Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization and preliminary X-ray analysis of the *Streptomyces olivaceoviridis* NgcE binding protein of the ABC transporter for *N*-acetylglucosamine

Akihiro Saito,[a,b,c] Zui Fujimoto,[a] Eiichi Minami,[a] Hiroshi Mizuno,[a] Kiyotaka Miyashita,[a] Hildgund Schrempf[b] and Mitsuru Momma[a]*

[a]National Institute of Agrobiological Sciences, Kan-nondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan, [b]FB Biologie/Chemie, Universität Osnabrück, Barbarastrasse 11, D-49069 Osnabrück, Germany, and [c]Department of Bioresources Chemistry, Faculty of Horticulture, Chiba University, Matsudo 648, Matsudo, Chiba 271-8510, Japan

Correspondence e-mail: momma@nias.affrc.go.jp

Received 2 July 2004
Accepted 19 October 2004

The NgcE protein binds *N*-acetylglucosamine (GlcNAc) as well as *N*,*N*′-diacetylchitobiose and is a component of the ABC transporter Ngc for GlcNAc uptake in *Streptomyces olivaceoviridis*. After cloning the corresponding gene in an *Escherichia coli* host, the NgcE protein was overproduced in a soluble form within the cytoplasm and purified to homogeneity by four consecutive chromatographic processes. Crystals of NgcE that grew in the presence of 1 m*M* GlcNAc, 20%(*w*/*v*) PEG MME 2000 and 100 m*M* Tris–HCl pH 8.5 had a plate-like shape and belonged to either space group $P2_12_12$ (unit-cell parameters $a = 59.9$, $b = 153.0$, $c = 41.7$ Å) or $P2_12_12_1$ ($a = 58.1$, $b = 96.3$, $c = 151.7$ Å). The former crystals diffracted to 1.8 Å resolution and the latter to 2.2 Å. Selenomethionine-containing crystals were generated under the same conditions and belonged to space group $P2_12_12_1$ with unit-cell parameters $a = 58.4$, $b = 96.6$, $c = 152.5$ Å, and diffracted to 2.0 Å resolution.

1. Introduction

ABC (ATP-binding cassette) transporters comprise one of the largest of all paralogous protein families; most of them are involved in uptake and efflux (for a review, see Higgins, 2001). The transporters play many physiological roles in nutrient uptake, elimination of waste products and toxins, resistance to antifungal agents, export of cellular components, antibiotic production and protein secretion. ABC transporter systems comprise two transmembrane and two ATP-binding domains. In bacteria and archea, a solute-binding protein is often included as an auxiliary component and is involved in capturing a ligand. These systems have been named binding-protein-dependent ABC transporters and are classified into 18 families (for reviews, see Saier, 2000; Schneider, 2001). The solute-binding proteins of these transporters are present within the periplasmic space of Gram-negative bacteria, while they are usually bound to the cytoplasmic membrane *via* a lipid anchor in Gram-positive bacteria and archea. The individual ligand is initially captured by the solute-binding protein, with the transmembrane domains of the corresponding transporter required for entering the cell (for a review, see Higgins, 2001).

Of the binding-protein-dependent ABC transporters, the maltose-uptake system in *Escherichia coli* has been the most extensively investigated by genetic, biochemical and crystallographic studies (Schneider, 2003). The positioning of the ligands within the binding protein MalE have been determined (Quiocho *et al.*, 1997) and a model has been deduced implying that maltose uptake involves a series of conformational changes (Chen *et al.*, 2001; Schneider, 2003).

Recently, the first ABC transporter Ngc for *N*-acetylglucosamine (GlcNAc) was discovered in the chitin-degrader *Streptomyces olivaceoviridis*. The corresponding substrate-binding protein NgcE exhibits high affinity for GlcNAc and *N*,*N*′-diacetylchitobiose (Xiao *et al.*, 2002). The Ngc system belongs to the transporter-1 (CUT-1) family of binding-protein-dependent transporters. NgcE is so far the only known binding protein that interacts with the monosaccharide GlcNAc as well as the disaccharide *N*,*N*′-diacetylchitobiose. It would be interesting to identify its crystal structure and compare its features with other members of the CUT-1 family. As a prerequisite for future structural studies, we describe the crystallization and the preliminary X-ray crystallographic analysis of the NgcE protein purified from an *E. coli* host.

2. Experimental and results

2.1. Overexpression and purification

In order to obtain an *E. coli* transformant carrying and overexpressing the *ngcE* gene on a plasmid, the following procedure was designed. Part of the *S. olivaceoviridis ngcE* gene without the region encoding the signal sequence (amino acids 1–44) was amplified by polymerase chain reaction (PCR) using the plasmid pSKR5 (Saito & Schrempf, 2004) containing the whole *ngcEFG* gene cluster as a

© 2004 International Union of Crystallography
Printed in Denmark – all rights reserved

template. The oligonucleotides ngcf2 (5′-GG**GCATGC**ACGAGCAGGAGAAGGC-GAAGGGC-3′) and ngcEPr (5′-GG**AAGCTT**TCACTGGTGCTTGTAGTGCTTC-3′), which contain *Sph*I and *Hin*dIII sites (in bold), respectively, were used as a primers. The amplified fragment was digested with *Sph*I and *Hin*dIII and inserted into the corresponding site in the *E. coli* expression plasmid vector pQE70 (Qiagen) to obtain plasmid pQE401. Plasmid pQE401 carries a gene encoding the NgcE protein which lacks the endogenous signal peptide and the His tag originally encoded in pQE70. *E. coli* XL1-Blue(pQE401) was cultivated in 2 l LB medium containing 100 µg ml^{-1} ampicillin at 310 K with shaking at 160 rev min^{-1}. 0.5 m*M* IPTG was added during the mid-logarithmic phase. After a further 3 h of cultivation, cells were harvested by centrifugation (6000*g*, 10 min, 277 K).

To overexpress the selenomethionine derivative of NgcE, *E. coli* B-834(DE3)-pLysS, which is auxotrophic for methionine, was transformed with plasmid pQE401 and the resulting *E. coli* transformant was cultivated in a minimal medium (LeMaster & Richards, 1985) supplemented with 25 mg l^{-1} selenomethionine. Prior to collecting cells by centrifugation (6000*g*, 10 min, 277 K), the transformant was further cultivated for 11 h at 310 K after the addition of 0.5 m*M* IPTG at an OD_{600} of 0.6.

E. coli XL1-Blue(pQE401) or B-834(DE3)-pLysS(pQE401) cells were collected by centrifugation (6000*g*, 10 min, 277 K), washed once with ice-cold 20 m*M* Tris–HCl pH 8.0 and then broken (on ice) by sonication. Following centrifugation (16 000*g*, 15 min, 277 K) of the cell lysate, the supernatant (50 ml) was collected as a fraction containing the soluble form of the NgcE protein and was subjected to anion-exchange chromatography using a column (2.5 × 10 cm) filled with DEAE-cellulose (Wako Pure Chemical). The protein was eluted (in 50 ml) with 20 m*M* Tris–HCl pH 8.0 containing 200 m*M* NaCl. 3 *M* $(NH_4)_2SO_4$ in 20 m*M* Tris–HCl pH 8.0 was added to the solution (to obtain a final concentration of 1.0 *M*) and the solution was kept at 277 K overnight. The supernatant was then obtained by centrifugation at 16 000*g* for 30 min at 277 K and was subjected to hydrophobic interaction chromatography using a phenyl Sepharose (Amersham-Pharmacia Biotech) column (1.5 × 5 cm). The portion which did not bind to the column in the presence of 1 *M* $(NH_4)_2SO_4$, 20 m*M* Tris–HCl pH 8.0 was collected, dialyzed against 20 m*M* Tris–HCl pH 8.0 and subjected to further anion-exchange chromatography using a MonoQ 5/50 GL column (Amersham-Pharmacia Biotech). NgcE was eluted with a 0–200 m*M* NaCl gradient in 20 m*M* Tris–HCl pH 8.0. The fractions containing NgcE were detected from its approximate molecular weight (∼45 kDa) on SDS–PAGE (Fig. 1). The main fractions containing NgcE were then subjected to a Superdex 75 10/300 GL gel-filtration column (Amersham-Pharmacia Biotech) in 20 m*M* Tris–HCl pH 8.0, 150 m*M* NaCl. The purity of the obtained NgcE protein was determined by SDS–PAGE (Fig.1). A single N-terminal amino-acid sequence was detected in the purified protein by Edman degradation using an automated protein sequencer. The protein concentration was measured as reported by Bradford (1976). After several rounds of optimization, 5 mg homogenous NgcE protein was reproducibly gained from a 1 l *E. coli* culture. The final protein solution was concentrated to 13 mg ml^{-1} using an Amicon Ultrafiltration system (Millipore).

2.2. Crystallization

Conditions for crystallization of the NgcE protein in the presence of 1 m*M* *N*-acetylglucosamine were initially investigated using solution systems: Index, Crystal Screen, Crystal Screen II, SaltRx (Hampton Research), Wizards I and II and Cryo I and II (Emerald Biostructures). The trials were carried out using the sitting-drop vapour-diffusion technique by mixing 0.2 µl protein solution (13 mg ml^{-1}) with an equal volume of precipitant solution on plates and equilibrating against 100 µl reservoir solutions. After 10 d incubation at 293 K, crystals were obtained using 20–30%(*w*/*v*) PEG 1000, PEG 1500 or PEG MME 2000 as precipitant at pH 8.0 or 8.5. After refinement of the crystallization conditions, solutions containing 20%(*w*/*v*) PEG MME 2000 and 0.1 *M* Tris–HCl pH 8.5 were shown to be optimal for crystallization of both the native NgcE protein and its selenomethionine derivative. Plate-like crystals (approximate crystal dimensions of 0.15 × 0.15 × 0.02 mm) were obtained after 3 d incubation at 293 K (Fig. 2).

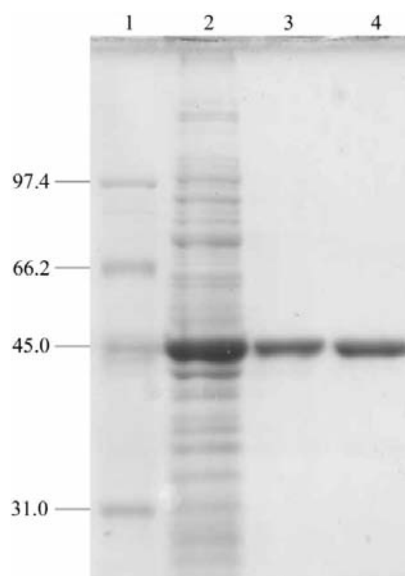

Figure 1
Analysis of the NgcE protein during the course of purification. The protein was isolated as outlined in the text. Samples were subjected to SDS–PAGE and stained with Coomassie Brilliant Blue. Lane 1, molecular-weight markers (kDa); lane 2, soluble protein fraction prepared from the cell lysate of the *E. coli* transformant; lane 3, purified NgcE protein; lane 4, purified selenomethionine derivative of NgcE.

2.3. Data collection and analysis

Diffraction data were collected using a Micromax R-AXIS VII imaging-plate X-ray diffractometer (Rigaku) or using a Quantum 210 CCD detector (ADSC) at beamline NW-12 of the Photon Factory Advanced Ring (PF-AR), Tsukuba, Japan. Crystals were soaked in the corresponding crystallization solution supplemented with 15%(*w*/*v*) glycerol as a cryoprotectant and then picked up in a fibre loop and flash-cooled in a stream of nitrogen gas at 95 K. All data sets were processed and scaled using *DENZO* and *SCALEPACK* from the *HKL*2000 program package (Otwinowski, 1993).

Crystals of the NgcE protein belonged to space group $P2_12_12$, with unit-cell parameters a = 59.9, b = 153.0, c = 41.7 Å, and diffracted to better than 2.2 Å (Table 1). Assuming the presence of one protein molecule in the asymmetric unit, the V_M

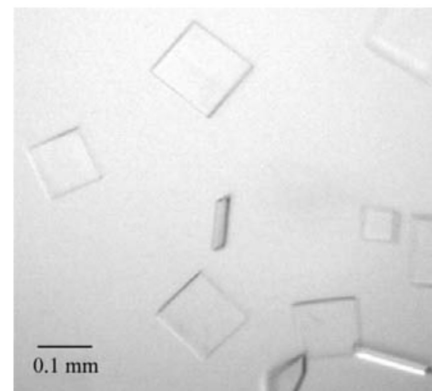

Figure 2
Plate-like crystals of NgcE protein. The crystals were generated in a buffer system containing 1 m*M* *N*-acetylglucosamine (the main ligand for NgcE), 20%(*w*/*v*) PEG MME 2000 and 0.1 *M* Tris–HCl pH 8.5.

Table 1
Parameters and statistics of data collection for the investigated crystals.

Values in parentheses are for the highest resolution shell.

Data	Native1	Native2	SeMet Peak	SeMet Edge	SeMet Remote
Space group	$P2_12_12$	$P2_12_12_1$	$P2_12_12_1$		
Unit-cell parameters (Å)					
a	59.9	58.1	58.4		
b	153.0	96.3	96.6		
c	41.7	151.7	152.5		
Z	4	8	8		
X-ray source	Cu $K\alpha$	PF-AR	PF-AR	PF-AR	PF-AR
Wavelength (Å)	1.5418	1.0000	0.9790	0.9792	0.9770
Resolution (Å)	50–2.2 (2.28–2.20)	20–1.8 (1.86–1.80)	50–2.0 (2.07–2.00)	50–2.2 (2.28–2.20)	50–2.4 (2.49–2.40)
No. reflections	140702	428940	846349	651614	505293
No. unique reflections	37531 (3744)	74967 (7004)	59077 (5819)	44778 (4363)	34791 (3410)
Completeness (%)	99.1 (99.1)	94.0 (89.0)	100.0 (100.0)	100.0 (100.0)	100.0 (100.0)
Multiplicity	3.8 (3.7)	5.8 (5.8)	14.3 (11.8)	14.6 (14.1)	14.5 (14.1)
$R_{\rm merge}$ (%)	16.6 (28.9)	8.0 (30.1)	10.1 (32.8)	8.8 (32.5)	8.9 (34.0)
Average $I/\sigma(I)$	4.0 (2.1)	18.4 (3.3)	40.8 (6.0)	41.2 (6.9)	41.4 (6.6)

value has been calculated to be 2.0 $Å^3$ Da^{-1} (Matthews, 1968), indicating a solvent content of approximately 38%. Another type of crystal grown in the same crystallization conditions belonged to space group $P2_12_12_1$ (unit-cell parameters a = 58.1, b = 96.3, c = 151.7 Å), with a $V_{\rm M}$ value of 2.2 $Å^3$ Da^{-1}, two molecules in the asymmetric unit and an estimated solvent content of 44%, and exhibited an even higher resolution of 1.8 Å (Table 1). The data demonstrated that the NgcE molecules can be packed in two different manners during the crystallization process under identical conditions. In order to solve the structure by the multiple-wavelength anomalous dispersion (MAD) method, we prepared a selenomethionine derivative (space group $P2_12_12_1$) and collected full data sets at three different wavelengths around 2.0 Å resolution.

3. Discussion

The data presented show that the NgcE protein and its selenomenthionine derivative can reproducibly form crystals in the presence of the ligand GlcNAc. As the analyzed crystals diffract in the range 1.8–2.2 Å resolution, it can be predicted that they will be suitable for the collection of further data in order to establish the three-dimensional structure of the NgcE protein in the presence of ligands. It is planned to also grow crystals in the presence of N,N'-diacetylchitobiose, which has also been shown to interact with NgcE (Xiao *et al.*, 2002; Saito & Schrempf, 2004). Previous plasmon surface-resonance studies have revealed the additional high affinity of NgcE for several chitooligomers (Xiao *et al.*, 2002); thus, in the future it will be rewarding to assign their binding coordinates within the protein. Furthermore, it will be interesting to compare the highly resolved features of the NgcE protein with those of the best studied members of the CUT-1 family. These include the MalE protein from *E. coli* (Quiocho *et al.*, 1997) binding maltose, maltotriose and maltotetraose, as well as the homologues *Pfu*MBP from *Pyrococcus furiosus* (Evdokimov *et al.*, 2001), TMBP from *Thermococcus litralis* (Diez *et al.*, 2001) and *Acy*MBP from *Alicyclobacillus acidocaldarius* (Schäfer *et al.*, 2004).

We are grateful for access to the synchrotron facilities of Photon Factory, the High Energy Accelerator Research Organization, Japan. AS appreciates the support of A. Schiefner and K. Diederichs in the research group headed by W. Welte at Universität Konstanz during the initial phase of the crystallization work and gratefully acknowledges an initial postdoctoral fellowship from the Alexander von Humboldt foundation (AvH) and a subsequent research fellowship from Japan Society for the Promotion of Science (JSPS). This work was partially supported by a Grant-in-Aid for JSPS Fellows to AS.

References

Bradford, M. M. (1976). *Anal. Biochem.* **72**, 248–254.
Chen, J., Scharma, S., Quiocho, F. A. & Davidson, A. L. (2001). *Proc. Natl Acad. Sci. USA*, **98**, 1525–1530.
Diez, J., Diederichs, K., Greller, G., Horlacher, R., Boos, W. & Welte, W. (2001). *J. Mol. Biol.* **305**, 905–915.
Evdokimov, A. G., Anderson, D. E., Routzahn, K. M. & Waugh, D. S. (2001). *J. Mol. Biol.* **305**, 891–904.
Higgins, C. F. (2001). *Res. Microbiol.* **152**, 205–210.
LeMaster, D. M. & Richards, F. M. (1985). *Biochemistry*, **24**, 7263–7268.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Otwinowski, Z. (1993). *Proceedings of the CCP4 Study Weekend. Data Collection and Processing*, edited by L. Sawyer, N. Isaacs & S. Bailey, pp. 56–62. Warrington: Daresbury Laboratory.
Quiocho, F. A., Spurlino, J. C. & Rodseth, L. E. (1997). *Structure*, **5**, 997–1015.
Saier, M. H. Jr (2000). *Mol. Microbiol.* **35**, 699–710.
Saito, A. & Schrempf, H. (2004). *Mol. Genet. Genomics*, **271**, 545–553.
Schäfer, K., Magnusson, U., Scheffel, F., Schiefner, A., Sandgren, M. O. J., Diedrechs, K., Welte, W., Hülsmann, A., Schneider, E. & Mowbray, S. L. (2004). *J. Mol. Biol.* **335**, 261–274.
Schneider, E. (2001). *Res. Microbiol.* **152**, 303–310.
Schneider, E. (2003). *ABC Proteins: From Bacteria to Man*, edited by I. B. Holland, S. P. C. Cole, K. Kuchler & C. F. Higgins, pp. 157–185. New York: Academic Press.
Xiao, X., Wang, F., Saito, A., Majka, J., Schlösser, A. & Schrempf, H. (2002). *Mol. Genet. Genomics*, **267**, 429–439.