-230X

Serendipitous crystallization and structure determination of cyanase (CynS) from *Serratia proteamaculans*

Agata Butryn, Gabriele Stoehr, Christian Linke-Winnebeck and Karl-Peter Hopfner*

Gene Center and Department of Biochemistry, Ludwig-Maximilians-University Munich, Feodor-Lynen-Strasse 25, 81377 Munich, Germany. *Correspondence e-mail: hopfner@genzentrum.lmu.de

Received 18 February 2015
Accepted 10 March 2015

Edited by W. N. Hunter, University of Dundee, Scotland

Keywords: CynS; cyanase; cyanate hydratase.

PDB reference: cyanase, 4y42

Supporting information: this article has supporting information at journals.iucr.org/f

Cyanate hydratase (CynS) catalyzes the decomposition of cyanate and bicarbonate into ammonia and carbon dioxide. Here, the serendipitous crystallization of CynS from *Serratia proteamaculans* (SpCynS) is reported. SpCynS was crystallized as an impurity and its identity was determined using mass-spectrometric analysis. The crystals belonged to space group *P*1 and diffracted to 2.1 Å resolution. The overall structure of SpCynS is very similar to a previously determined structure of CynS from *Escherichia coli*. Density for a ligand bound to the SpCynS active site was observed, but could not be unambiguously identified. Additionally, glycerol molecules bound at the entry to the active site of the enzyme indicate conserved residues that might be important for the trafficking of substrates and products.

1. Introduction

Cyanate hydratase (CynS; EC 4.2.1.104) decomposes cyanate according to the reaction $NCO^- + HCO_3^- + 2H^+ \rightarrow NH_3 + 2CO_2$. It is an inducible enzyme found in bacteria (Johnson & Anderson, 1987; Sung *et al.*, 1987; Walsh *et al.*, 2000), fungi (Elleuche & Pöggeler, 2008) and plants (Qian *et al.*, 2011). In bacteria, CynS is part of the *cyn* operon, which also encodes the regulator CynR, the carbonic anhydrase CynT and the cyanate transporter CynX. Cyanate is a toxic compound, but some bacteria such as *Escherichia coli* are also able to use cyanate as the sole source of nitrogen through the action of CynS (Guilloton & Karst, 1987; Johnson & Anderson, 1987).

The crystal structure of *E. coli* CynS (EcCynS) revealed that this protein assembles into a functional decamer composed of ten protomers organized into five identical dimers (Walsh *et al.*, 2000). Each protomer comprises an N-terminal α-helix bundle, which structurally resembles the DNA-binding domain of helix–turn–helix transcriptional regulators, and a C-terminal catalytic cyanase domain of a unique fold. The dimer is formed through an intricate interaction of two C-terminal cyanase domains.

Analogues of the EcCynS substrates or its intermediate products such as low-molecular-weight dicarboxylic acids and monoanions have previously been shown to inhibit EcCynS by binding to the adjacent cyanate and bicarbonate binding sites (Anderson *et al.*, 1987). In the EcCynS co-structures, oxalate and chloride bind to a site formed by residues from four EcCynS protomers (Walsh *et al.*, 2000). These highly conserved residues, Arg96, Glu99 and Ser122, have been

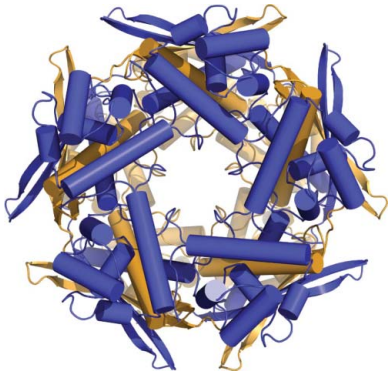

Table 1
Data-collection and refinement statistics.

Values in parentheses are for the outer shell.

Data collection	
Diffraction source	X06SA, SLS
Wavelength (Å)	1.00
Detector	PILATUS 6M-F
Space group	*P*1
a, *b*, *c* (Å)	82.55, 82.46, 87.98
α, β, γ (°)	98.56, 117.38, 110.46
Mosaicity (°)	0.182
Solvent content (%)	53.4
Molecules per asymmetric unit	10
Matthews coefficient (Å^3 Da^{-1})	2.64
Resolution range (Å)	47.05–2.09 (2.22–2.09)
Total No. of reflections	177433 (26673)
No. of unique reflections	98524 (14907)
Completeness (%)	92.5 (86.8†)
Multiplicity	1.8 (1.8)
$\langle I/\sigma(I)\rangle$	11.4 (2.3)
$R_{\rm meas}$	0.05 (0.32)
Overall *B* factor from Wilson plot (Å^2)	28.6
Refinement	
Resolution range (Å)	47.05–2.09 (2.11–2.09)
No. of reflections	98486
Final $R_{\rm work}$	0.15 (0.25)
Final $R_{\rm free}$	0.20 (0.33)
No. of non-H atoms	
Protein	11975
Ligand	24
Water	1414
R.m.s. deviations	
Bonds (Å)	0.008
Angles (°)	1.1
Average *B* factors (Å^2)	
Protein	30.1
Ligand	50.3
Water	40.6
Ramachandran plot	
Favoured regions (%)	99.0
Additionally allowed (%)	1.0
Outliers (%)	0.0

† Owing to low crystal symmetry and data redundancy.

proposed to constitute the active centre of CynS, but have not yet been tested experimentally for their actual role in cyanate decomposition.

Here, the crystal structure of a CynS homologue from *Serratia proteamaculans* (SpCynS) is reported. This structure is the result of the serendipitous crystallization of SpCynS during crystallization screens for an unrelated eukaryotic complex and extensive analyses of these crystals.

2. Materials and methods

2.1. Protein purification and crystallization

SpCynS was inadvertently copurified with a eukaryotic multiprotein–DNA complex. This complex was assembled from components expressed in both *E. coli* Rosetta DE3 (Merck Millipore) and *Trichoplusia ni* High Five cells (Life Technologies). In brief, each of the N-terminally His_6-tagged complex subunits were separately expressed and purified using Ni^{2+}–NTA agarose (Qiagen). The proteins were further purified using ion-exchange chromatography (HiTrap SP HP or HiTrap Q HP 5 ml columns, GE Healthcare) and size-exclusion chromatography (HiLoad 16/60 Superdex 75 or 200 pg columns, GE Healthcare). Finally, the DNA–protein complex was reconstituted and isolated by size-exclusion chromatography (Superdex 200 10/300 GL column, GE Healthcare) using SEC buffer (20 m*M* MES potassium salt pH 6.5, 60 m*M* KCl, 5 m*M* $MgCl_2$, 2 m*M* DTT). The DNA–protein complex eluted as a single symmetric peak and was >90% pure as estimated by SDS–PAGE analysis.

The DNA–protein complex was screened for crystallization using several commercial screens. Using a Phoenix Crystallization robot (Art Robbins), 500 nl of the sample at a protein concentration of 5 mg ml^{-1} in SEC buffer was mixed with 500 nl reservoir solution and equilibrated against 55 µl reservoir solution in a 96-well sitting-drop format (MRC Crystallization Plate, Molecular Dimensions). After four months at 20°C, a few thin, plate-like crystals of rhomboidal shape and with dimensions of ∼150 × 150 µm were observed in two different conditions and were used directly for X-ray data collection. The best data set used during refinement was collected from a crystal grown in 0.1 *M* HEPES pH 7.5, 2%(*v*/*v*) 2-propanol, 0.1 *M* sodium acetate, 18%(*w*/*v*) PEG 8000 (JBScreen Classic HTS I, Jena Bioscience).

2.2. Data collection and processing

Crystals were soaked in cryoprotecting solution consisting of mother liquor containing 25%(*v*/*v*) glycerol, mounted using a CryoLoop (Hampton Research) and flash-cooled in liquid nitrogen. A small number of crystals were tested for X-ray diffraction on the X06SA beamline at the Swiss Light Source, Villigen, Switzerland. A 180° data set was collected at a wavelength of 1.00 Å and −173°C with 0.25° per image (exposure time 1 s per 0.25°). The data set was indexed, integrated and scaled in space group *P*1 to 2.1 Å resolution using *XDS* (Kabsch, 2010). Data-collection statistics are listed in Table 1. The self-rotation function was analyzed using *MOLREP* (Supplementary Fig. S1*a*; Vagin & Teplyakov, 2010).

2.3. Analysis of the crystal content

To identify the crystal content, protein was recovered from several crystallization drops from the original crystallization plate. Protein samples were analyzed by SDS–PAGE (Supplementary Fig. S1*b*), which revealed that the original components were degraded except for a species migrating at 14 kDa. The band was excised, subjected to tryptic in-gel digestion and analyzed by liquid chromatography–tandem mass spectrometry (LC-MS/MS) using an EASY-nLC 1000 coupled online to an LTQ-Orbitrap Elite system (both from Thermo Fisher Scientific). `*.raw` data were analyzed using the *MaxQuant* software suite (v.1.5.0.0; Cox & Mann, 2008). An Andromeda database search (Cox *et al.*, 2011) was performed against the proteome databases of *E. coli*, *Spodoptera frugiperda*, *Bombyx mori* and *Drosophila melanogaster* sequences as available in UniProt and against the sequences of the recombinantly expressed proteins. In the second analysis, the

data were re-analyzed with *MaxQuant* applying a database search against the *E. coli* and *S. proteamaculans* CynS sequences.

2.4. Structure solution and refinement

No molecular-replacement solution was found using the target DNA–protein complex or its components as a search model. The EcCynS crystal structure (PDB entry 1dw9; Walsh *et al.*, 2000) was used as a search model for molecular replacement using *Phaser* (McCoy *et al.*, 2007). The SpCynS model was refined in iterative cycles of manual building using *Coot* (Emsley *et al.*, 2010) and maximum-likelihood refinement using *PHENIX* applying NCS restraints (Adams *et al.*, 2010). Refinement statistics are reported in Table 1. The coordinates of the final model were deposited in the Protein Data Bank with accession code 4y42.

Figures were created using the *PyMOL* molecular-graphics system (v.1.5.0.2; Schrödinger).

3. Results and discussion

3.1. Crystallization and structure solution

Crystallization screens for a eukaryotic protein–DNA complex, the components of which had been expressed in both *E. coli* and *T. ni* High Five cells, resulted in well diffracting crystals after an incubation time of four months at 20°C. However, no molecular-replacement solutions (as judged from *Z*-scores) with search models derived from previously determined parts of the protein complex could be found. The analysis of the self-rotation function indicated the presence of five twofold and one perpendicular fivefold symmetry axes (Supplementary Fig. S1*a*), although neither the complex nor its components had been expected to form pentamers. When the content of the protein crystallization drops was analyzed, it became clear that all of the target protein had degraded over the four months of incubation time. Only one prominent species running at an apparent size of 14 kDa was observed (Supplementary Fig. S1*b*). The identification of the source of this band was accomplished using in-gel digestion in combination with LC-MS/MS analysis searching against the sequences of the complex components and the proteomes from *E. coli* and, in the absence of genomic data for *T. ni*, from lepidopterans. The most intense hit was for *E. coli* cyanate hydratase (EcCynS), with three unique peptides covering 18% of the EcCynS sequence. EcCynS forms pentamers of homodimers (Walsh *et al.*, 2000), which is consistent with the fivefold symmetry axis observed in the self-rotation function. Indeed, using the EcCynS structure with PDB code 1dw9 as search model, a convincing molecular-replacement solution was found.

However, close inspection of the resulting electron-density map revealed that the density of many side chains, particularly in the N-terminal region, did not correlate well with the sequence of the EcCynS search model. Therefore, the structure was initially built as a polyalanine model. Side chains were assigned based on the $2F_o - F_c$ and $F_o - F_c$ electron density, which for the most part was quite unambiguous. The sequence derived from this procedure was used to consult the NCBI protein database using *BLAST* (Altschul *et al.*, 1997) and was unambiguously identified as CynS from *S. proteamaculans* (SpCynS). To test the sequence assignment, the mass-spectrometric data for the gel band were re-analyzed against the *S. proteamaculans* CynS sequence. Indeed, 15 peptides (13 unique) were identified, which in total covered 98% of the SpCynS sequence (Supplementary Fig. S1*c* and Supplementary Table S1). Thus, the crystal structure was further refined using the sequence of SpCynS.

Although the exact source of the SpCynS protein is unclear, it seems likely that this contaminant (homodecamer of 172 kDa) was initially copurified with the 144 kDa complex subunit expressed in High Five cells and co-eluted with the reconstituted DNA–protein complex (approximately 220 kDa) during the final size-exclusion chromatography step. *S. proteamaculans* belongs to the Enterobacteriaceae family and has been isolated from soil, plants and animals (Grimont *et al.*, 1981; Bollet *et al.*, 1993; Ashelford *et al.*, 2002). We deem

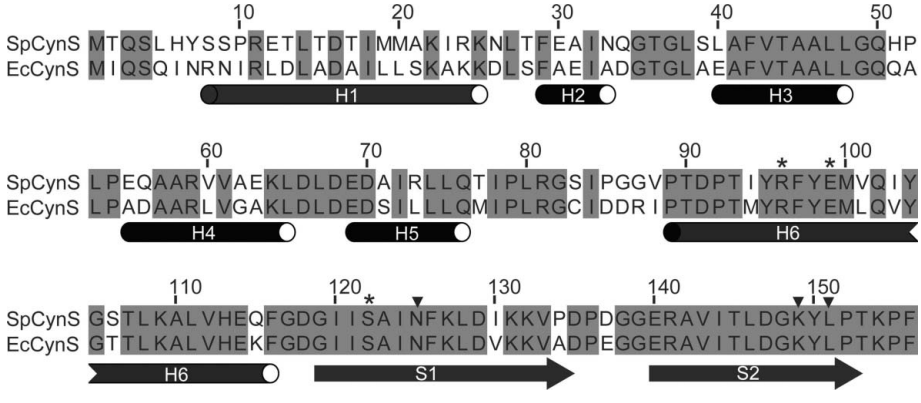

Figure 1
Amino-acid sequence alignment and secondary-structure assignment of *S. proteamaculans* and *E. coli* CynS. Invariant residues are shaded in grey. Key residues localized in the centre of the active site (Arg96, Glu99 and Ser122) are marked with asterisks. Residues potentially involved in the substrate/product trafficking identified in this study (Asn125 and Lys149) and Leu151 are marked with black triangles.

it likely that *S. proteamaculans* may have contaminated our High Five cells in this particular culture rather than the *E. coli* expression cultures, as the former contain only low amounts of selective antibiotics and were incubated for 3 d at 27.5°C.

3.2. Overall structure of SpCynS

The only structurally characterized homologue of SpCynS is EcCynS, which was used for molecular replacement. Both EcCynS and SpCynS protomers consist of 156 amino acids and share 58 and 87% sequence identity in the N-terminal and the C-terminal catalytic domains, respectively (overall sequence identity of 70%; Fig. 1). The root-mean-square deviation between protomers of SpCynS and EcCynS is 0.66 Å over 152 C^{α} atoms (Maiti *et al.*, 2004). SpCynS protomers form dimers through intricate interactions of their C-terminal domains. Both protomers contribute two β-strands each to one continuous antiparallel β-sheet, and the two β-strands of one protomer are packed against a long α-helix from the other protomers within the dimer (Fig. 2*a*). Like EcCynS, SpCynS assembles into a ring-like decamer with 5/2 symmetry of approximate dimensions 85 × 85 × 70 Å (Fig. 2*b*). No larger, substantial differences between the structures of SpCynS and EcCynS could be detected.

3.3. Identification of the ligand at the active site of *S. proteamaculans* CynS

In analogy to EcCynS, the active site of SpCynS is expected to be formed by Arg96, Glu99, Ser122 and their NCS-related equivalents at the interface of two adjacent dimers (Fig. 3*a*). There are no significant changes compared with the active site of EcCynS, which was crystallized in the presence of inhibitors (Fig. 3*b*). However, as clearly defined in $2F_o - F_c$ and $F_o - F_c$ electron-density maps, all five SpCynS active sites are occupied by ligands. Their density does not clearly match any components of the crystallization or sample buffer. These

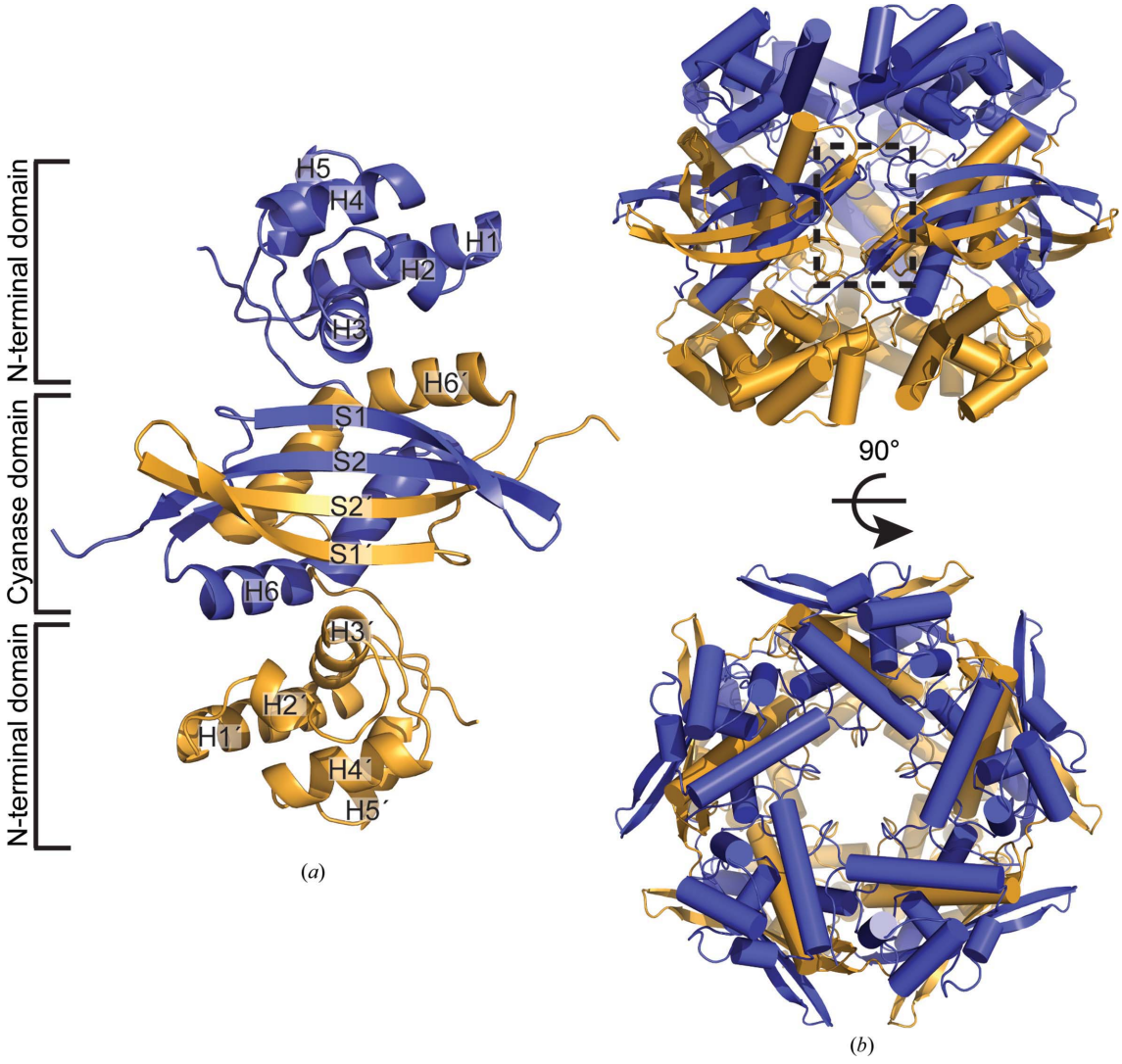

Figure 2
Structural organization of the *S. proteamaculans* CynS decamer. (*a*) The SpCynS dimer with its protomeric units coloured yellow and blue. Secondary-structure elements (H for α-helices and S for β-sheets) are labelled. (*b*) Side and top view of the SpCynS decamer. A dashed rectangle highlights the location of the active site formed at the interface of two adjacent dimers.

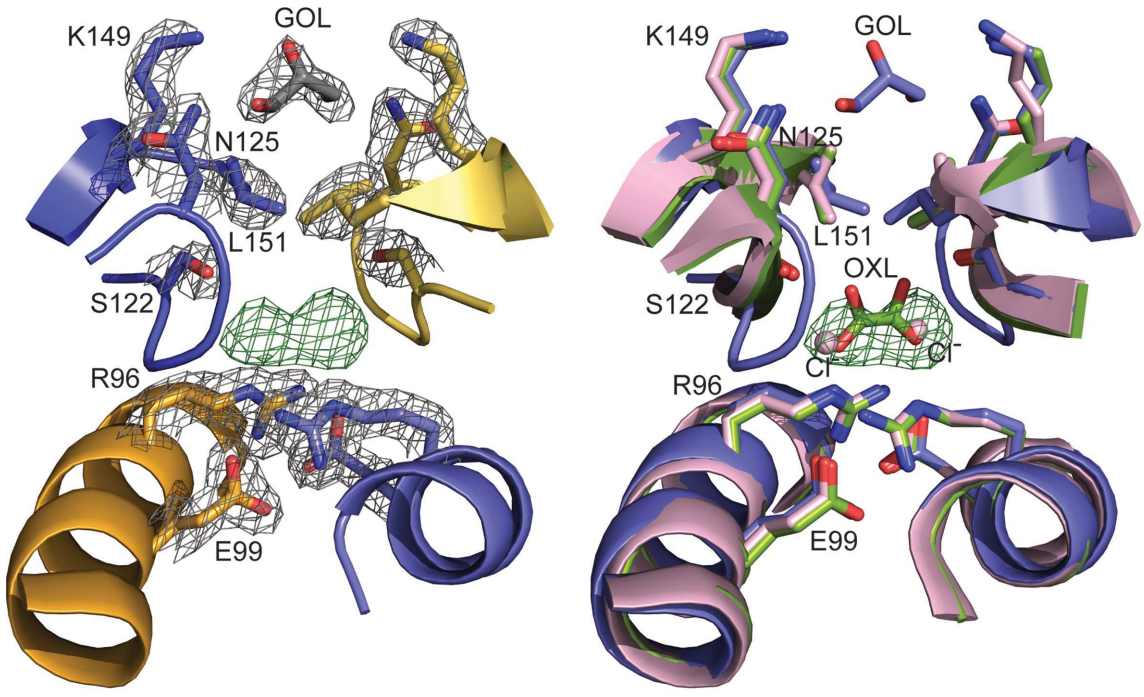

Figure 3
The active site of CynS. (*a*) Organization of the SpCynS catalytic centre formed by the residues belonging to four adjacent protomers. Key residues and the glycerol molecule bound at the entry channel are marked. The $2F_o - F_c$ map is contoured at 1.5σ (grey mesh). $F_o - F_c$ difference density map corresponding to the ligand bound in the active site of SpCynS is contoured at 4σ (green mesh). (*b*) Superimposition of the active site of SpCynS (blue), oxalate (OXL)-bound EcCynS (green; PDB entry 1dwk) and chloride (Cl^-)-bound EcCynS (pink; PDB entry 1dw9).

molecules may be derived from *S. proteamaculans* or the culture medium.

The additional density has twofold symmetry with two peaks >12σ, which suggest the presence of two electron-rich atoms. The positions of these peaks correspond well to two chloride ions found in the EcCynS crystal structure (PDB entry 1dw9; Walsh *et al.*, 2000). As chloride was also present in the crystallization buffer for SpCynS, we tried to model two chlorides per SpCynS active site. Potential interactions between the modelled chlorides and SpCynS agree with expected bond lengths (Davey *et al.*, 2002), but further difference density is observed after their refinement. Thus, two chlorides do not appear to fully explain the observed additional electron density.

The position, the environment and the symmetry of the ligand density would in principle also accommodate dicarboxylates such as oxalate or malonate. Oxalate [$(COO^-)_2$] is present in the active sites of one of the EcCynS crystal structures (PDB entry 1dwk; Walsh *et al.*, 2000). However, oxalate does not seem to fit well within the additional density in SpCynS, with malonate [$C(COO^-)_2$] a more likely candidate. Alternatively, a smaller molecule such as acetate [$CH_2(COO^-)$] bound to either substrate-binding site with half occupancy may not be excluded. Sodium acetate was present in the crystallization buffer, but the additional ligand density in a SpCynS crystal grown in an acetate-free condition harboured the same features (data not shown). Therefore, it is likely that the additional density in each of the active sites is a compilation of the electron density corresponding to two chloride ions (present in the crystallization buffer) and a small organic molecule, all with partial occupancies. However, since the identity of the ligands could not be determined unambiguously, no ligands were included in the final model.

3.4. Gating mechanism of the substrate/product trafficking

EcCynS Leu151, localized in the solvent-exposed site of the active-site cavity, was suggested to provide a gating mechanism for the substrate entry and/or product exit (Walsh *et al.*, 2000). In agreement with this, the $2F_o - F_c$ and $F_o - F_c$ electron-density maps of SpCynS Leu151 point towards the presence of alternative side-chain conformations which may affect the passage of molecules through this channel. Additionally, glycerol molecules are found at the entry of four out of five active-site cavities. Distal glycerol hydroxyl groups are coordinated by two symmetry-related Asn125 residues, whereas the central hydroxyl group is coordinated by Lys149. Since glycerol bears some structural resemblance to CynS inhibitors, its localization points to highly conserved residues that could potentially accommodate negative charges and be involved in passage into/out of the active site. Mutational analysis would be required to experimentally test the role of Asn125 and Lys149 in CynS activity.

Acknowledgements

We would like to thank Dr Thomas Fröhlich for discussion of the sample analysis. We thank the Swiss Light Source synchrotron facility for beam time and excellent on-site support. AB acknowledges support from the International

Max Planck Research School for Molecular and Cellular Life Sciences.

References

Adams, P. D. *et al.* (2010). *Acta Cryst.* D**66**, 213–221.
Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997). *Nucleic Acids Res.* **25**, 3389–3402.
Anderson, P. M., Johnson, W. V., Endrizzi, J. A., Little, R. M. & Korte, J. J. (1987). *Biochemistry*, **26**, 3938–3943.
Ashelford, K. E., Fry, J. C., Bailey, M. J. & Day, M. J. (2002). *Int. J. Syst. Evol. Microbiol.* **52**, 2281–2289.
Bollet, C., Grimont, P., Gainnier, M., Geissler, A., Sainty, J.-M. & De Micco, P. (1993). *J. Clin. Microbiol.* **31**, 444–445.
Cox, J. & Mann, M. (2008). *Nature Biotechnol.* **26**, 1367–1372.
Cox, J., Neuhauser, N., Michalski, A., Scheltema, R. A., Olsen, J. V. & Mann, M. (2011). *J. Proteome Res.* **10**, 1794–1805.
Davey, C. A., Sargent, D. F., Luger, K., Maeder, A. W. & Richmond, T. J. (2002). *J. Mol. Biol.* **319**, 1097–1113.
Elleuche, S. & Pöggeler, S. (2008). *Fungal Genet. Biol.* **45**, 1458–1469.
Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.
Grimont, P. A. D., Grimont, F. & Starr, M. P. (1981). *Curr. Microbiol.* **5**, 317–322.
Guilloton, M. & Karst, F. (1987). *J. Gen. Microbiol.* **133**, 655–665.
Johnson, W. V. & Anderson, P. M. (1987). *J. Biol. Chem.* **262**, 9021–9025.
Kabsch, W. (2010). *Acta Cryst.* D**66**, 125–132.
Maiti, R., Van Domselaar, G. H., Zhang, H. & Wishart, D. S. (2004). *Nucleic Acids Res.* **32**, W590–W594.
McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
Qian, D., Jiang, L., Lu, L., Wei, C. & Li, Y. (2011). *PLoS One*, **6**, e18300.
Sung, Y. C., Parsell, D., Anderson, P. M. & Fuchs, J. A. (1987). *J. Bacteriol.* **169**, 2639–2642.
Vagin, A. & Teplyakov, A. (2010). *Acta Cryst.* D**66**, 22–25.
Walsh, M. A., Otwinowski, Z., Perrakis, A., Anderson, P. M. & Joachimiak, A. (2000). *Structure*, **8**, 505–514.