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

Preliminary X-ray analysis of twinned crystals of sarcosine dimethylglycine methyltransferase from *Halorhodospira halochoris*

Juha Pekka Kallio,[a] Janne Jänis,[a] Antti Nyyssölä,[b,c] Nina Hakulinen[a] and Juha Rouvinen[a]*

[a]Department of Chemistry, University of Joensuu, PO Box 111, 80101 Joensuu, Finland, [b]VTT Technical Research Centre of Finland, PO Box 1000, 02044 VTT, Finland, and [c]Laboratory of Bioprocess Engineering, Helsinki University of Technology, PO Box 6100, 02015 Espoo, Finland

Correspondence e-mail: juha.rouvinen@joensuu.fi

Received 16 April 2009
Accepted 6 July 2009

Sarcosine dimethylglycine methyltransferase (EC 2.1.1.157) is an enzyme from the extremely halophilic anaerobic bacterium *Halorhodospira halochoris*. This enzyme catalyzes the twofold methylation of sarcosine to betaine, with *S*-adenosylmethionine (AdoMet) as the methyl-group donor. This study presents the crystallization and preliminary X-ray analysis of recombinant sarcosine dimethylglycine methyltransferase produced in *Escherichia coli*. Mass spectroscopy was used to determine the purity and homogeneity of the enzyme material. Two different crystal forms, which initially appeared to be hexagonal and tetragonal, were obtained. However, on analyzing the diffraction data it was discovered that both crystal forms were pseudo-merohedrally twinned. The true crystal systems were monoclinic and orthorhombic. The monoclinic crystal diffracted to a maximum of 2.15 Å resolution and the orthorhombic crystal diffracted to 1.8 Å resolution.

1. Introduction

Glycine betaine (betaine) is a quaternary amine which is accumulated in cells in response to low water activity. Betaine protects a wide variety of cells and enzymes against many different types of stresses (Galinski & Trüper, 1994; Gorham, 1995; Yancey *et al.*, 1982). Therefore, an interesting application for enzymes that catalyze betaine synthesis is in the metabolic engineering of stress-tolerant plants and microorganisms (McCue & Hanson, 1990).

In bacteria, the *de novo* synthesis of betaine proceeds *via* the threefold methylation of glycine at the N-position (Nyyssölä *et al.*, 2000; Roberts *et al.*, 1992; Waditee *et al.*, 2003). One of the first microorganisms in which the methyltransferases catalyzing this pathway were characterized was the extremely halophilic anaerobic bacterium *Halorhodospira halochloris*. *H. halochloris* has two enzymes with partly overlapping substrate specifities for glycine methylation. Glycine sarcosine methyltransferase catalyzes the methylation of glycine to sarcosine (*N*-monomethylglycine) and of sarcosine to dimethylglycine. Sarcosine dimethylglycine methyltransferase (SDMT) catalyzes the steps from sarcosine to dimethylglycine and from dimethylglycine to betaine. *S*-Adenosylmethionine (AdoMet) is the methyl-group donor in these reactions (Nyyssölä *et al.*, 2000, 2001). The enzyme-catalyzed reaction is presented in Fig. 1.

Despite the wide occurrence and importance of betaine, no structural information on the enzymes that catalyze microbial betaine synthesis is available at present. This is the first report on the crystallization of a microbial methyltransferase involved in the synthesis of betaine from glycine. However, the coordinates for a putative sarcosine dimethylglycine methyltransferase from *Galdieria sulphuraria* are available in the PDB. These coordinates result from a structural genomics project, but no publication covering the structural data exists.

2. Materials and methods

2.1. Production and purification

The sarcosine dimethylglycine methyltransferase was produced recombinantly in *Escherichia coli* and purified to homogeneity by

Table 1
Data-collection statistics for the SDMT crystals.

	sdmt1	sdmt2
Wavelength (Å)	1.000	0.979
Space group	*C*2	$P2_12_12_1$
Unit-cell parameters (Å, °)	a = 84.9, b = 49.4, c = 179.8, $\alpha = \gamma = 90$, $\beta = 89.9$	a = 50.8, b = 51.2, c = 197.1, $\alpha = \beta = \gamma = 90$
Resolution (Å)	19.3–2.15 (2.3–2.15)	49.3–1.8 (1.9–1.8)
No. of observed reflections	146895 (24825)	390988 (48221)
No. of unique reflections	40587 (7104)	48572 (7169)
Redundancy	3.6	8.1
Completeness (%)	98.9 (95.8)	99.7 (99.4)
R_{obs}† (%)	4.4 (24.2)	8.6 (28.2)
R_{meas}‡ (%)	5.2 (28.6)	9.1 (30.5)
$I/\sigma(I)$	18.9 (5.2)	17.3 (6.9)
Mosaicity (from *XDS*)	0.26	0.15
V_M (Å^3 Da^{-1})	1.95	1.99
Solvent content (%)	37	38

† $R_{obs} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$. ‡ R_{meas} is the redundancy-independent R factor on intensities.

ammonium sulfate fractionation, hydrophobic interaction chromatography and ion-exchange chromatography. The expression and purification have been reported in more detail in a previous study (Nyyssölä *et al.*, 2001).

2.2. Mass spectrometry

All experiments were performed on a 4.7 T Bruker APEX-Qe hybrid quadrupole–Fourier transform ion cyclotron resonance (Q–FT-ICR) mass spectrometer (Bruker Daltonics, Billerica, Massachusetts, USA) equipped with an external Apollo-II electrospray ionization (ESI) source. The sample solutions were directly infused at a flow rate of 1.5 µl min^{-1}. ESI-generated ions were externally accumulated in the hexapole ion trap for 500 ms. The quadrupole was operated in the RF-only mode, transmitting ions at m/z 700–3000. The ions were then passed through high-voltage ion optics to the ICR cell for 'Sidekick' trapping, conventional 'RF-chirp' excitation and direct broadband detection. A total of 2000 co-added 512 kword time-domain transients were acquired and zero-filled twice, followed by magnitude calculation, fast Fourier transform and external frequency-to-m/z calibration with respect to the ions of an ES Tuning Mix (Agilent Technologies, Santa Clara, California, USA). All data were acquired and processed using Bruker *XMASS* 7.0.8 software.

2.3. Crystallization

Screening for crystallization conditions was performed using the hanging-drop vapour-diffusion method. The droplet size was 2 + 2 µl with a reservoir volume of 500 µl. The sample solution concentrations were 13.5 mg ml^{-1} protein, 20 m*M* Tris–HCl [2-amino-2-(hydroxymethyl)-1,3-propanediol hydrochloride] buffer pH 7.5, 2 m*M* DTT (*threo*-1,4-dimercapto-2,3-butanediol) and 0.2% azide. The tested protein concentrations were between 5 and 13.5 mg ml^{-1}. Since *S*-adenosylhomocysteine (AdoHcy) has been reported to be a highly efficient competitive inhibitor of the methylation reaction (Nyyssölä *et al.*, 2000), it was added to the protein solution to aid crystallization (2–5 m*M*). The exact concentration of AdoHcy cannot be reported owing to its extremely low solubility in water. Three different crystallization kits were used: PEG/Ion Screen, Crystal Screen 1 (both from Hampton Research) and a complex screen (Radaev & Sun, 2002) that was made in our laboratory.

2.4. Data collection and analysis

Data were collected using synchrotron radiation on beamline X12 at EMBL/DESY, Hamburg and on beamline ID29 at ESRF, Grenoble. The two data sets from different crystal forms were named sdmt1 (monoclinic) and sdmt2 (hexagonal). The wavelengths were 1.0 Å for sdmt1 and 0.979 Å for sdmt2. The diffraction images were collected with an oscillation angle of 0.25° for sdmt1 and of 0.2° for sdmt2. For the sdmt2 data, we also collected a low-resolution data set with an increased crystal-to-detector distance and with an oscillation angle of 0.5° owing to the overlapping of reflections in the low-resolution area of the higher resolution set. A temperature of 100 K was used for all measurements. Prior to data collection, the crystals were briefly soaked in a cryoprotectant solution consisting of the crystallization solution with a concentrated volume of polyethylene glycol [35%(*v*/*v*)]. After mounting in loops, the crystals were flash-frozen in liquid nitrogen.

The data were processed using *XDS* and scaled with *XSCALE* (Kabsch, 1993). Data-processing statistics are presented in Table 1.

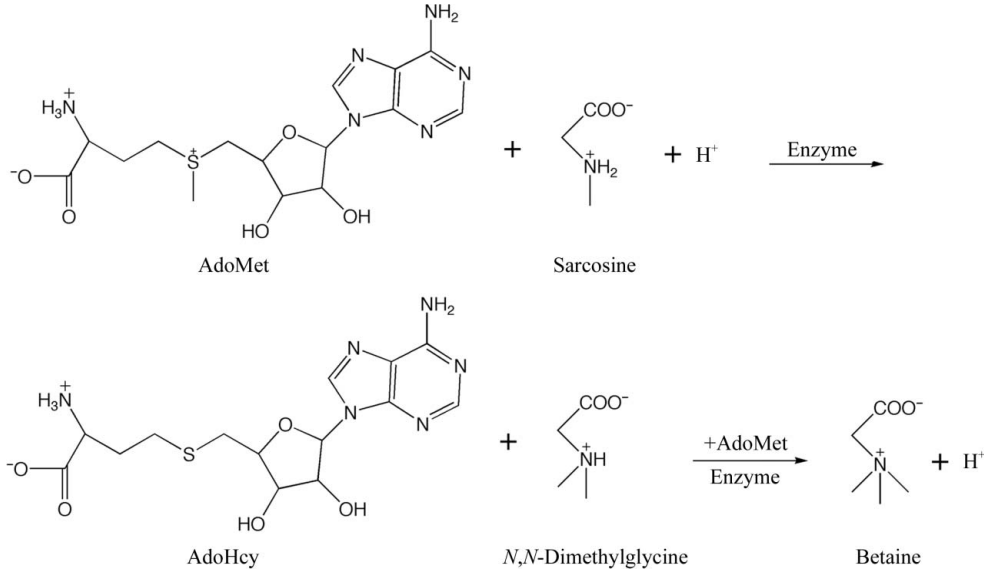

Figure 1
Schematic view of the reaction catalyzed by the SDMT enzyme.

The *phenix.xtriage* module of the *PHENIX* program (Adams *et al.*, 2002) was used to analyse the collected data for twinning, which was first suggested by the intensity distribution of the processed data.

3. Results and discussion

Fig. 2 presents the ESI Q–FT-ICR mass spectrum of SDMT measured in acetonitrile–water–acetic acid (49.5:49.5:1.0 by volume) solution. The mass spectrum exhibited a major protein-ion ($[M + zH]^{z+}$) charge-state distribution from z = 12+ to 32+ centered around 21+ (unlabelled signals). The most abundant isotopic mass of this protein form, averaged over the charge-state distribution, was determined to be 32 094.71 ± 0.09 Da. The results suggest that the N-terminal methionine had been cleaved and that the C-terminal glycine was in its amidated form (32 094.65 Da calculated from the amino-acid sequence). Another protein-ion charge-state distribution (signals labelled with filled circles) was observed at considerably higher charge states (centred around 25+), with the most abundant isotopic mass determined being 32 215.7 ± 0.4 Da, which was consistent with a mass increase of ∼121 Da compared with the major protein form. However, the mass difference of 121 Da does not correspond to any of the known post-translational modifications, nor can it result from partial incorporation of the protein signal sequence. A third and less abundant protein-ion charge-state distribution (signals labelled with unfilled circles) was also observed, but owing to low resolution only an estimation of the average mass (29 420.1 Da) could be made. This protein is most probably an impurity in the sample, since the mass was ∼2.5 kDa less than the masses of the two protein forms of SDMT.

In the crystal screening, the initial crystallization condition was found using the complex screen solution No. 11 {20% polyethylene glycol 3350, 0.1 *M* $MgCl_2$, 0.1 *M* MOPS [3-(*N*-morpholino)butanesulfonic acid] pH 7}. From the screening (146 conditions), we discovered one common denominator. Tiny stick-shaped and plate-shaped crystals were detected in droplets containing the second-group elements magnesium and calcium. The crystallization conditions were optimized by altering the pH and the buffer composition and by changing the concentrations of the precipitants and salts. Different compositions of salts of the second-group elements magnesium, calcium, strontium and barium were also tested (about 50 different conditions). Crystals were obtained with all of these metal cations except barium. The crystallization solution for the first crystal form contained 15%(*w*/*v*) polyethylene glycol 3350, 0.1 *M* calcium acetate and 0.1 *M* HEPES [*N*-(2-hydroxyethyl)piperazine-*N*′-(2-ethanesulfonic acid)] pH 7 (sdmt1; Fig. 3*a*). Despite the optimization, the crystals grew in bundles. Nevertheless, the crystals were separable for X-ray measurements. The crystals grew to dimensions of 0.4 × 0.3 × 0.05 mm within a few days. For the other crystal form obtained, only the composition of the salt was altered in the crystallization conditions. From the conditions 15%(*w*/*v*) polyethylene glycol 3350, 0.1 *M* $MgCl_2$ and 0.1 *M* HEPES pH 7.5, we obtained a single crystal with dimensions of 0.5 × 0.3 × 0.1 mm after four months (sdmt2; Fig. 3*b*). Crystals of this crystal form again mostly grew in bundles, but some single crystals were also obtained. The best protein concentration for crystallization experiments was approximately 9 mg ml^{-1}. Polystyrene nanoparticles (Kallio *et al.*, 2009) were tested in some crystallization experiments as additives in the crystallization droplet. Nanospheres had a positive effect on crystal growth and we were able to grow large crystals in less time (1–7 d). An attempt to improve the quality of the crystals by streak-seeding did not prove succesful. Nanospheres were not used in these streak-seeding experiments.

The sdmt1 crystal diffracted to a resolution of 2.15 Å. The most likely space group was *C*2, with unit-cell parameters a = 84.9, b = 49.4, c = 179.8 Å, β = 89.9°. In the monoclinic crystal system, when the β angle is close to 90°, it is possible to detect pseudo-merohedral twinning. The sdmt2 crystal diffracted to a resolution of 1.8 Å and the most likely space group was $P2_12_12_1$, with unit-cell parameters a = 50.8, b = 51.2, c = 197.1 Å. In the orthorhombic system, when two unit-cell axes are about the same (in this case $a \simeq b$), the existence of pseudo-merohedral twinning is again possible. Crystal systems and space groups were defined based on processing statistics and the utility programs *XPREP* (Sheldrick, 1991) and *phenix.xtriage* (Adams *et al.*, 2002).

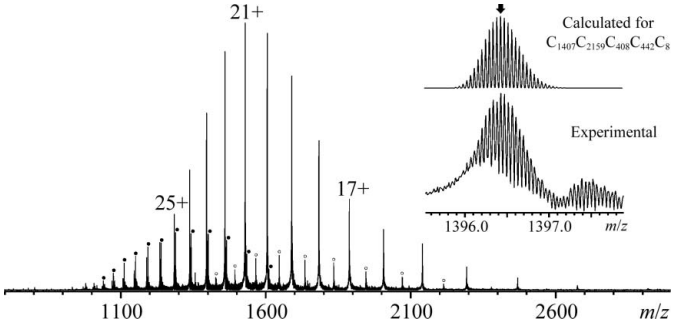

Figure 2
ESI Q–FT-ICR mass spectrum of 7 µ*M* SDMT measured in an acetonitrile–water–acetic acid (49.5:49.5:1.0 by volume) solution. Numbers denote protein-ion charge states (z+). The inset shows a magnified view of the isotopically resolved charge state 23+ (both experimental and calculated isotopic distributions), with the peak corresponding to the most abundant isotopic mass assigned with an arrow. Signals representing minor protein forms are labelled with filled and and unfilled circles.

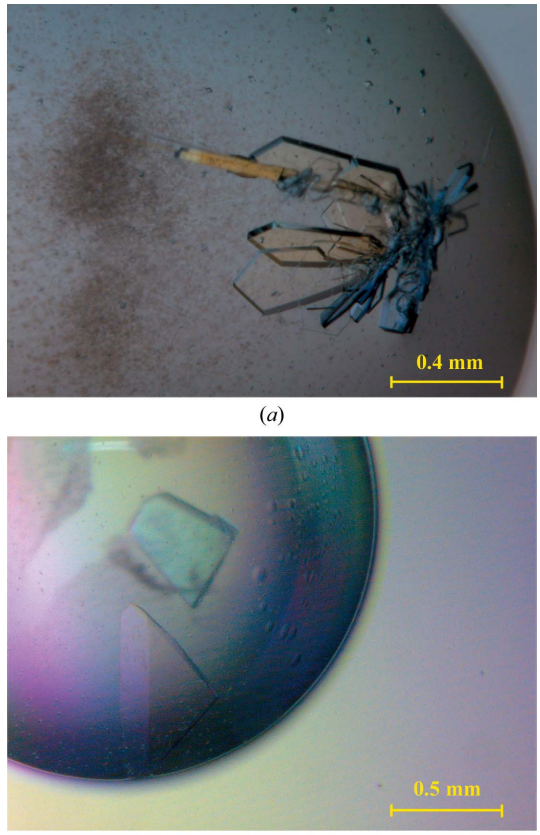

Figure 3
SDMT crystals: (*a*) sdmt1, (*b*) sdmt2.

Table 2
Results of the twinning tests for the SDMT data.

The expected values for twinning tests are $\langle I^2\rangle/\langle I\rangle^2$ untwinned = 2.000, perfect twin = 1.500; $\langle F\rangle^2/\langle F^2\rangle$ untwinned = 0.785, perfect twin = 0.885; $\langle|E^2-1|\rangle$ untwinned = 0.736; perfect twin = 0.541.

	sdmt1	sdmt2
Apparent crystal system	Hexagonal	Tetragonal
Likely (true) space group	*C*2	$P2_12_12_1$
Twinning tests		
For acentric data		
$\langle I^2\rangle/\langle I\rangle^2$	1.803	1.763
$\langle F\rangle^2/\langle F^2\rangle$	0.839	0.841
$\langle\|E^2-1\|\rangle$	0.645	0.638
Multivariate *Z* score *L*-test	5.864	6.295
Possible twin operator(s)	$-h, -k, l$	$k, h, -l$
	$-1/2h + 3/2k, 1/2h + 1/2k, -l$	
	$-1/2h - 3/2k, -1/2h + 1/2k, -l$	
	$1/2h + 3/2k, 1/2h - 1/2k, -l$	
	$1/2h - 3/2k, -1/2h - 1/2k, -l$	

In both cases, the symmetry of the crystals seemed at first glance to be higher. However, the intensity statistics from *XDS* showed the first indications of twinning. Twinning tests were performed using the *phenix.xtriage* module of the *PHENIX* program (Adams *et al.*, 2002). The multivariate *Z* scores from the *L*-tests of *PHENIX* were well above the expected value of 3.5% for both data sets. A high *Z*-score value indicates that the data differ significantly from what is expected for untwinned data. The results of the twinning tests and possible twinning laws are shown in Table 2. Preliminary data analysis was continued with the sdmt2 data, which were of higher resolution than the sdmt1 data. The most probable content of the asymmetric unit was two SDMT monomers, with a solvent content of 38% and a Matthews coefficient of 1.99 $Å^3$ Da^{-1}. Based on sequence identity (44%), the phases for the structure determination could potentially be estimated using the available coordinates of a putative sarcosine dimethylglycine methyltransferase from *G. sulphuraria* (PDB code 2o57).

The authors thank Reetta Kallio-Ratilainen and Ritva Romppanen for their skilled technical assistance. We gratefully acknowledge the European Synchrotron Radiation Facility for provision of synchrotron access. We also gratefully acknowledge access to the EMBL beamline X12 at the DORIS storage ring, Hamburg, Germany. Support from the European Community Research Infrastructure Action under the FP6 'Structuring the European Research Area' Programme contract No. II3-CRT-2004-506008 is also acknowledged.

References

Adams, P. D., Grosse-Kunstleve, R. W., Hung, L.-W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W., Read, R. J., Sacchettini, J. C., Sauter, N. K. & Terwilliger, T. C. (2002). *Acta Cryst.* D**58**, 1948–1954.
Galinski, E. A. & Trüper, H. G. (1994). *FEMS Microbiol. Rev.* **15**, 95–108.
Gorham, J. (1995). *Amino Acids and Their Derivatives in Higher Plants*, edited by R. M. Wallsgrove, pp. 173–203. Cambridge University Press.
Kabsch, W. (1993). *J. Appl. Cryst.* **26**, 795–800.
Kallio, J. M., Hakulinen, N., Kallio, J. P., Niemi, M. H., Kärkkäinen, S. & Rouvinen, J. (2009). *PLoS ONE*, **4**, e4198.
McCue, K. F. & Hanson, A. D. (1990). *Trends Biotechnol.* **8**, 358–362.
Nyyssölä, A., Kerovuo, J., Kaukinen, P., von Weymarn, N. & Reinikainen, T. (2000). *J. Biol. Chem.* **275**, 22196–22222.
Nyyssölä, A., Reinikainen, T. & Leisola, M. (2001). *Appl. Environ. Microbiol.* **67**, 2044–2050.
Radaev, S. & Sun, P. D. (2002). *J. Appl. Cryst.* **35**, 674–676.
Roberts, M. F., Lai, M.-C. & Gunsalus, R. P. (1992). *J. Bacteriol.* **174**, 6688–6693.
Sheldrick, G. (1991). *XPREP. Space Group Determination and Reciprocal Space Plots.* Siemens Analytical X-ray Instruments, Madison, Wisconsin, USA.
Waditee, R., Tanaka, Y., Aoki, K., Hibino, T., Jikuya, H., Takano, J., Takabe, T. & Takabe, T. (2003). *J. Biol. Chem.* **278**, 4932–4942.
Yancey, P. H., Clark, M. E., Hand, S. C., Bowlus, R. D. & Somero, G. N. (1982). *Science*, **217**, 1214–1222.