Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Purification, crystallization and preliminary X-ray diffraction analysis of an archaeal ABC-ATPase

Grégory Verdon,[a] Sonja-V. Albers,[b] Bauke W. Dijkstra,[a] Arnold J. M. Driessen[b] and Andy-Mark W. H. Thunnissen[a]*

[a]Laboratory of Biophysical Chemistry, Department of Chemistry, University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands, and [b]Department of Microbiology, University of Groningen, Kerklaan 30, 9751 NN Haren, The Netherlands

Correspondence e-mail: a.thunnissen@chem.rug.nl

Received 16 October 2001
Accepted 3 December 2001

In the archaeon *Sulfolobus solfataricus* glucose uptake is mediated by an ABC transport system. The ABC-ATPase of this transporter (GlcV) has been overproduced in *Escherichia coli* and purified. Crystals of GlcV suitable for data collection were obtained in the absence of nucleotide by microseeding combined with vapour diffusion from a mixture of PEG polymers and NaCl. Appearing under identical conditions, two crystal forms have been characterized by X-ray diffraction. Both forms diffract to high resolution using synchrotron radiation and both belong to space group $P2_12_12_1$. The related crystal forms *A* (unit-cell parameters a = 47.0, b = 48.2, c = 182.1 Å) and *B* (a = 47.0, b = 146.6, c = 178.5 Å) feature one and three GlcV molecules in the asymmetric unit, respectively, with a solvent content of about 50%. Crystals have also been obtained in the presence of sodium iodide. From single-wavelength anomalous diffraction data extending to 2.1 Å resolution, an iodide substructure could be resolved.

1. Introduction

ATP-binding cassette (ABC) transporters are a large family of proteins involved in the specific and active translocation of molecules across cellular membranes (Holland & Blight, 1999). Their activity is required for several processes essential to cells, such as the uptake of nutrients and protection from a wide range of noxious compounds (*e.g.* multi-drug resistance). All ABC transporters have a similar organization that consists minimally of two membrane-embedded modules forming a channel, in complex with two cytosolic ABC-ATPase domains (or ATP-binding cassettes) energizing the translocation process. Eukaryotic systems are in majority built up from one polypeptide chain, while their prokaryotic counterparts are composed of distinct subunits. Prokaryotic uptake systems include an additional protein for capturing the compound to be transported.

The ABC-ATPases are the most conserved components of the ABC transport systems and their primary sequences feature a characteristic set of motifs, including most notably the Walker A, Walker B and ABC signature motifs. The Walker A (or P-loop) and Walker B motifs are involved in the binding of the substrate ATP and the catalytic cofactor Mg^{2+} (Walker *et al.*, 1982). The ABC signature motif (LSGGQQ) is strictly specific to the ABC-ATPases. Mutants in this pattern are defective in ATP hydrolysis, while retaining the ATP-binding capability (Schmees *et al.*, 1999).

The high sequence similarity among the ABC-ATPases implies they share a common three-dimensional fold, which has been confirmed by the crystal structures of the ABC-ATPases HisP from *Salmonella typhimurium*, MalK from *Thermococcus litoralis*, MJ1267 and MJ0796 from *Methanococcus jannaschii* (Hung *et al.*, 1998; Diederichs *et al.*, 2000; Yuan *et al.*, 2001; Karpowich *et al.*, 2001). Moreover, the crystal structures of the DNA-maintenance ATPases Rad50 from *Pyrococcus furiosus*, MutS from *Escherichia coli* and SMC from *Thermotoga maritima* (Hopfner *et al.*, 2000; Lamers *et al.*, 2000; Löwe *et al.*, 2001) showed that ABC-ATPase domains are also involved in other mechanochemical systems, emphasizing their uniform role in transducing the energy of ATP hydrolysis to drive a mechanical process. The precise nature of this mechanochemical coupling mechanism remains largely unknown. The structural characterization of an ATP-like bound state of Rad50 revealed that dimerization of the ABC-ATPase domains is part of an ATP-induced conformational switch mechanism. Interestingly, the highly conserved ABC signature motif from one monomer was found to complement the binding of ATP in the active site of the other monomer. It seems likely that such a dimerization of the ABC-ATPases is a conserved feature of all ABC-ATPase containing systems (Hopfner *et al.*, 2000). However, so far this type of dimer has not been identified in the crystal structures of ABC-ATPases involved in transport.

In the hyperthermoacidophilic *Sulfolobus solfataricus*, an ABC transporter dedicated to glucose uptake has been identified (Albers *et*

al., 1999). This system is built up from four different gene products: a membrane-anchored glucose-binding protein (GlcS), two sugar permeases (GlcT and GlcU) and the ABC-ATPase GlcV. Sequence analysis indicates that the 40 kDa GlcV protein (353 residues) contains a typical ABC-ATPase domain of about 225 residues and a C-terminal extension of 125 residues of unknown function. GlcV is most closely related to MalK from *Thermococcus litoralis*, with a sequence identity of 43%.

Our work aims at elucidating the structural basis of the coupling between the ABC-ATPase catalytic cycle and the transport process using X-ray crystallography and biochemical methods. Here, we present the overproduction, purification and crystallization procedures of GlcV and a preliminary X-ray analysis of the crystals obtained.

2. Material and methods

2.1. Expression and purification

The *glcV* gene was amplified by PCR from *S. solfataricus* P2 genomic DNA using two primers (FP, 5′-CGCGCC**ATGGTT**AGGATTATTG-3′; RP, 5′-CGGC**GGATCC**TTATTTTTTTTCAA-3′) featuring an *Nco*I and a *Bam*HI endonuclease site, respectively. Ligation of the *Nco*I–*Bam*HI digested PCR product and expression vector pET-15b (Novagen) yielded pET2150 (Amp^r), which was introduced into *E. coli* C43 (DE3) cells (Miroux & Walker, 1996) together with p1244 (Spc^r) coding for rare *E. coli* tRNAs (Kim *et al.*, 1998). Fermentation and expression were performed at 303 K in LB$^+$ medium (LB medium plus 5 g l^{-1} of yeast extract) supplemented with carbenicillin and spectinomycin (80 µg l^{-1} each). Expression of the *glcV* gene was induced with 0.8 m*M* isopropyl-D-thiogalactoside at an OD_{600} of 1.0. Cells were harvested by centrifugation after 6 h of induction and stored at 193 K.

All purification steps were performed at 285 K. Cell extracts were prepared by sonication in 50 m*M* MES buffer pH 6.5, 100 m*M* NaCl, 20% glycerol with protease inhibitors. After centrifugation at 40 000*g* for 30 min, the supernatant was diluted with buffer *A* (30 m*M* MES buffer pH 6.5, 20% glycerol), loaded onto a SP Sepharose FF column (Pharmacia) and eluted with a linear gradient of 0–0.6 *M* NaCl in buffer *A*. The fractions containing GlcV, as identified by SDS–PAGE, were pooled and, after raising the NaCl concentration to 3 *M*, were loaded on a butyl Sepharose FF column (Pharmacia) equilibrated with 3 *M* NaCl in buffer *B* (30 m*M* HEPES buffer pH 7.0, 5% glycerol). GlcV was eluted using a linear gradient of 3–0 *M* NaCl in buffer *B*. The GlcV solution was then diluted ten times with water and loaded onto a Red-dye agarose column (Reactive red 120 type 3000CL, Sigma) equilibrated in 30 m*M* HEPES buffer pH 7.5, 20% glycerol (buffer *C*) and 0.3 *M* NaCl. A first gradient was performed from 0.3 to 0.8 *M* NaCl to remove weakly bound contaminants and GlcV was eluted with a step gradient at 2 *M* NaCl in buffer *C*. Finally, the GlcV pool was concentrated and further equilibrated in 20 m*M* MES buffer pH 6.5, 150 m*M* NaCl, 5% glycerol for storage at 253 K.

To test the activity of the purified protein, GlcV (3 µg) was pre-heated at 343 K for 2 min in 100 µl of assay buffer (storage solution plus 5 m*M* $MgCl_2$) and ATP was added at a concentration of 1 m*M*. The reaction was stopped after 5 min by freezing in liquid nitrogen and the amount of released inorganic phosphate was determined colorimetrically (Lanzetta *et al.*, 1979). Data were corrected for non-enzymatic hydrolysis of ATP at 343 K.

Table 1
Crystallographic data and processing statistics.

Values in parentheses refer to the highest resolution shell.

Data	Crystal form *A*	Crystal form *B*	Iodide crystal form *A*
Beamline	I711, Maxlab	ID14-EH4, ESRF	BW7A, DESY
Detector	MAR 345 mm IP	ADSC Q4 CCD	MAR 165 mm CCD
Temperature (K)	100	100	100
Wavelength (Å)	1.02	1.0	1.7
No. of images	230	200	300
Oscillation angle per frame (°)	0.5	0.5	0.5
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Unit-cell parameters (Å, °)	a = 47.0, b = 48.2, c = 182.1, $\alpha = \beta = \gamma$ = 90	a = 47.0, b = 146.6, c = 178.5, $\alpha = \beta = \gamma$ = 90	a = 46.1, b = 48.2, c = 183.0, $\alpha = \beta = \gamma$ = 90
No. of molecules per asymmetric unit	1	3	1
No. of unique reflections	30914	69924	24485
Redundancy	4–5	4–5	5–6
Resolution range (Å)	40.0–1.92 (2.0–1.92)	46.2–2.1 (2.2–2.1)	37.0–2.1 (2.2–2.1)
Completeness (%)	94.9 (85.9)	95.4 (71.0)	98.7 (93.8)
$\langle I/\sigma(I)\rangle$	31.8 (5.0)	20.5 (3.0)	14.6 (5.7)
R_{sym}† (%)	3.3 (19.8)	5.7 (33.6)	10.0 (27.0)

† $R_{sym} = \sum_h \sum_i I_{(h,i)} - \langle I\rangle_h| / \sum_h \sum_i \langle I\rangle_h$.

2.2. Crystallization and data collection

Crystallization trials were performed in the absence of nucleotide with Crystal Screens I and II (Hampton Research) using the hanging-drop vapour-diffusion method. One condition resulting in a crystalline precipitate (30% PEG 4000, 0.1 *M* Tris buffer pH 8.5, 200 m*M* $MgCl_2$) could be optimized and crystals were grown at 293 K using microseeding in hanging-drop vapour-diffusion setups. The drops (4 µl) consisted of equal volumes of protein solution (4.25 mg ml^{-1} GlcV in storage buffer) and reservoir solution containing 15% PEG 3350, 15% PEG 400, 0.1 *M* Tris buffer pH 8.3, 0.5 *M* NaCl, 5% glycerol. Iodide-derivatized crystals could be obtained under the same conditions, except that NaI was substituted for NaCl. Because of the presence of PEG 400 and glycerol, crystals could be directly frozen in liquid nitrogen for data collection at 100 K. The crystallographic data were collected using synchrotron radiation (Table 1) and processed using the programs *DENZO* and *SCALEPACK* (Otwinowski & Minor, 1997).

3. Results and discussion

GlcV could be overproduced in *E. coli* as a soluble protein and was isolated using

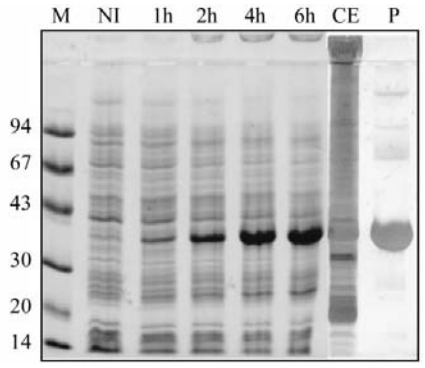

Figure 1
SDS–PAGE analysis of the time-course production of GlcV in *E. coli* C43 (DE3) cells and its purification: M, molecular-weight markers (kDa); NI, non-induced; 1h, 2h, 4h and 6h, 1, 2, 4 and 6 h after induction, respectively; CE, crude extract; P, purifed GlcV protein (20 µg). CE and P were revealed by silver-staining and the other lanes developed using Coomassie Blue.

cation-exchange, hydrophobic interaction and dye-ligand chromatography. From 16 g of cells (2 l of culture), 75 mg of protein could be purified to near electrophoretic homogeneity as judged from silver-stained SDS–PAGE (Fig. 1). Optimal ATPase activity was observed at pH 6.5 and 343 K, with a K_m for ATP of about 290 μM. Such a low affinity for ATP has also been measured for the ABC-ATPases HisP from *S. typhimurium* and MalK from *T. litoralis* (Nikaido *et al.*, 1997; Greller *et al.*, 1999). In the absence of nucleotide, GlcV is monomeric in solution as interpreted from size-exclusion chromatography and dynamic light-scattering experiments (data not shown).

Crystals of GlcV could be grown by microseeding in hanging-drop vapour-diffusion setups. The crystals appear overnight and grow to final dimensions of $0.1 \times 0.05 \times 0.45$ mm in two weeks at 293 K (Fig. 2). In the early stages of this study, diffraction experiments were severely hampered by crystals showing 'multiple' diffraction patterns. Since the purification of GlcV as described was originally performed with a phosphate buffer in the second chromatographic step, we considered the possibility that phosphate ions remaining from the purification could be bound in the active sites of some GlcV molecules, thus introducing some heterogeneity. By replacing the phosphate buffer with a HEPES buffer, the crystals that were subsequently obtained always showed a single diffraction pattern (Fig. 3).

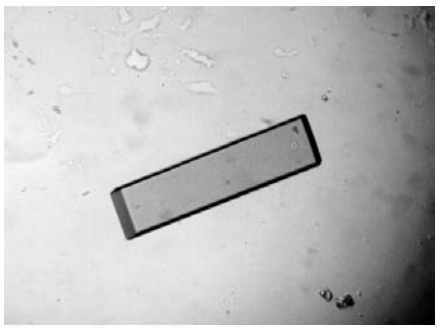

Figure 2
Single orthorhombic crystal of GlcV in either form *A* or *B*: from visual inspection it is impossible to distinguish the two forms.

Figure 3
Diffraction image from a GlcV form *A* crystal collected using a 0.5° oscillation and a 10 s exposure on a MAR 345 mm image-plate detector (beamline I711, Maxlab, Lund). The resolution at the edge of the detector is 1.9 Å.

Appearing under identical conditions and indistinguishable by shape, two crystal forms have been characterized by X-ray diffraction (Fig. 3, Table 1). Both forms diffract to at least 2.1 Å resolution using synchrotron radiation and both belong to space group $P2_12_12_1$. Crystal form *A* (unit-cell parameters a = 47.0, b = 48.2, c = 182.1 Å) and form *B* (a = 47.0, b = 146.6, c = 178.5 Å) have related unit cells with similar a and c unit-cell parameters, but with the b cell axis three times longer in crystal form *B* than in crystal form *A*. A self-Patterson function calculated with the data of crystal form *B* showed very high peaks at 1/3 and 2/3 along v, while self-rotation functions calculated for both crystal forms did not reveal any difference. Therefore, we expect crystal form *B* to be related to crystal form *A* by a symmetry breakdown along the crystallographic twofold screw axis in the b direction. From the calculated Matthews coefficients we assume that crystal form *A* contains one monomer of GlcV per asymmetric unit and crystal form *B* three, with a solvent content of about 50%. The cause of the polymorphism remains unclear, but interestingly the appearance of form *B* crystals is favoured by a higher concentration of glycerol.

For structure determination, molecular replacement using models based on the HisP coordinates (the only available ABC-ATPase structure at that time) and crystallization of selenomethionine-substituted GlcV protein were without success. As an alternative approach, iodide-derivatized crystals were prepared by replacing sodium chloride in the crystallization solution with sodium iodide. As shown recently, the anomalous scattering of halide ions bound to the protein surface can be used as a unique source of phase information for protein-structure determination (Dauter *et al.*, 2000). Following this method, single-wavelength anomalous diffraction data were collected to 2.1 Å resolution on an iodide-derivatized GlcV crystal (Table 1). From the anomalous differences, five iodide sites were clearly identified by direct methods with the program *Shake-and-Bake* (Weeks & Miller, 1999). Refinement against the anomalous data using the program *SHARP* (de La Fortelle & Bricogne, 1997) led to the characterization of 12 partially occupied iodide sites (0.8–0.2). The overall figure of merit for the phases was 0.32 (37.0–2.1 Å). After a solvent-flipping procedure performed with the program *SOLOMON* (Abrahams & Leslie, 1996), the electron density was already of sufficient quality to locally identify some helical features, although the overall accuracy of the phases was still too low to allow extensive building.

However, new data collected to 1.65 Å resolution on a similar iodide-derivatized GlcV crystal (ID14-EH2, ESRF, Grenoble) in combination with the preliminary set of phases allowed the complete structure determination of GlcV using the phase-improvement and model building/refinement routines implemented in *ARP*/*wARP* (Perrakis *et al.*, 1997, 1999).

The authors thank Wil N. Konings for his interest in this project and John E. Walker (MRC, Cambridge) and Sung-Hou Kim (University of California at Berkeley, USA) for the *E. coli* C43 (DE3) strain and p1244 vector, respectively. We are grateful to staff scientists and local contacts at the beamlines I711 (Maxlab, Lund), ID14-EH4 (ESRF, Grenoble) and BW7A (DESY, Hamburg) for help during data collections. This work was supported by an Ubbo Emmius PhD fellowship from the University of Groningen to GV and a TMR fellowship (ERBFMBIC971980) to SVA.

References

Abrahams, J. P. & Leslie, A. G. W. (1996). *Acta Cryst.* D**52**, 30–42.
Albers, S. V., Elferink, M. G. L., Charlebois, R. L., Sensen, C. W., Driessen, A. J. M. & Konings, W. N. (1999). *J. Bacteriol.* **181**, 4285–4291.
Dauter, Z., Dauter, M. & Rajashankar, K. R. (2000). *Acta Cryst.* D**56**, 232–237.
Diederichs, K., Diez, J., Greller, G., Müller, C., Breed, J., Schnell, C., Vonrheim, C., Boos, W. & Welte, W. (2000). *EMBO J.* **19**, 5951–5961.
Greller, G., Horlacher, R., DiRuggerio, J. & Boos, W. (1999). *J. Biol. Chem.* **274**, 20259–20264.
Holland, I. B. & Blight, M. A. (1999). *J. Mol. Biol.* **293**, 381–399.
Hopfner, K.-P., Karcher, A., Shin, D. S., Craig, L., Arthur, M. L., Carney, J. P. & Tainer, J. A.

(2000). *Cell*, **101**, 789–800.
Hung, L. W., Wang, I. X., Nikaido, K., Liu, P. Q., Ames, G. F. L. & Kim, S.-H. (1998). *Nature (London)*, **396**, 703–707.
Karpowich, N., Martsinkevich, O., Millen, L., Yuan, Y., Dai, P. L., Macvey, K., Thomas, P. J. & Hunt, J. F. (2001). *Structure Fold. Des.* **9**, 571–586.
Kim, R., Sandler, S. J., Goldman, S., Yokota, H., Clark, A. J. & Kim, S.-H. (1998). *Biotech. Lett.* **20**, 207–210.
La Fortelle, E. de & Bricogne, G. (1997). *Methods Enzymol.* **276**, 472–494.
Lamers, M. H., Perrakis, A., Enzlin, J. H., Winterwerp, H. H., de Wind, N. & Sixma, T. K. (2000). *Nature (London)*, **407**, 711–717.
Lanzetta, P. A., Alvarez, L. J., Reinach, P. S. & Candia, O. A. (1979). *Anal. Biochem.* **100**, 95–97.
Löwe, J., Cordell, S. C. & van den Ent, F. (2001). *J. Mol. Biol.* **306**, 25–35.
Miroux, B. & Walker, J. E. (1996). *J. Mol. Biol.* **260**, 289–298.
Nikaido, K., Liu, P.-Q. & Ames, G. F.-L. (1997). *J. Biol. Chem.* **272**, 27745–27752.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Perrakis, A., Morris, R. J. & Lamzin, V. S. (1999). *Nature Struct. Biol.* **6**, 458–463.
Perrakis, A., Sixma, T. K., Wilson, K. S. & Lamzin, V. S. (1997). *Acta Cryst.* D**53**, 448–455.
Schmees, G., Stein, A., Hunke, S., Landmesser, H. & Schneider, E. (1999). *Eur. J. Biochem.* **266**, 420–430.
Walker, J. E., Saraste, M., Runswick, M. J. & Gay, N. J. (1982). *EMBO J.* **1**, 945–951.
Weeks, C. M. & Miller, R. (1999). *J. Appl. Cryst.* **32**, 120–124.
Yuan, Y.-R., Blecker, S., Martsinkevich, O., Millen, L., Thomas, P. J. & Hunt, J. F. (2001). *J. Biol. Chem.* **276**, 32313–32321.