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

Overexpression, purification and crystallization of a choline-binding protein CbpI from *Streptococcus pneumoniae*

Neil G. Paterson,[a]* Alan Riboldi-Tunicliffe,[a] Timothy J. Mitchell[b] and Neil W. Isaacs[a]

[a]Department of Chemistry and WestCHEM, Glasgow Biomedical Research Centre (GBRC), University of Glasgow, 120 University Place, Glasgow G12 8TA, Scotland, and [b]Division of Infection and Immunity (IBLS), Glasgow Biomedical Research Centre (GBRC), University of Glasgow, 120 University Place, Glasgow G12 8TA, Scotland

Correspondence e-mail: neison@chem.gla.ac.uk

Received 26 April 2005
Accepted 12 May 2006

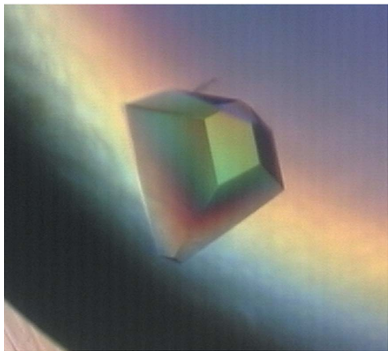
The choline-binding protein CbpI from *Streptococcus pneumoniae* is a 23.4 kDa protein with no known function. The protein has been successfully purified initially using Ni–NTA chromatography and to homogeneity using Q-Sepharose ion-exchange resin as an affinity column. CbpI was crystallized using PEG 3350 as a precipitant and X-ray crystallographic analysis showed that the crystals belonged to the tetragonal space group $P4$, with unit-cell parameters $a = b = 83.31$, $c = 80.29$ Å, $\alpha = \beta = \gamma = 90^\circ$. The crystal contains two molecules in the asymmetric unit with a solvent content of 55.7% ($V_M = 2.77$ Å^3 Da^{-1}) and shows a diffraction limit of 3.5 Å.

1. Introduction

Choline-binding proteins are a family of between ten and 15 surface proteins (dependent on strain; Bergmann & Hammerschmidt, 2006) from *Streptococcus pneumoniae* that share a common surface-attachment domain. This domain facilitates attachment to multiple choline head units of the teichoic and lipoteichoic acids that are found throughout the pneumococcal cell wall (Brundish & Baddiley, 1968; Tomasz, 1967) and is comprised of a highly conserved ∼20-amino-acid repeat with between two and ten repeats forming the domain (Yother & White, 1994). Members of the family show diverse functionality and are often strongly implicated in organism virulence (Gosink *et al.*, 2000). There are also several family members for which no function is available and no sequence similarity exists and these may also be important pneumococcal virulence factors.

CbpI was chosen as part of a pilot structural genomics project at the University of Glasgow studying potential surface virulence factors from *S. pneumoniae* (Wisemann *et al.*, 2001). CbpI is a 211-amino-acid 23.4 kDa protein of unknown function that contains six choline-binding repeats (Garcia *et al.*, 1988) located towards the C-terminus of the protein as identified by *PROSITE* (Gattiker *et al.*, 2002). There is no significant sequence similarity between the N-terminal region and any previously identified sequence.

We report here the purification and crystallization of full-length CbpI from *S. pneumoniae*. A crystal structure of CbpI may elucidate a function for this member of a pathogenically important extracellular protein family.



2. Materials and methods

2.1. Cloning and expression

A PCR product containing the coding region for full-length CbpI was cloned between the *Bam*HI and *Hin*dIII sites of vector pQE-10 (Qiagen) with an in-frame non-cleavable N-terminal His_6 tag and linker (MRGSHHHHHHTDP). Transformation was carried out into *Escherichia coli* strain BL21 (DE3) for expression in the cytoplasm and the cells were grown overnight on LB-agar plates containing 50 µg ml^{-1} ampicillin at 310 K. A single colony was picked and grown in 10 ml LB containing 50 µg ml^{-1} ampicillin overnight at 310 K. Cultures (8 × 1 l) for induction were inoculated using 1 ml of this overnight culture and were grown in LB media containing the selective antibiotic at 293 K. This temperature was essential for

Table 1
Summary of data-collection statistics for CbpI.

Values in parentheses are for the highest resolution shell (3.63–3.50 Å).

X-ray source	ID 14.2, ESRF
Wavelength (Å)	0.9340
Space group	*P*4
Unit-cell parameters	
$a = b$ (Å)	83.31
c (Å)	80.29
$\alpha = \beta = \gamma$ (°)	90.0
Resolution limits (Å)	28.91–3.50
No. of observations	23177
No. of unique observations	12191
Average redundancy	1.90 (1.94)
Completeness (%)	89.7 (90.3)
$\langle I/\sigma(I)\rangle$	6.8 (2.1)
$R_{\rm merge}$	0.091 (0.281)

expression, as no expression was observed at 310 K. Cells were grown to an A_{600} of 0.6, at which point isopropyl thio-β-D-galactoside (IPTG) was added to a final concentration of 1 m*M*. Cultures were grown to an A_{600} of 1.0 and then harvested by centrifugation at 3500*g* and 277 K for 15 min.

2.2. Purification

Cells were resuspended in lysis buffer [50 m*M* Tris–HCl pH 7.5, 500 m*M* NaCl, one EDTA-free protease-inhibitor tablet (Roche) per 25 ml] and lysed by sonication (Status US200 with TT13 tip, 10 × 30 s bursts at 100% power on ice). Cell debris was pelleted by centrifugation at 8000*g* and 277 K for 15 min and any residual cell debris was removed by further centrifugation of the supernatant at 40 000*g* and 277 K for 20 min. The supernatant was loaded onto a 15 ml Ni–NTA (Qiagen) column previously equilibrated with 50 m*M* Tris–HCl pH 7.5, 500 m*M* NaCl and washed with this buffer until a constant A_{280} was observed. Protein was eluted with a gradient to 50 m*M* Tris–HCl pH 7.5, 500 m*M* NaCl, 250 m*M* imidazole and fractions containing CbpI were pooled.

The charged head unit of Q-Sepharose ion-exchange resin (Amersham) is a choline-like moiety (Fig. 1) and it proved possible to use this resin for affinity chromatography in a similar manner to the approach previously reported using DEAE resin (Sanchez-Puelles *et al.*, 1990). Protein previously purified by Ni–NTA affinity was dialysed into 50 m*M* Tris–HCl pH 7.5, 150 m*M* NaCl and loaded onto a column packed with 50 ml Q-Sepharose resin. This column was then washed with a gradient to 50 m*M* Tris–HCl pH 7.5, 1.0 *M* NaCl to remove contaminant proteins. CbpI was eluted by a step to 50 m*M* Tris–HCl pH 7.5, 500 m*M* NaCl, 150 m*M* choline chloride. This resulted in protein that was judged by eye to be over 95% pure by SDS–PAGE analysis (Fig. 2).

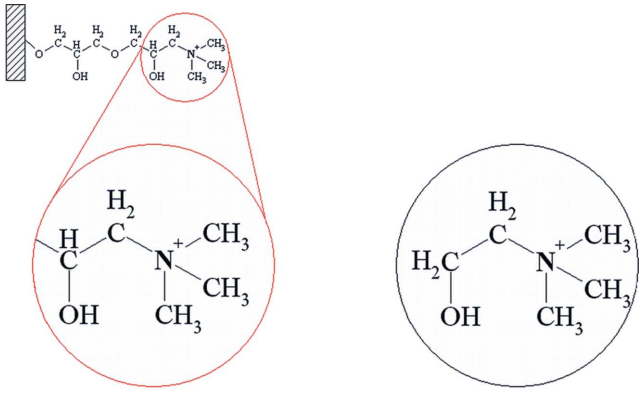


Figure 1
Chemical similarity between Q-Sepharose ion-exchange resin and choline.

A final polishing step and buffer exchange was carried out using gel filtration on a 100 ml Superdex 75 column (Amersham) with 50 m*M* Tris–HCl pH 7.5, 500 m*M* NaCl, 50 m*M* choline chloride and 0.4 m*M* DTT. Purified CbpI was then concentrated to 7 mg ml^{-1} (Amicon, 10 kDa molecular-weight cutoff, calculated from a theoretical extinction coefficient of 89 840 M^{-1} cm^{-1}; Gasteiger *et al.*, 2005). Approximately 45 mg purified protein was obtained from 8 l culture.

2.3. Crystallization and data collection

Crystallization trials were performed using the sitting-drop vapour-diffusion method at 295 K with a protein concentration of 7 mg ml^{-1} in 10 m*M* Tris–HCl pH 7.5, 500 m*M* NaCl, 50 m*M* choline chloride. Initial screening was conducted using Hampton 24-well Cryschem Plates, a drop volume of 1.5 µl protein solution and 1.5 µl reservoir solution and a reservoir volume of 750 µl. Drops were mixed by aspiration. Screening was conducted using Hampton Research Crystal Screens 1 and 2 (Jancarik & Kim, 1991), Emerald Biosystems Wizard 1 and 2 and Cryo 1 and 2, Molecular Dimensions Structure Screen 1 and 2 and a wide range of in-house conditions. Initial screening revealed a single crystal in an in-house condition comprising 20%(*w*/*v*) PEG 3350, 0.2 *M* tetraethylammonium bromide that grew after three months (Fig. 3).

The crystal was flash-cooled to 100 K using dry paraffin oil as the cryoprotectant (Riboldi-Tunnicliffe & Hilgenfeld, 1999) and exposed to synchrotron X-rays at ESRF beamline ID14.2. Diffraction was observed to 3.5 Å (Fig. 4). A data set was collected with an oscillation angle of 0.5° and a summary of data-collection statistics is shown in Table 1. Data were indexed with *LABELIT* (Sauter *et al.*, 2004) and then processed and scaled using the *d*TREK* package (Pflugrath, 1999) in the tetragonal space group *P*4, with unit-cell parameters $a = b = 83.31$, $c = 80.29$ Å, $\alpha = \beta = \gamma = 90.0°$. The data exhibited high

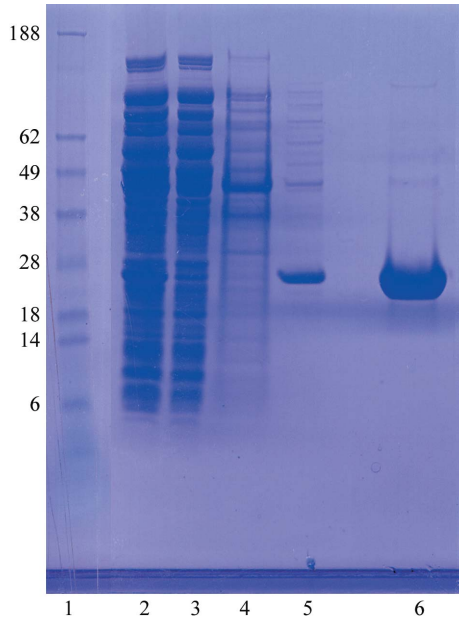


Figure 2
NuPAGE 4–12% Bis-Tris gel (Invitrogen) run with NuPAGE MES buffer. Lane 1, SeeBlue PreStained standard with molecular weights (kDa) given on the left; lane 2, soluble cell fraction; lane 3, Ni–NTA flowthrough; lane 4, Ni–NTA wash; lane 5, Ni–NTA peak elution/Q-Sepharose load; lane 6, Q-Sepharose elution.

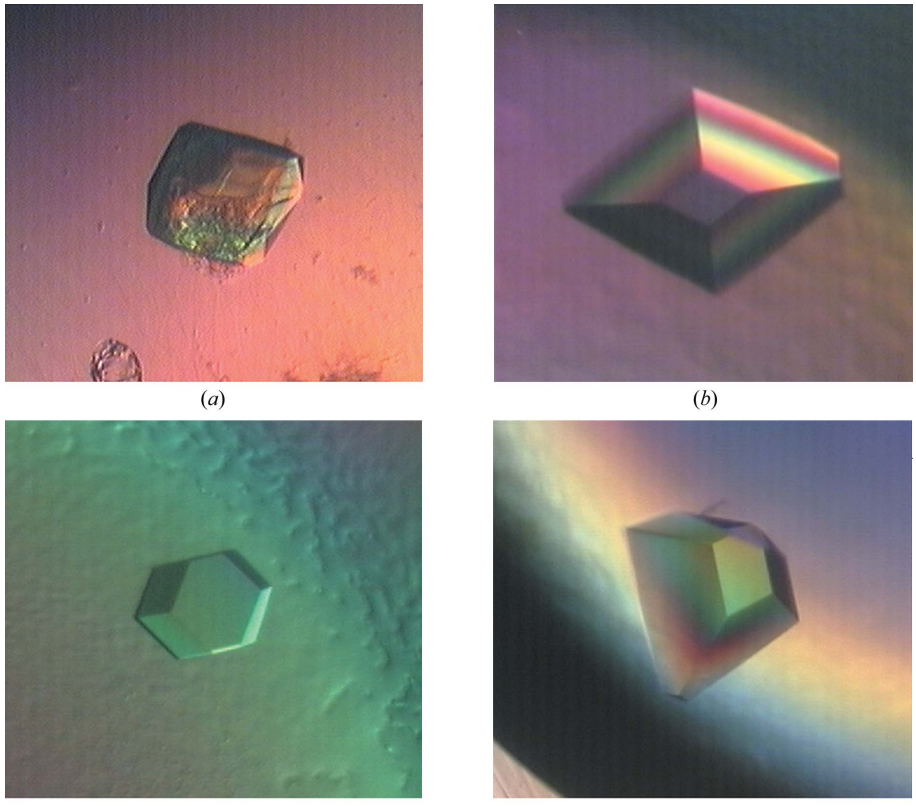


Figure 3
Crystals of CbpI. (*a*) Initial condition; (*b*), (*c*) and (*d*) pH/precipitant variations.

mosaicity, with a value of around 3.5° indicated by *d*TREK*, although data-processing and scaling statistics were reasonable. The crystal contains two molecules in the asymmetric unit, with a solvent content of 55.7% (V_M = 2.77 Å^3 Da^{-1}).

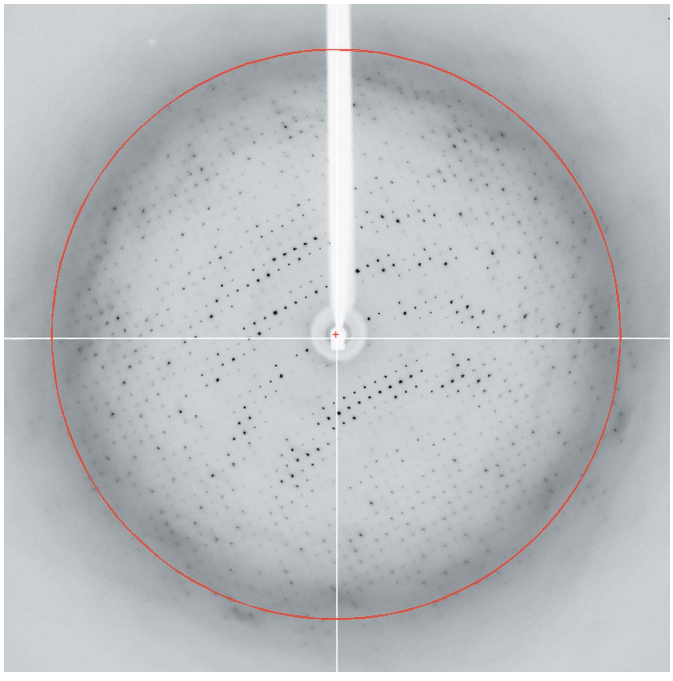


Figure 4
Diffraction image from CbpI. The red circle indicates a resolution of 3.5 Å.

Slight variation of the initial crystallization precipitant concentration and reduction of the pH by addition of 0.1 *M* Tris–HCl pH 7.0 to the crystallization solution yielded crystals of improved appearance with no reduction in growth time (Fig. 3). The diffraction quality of these new crystals was poor, with resolution limits of around 5 Å and apparent high mosaicity. Indexing the data gathered from these new crystals proved impossible.

3. Concluding remarks

Full-length CbpI from *S. pneumoniae* has been successfully expressed, purified and crystallized in space group *P*4 with a diffraction limit of 3.5 Å. Several structures of other choline-binding proteins have already been determined (Fernandez-Tornero *et al.*, 2001; Garau *et al.*, 2005) and work is under way to determine the structure by molecular replacement using choline-binding domain fragments from these structures. Work is also in progress to produce a cleavable His_6 tag construct of the protein with the aim of obtaining crystals of better diffraction quality.

We would like to thank the beamline staff at station ID14.2, ESRF for help and support during data collection. This work was made possible by a studentship awarded to NGP by the BBSRC.

References

Bergmann, S. & Hammerschmidt, S. (2006). *Microbiology*, **152**, 295–303.
Brundish, D. E. & Baddiley, J. (1968). *Biochem. J.* **110**, 573–582.
Fernandez-Tornero, C., Lopez, R., Garcia, E., Gimenez-Gallego, G. & Romero, A. (2001). *Nature Struct. Biol.* **8**, 1020–1024.

Garau, G., Lemaire, D., Vernet, T., Dideberg, O. & Di Guilmi, A. M. (2005). *J. Biol. Chem.* **280**, 28591–28600.
Garcia, E., Garcia, J. L., Garcia, P., Arraras, A., Sanchez-Puelles, J. M. & Lopez, R. (1988). *Proc. Natl Acad. Sci. USA*, **85**, 914–918.
Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D. & Bairoch, A. (2005). *The Proteomic Protocols Handbook*, edited by J. M. Walker, pp. 571–607. Totowa, NJ, USA: Humana Press.
Gattiker, A., Gasteiger, E. & Bairoch, A. (2002). *Appl. Bioinformatics*, **1**, 107–108.
Gosink, K. K., Mann, E. R., Guglielmo, C., Tuomanen, E. I. & Masure, H. R. (2000). *Infect. Immun.* **68**, 5690–5695.
Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.
Pflugrath, J. W. (1999). *Acta Cryst.* D**55**, 1718–1725.
Riboldi-Tunnicliffe, A. & Hilgenfeld, R. (1999). *J. Appl. Cryst.* **32**, 1003–1005.
Sanchez-Puelles, J. M., Sanz, J. M., Garcia, J. L. & Garcia, E. (1990). *Gene*, **89**, 69–75.
Sauter, N. K., Grosse-Kunstleve, R. W. & Adams, P. D. (2004). *J. Appl. Cryst.* **37**, 399–409.
Tomasz, A. (1967). *Science*, **157**, 694–697.
Wisemann, T. M., Heinrichs, J. H., Adamou, J. E., Erwin, A. L., Kunsch, C., Choi, G. H., Barash, S. C., Rosen, C. A., Masure, H. R., Tuomanen, E., Gayle, A., Brewah, Y. A., Walsh, W., Barren, P., Lathigra, R., Hanson, M., Langermann, S., Johnson, S. & Koenig, S. (2001). *Infect. Immun.* **69**, 1593–1598.
Yother, J. & White, J. M. (1994). *J. Bacteriol.* **176**, 2976–2985.