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

Crystallization and preliminary X-ray crystallographic studies of an oligomeric species of a refolded C39 peptidase-like domain of the *Escherichia coli* ABC transporter haemolysin B

Christian K. W. Schwarz,[a] Britta Tschapek,[a] Thorsten Jumpertz,[a] Stefan Jenewein,[a] Justin Lecher,[b,c] Dieter Willbold,[b,c] Santosh Panjikar,[d] I. Barry Holland,[e] Sander H. J. Smits[a] and Lutz Schmitt[a]*

[a]Institute of Biochemistry, Heinrich Heine University Düsseldorf, Universitätsstrasse 1, D-40225 Düsseldorf, Germany, [b]Institute of Physical Biology, Heinrich Heine University Düsseldorf, Universitätsstrasse 1, D-40225 Düsseldorf, Germany, [c]Institute of Structural Biochemistry (ICS-6), Research Centre Jülich GmbH, Jülich, Germany, [d]EMBL Hamburg Outstation, c/o DESY, Notkestrasse 85, D-22603 Hamburg, Germany, and [e]Institute of Microbiology and Genetics, University of Paris-Sud, Orsay, France

Correspondence e-mail: lutz.schmitt@hhu.de

Received 28 January 2011
Accepted 23 March 2011

The ABC transporter haemolysin B (HlyB) from *Escherichia coli* is part of a type I secretion system that translocates a 110 kDa toxin in one step across both membranes of this Gram-negative bacterium in an ATP-dependent manner. Sequence analysis indicates that HlyB contains a C39 peptidase-like domain at its N-terminus. C39 domains are thiol-dependent peptidases that cleave their substrates after a GG motif. Interestingly, the catalytically invariant cysteine is replaced by a tyrosine in the C39-like domain of HlyB. Here, the overexpression, purification and crystallization of the isolated C39-like domain are described as a first step towards obtaining structural insights into this domain and eventually answering the question concerning the function of a degenerated C39 domain in the ABC transporter HlyB.

1. Introduction

The ATP-binding cassette (ABC) transporter HlyB is a central element of the type I secretion machinery of *Escherichia coli*. Together with the inner membrane protein HlyD, which interacts with the outer membrane protein TolC, HlyB forms a continuous tunnel from the cytoplasm to the exterior to secrete the 110 kDa toxin HlyA without any periplasmic intermediates (Balakrishnan *et al.*, 2001; Mackman *et al.*, 1986, 1987; Benabdelhak *et al.*, 2003; Gray *et al.*, 1986).

In general, a functional ABC transporter consists of two nucleotide binding domains (NBDs) and two transmembrane domains (TMDs). However, and in contrast to this general blueprint, HlyB contains an additional N-terminal domain of about 150 amino acids that is conserved in bacteriocin ABC transporters (Wu & Tai, 2004; Ishii *et al.*, 2010). A sequence alignment classifies the N-terminal region of HlyB as a C39 domain. This family of cysteine peptidases cleave their substrates after a GG motif using a catalytic dyad consisting of a conserved cysteine and histidine. In HlyB, however, the catalytic dyad is degenerated: the highly conserved cysteine residue is mutated to tyrosine (Ishii *et al.*, 2010; Wu & Tai, 2004). Thus, this domain is very likely to be inactive, giving rise to the name 'C39-like domain'.

Curiously, the C39-like domain of HlyB appears to be essential for secretion of the toxin HlyA. In a system in which HlyB lacked the C39-like domain, secretion of HlyA was completely abolished (data not shown) and therefore the question arises as to the specific role that this defective protease may play during substrate secretion. We expressed and purified the C39-like domain in *E. coli* with the objective of determining its structure using X-ray crystallography. Here, we describe the preliminary results of the crystallization of the C39-like domain of HlyB from *E. coli*.

2. Materials and methods

2.1. Cloning, expression and purification

E. coli strain BL21 (DE3) pLysS was transformed with the vector pET28b harbouring the HlyB-C39 gene (gene ID Q1R2T6) for overexpression of the C39-like domain (residues 2–144 of HlyB). Since the original sequence of the C39-like domain contained no

methionine, two isoleucines (Ile16 and Ile30) were exchanged for methionines in order to facilitate the possibility of introducing selenomethionine and providing a tool for structure determination by anomalous diffraction. Additionally, the construct contained an N-terminal His_{10} tag and a linker (primary sequence SENLYFQ-GAMANS). Cells were grown at 310 K and 180 rev min^{-1} in rich medium (2×YT) to an OD_{600} of 0.6 before expression of the C39-like domain was induced by the addition of 1 m*M* isopropyl *β*-D-1-thio-galactopyranoside (IPTG). All subsequent steps were performed at 277 K. Cells from 4 l bacterial cell culture were harvested by centrifugation 5 h after induction and resuspended in 60 ml 50 m*M* HEPES pH 7.0, 150 m*M* NaCl, 10%(*v*/*v*) glycerol. After addition of DNAse (0.1 mg ml^{-1}), the cells were incubated for 20 min and disrupted three times at 2.5 MPa using a cell disrupter (Constant Systems). The homogenate was centrifuged for 30 min at 30 000*g*. Expression of the C39-like domain resulted in the formation of inclusion bodies. The pellet of inclusion bodies was washed once with 50 m*M* Tris–HCl pH 7.0, 0.5% Triton X-100, 1 m*M* dithiothreitol (DTT), 0.05% NaN_3 and once with 50 m*M* Tris–HCl pH 7.0, 1 m*M* EDTA, 0.05% NaN_3. Following the final centrifugation step, the inclusion bodies were resuspended in 20 ml buffer (6 *M* guanidinium hydrochloride, 10 m*M* Tris–HCl pH 7.0, 100 m*M* KCl, 10 m*M* DTT) until the pellet was dissolved. The protein suspension was centrifuged for 1 h at 50 000*g* and the protein concentration of the supernatant was determined using a NanoDrop device (PeqLab Biotechnology GmbH) using the molecular mass of the C39-like domain (19.46 kDa including the His_{10} tag) and its calculated extinction coefficient (25 440 M^{-1} cm^{-1}; http://expasy.org/tools/protparam.html). The resulting supernatant was diluted to 2 mg ml^{-1} in resuspension buffer and dialyzed overnight in a 1:125 ratio against renaturation buffer consisting of 50 m*M* Tris–HCl pH 8.0, 100 m*M* NaCl, 1 m*M* EDTA, 10%(*v*/*v*) glycerol, 2 m*M* DTT, 0.5%(*w*/*v*) PEG 3350 using dialysis tubing with a molecular-weight cutoff (MWCO) of 6–8 kDa (Spektra/Por, Spectrum Laboratories, California, USA). After dialysis, any precipitate was separated from soluble protein by centrifugation for 10 min at 21 000*g*. The supernatant was applied onto an immobilized metal-ion affinity chromatography (IMAC) column for purification of the C39-like domain. However, the C39-like domain precipitated after elution. Therefore, we established a purification procedure without IMAC. The supernatant was applied onto a cation-exchange (cIEX) chromatography column (5 ml SP HP column, GE Healthcare) pre-equilibrated with buffer *A* (10 m*M* HEPES pH 7.0, 1 m*M* DTT). After washing with five column volumes (CV) of buffer *A*, proteins were eluted with a linear gradient of 40 CV to 100% buffer *B* (buffer *A* containing 500 m*M* NaCl). Proteins eluting at 190 m*M* NaCl (38% buffer *B*) and 375 m*M* NaCl (75% buffer *B*) were concentrated using an Amicon Ultra Centrifugal Filter device (10 kDa molecular-weight cutoff, Millipore) and analyzed by SDS–PAGE and Western blotting with an antibody recognizing the His_{10} tag (according to the protocol of the manufacturer, Qiagen). However, the protein fraction that eluted from the column at 375 m*M* NaCl could not be concentrated above 2 mg ml^{-1}. Nevertheless, both species were applied onto a size-exclusion chromatography (SEC) column (Superdex 200 16/60, GE Healthcare) equilibrated with 50 m*M* Tris–HCl pH 8.0, 150 m*M* NaCl, 1 m*M* DTT.

2.2. Buffer screen

A buffer screen was performed with the C39-like domain fraction that eluted at 375 m*M* NaCl during cIEX chromatography. The protein (2 mg ml^{-1}) was mixed in a 1:1 ratio with commercially available crystallization buffers (Basic Crystallography Kit, Sigma–Aldrich). Buffer conditions with no visible precipitation were analyzed further. 1 ml of the C39-like domain (2 mg ml^{-1}) was dialyzed against 1 l of each buffer positively evaluated in the 96-well format using Slide-A-Lyzer dialysis cassettes (10 kDa molecular-weight cutoff, Thermo Scientific) and subsequently concentrated using filter devices (see above) until precipitation occurred. After purification and dialysis against 50 m*M* citrate pH 6.5 and 20 m*M* LiCl, the C39-like domain could be concentrated to 20 mg ml^{-1}.

2.3. SEC–MALS

SEC–MALS (SEC–multi-angle light scattering) experiments were performed using an analytical Superdex 200 10/300 column (GE Healthcare) equilibrated with 25 m*M* Tris pH 8.0 and 150 m*M* NaCl on an ÄKTA Purifier (GE Healthcare) connected to a triple-angle light-scattering detector (miniDAWN TREOS, Wyatt Technology) followed by a differential refractive-index detector (Optilab rEX, Wyatt Technology). 100 µl purified C39-like domain at a protein concentration of 1 mg ml^{-1} was loaded onto the analytical Superdex 200 10/300 column. Data were analyzed with the *ASTRA* software package (Wyatt Technology).

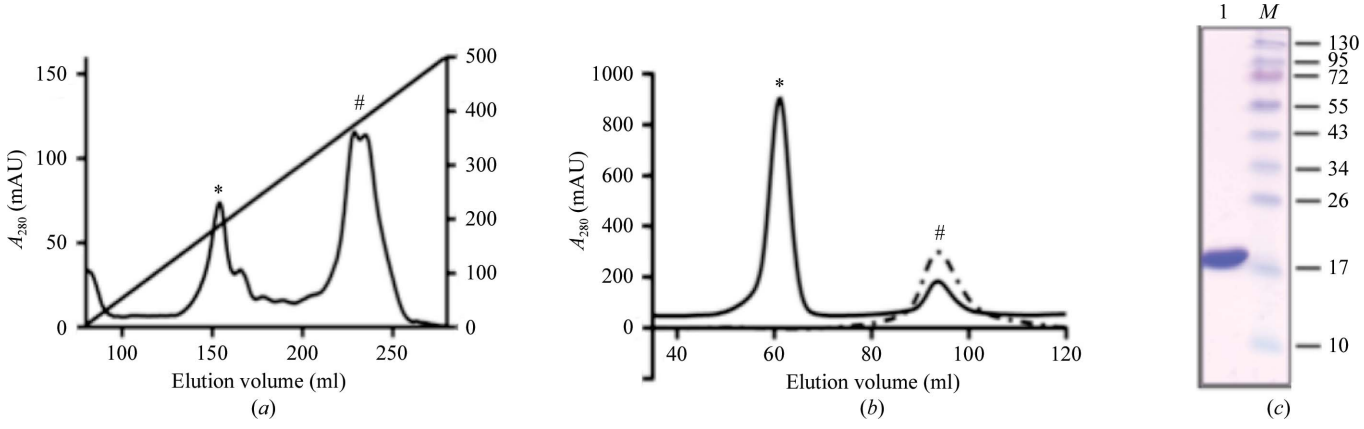

Figure 1
(*a*) cIEX column elution profile of the refolded C39-like domain. One low-salt peak (*) and one high-salt peak (#) were observed. (*b*) Size-exclusion elution profile (Superdex 200 16/60, GE Healthcare) of the low-salt (black line) and high-salt (dashed line) species. The peaks corresponding to oligomeric (*) and monomeric (#) C39-like domain are marked. (*c*) A Coomassie-stained SDS–PAGE gel of the C39-like domain after purification.

2.4. Crystallization and preliminary X-ray analysis

Crystallization conditions were screened using the sitting-drop vapour-diffusion method in combination with commercially available buffers (Sigma Basic Screen I and II) using the monomeric as well as the oligomeric form of the C39-like domain at a protein concentration of 20 mg ml^{-1}. Crystals appeared when samples of protein were mixed with an equal volume of precipitant solution containing 3.6–4 *M* formate. Using the 'no-fail' cryoprotection method (http://capsicum.colgate.edu/chwiki/), 37.5% glucose in mother liquor was added gradually to the crystal solutions to give a final concentration of 20% glucose before flash-cooling the crystals in liquid nitrogen. A high-resolution data set was collected from a single crystal at 100 K at a wavelength of 0.87260 Å using a MAR 225 detector (MAR Research) on ESRF beamline ID23eh1 using 0.75° rotation per frame. The data set was processed using the *XDS* package (Kabsch, 2010) and scaled using *XSCALE*.

3. Results and discussion

The C39-like domain of HlyB was expressed as inclusion bodies in *E. coli*. After preparation and crude purification of the inclusion bodies, proteins were denatured and refolded. A cIEX step was used to purify the refolded C39-like domain (see §2). During elution with a salt gradient, two species eluted at 190 m*M* NaCl (38% buffer *B*) and 375 m*M* NaCl (75% buffer *B*) (Fig. 1). SDS–PAGE and Western-blot analysis using an anti-His-tag antibody showed that both eluted fractions corresponded to the C39-like domain (data not shown). Therefore, both species were concentrated and applied onto an SEC column. However, the resulting elution chromatograms were different. The low-salt species (190 m*M* NaCl) gave two peaks (61 and 93 ml) corresponding to an oligomeric and a monomeric species, respectively. Based on a calibration curve, the molecular weights were calculated as approximately 200 and 20 kDa, respectively. The higher oligomeric fraction of the C39-like domain yielded approximately six times more protein: 9 mg per litre of cells for the oligomer compared with 1.5 mg per litre of cells for the monomeric species. For further investigation of the molecular weight of the low-salt oligomeric species, more precise MALS experiments were performed and the molecular weight was determined to be 233.4 ± 1.4 kDa, corresponding to a dodecamer of the C39-like domain (Fig. 2). Analysis of the high-salt species (375 m*M* NaCl) by size-exclusion chromatography revealed that the species corresponded to a monomeric C39-like domain only. After purification, the C39-like domain was obtained without any visible impurities on an SDS–PAGE gel (Fig. 1*c*).

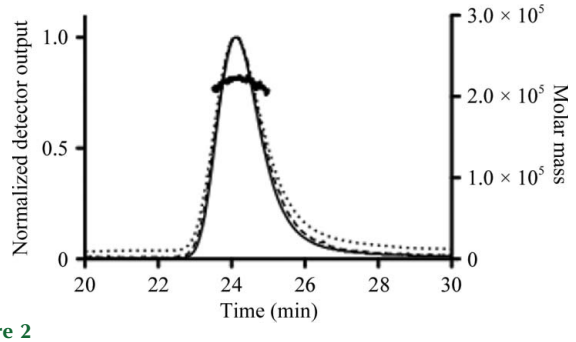

Figure 2
The oligomeric state of the C39-like domain was analyzed by SEC–MALS. The elution profile from a Superdex 200 10/300 column is shown. Straight line, normalized UV signal; dashed line: normalized LS (90°) signal; dotted line, refractive index signal. Filled circles indicate the corresponding molar masses.

Table 1
Data-collection statistics.

Values in parentheses are for the highest resolution shell.

No. of crystals	1
Beamline	ID23eh1, ESRF
Wavelength (Å)	0.87260
Detector	MAR Mosaic 225 mm
Crystal-to-detector distance (mm)	200
Rotation range per image (°)	0.75
Total rotation range (°)	230
Exposure time per image (s)	1
Resolution range (Å)	20–2.1 (2.2–2.1)
Space group	$P2_12_12$ or $P2_12_12_1$
Unit-cell parameters	
a (Å)	117.6
b (Å)	134.2
c (Å)	139.4
$\alpha = \beta = \gamma$ (°)	90.0
Mosaicity (°)	0.65
Total No. of measured intensities	720832
Unique reflections	240680
Multiplicity	2.9
Completeness (%)	95.6 (97.3)
$R_{\rm merge}$ (%)†	9.0 (21.9)
$\langle I/\sigma(I)\rangle$	9.4 (4.6)
Overall B factor from Wilson plot (Å^2)	24.5

† $R_{\rm merge}$ is defined as $\sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle|/\sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the ith observation of reflection hkl and $\langle I(hkl)\rangle$ is the weighted average intensity for all observations i of reflection hkl.

The monomeric species precipitated at the relatively low protein concentration of 2 mg ml^{-1}. In our hands, this concentration was too low to obtain any protein crystals since more than 90% of the drops in the crystallization experiments did not show any precipitation at such a low protein concentration. The buffer screen performed revealed a buffer condition (20 m*M* citrate pH 6.5 and 20 m*M* $NiCl_2$) in which the protein was very stable (to above 20 mg ml^{-1}). However, despite extensive efforts no well diffracting crystals of the C39-like domain could be obtained.

Therefore, we went back to the initial oligomeric species and crystallized this form of the C39-like domain containing a His_{10} tag. Crystals were observed in various screening conditions. The best diffracting crystals (Fig. 3) were obtained after 30 d at 285 K in 3.8 *M* sodium formate. For diffraction analysis, crystals were cryoprotected using the 'no-fail' cryoprotection method (http://capsicum.colgate.edu/chwiki/). Here, crystal-containing drops were substituted step by step with small volumes of 37% glucose dissolved in mother liquor to give a final glucose concentration of 20%. Using this method, well diffracting crystals could be obtained and a native data set was collected.

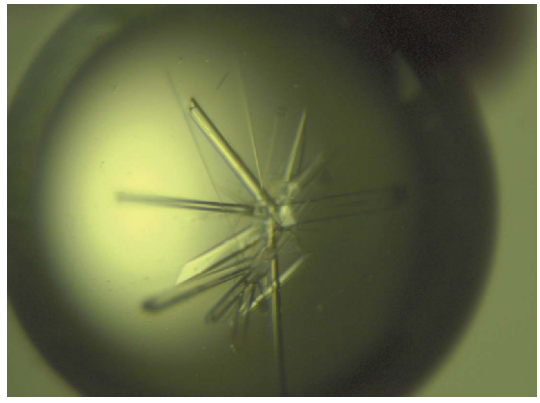

Figure 3
Crystal of the C39-like domain. The crystal has dimensions of 80 × 80 × 100 µm.

Preliminary analysis of the data set using the *XDS* package revealed that the crystals belonged to a primitive orthorhombic space group (either $P2_12_12$ or $P2_12_12_1$) with large unit-cell parameters (a = 117.6, b = 134.2, c = 139.4 Å) for a protein of this size. At the moment we cannot distinguish between the two space groups since the systematic absences observed are not conclusive. A self-rotation function was calculated and revealed three twofold axes perpendicular to each other, indicating an orthorhombic space group. The presence of substantial peaks on the $\chi = 180^\circ$ and $\chi = 120^\circ$ sections of the self-rotation function indicated the presence of twofold and threefold rotational symmetry in the expected dodecamer of the asymmetric unit. Additionally, no significant peaks could be detected in the $\chi = 60^\circ$ or $\chi = 90^\circ$ sections, indicating a lack of fourfold or sixfold symmetry. The overall data set was cut off at a resolution of 2.1 Å with an $R_{\rm merge}$ value in the highest resolution shell of below 30% (see Table 1 for data statistics). Consideration of the unit-cell volume (2.2×10^6 Å^3) and the molecular weight of 19 460 Da suggests that the asymmetric unit contains one copy of the dodecamer, corresponding to a $V_{\rm M}$ value of 2.36 Å^3 Da^{-1} with an estimated solvent content of 47.61% (Matthews, 1968).

This suggests that the initially observed oligomeric species is likely to be the state in which the C39-like domain is sufficiently stable and homogenous to crystallize under the conditions used in this study.

An attempt to solve the structure of the C39-like domain taking advantage of the introduced selenomethionine residues by using multiple anomalous dispersion (MAD) phasing is currently under way. The result of this structural analysis should provide an insight into the function of the C39-like domain of HlyB and its role in the secretion of HlyA.

We would like to acknowledge Solvej Siedler for important assistance during the initial stages of this project. This work was funded by the EDICT (European Initiative on Channels and Transporters) consortium FP7 Theme (Health-2007-2.1.1-5) to BT and LS. We thank the Ministry of Innovation, Science and Research of the German Federal State North Rhine-Westphalia (NRW) and the Heinrich Heine University Düsseldorf (scholarship from the CLIB-Graduate Cluster Industrial Biotechnology to CKWS).

References

Balakrishnan, L., Hughes, C. & Koronakis, V. (2001). *J. Mol. Biol.* **313**, 501–510.
Benabdelhak, H., Kiontke, S., Horn, C., Ernst, R., Blight, M. A., Holland, I. B. & Schmitt, L. (2003). *J. Mol. Biol.* **327**, 1169–1179.
Gray, L., Mackman, N., Nicaud, J. M. & Holland, I. B. (1986). *Mol. Gen. Genet.* **205**, 127–133.
Ishii, S., Yano, T., Ebihara, A., Okamoto, A., Manzoku, M. & Hayashi, H. (2010). *J. Biol. Chem.* **285**, 10777–10785.
Kabsch, W. (2010). *Acta Cryst.* D**66**, 125–132.
Mackman, N., Baker, K., Gray, L., Haigh, R., Nicaud, J. M. & Holland, I. B. (1987). *EMBO J.* **6**, 2835–2841.
Mackman, N., Nicaud, J. M., Gray, L. & Holland, I. B. (1986). *Curr. Top. Microbiol. Immunol.* **125**, 159–181.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Wu, K.-H. & Tai, P. C. (2004). *J. Biol. Chem.* **279**, 901–909.