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

Crystallization and preliminary X-ray diffraction studies of choline-binding protein F from *Streptococcus pneumoniae*

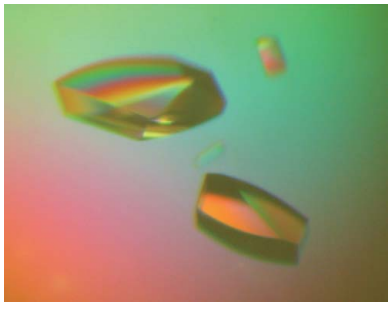
Rafael Molina,[a] Ana González,[b] Miriam Moscoso,[b] Pedro García,[b] Meike Stelter,[c]‡ Richard Kahn[c] and Juan A. Hermoso[a]*

[a]Grupo de Cristalografía Macromolecular y Biología Estructural, Instituto Química Física Rocasolano, CSIC, Serrano 119, 28006 Madrid, Spain, [b]Departamento de Microbiología Molecular, Centro de Investigaciones Biológicas, CSIC, Ramiro de Maeztu 9, 28040 Madrid, Spain, and [c]Institut de Biologie Structurale J.-P. Ebel CEA CNRS UJF, Laboratoire de Cristallographie Macromoléculaire, 41 Rue Jules Horowitz, 38027 Grenoble CEDEX 1, France

‡ Present address: Instituto de Tecnologia Química e Biológica, Universidade Nova de Lisboa, 2781-901 Oeiras, Portugal.

Correspondence e-mail: xjuan@iqfr.csic.es

Received 18 May 2007
Accepted 23 July 2007



Choline-binding protein F (CbpF) is a modular protein that is bound to the pneumococcal cell wall through noncovalent interactions with choline moieties of the bacterial teichoic and lipoteichoic acids. Despite being one of the more abundant proteins on the surface, along with the murein hydrolases LytA, LytB, LytC and Pce, its function is still unknown. CbpF has been crystallized using the hanging-drop vapour-diffusion method at 291 K. Diffraction-quality orthorhombic crystals belong to space group $P2_12_12$, with unit-cell parameters a = 49.13, b = 114.94, c = 75.69 Å. A SAD data set from a Gd-HPDO3A-derivatized CbpF crystal was collected to 2.1 Å resolution at the gadolinium L_{III} absorption edge using synchrotron radiation.

1. Introduction

Streptococcus pneumoniae is a human pathogen responsible for infectious diseases that have some of the highest indexes of morbidity and mortality worldwide. This Gram-positive bacterium exhibits an absolute requirement for choline for growth. The presence of choline as a component of teichoic acid (TA) in the pneumococcal cell wall seems to have selected for the acquisition of a specialized domain, the choline-binding module (CBM), by certain surface proteins. This domain allows these proteins to specifically anchor themselves to the cell wall using choline residues (Yother *et al.*, 1992; Yother & White, 1994; López & García, 2004, and references therein). Moreover, many host interactions appear to be mediated by choline residues being recognized by components of the host response (Cundell *et al.*, 1995; Pepis & Hirschfield, 2003), such as human C-reactive protein and the PAF receptor (Hermoso *et al.*, 2005 and references therein). Choline-binding proteins (CBPs) are an important family of pneumococcal surface proteins and include the murein hydrolases LytA, LytB, LytC and Pce. These enzymes break various bonds in peptidoglycan or in teichoic acid (López & García, 2004). CBPs have been demonstrated to participate in a series of important biological functions such as cell adhesion, division and virulence (Yother *et al.*, 1992; Hammerschmidt *et al.*, 1997; Rosenow *et al.*, 1997; Sánchez-Beato *et al.*, 1998; Brooks-Walter *et al.*, 1999). Choline-binding protein F (CbpF) appears on the surface of the pneumococcal cell wall, bound to the characteristic choline moieties of the teichoic and lipoteichoic acids. The function of CbpF is still unknown; the protein carries an N-terminal signal sequence of 28 amino acids, with the mature form of the protein (311 amino acids; 36 288 Da) containing a putative functional module localized at the N-terminus and a CBM at the C-terminus which attaches the enzyme to the bacterial cell wall.

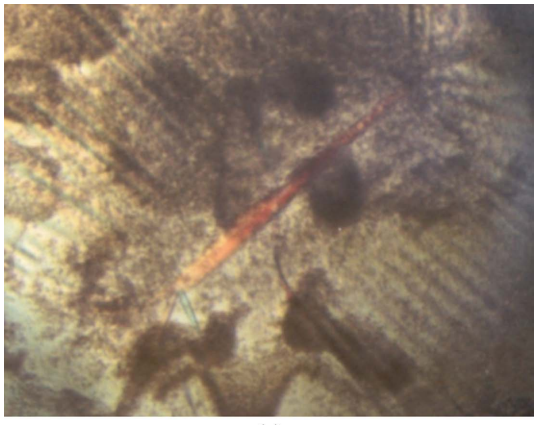
Comparison of the sequence of the putative functional module reveals no similarity to any protein other than the CBMs of other CBPs. In this sense, this N-terminal module of CbpF seems to have evolved away from the canonical choline-binding repeats (the homologous repeating units of about 20 amino acids forming the CBM of the CBPs). Taking into account the fact that CbpF, together with LytA, LytB, LytC and Pce (García *et al.*, 1999), is one of the more abundant proteins in the pneumococcal cell envelope, it seems clear

that CbpF must play a relevant role in pneumococcal physiology or virulence or both. Structure determination should help to understand the function of this protein and could open new therapeutic prospects for treating pneumococcal diseases. Here, we present the preliminary results obtained by X-ray crystallography on the complete CbpF from *S. pneumoniae*.

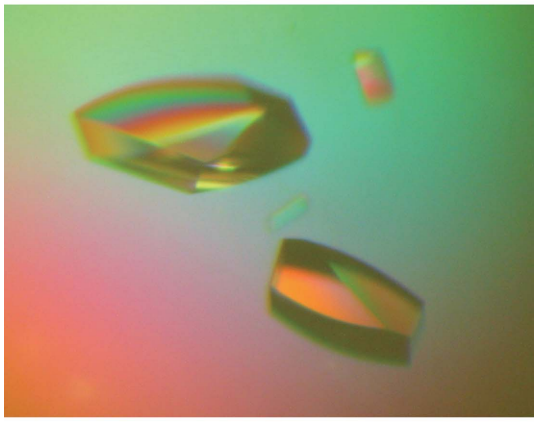
2. Experimental procedures

2.1. CbpF production and purification

The gene *spr0337* from the pneumococcal R6 strain, encoding CbpF protein, was cloned into plasmid pIN-III-A3 (Masui *et al.*, 1984). The corresponding PCR fragment was amplified with the oligonucleotides 5′-ggaattc**atg**AATACCACAGGTGGCCGATTTG-3′ and 5′-cgggatccCGCAATCGCTTCTT**TCATTA**TTG-3′, which are complementary to the 5′ and 3′ ends, respectively, of the DNA sequence coding for the mature CbpF protein. The lower case letters in the oligonucleotides indicate extensions to introduce the appropriate restriction sites (*Eco*RI and *Bam*HI). The start codon and the sequence complementary to the stop codons are shown in bold. The resulting recombinant plasmid (pAPM40) harboured the *cbpF* gene and its DNA sequence was confirmed using an automated Abi Prism 3700 DNA sequencer (Applied Biosystems). All primers were synthesized on a Beckman Oligo 1000M synthesizer. Restriction enzymes and other DNA-modifying enzymes were purchased from Amersham. For the purification of the mature protein, the recombinant *Escherichia coli* DH5α (pAPM40) strain was cultured in Terrific Broth (Sambrook *et al.*, 1989) containing ampicillin (100 µg ml^{-1}) at 310 K with aeration; when the culture reached an OD_{600} of about 0.8, protein expression was induced with 50 µ*M* isopropyl β-D-thiogalactopyranoside and incubation was continued at the lower temperature of 298 K for 16 h in order to minimize the presence of inclusion bodies. Cells were collected by centrifugation, suspended in 20 m*M* sodium phosphate buffer pH 6.9 and disrupted by passage twice through a French press (7.6 MPa). The cell lysate was centrifuged and nucleic acids were separated from protein using streptomycin sulfate precipitation (Weissborn *et al.*, 1994). The soluble fraction was loaded onto a DEAE-cellulose column equilibrated with 20 m*M* sodium phosphate buffer pH 6.9 in order to purify CbpF protein by a one-step purification procedure according to Sánchez-Puelles *et al.* (1990). The homogeneity of the protein preparations was confirmed by SDS–PAGE and mass spectroscopy. Pure and highly concentrated fractions were pooled and extensively dialyzed against 20 m*M* Tris buffer pH 8.0 containing 140 m*M* choline chloride. This high concentration of choline (140 m*M*) in the dialysis buffer is essential; at lower choline concentrations the protein precipitated rapidly. The enzyme was then concentrated at 277 K with a 10 kDa cutoff protein concentrator (Amicon, YM-10) to approximately 8 mg ml^{-1}. The final protein concentration was determined by UV spectrophotometry assuming a molar absorption coefficient of 152 640 M^{-1} cm^{-1} at 280 nm.



(*a*)



(*b*)

Figure 1
The two CbpF crystal habits obtained after crystallization trials. (*a*) Crystal habit I. These crystals were grown at 291 K in 0.1 *M* sodium cacodylate pH 6.5, 0.2 *M* ammonium sulfate, 30% PEG 8000. (*b*) Crystal habit II was obtained at the same temperature with 0.1 *M* Tris pH 8.5, 0.01 *M* $NiCl_2$, 20% PEG MME 2000.

2.2. Mass spectrometry

MALDI–TOF measurements were performed using a Voyager DE-PRO mass spectrometer from Applied Biosystems equipped with a pulsed nitrogen laser (λ = 337 nm, 3 ns pulse width, 20 Hz frequency) and a delayed-extraction ion source. Ions generated by laser desorption were introduced into the flight tube (1.3 m flight path) with an acceleration voltage of 25 kV in the linear positive-ion mode. All mass spectra were collected by averaging the signals from 5000 laser shots.

2.3. Crystallization

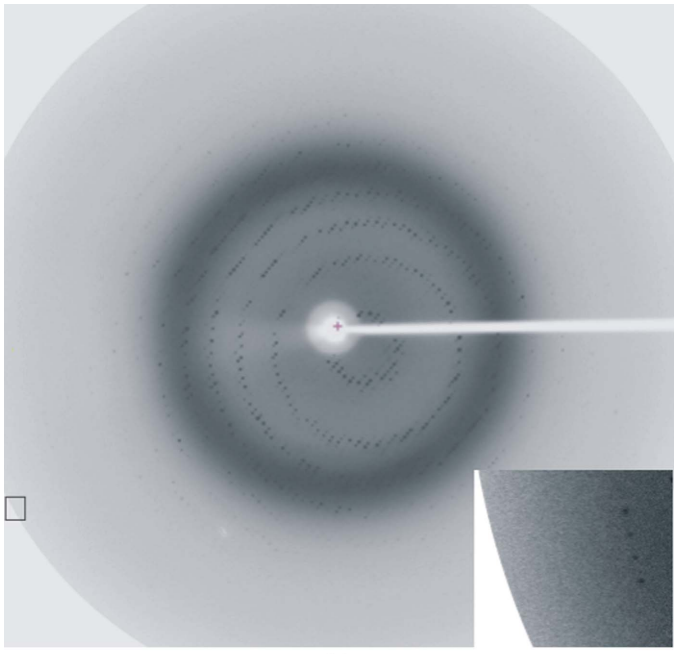
The initial crystallization conditions were established utilizing the sparse-matrix sampling technique (Jancarik & Kim, 1991) and the hanging-drop vapour-diffusion method at 291 K using commercial screens (Index Screen and Crystal Screens I and II) from Hampton Research. Initial drops consisted of 1 µl protein solution (7.8 mg ml^{-1} in 20 m*M* Tris–HCl pH 8.0, 140 m*M* choline chloride) and 1 µl well solution and were equilibrated against 500 µl well solution. Two crystal habits (I and II) were obtained using different crystallization conditions. Habit I crystals were grown in 0.2 *M* ammonium sulfate, 30% PEG 8000 and 0.1 *M* sodium cacodylate buffer pH 6.5 (Fig. 1*a*) and habit II crystals were produced in 0.01 *M* $NiCl_2$, 20% PEG MME 2000, 0.1 *M* Tris pH 8.5 (Fig. 1*b*). Considering the superior diffraction of the habit II crystals (see *Results and discussion*), the crystallization conditions of habit II crystals were optimized by mixing 3 µl protein solution (3.9 mg ml^{-1} in 20 m*M* Tris–HCl pH 8.0, 140 m*M* choline chloride) and 1 µl well solution, producing crystals with maximum dimensions of 0.7 × 0.4 × 0.3 mm in 3 d that diffracted X-rays to 1.67 Å resolution.

2.4. X-ray diffraction experiments

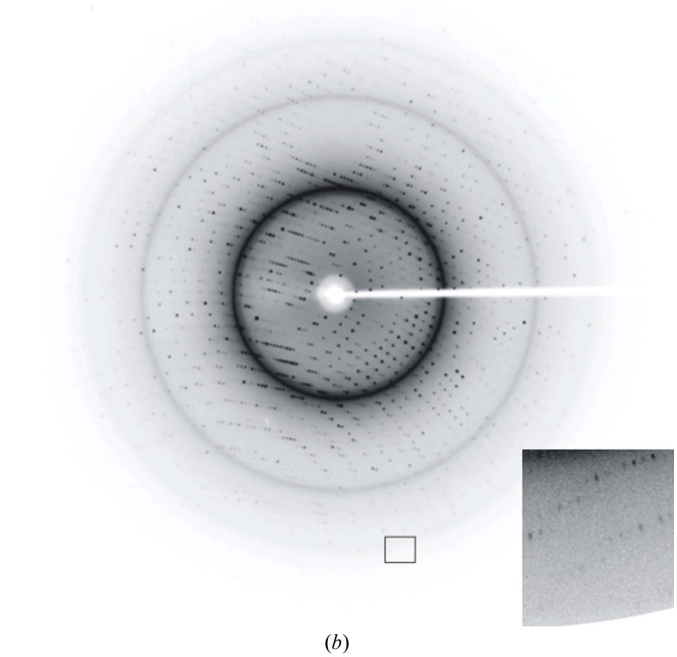
Diffraction data sets from native crystals of both habits were collected on a MAR 345 image-plate detector using Cu *Kα* X-rays

generated by an in-house Enraf–Nonius rotating-anode generator equipped with a double-mirror focusing system and operated at 40 kV and 90 mA. Crystals were cryoprotected by a quick soak (10 s) in reservoir solution containing 17%(*v*/*v*) glycerol.

An isomorphous Gd-derivative crystal was obtained by cocrystallization using the mother liquor used for native crystals (habit II) and 50 m*M* of the neutral gadolinium complex Gd-HPDO3A (Girard *et al.*, 2003), which produced derivative crystals after 7 d. The optimal wavelength for data collection was taken as the white line



(*a*)



(*b*)

Figure 2
X-ray diffraction patterns from CbpF crystals. (*a*) Crystal habit I (oscillation range 1°). (*b*) Crystal habit II (oscillation range 1°). The inset in (*a*) shows reflections at 2.0 Å resolution; in (*b*) reflections reach the edge of the plate, which corresponds to 1.67 Å resolution.

Table 1
Data-collection statistics for native CbpF and Gd-HPDO3A-derivatized CbpF crystals.

Values in parentheses are for the highest resolution shell.

	Native, habit I	Native, habit II	Gd derivative
Space group	$P2_12_12$	$P2_12_12$	$P2_12_12$
Unit-cell parameters			
a (Å)	52.40	49.15	49.54
b (Å)	115.79	115.01	115.10
c (Å)	72.98	75.79	76.29
Data collection			
Temperature (K)	120	120	100
Wavelength (Å)	1.5418	1.5418	1.7106
Resolution (Å)	26.92–2.10 (2.14–2.10)	21.07–1.67 (1.76–1.67)	49.39–2.13 (2.28–2.13)
Total reflections	102654	625729	153121
Unique reflections	25934	50379	24209
Redundancy	3.9 (3.9)	12.4 (11.5)	6.3 (4.3)
Completeness (%)	99.9 (100)	99.1 (96.7)	97.5 (87.0)
$I/\sigma(I)$	11.0 (2.3)	33.1 (7.0)	16.3 (7.8)
R_{sym}†	0.10 (0.56)	0.05 (0.34)	0.07 (0.14)

† $R_{sym} = \sum I - I_{av}|/\sum I$, where the summation is over symmetry-equivalent reflections.

(λ = 1.7110 Å) in the L_{III} absorption edge of Gd as determined from an X-ray fluorescence spectrum collected from the derivative crystal. A SAD data set was collected at 100 K using a CCD detector on beamline ID29 at the European Synchrotron Radiation Facility (ESRF). The crystal-to-detector distance was fixed at 110 mm. All data sets were processed and scaled using *XDS* (Kabsch, 1988) and *SCALA* from the *CCP*4 package (Collaborative Computational Project, Number 4, 1994). Data-collection statistics for native crystals (habits I and II) and the Gd-HPDO3A derivative of CbpF are summarized in Table 1.

3. Results and discussion

Both crystal habits belong to the orthorhombic space group $P2_12_12$, but the resolution limits and unit-cell parameters are different. While crystals with habit I (unit-cell parameters *a* = 52.40, *b* = 115.79, *c* = 72.98 Å) diffract X-rays to 1.95 Å (Fig. 2*a*), crystals of habit II (unit-cell parameters *a* = 49.13, *b* = 114.94, *c* = 75.69 Å) diffract to

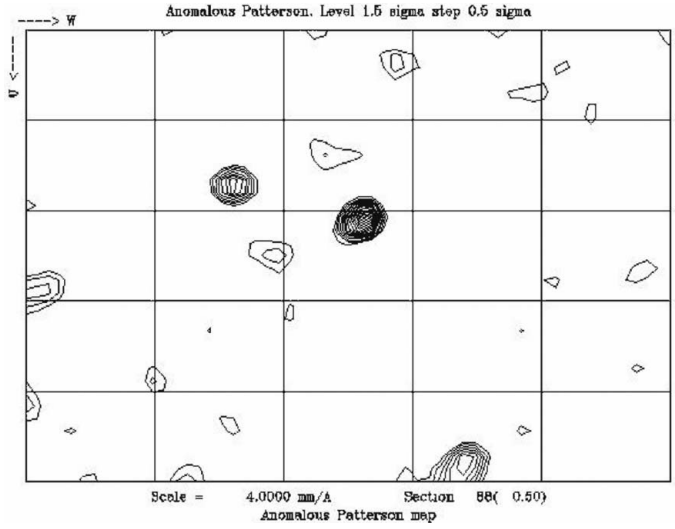


Figure 3
Plot of the *v* = 1/2 Harker section of the anomalous Patterson map for the Gd-HPDO3A derivative. Levels are contoured in 0.5σ steps starting at 1.5σ.

1.67 Å resolution (Fig. 2*b*). Considering the superior resolution of the diffraction pattern of the habit II crystal, these crystals were used in the following steps. Considering the molecular weight of CbpF and the unit-cell volume, a Matthews coefficient of 2.94 $Å^3\ Da^{-1}$ (Matthews, 1968) for a monomer in the asymmetric unit and a solvent content of 58% are obtained. Cocrystallization of CbpF in the presence of the neutral gadolinium complex Gd-HPDO3A (Girard *et al.*, 2003) led to derivative crystals of high quality that diffracted X-rays to 2.13 Å resolution. An anomalous Patterson map was calculated using data in the entire resolution range. The resulting $v = 1/2$ Harker section is shown in Fig. 3. Peaks corresponding to Gd-binding sites were clearly observed in anomalous Patterson maps and four sites were located using *SHELX* (Sheldrick, 1998). Initial phases were calculated at 2.13 Å resolution using the program *SHARP* (de La Fortelle & Bricogne, 1997). Model building of the complete modular CbpF is currently in progress.

The authors thank Bracco Imaging (Milan) for kindly providing a sample of Gd-HPDO3A, the ID29 beamline staff at ESRF for help and Douglas Laurents for critical reading of the manuscript. This work was supported by grants from the Spanish Ministry of Science and Technology (BFU2005-01645 and CONSOLIDER-INGENIO 2010 CSD2006-00015). RM holds a fellowship from the Spanish Ministry of Science and Technology.

References

Brooks-Walter, A., Briles, D. E. & Hollingshead, S. K. (1999). *Infect. Immun.* **67**, 6533–6542.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Cundell, D. R., Gerard, N. P., Gerard, C., Idanpaan-Heikkila, I. & Tuomanen, E. I. (1995). *Nature (London)*, **377**, 435–438.
García, P., González, M. P., García, E., García, J. L. & López, R. (1999). *Mol. Microbiol.* **33**, 128–138.
Girard, É., Stelter, M., Vicat, J. & Kahn, R. (2003). *Acta Cryst.* D**59**, 1914–1922.
Hammerschmidt, S., Talay, S. R., Brandtzaeg, P. & Chhatwai, G. S. (1997). *Mol. Microbiol.* **25**, 1113–1124.
Hermoso, J. A., Lagartera, L., González, A., Stelter, M., García, P., Martínez-Ripoll, M., García, J. L. & Menéndez, M. (2005). *Nature Struct. Mol. Biol.* **12**, 533–538.
Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.
Kabsch, W. (1988). *J. Appl. Cryst.* **21**, 916–924.
La Fortelle, E. de & Bricogne, G. (1997). *Methods Enzymol.* **276**, 472–494.
López, R. & García, E. (2004). *FEMS Microbiol. Rev.* **28**, 553–580.
Masui, Y., Mizuno, T. & Inouye, M. (1984). *Biotechnology (N.Y.)*, **2**, 81–85.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Pepis, M. B. & Hirschfield, G. M. (2003). *J. Clin. Invest.* **111**, 1808–1812.
Rosenow, C., Ryan, P., Weiser, J. N., Johnson, S., Fontan, P., Ortqvist, A. & Masure, H. R. (1997). *Mol. Microbiol.* **25**, 819–829.
Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*, 2nd ed. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.
Sánchez-Beato, A. R., López, R. & García, J. L. (1998). *FEMS Microbiol. Lett.* **164**, 207–214.
Sánchez-Puelles, J. M., Sanz, J. M., García, J. L. & García, E. (1990). *Gene*, **89**, 69–75.
Sheldrick, G. M. (1998). *Direct Methods for Solving Macromolecular Structures*, edited by S. Fortier, pp. 401–411. Dordrecht: Kluwer Academic Publishers.
Weissborn, A. C., Liu, Q., Rumley, M. K. & Kennedy, E. P. (1994). *J. Bacteriol.* **176**, 2611–2618.
Yother, J., Handsome, G. L. & Briles, D. E. (1992). *J. Bacteriol.* **174**, 610–618.
Yother, J. & White, J. M. (1994). *J. Bacteriol.* **176**, 2976–2985.