

# The A component (SmhA) of a tripartite pore-forming toxin from *Serratia marcescens*: expression, purification and crystallographic analysis

Alicia M. Churchill-Angus,<sup>a</sup> Svetlana E. Sedelnikova,<sup>a</sup> Thomas H. B. Schofield<sup>a,b</sup> and Patrick J. Baker<sup>a\*</sup>

Received 27 August 2020

Accepted 17 October 2020

Edited by M. W. Bowler, European Molecular Biology Laboratory, France

**Keywords:** pore-forming toxin; crystallization; ClyA family; *Serratia*.

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

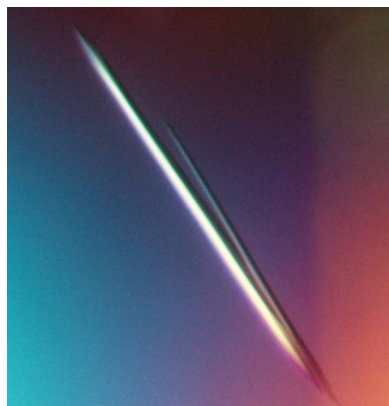
<sup>a</sup>Department of Molecular Biology and Biotechnology, The University of Sheffield, Western Bank, Sheffield S10 2TN, United Kingdom, and <sup>b</sup>Astbury Centre for Structural Molecular Biology, University of Leeds, Leeds LS2 9JT, United Kingdom. \*Correspondence e-mail: p.baker@sheffield.ac.uk

Tripartite  $\alpha$ -pore-forming toxins are constructed of three proteins (A, B and C) and are found in many bacterial pathogens. While structures of the B and C components from Gram-negative bacteria have been described, the structure of the A component of a Gram-negative  $\alpha$ -pore-forming toxin has so far proved elusive. SmhA, the A component from the opportunistic human pathogen *Serratia marcescens*, has been cloned, overexpressed and purified. Crystals were grown of selenomethionine-derivatized protein and anomalous data were collected. Phases were calculated and an initial electron-density map was produced.

## 1. Introduction

Tripartite  $\alpha$ -pore-forming toxins ( $\alpha$ -PFTs) are members of the ClyA  $\alpha$ -pore-forming toxin family (Fagerlund *et al.*, 2008; Wilson *et al.*, 2019); however, unlike ClyA, where the pore is formed from an oligomer of a single protomer, three proteins (A, B and C) are involved in active pore formation (Sastalla *et al.*, 2013; Lindbäck *et al.*, 2004; Wilson *et al.*, 2019; Beecher & Macmillan, 1991). In the ClyA family the active pore is formed when the soluble protein(s) undergo a large-scale conformational change to expose the membrane-binding regions, with protomers assembling into a hydrophilic-lined pore (Benke *et al.*, 2015; Roderer & Glockshuber, 2017). It is proposed that each protein of the tripartite  $\alpha$ -PFT (A, B and C) fulfils a role in the active pore that is provided by different regions of the ClyA protomer (Wilson *et al.*, 2019). The C component makes the first attachment to the target cell, binding to a single leaflet of the membrane, and is equivalent in function to the  $\beta$ -tongue region of soluble ClyA. In ClyA the pore is completed by the N-terminal amphipathic helix of each protomer assembling to construct the membrane-spanning, hydrophilic-lined pore of the oligomer (Roderer & Glockshuber, 2017; Wallace *et al.*, 2000; Benke *et al.*, 2015). In the tripartite  $\alpha$ -PFTs the A and B components are functionally equivalent to this region of ClyA. The B component acts as the pore-forming unit, using two hydrophobic helices to span the membrane, with the A component proposed to provide amphipathic helices that produce the hydrophilic interior lining of the oligomeric pore (Wilson *et al.*, 2019; Mueller *et al.*, 2009; Benke *et al.*, 2015).

Tripartite  $\alpha$ -PFTs were first identified in the pathogenic Gram-positive bacterium *Bacillus cereus*, when the Hbl



OPEN ACCESS

system, and later the NheABC system, were identified as vital toxins in its pathogenicity and the cause of a major food-poisoning outbreak in Norway (Thompson *et al.*, 1984; Lund & Granum, 1996; Beecher *et al.*, 1995). The tripartite  $\alpha$ -PFT family has recently been expanded into a large number of clinically and economically important Gram-negative bacteria, including the fish and opportunistic human pathogen *Aeromonas hydrophila* (Wilson *et al.*, 2019). The  $\alpha$ -PFT toxin AhlABC from *A. hydrophila*, like NheABC, has been shown to be lytic to mammalian cells and forms pores in membranes (Wilson *et al.*, 2019; Lindbäck *et al.*, 2004).

Structures of soluble AhlB and AhlC, and also a pore structure of AhlB, have been solved by X-ray crystallography; however, a structure of the A component from a Gram-negative  $\alpha$ -PFT has yet to be determined (Wilson *et al.*, 2019) and thus the structural role of this protein in the active pore is as yet unknown. Within the Gram-positive *B. cereus*  $\alpha$ -PFTs, only the structures of NheA (PDB entry 4k1p; Ganash *et al.*, 2013) and HblB (PDB entry 2nrj; Madegowda *et al.*, 2008) have been determined. HblB is functionally equivalent to AhlC, yet these two proteins share less than 10% sequence identity and their structures are significantly different (Wilson *et al.*, 2019). Similarly, NheA and AhlA share only 6% sequence identity, and thus structures of the A component from the Gram-negative bacterial  $\alpha$ -PFT systems may also vary substantially from that of NheA.

*Serratia marcescens* is a nosocomial human-pathogenic Gram-negative bacteria (Su *et al.*, 2003; Kurz *et al.*, 2003; Iguchi *et al.*, 2014). Genomic analysis has shown that it possesses an  $\alpha$ -PFT with three proteins (SmhABC) homologous to the AhlABC proteins (Wilson *et al.*, 2019).

In this paper, we present the overexpression, purification and crystallization of SmhA and show the first electron-density map for an A component of a Gram-negative tripartite  $\alpha$ -PFT.

## 2. Materials and methods

### 2.1. Macromolecule production

**2.1.1. Cloning and overexpression.** The open reading frame for SmhA from *S. marcescens* MSU97 (NCBI accession No. OKB64935.1) was synthesized and cloned into the pET-21a expression vector by GenScript to contain a C-terminal His<sub>6</sub> tag.

The plasmid was transformed into an *Escherichia coli* BL21 (DE3) expression cell line (NEB). One colony was used to inoculate a 250 ml flask containing 50 ml Luria–Bertani (LB) broth supplemented with 100  $\mu$ g ml<sup>-1</sup> ampicillin and was grown overnight at 37°C. 10 ml of this overnight culture was then used to inoculate 500 ml LB broth supplemented and incubated as described above until an OD<sub>600</sub> of 0.6 was reached, at which point protein expression was induced by the addition of 1 mM isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG). Protein expression was carried out overnight at 16°C.

To prepare selenomethionine-incorporated SmhA, 2  $\times$  500 ml of cells were grown as described above and harvested

**Table 1**  
Macromolecule-production information.

Source organism	<i>S. marcescens</i> MSU97
Restriction sites	NdeI/XhoI
Cloning vector	pET-21a(+)
Expression vector	pET-21a(+)
Expression host	<i>E. coli</i> BL21 (DE3)
Complete amino-acid sequence of the construct produced	MNNLTSIDLSPQTLMMHHSISSQALLNQS YNNLLSQLLTSQSM DPGLTVKIKAYQ NQLRQQAQVFKQNTVAELIGLYTKASNF AALVNAVNALYSTEDPQVSQKGAEMVAA LSDVAQHYQAAAQAVHTQLQAKREMLEP LMGNFLNVIDAIEQGLNABEAKQQAQTIA ELNEAIKNIQSIADAGFKAGEGVVQLG QSIVAAVPLGPTDKKPKAPTAPPKPLS DQASYMISGIQAI SAGASGAQQAVNELK ANYAKLAVAYRALATANALLSVAKSVA QAQLFVDTYVLTEQRMALLPTEWGVKVAE AYLTAAPITNQAGSAAEIKQAKQIISLN AEKWQLFSKSIDNAKANYAGNNILPEVL EHHHHHH

prior to induction. The cells were washed and resuspended in selenomethionine minimal medium [10.5 g l<sup>-1</sup> K<sub>2</sub>HPO<sub>4</sub>, 1.0 g l<sup>-1</sup> (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 4.5 g l<sup>-1</sup> KH<sub>2</sub>PO<sub>4</sub>, 0.5 g l<sup>-1</sup> trisodium citrate-2H<sub>2</sub>O, 5.0 g l<sup>-1</sup> glycerol and 0.5 g l<sup>-1</sup> each of adenine, guanosine, thymine and uracil; medium A] and added to 2  $\times$  500 ml of medium A supplemented with 1.0 g l<sup>-1</sup> MgSO<sub>4</sub>·7H<sub>2</sub>O, 4.0 mg l<sup>-1</sup> thiamine; 100 mg l<sup>-1</sup> each of L-lysine, L-phenylalanine and L-threonine; 50 mg l<sup>-1</sup> each of L-isoleucine, L-leucine and L-valine; and 40 mg l<sup>-1</sup> seleno-L-methionine. Growth was continued until an OD<sub>600</sub> of 0.6 was reached before induction with 1 mM IPTG. The protein was expressed overnight at 16°C. The cells were harvested and pelleted before storage at -25°C.

**2.1.2. Purification.** Harvested cells of either native or selenomethionine-derivatized (SeMet) SmhA were defrosted, resuspended in lysis buffer (50 mM Tris pH 8.0) and lysed by sonication (3  $\times$  20 s bursts at 16  $\mu$ m amplitude). Insoluble material was removed by centrifugation at 40 000g for 15 min. The supernatant was applied onto a 5 ml nickel HiTrap column (GE Healthcare) in binding buffer (50 mM Tris pH 8.0, 0.5 M NaCl). The protein was eluted with a linear gradient of 0–1 M imidazole in binding buffer and fractions containing protein were pooled, concentrated and buffer-exchanged into 50 mM Tris pH 8.0, 10 mM NaCl for crystallization; the purification was analysed by SDS–PAGE. Macromolecule-production information is summarized in Table 1.

### 2.2. Crystallization

Purified SmhA was concentrated to 7 mg ml<sup>-1</sup> for crystallization using a Vivaspin 30 kDa molecular-weight cutoff concentrator (Sartorius). The concentrated protein was used to set up 96-well sitting-drop crystallization trials using a TTP LabTech Mosquito LCP robot, with both 200 nl:200 nl and 200 nl:100 nl well solution:protein solution drops, and stored at 7°C. Crystallization-condition suites used for preliminary screens included JCSG+, PACT *premier*, MPD, Morpheus, ProPlex and AmSO4 (Qiagen and Molecular Dimensions).

Initial crystals of both native and SeMet SmhA grew in PACT *premier* condition B11 (0.2 M MES pH 6, 0.2 M CaCl<sub>2</sub>, 20% PEG 6000). Optimization (using a Formulatrix Formulator robot) of the SeMet SmhA crystals around PACT *premier* condition B11 in a 96-well sitting-drop plate with 200 nl:200 nl drops gave larger more defined crystals from 0.1 M MES pH 6.1, 0.14 M CaCl<sub>2</sub>, 21% PEG 6000 (Fig. 1). Crystallization information is summarized in Table 2.

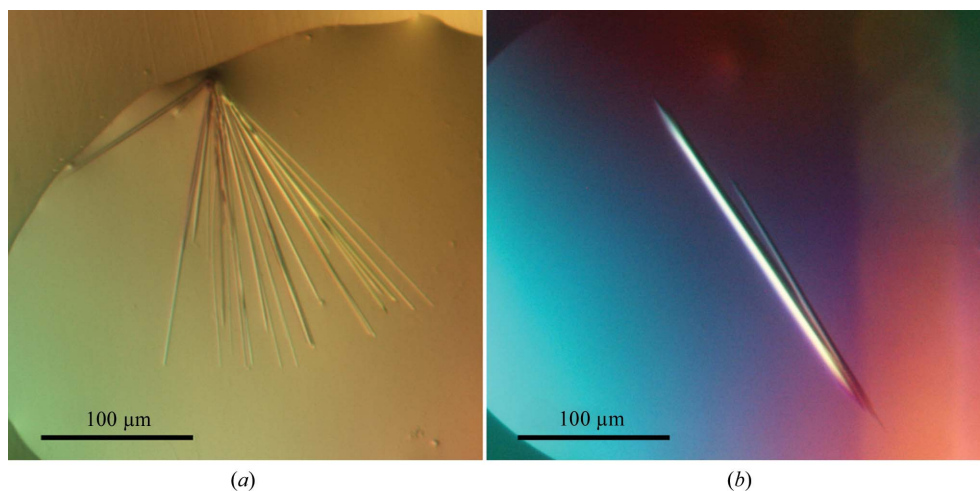
### 2.3. Data collection and processing

A single SeMet SmhA crystal was flash-cooled in liquid nitrogen using a cryoprotectant consisting of 20% ethylene

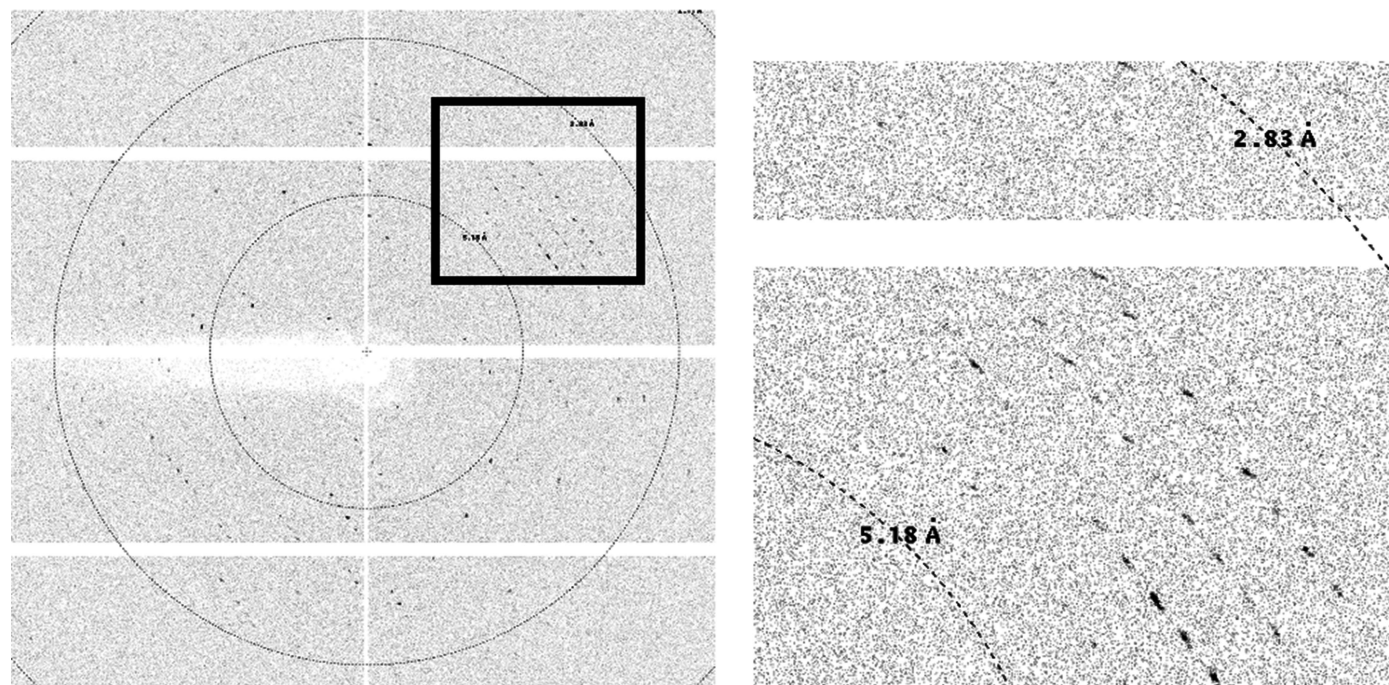
**Table 2**  
Crystallization.

Method	Sitting-drop
Plate type	96-well sitting drop
Temperature (K)	280
Protein concentration (mg ml <sup>-1</sup> )	7
Buffer composition of protein solution	50 mM Tris pH 8.0, 10 mM NaCl
Composition of reservoir solution	0.2 M MES pH 6, 0.2 M CaCl <sub>2</sub> , 20% PEG 6000
Volume and ratio of drop	200 nl:200 nl
Volume of reservoir (μl)	50

glycol, 0.2 M MES pH 6, 0.2 M CaCl<sub>2</sub>, 20% PEG 6000 and data were collected at the selenium absorption edge (0.9792 Å) on beamline I03 at Diamond Light Source (DLS; Fig. 2). Data



**Figure 1**  
(a) Native SmhA crystals grown in PACT *premier* condition B11 (0.2 M MES pH 6, 0.2 M CaCl<sub>2</sub>, 20% PEG 6000). (b) SeMet SmhA crystals grown in optimized conditions based on PACT *premier* condition B11 (0.1 M MES pH 6.1, 0.14 M CaCl<sub>2</sub>, 21% PEG 6000).



**Figure 2**  
A representative 0.1° oscillation image from an SeMet SmhA crystal collected using an EIGER2 XE 16M detector on beamline I03 at Diamond Light Source. An enlarged view of the region highlighted by the square shows that diffraction extends to around 3.3 Å resolution.

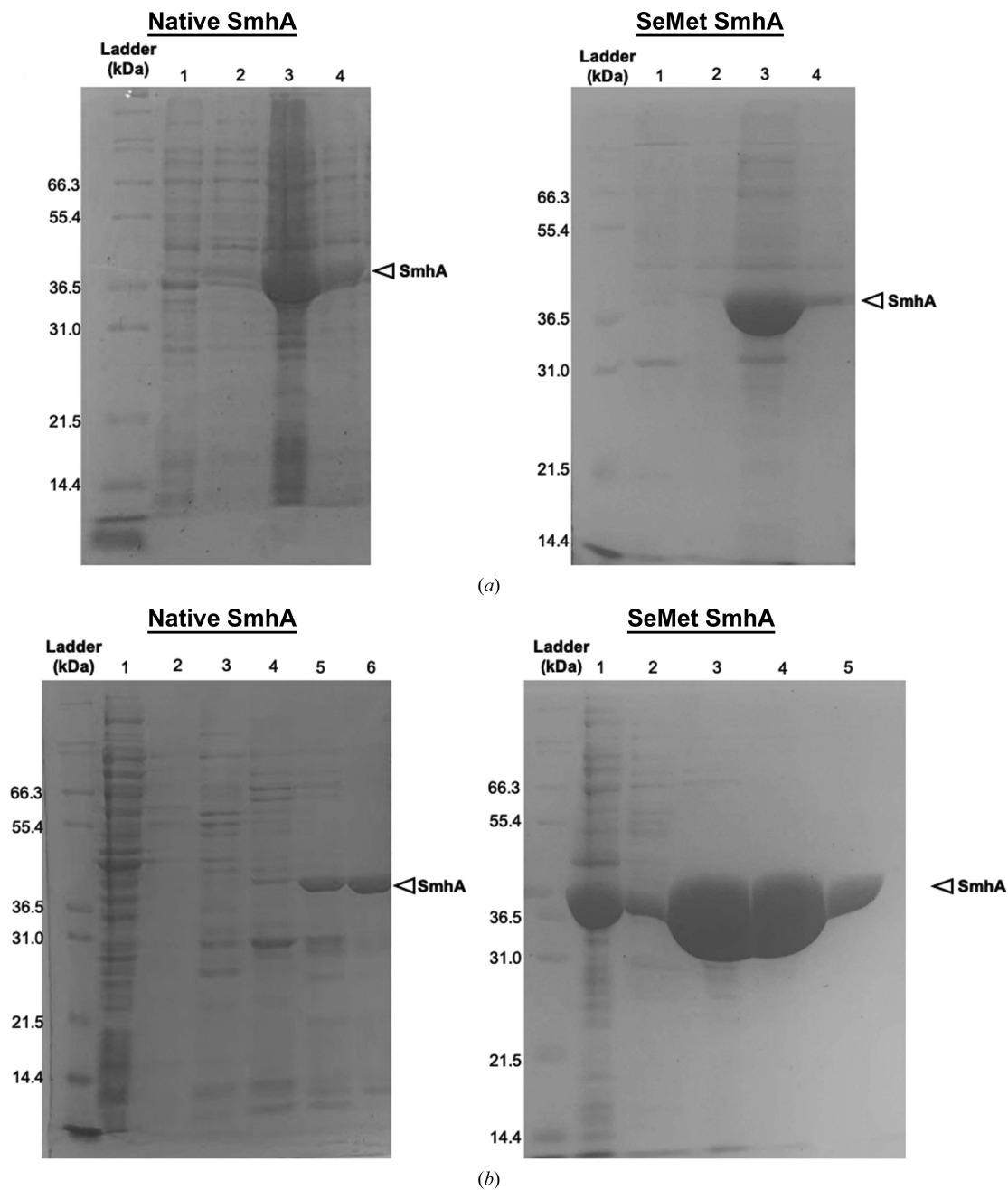
were processed to 3.3 Å resolution using the *xia2/DIALS* pipeline (Winter *et al.*, 2018) and showed that the crystal belonged to space group  $P4_2$ , with unit-cell parameters  $a = b = 151.8$ ,  $c = 134.0$  Å (Table 3).

### 3. Results and discussion

#### 3.1. Construct design

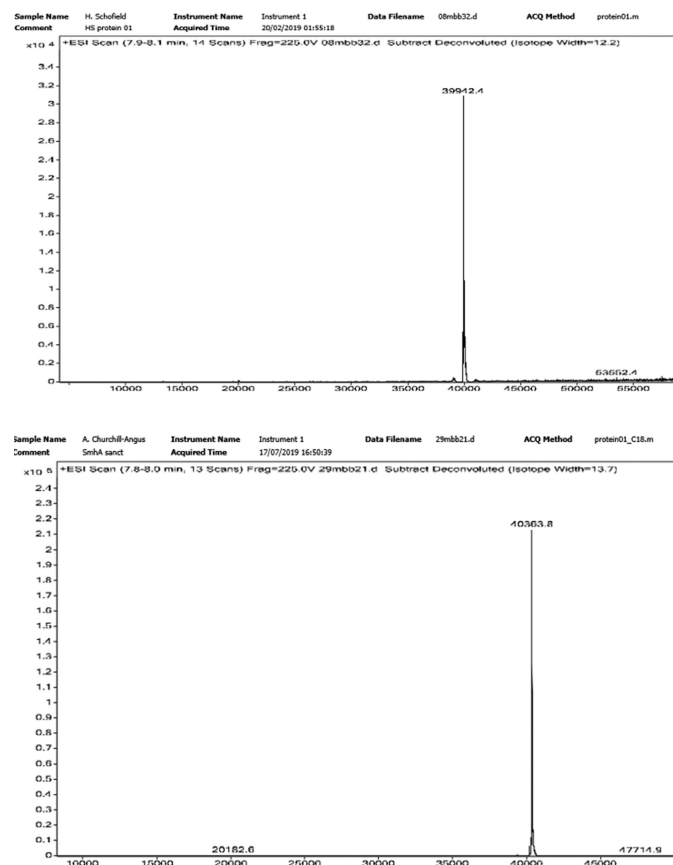
Structural studies of the tripartite toxins have been hampered in part by difficulties in producing large quantities

of stable protein. The expression of *A. hydrophila* AhlA using constructs generated from genomic DNA in *E. coli* BL21 cells (Wilson *et al.*, 2019) produced protein for assays, but the yield was low and insufficient for crystallization. Pairwise sequence alignment of AhlA with *S. marcescens* SmhA shows 43% identity and 53% similarity (Supplementary Fig. S1), identifying SmhA as a good candidate for structural studies of the A component from a Gram-negative bacterium. As with AhlA, initial attempts to overexpress SmhA using constructs from genomic DNA also proved unsuccessful. To try to improve expression in *E. coli* BL21, the SmhA gene (NCBI accession

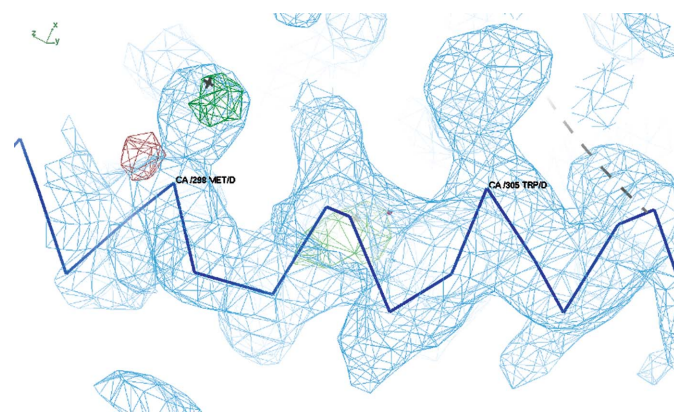


**Figure 3**  
 (a) SDS-PAGE gels showing overexpression of native and SeMet SmhA. Lane 1, pre-induction insoluble fraction; lane 2, pre-induction soluble fraction; lane 3, post-induction insoluble fraction; lane 4, post-induction soluble fraction. (b) SDS-PAGE gels showing nickel HiTrap column purification of native and SeMet SmhA. Lane 1, cell-free extract; lanes 2–5/6, elution fractions from the nickel HiTrap column. Fraction 6 and fractions 4 and 5 (native and SeMet, respectively) were >90% pure and were used for crystallization.

No. OKB64935.1) was synthesized and optimized in both GC content and codon usage for expression in *E. coli* (GenScript). This resulted in 38% of the codons being altered (Supplementary Fig. S2). This new construct was successfully used to express SmhA in *E. coli* with a C-terminal 6×His tag and,



**Figure 4**  
Mass spectrum for native SmhA (top) and SeMet SmhA (bottom) as used for crystallization. A molecular weight of 40 363.8 Da for SeMet SmhA (the molecular weight of native SmhA with a His<sub>6</sub> tag is 39 942.4 Da) shows the incorporation of selenium at all nine methionine sites.



**Figure 5**  
Initial electron-density map contoured at  $1.0\sigma$  (blue) and an anomalous difference map (positive, green) showing a helical section of SmhA. Density can be seen for the side chains of Trp305 and Met298, with positive difference for the Se atom (shown as a cross) in Met298.

**Table 3**  
Data collection and processing.

Values in parentheses are for the outer shell.

Diffraction source	I03, DLS
Wavelength (Å)	0.9792
Temperature (K)	100
Detector	EIGER2 XE 16M
Rotation range per image (°)	0.1
Total rotation range (°)	360
Exposure time per image (s)	0.008
Space group	<i>P</i> <sub>4</sub> <sub>2</sub>
<i>a</i> , <i>b</i> , <i>c</i> (Å)	151.8, 151.8, 134.0
$\alpha$ , $\beta$ , $\gamma$ (°)	90, 90, 90
Mosaic spread (°)	0.19
Resolution range (Å)	67.9–3.34 (3.40–3.34)
Total No. of reflections	612546 (30358)
No. of unique reflections	44270 (2175)
Completeness (%)	100 (98.5)
Multiplicity	13.8 (14.0)
$\langle I/\sigma(I) \rangle^\dagger$	4.7 (1.0)
$R_{\text{r.i.m.}}$	0.118 (0.759)
Overall <i>B</i> factor from Wilson plot (Å <sup>2</sup> )	56.626

<sup>†</sup> The high-resolution cutoff for the data was automatically determined in the *DIALS* pipeline, which uses  $CC_{1/2} = 0.5$  as the limit of usable data. For the outer shell the mean  $\langle I/\sigma(I) \rangle$  is 2.0 at 3.7 Å resolution.

following nickel column purification, resulted in protein with >90% purity and a good yield (9.5 mg l<sup>-1</sup>; Fig. 3).

### 3.2. Data analysis of SmhA

Analysis of the Matthews coefficient for SeMet SmhA showed that the asymmetric unit most likely contained between six and ten molecules with a solvent content between 63% and 38%, respectively, with eight molecules being the most probable, with a  $V_M$  value of 2.39 Å<sup>3</sup> Da<sup>-1</sup> and a solvent content of 48% (Matthews, 1968; Kantardjieff & Rupp, 2003). Mass spectrometry showed that the molecular weight of SeMet SmhA was 40 363.8 Da, which is 421.4 Da more than the native sample, indicating full incorporation of Se atoms into the nine methionine residues of the protein and also indicating that the protein was of high purity (Fig. 4). In order to maximize the quality of the single-wavelength selenium anomalous signal, data-collection parameters were chosen to minimize radiation damage, whilst still providing good multiplicity, albeit at the expense of resolution. A beam size of 80 × 20 μm was selected to match the dimensions of the crystal, with a beam transmission of 20%, giving a flux of 7.44 × 10<sup>11</sup> photons s<sup>-1</sup>. 3600 images of 0.1° and an exposure of 0.008 s gave a data set with a half-set correlation coefficient of 1.0, an anomalous multiplicity of 7.0 and an anomalous correlation coefficient of 0.2 with no obvious signs of radiation damage. Selenium positions were calculated from these Se SAD data, and an initial density map and model were generated using the *CRANK2* pipeline (Skubák & Pannu, 2013; Fig. 5). A preliminary initial model of SmhA, placing 2809 residues assigned to 28 fragments, was automatically built into the electron density. Visual inspection of the map and model confirmed that eight molecules were present in the asymmetric unit, with the side-chain positions of the methionine residues clearly aligned with the positive density of the

anomalous difference map. A self-rotation function calculated using the data between 50 and 6 Å resolution showed the presence of a noncrystallographic twofold axis perpendicular to the crystallographic fourfold (peak of 88% of the origin at polar coordinates 90.0, 111.2, 180°). Inspection of the initial model showed that two subunits were related by this rotation axis, but were separated by 20 Å. In addition, a self-Patterson indicated the presence of noncrystallographic translational symmetry with a peak of 43% of the origin at 0, 0, 0.467 and the model showed that six of the eight molecules were related by this translational symmetry. However, despite these non-crystallographic symmetry relationships, no high-order oligomeric arrangement could be observed for the eight subunits, indicating that the structure of SmhA was of the monomeric soluble form of the protein, rather than an oligomeric structure assembled around a central rotation axis as required for the proposed pore form.

Further work to extend the resolution of the data and refine the SmhA structure is ongoing.

### Acknowledgements

We would like to thank Diamond Light Source for access to beamline I03 under proposal MX17773, which contributed to these results, and the beamline scientists for their assistance with data collection. Mass spectrometry was performed at the ChemMS Mass Spectrometry Centre at the University of Sheffield.

### Funding information

Funding for this research was provided by: University of Sheffield (scholarship to Alicia M. Churchill-Angus; studentship to Thomas H. B. Schofield); EPSRC (grant No. EP/L026872/1 to Professor M. D. Ward).

### References

Beecher, D. J. & Macmillan, J. D. (1991). *Infect. Immun.* **59**, 1778–1784.

Beecher, D. J., Schoeni, J. L. & Wong, A. C. (1995). *Infect. Immun.* **63**, 4423–4428.

Benke, S., Roderer, D., Wunderlich, B., Nettels, D., Glockshuber, R. & Schuler, B. (2015). *Nat. Commun.* **6**, 6198.

Fagerlund, A., Lindbäck, T., Storset, A. K., Granum, P. E. & Hardy, S. P. (2008). *Microbiology*, **154**, 693–704.

Ganash, M., Phung, D., Sedelnikova, S. E., Lindbäck, T., Granum, P. E. & Artymiuk, P. J. (2013). *PLoS One*, **8**, e74748.

Iguchi, A., Nagaya, Y., Pradel, E., Ooka, T., Ogura, Y., Katsura, K., Kurokawa, K., Oshima, K., Hattori, M., Parkhill, J., Sebahia, M., Coulthurst, S. J., Gotoh, N., Thomson, N. R., Ewbank, J. J. & Hayashi, T. (2014). *Genome Biol.* **6**, 2096–2110.

Kantardjieff, K. A. & Rupp, B. (2003). *Protein Sci.* **12**, 1865–1871.

Kurz, C. L., Chauvet, S., Andrès, E., Aurouze, M., Vallet, I., Michel, G. P. F., Uh, M., Celli, J., Filloux, A., De Bentzmann, S., Steinmetz, I., Hoffmann, J. A., Finlay, B. B., Gorvel, J. P., Ferrandon, D. & Ewbank, J. J. (2003). *EMBO J.* **22**, 1451–1460.

Lindbäck, T., Fagerlund, A., Rødland, M. S. & Granum, P. E. (2004). *Microbiology*, **150**, 3959–3967.

Lund, T. & Granum, P. E. (1996). *FEMS Microbiol. Lett.* **141**, 151–156.

Madegowda, M., Eswaramoorthy, S., Burley, S. K. & Swaminathan, S. (2008). *Proteins*, **71**, 534–540.

Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.

Mueller, M., Grauschopf, U., Maier, T., Glockshuber, R. & Ban, N. (2009). *Nature*, **459**, 726–730.

Roderer, D. & Glockshuber, R. (2017). *Philos. Trans. R. Soc. B*, **372**, 20160211.

Sastalla, I., Fattah, R., Coppage, N., Nandy, P., Crown, D., Pomerantsev, A. P. & Leppla, S. H. (2013). *PLoS One*, **8**, e76955.

Skubák, P. & Pannu, N. S. (2013). *Nat. Commun.* **4**, 2777.

Su, L.-H., Ou, J. T., Leu, H.-S., Chiang, P.-C., Chiu, Y.-P., Chia, J.-H., Kuo, A.-J., Chiu, C.-H., Chu, C., Wu, T.-L., Sun, C.-F., Riley, T. V., Chang, B.-J. & The Infection Control Group (2003). *J. Clin. Microbiol.* **41**, 4726–4732.

Thompson, N. E., Ketterhagen, M. J., Bergdoll, M. S. & Schantz, E. J. (1984). *Infect. Immun.* **43**, 887–894.

Wallace, A. J., Stillman, T. J., Atkins, A., Jamieson, S. J., Bullough, P. A., Green, J. & Artymiuk, P. J. (2000). *Cell*, **100**, 265–276.

Wilson, J. S., Churchill-Angus, A. M., Davies, S. P., Sedelnikova, S. E., Tzokov, S. B., Rafferty, J. B., Bullough, P. A., Bisson, C. & Baker, P. J. (2019). *Nat. Commun.* **10**, 2900.

Winter, G., Waterman, D. G., Parkhurst, J. M., Brewster, A. S., Gildea, R. J., Gerstel, M., Fuentes-Montero, L., Vollmar, M., Michels-Clark, T., Young, I. D., Sauter, N. K. & Evans, G. (2018). *Acta Cryst. D* **74**, 85–97.