

Christian Linke,<sup>a</sup> Nikolai  
 Siemens,<sup>b</sup> Martin J.  
 Middleditch,<sup>a</sup> Bernd  
 Kreikemeyer<sup>b</sup> and Edward N.  
 Baker<sup>a\*</sup>

<sup>a</sup>School of Biological Sciences and Maurice  
 Wilkins Centre, University of Auckland, Private  
 Bag 92019, Auckland 1142, New Zealand, and

<sup>b</sup>Institute of Medical Microbiology, Virology and  
 Hygiene, University of Rostock,  
 Schillingallee 70, 18057 Rostock, Germany

Correspondence e-mail:  
 en.baker@auckland.ac.nz

Received 20 April 2012  
 Accepted 5 May 2012

# Purification, crystallization and preliminary crystallographic analysis of the adhesion domain of Epf from *Streptococcus pyogenes*

The extracellular protein Epf from *Streptococcus pyogenes* is important for streptococcal adhesion to human epithelial cells. However, Epf has no sequence identity to any protein of known structure or function. Thus, several predicted domains of the 205 kDa protein Epf were cloned separately and expressed in *Escherichia coli*. The N-terminal domain of Epf was crystallized in space groups  $P2_1$  and  $P2_12_12_1$  in the presence of the protease chymotrypsin. Mass spectrometry showed that the species crystallized corresponded to a fragment comprising residues 52–357 of Epf. Complete data sets were collected to 2.0 and 1.6 Å resolution, respectively, at the Australian Synchrotron.

## 1. Introduction

The Gram-positive bacterium *Streptococcus pyogenes* is a common human pathogen that causes a broad range of conditions such as tonsillitis, skin diseases and the lethal toxic shock syndrome (Cunningham, 2000). In the colonization of human epithelia, a first step in infection, *S. pyogenes* depends on a number of cell-wall-attached proteins called adhesins which recognize and bind to specific receptors on the surface of human epithelia (Cunningham, 2000). These adhesins, which are generally encoded within pathogenicity islands, vary in their distribution in different strains. Definition of their roles and mechanisms of action is therefore highly relevant to our understanding of infection and disease.

One of these adhesins is the cell-wall-anchored 205 kDa protein called Epf, which is encoded within the ERES pathogenicity island found in serotypes M1, M4, M12, M28 and M49 of *S. pyogenes* (Kreikemeyer *et al.*, 2007). Preliminary experiments with recombinant Epf from *S. pyogenes* serotype M49 indicated that it binds to human plasminogen (Kreikemeyer *et al.*, 2007), and a parallel study on its close orthologue (71% sequence identity) from *S. pyogenes* serotype M1 showed that it is essential for virulence in the mouse model (Kwinn *et al.*, 2007). Epf belongs to a group of large cell-surface proteins that are widespread in Gram-positive bacteria (Kreikemeyer *et al.*, 2007; Clarke *et al.*, 2002; Smith *et al.*, 1993) and are also called MSCRAMMs (microbial surface components recognizing adhesive matrix molecules; Joh *et al.*, 1999; Patti *et al.*, 1992). These proteins typically comprise multiple domains, with a common pattern that includes a classical signal sequence at the N-terminus followed by a highly variable N-terminal domain that contains their binding or catalytic activities, a central region comprising a number of conserved repeat domains and a C-terminal domain that includes an LPXTG sortase recognition motif through which they are attached to the cell wall.

Epf comprises an N-terminal region that has no detectable sequence identity to other proteins followed by a large number of repeats of a domain referred to as DUF1542 (domain of unknown function 1542; Fig. 1). In *S. pyogenes* serotype M49 the N-terminal domain (EpfN) is followed by 18 DUF1542 repeat domains. The respective biological roles of these domains are not yet known, but we hypothesize that the N-terminal domain contains the adhesive functions of Epf, whereas the DUF1542 domains have a structural role, in conformity with the pattern seen in many other MSCRAMMs.



**Table 1**

Constructs and PCR primers used in this study.

Amino-acid residue numbers refer to NCBI accession No. ZP\_00366506. The restriction site of each primer is shown in bold.

Construct	Residues	Primer	Sequence	Restriction enzyme
EpfN	45–386	NtermIBA2for NtermIBA2rev	5'-CAT TCA GCG <b>AAT TCA</b> ATG GCG TGA TGG TCG TAA AG-3' 5'-CTG GTT TTT <b>CGT CGA</b> CCG GTT TTT CTG GTA GCC A-3'	<i>EcoRI</i> <i>Sall</i>
DUF1542 repeat 1	387–464	DUF1for DUF1rev	5'-CCA GAA <b>GAA TTC</b> GAA GGC GAA AAA CCA GTA CAA A-3' 5'-CAT TTA CTT TAT CGC TTA <b>GTC GAC</b> CCA GTG ATT TTT T-3'	<i>EcoRI</i> <i>Sall</i>
DUF1542 repeats 1–4	387–796	DUF1for DUF1–4rev	See above 5'-TGC TTT TTG CTT <b>GTC GAC</b> TGC AAG AAG TTT TTC-3'	<i>EcoRI</i> <i>Sall</i>
DUF1542 repeats 1–8	387–1118	DUF1for DUF1–8rev	See above 5'-CTT GAG CGA <b>CGT CGA</b> CAA GTT GCT TAC CAT-3'	<i>EcoRI</i> <i>Sall</i>
DUF1542 repeats 1–16	387–1766	DUF1for DUF1–16rev	See above 5'-CTT TAG CTA <b>CGT CGA</b> CAA GAT TAG CTT TAG CAG C-3'	<i>EcoRI</i> <i>Sall</i>

To further characterize Epf and its role in streptococcal pathogenicity, we cloned and purified various domains of Epf. Here, we report the successful crystallization of the N-terminal domain of Epf using *in situ* proteolysis.

## 2. Materials and methods

### 2.1. Cloning and expression

The primary structure of Epf (1879 amino acids; NCBI accession No. ZP\_00366506) from *S. pyogenes* strain 591 (serotype M49) was analysed manually and using the *SMART* server (Letunic *et al.*, 2009) to identify domain boundaries. This led to the choice of an N-terminal domain (EpfN; residues 45–386) and several constructs designed to include varying numbers of DUF1542 repeat domains (see Table 1).

Chromosomal DNA of the *S. pyogenes* wild-type strain was isolated using the DNeasy Blood & Tissue Kit (Qiagen) according to the manufacturer's instructions with minor modifications. Briefly, cells from a 10 ml overnight culture grown in THY (Todd Hewitt Broth) were harvested and resuspended in enzymatic lysis buffer [20 mM Tris-HCl, pH 8.0, 2 mM Na EDTA, 1.2% (v/v) Triton X-100, 20 mg ml<sup>-1</sup> lysozyme]. After 30 min incubation at 310 K, 50 µl Proteinase K solution (1 mg ml<sup>-1</sup>) was added to the suspension, followed by an additional incubation at 310 K for 30 min. The chromosomal DNA was purified from the lysis suspension according to the manufacturer's guidelines. The *epf* domains were amplified from genomic DNA of *S. pyogenes* strain 591 (see Table 1 for primers) with Phusion polymerase (NEB). The amplified fragments were digested with *EcoRI* and *Sall* and cloned into the vector pASK\_IBA2 (IBA GmbH). This expression vector confers an N-terminal *ompA* signal, which is removed during export into the periplasmic space, and adds a C-terminal Strep-tag onto the expressed peptide.

All Epf domain constructs (Table 1) were tested for expression and were subjected to initial purification experiments. EpfN was highly expressed and appeared to be stable over several weeks as judged

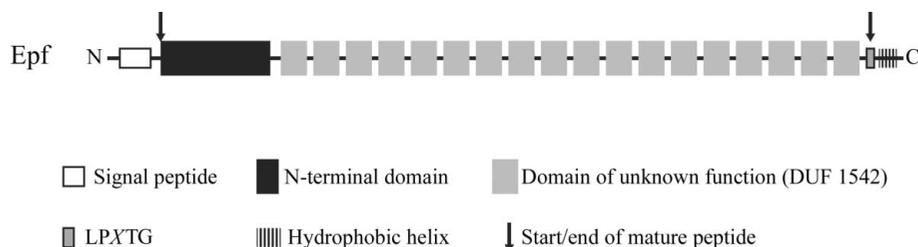
by SDS-PAGE. The construct comprising DUF1542 repeats 1–4 was also highly expressed but degraded within a week. The other constructs could only be expressed in a small-scale format (DUF1542 repeats 1–16) or not at all.

For heterologous overexpression of EpfN, *Escherichia coli* BL21 (DE3) cells were transformed with the pASK\_IBA2\_EpfN expression vector and cultured in M9 medium (0.25 l per 2 l baffled flask; 42 mM Na<sub>2</sub>HPO<sub>4</sub>, 22 mM KH<sub>2</sub>PO<sub>4</sub>, 19 mM NH<sub>4</sub>Cl, 8.6 mM NaCl, 2 mM MgSO<sub>4</sub>, 22 mM glucose, 0.15 µM thiamine and a mixture of trace metals) with 100 µg ml<sup>-1</sup> ampicillin at 310 K with shaking. When the optical density at a wavelength of 550 nm (OD<sub>550</sub>) reached 0.2–0.4, expression was induced with 200 ng ml<sup>-1</sup> anhydrotetracycline and the cells were incubated at 291 K with shaking for a further 18 h. The cells were then harvested by centrifugation (20–30 min, 3800g, 277 K).

### 2.2. Purification

To release EpfN from the periplasmic space, the cells from 1 l expression culture were resuspended in 10 ml chilled buffer *P* [100 mM Tris-HCl pH 8.0, 500 mM sucrose, 1 mM EDTA, 0.3 mM NaN<sub>3</sub>, Complete Protease Inhibitor Cocktail Mini Tablet EDTA-free (Roche), 10 mg ml<sup>-1</sup> hen egg-white lysozyme (Roche)] and incubated on ice for 30 min. Spheroblasts and debris were then removed by centrifugation (15 min, 13 000g, 277 K).

EpfN was purified by affinity chromatography using the Strep-tag derived from the expression vector. The cleared supernatant was applied onto Strep-Tactin Superflow High Capacity beads (IBA GmbH) pre-equilibrated in buffer *W* (100 mM Tris-HCl pH 8.0, 150 mM NaCl, 1 mM EDTA, 0.3 mM NaN<sub>3</sub>) in a gravity-flow column at 277 K. The bound protein was washed with five column volumes of buffer *W* and eluted with three column volumes of buffer *E* (100 mM Tris-HCl pH 8.0, 150 mM NaCl, 1 mM EDTA, 2.5 mM desthiobiotin, 0.3 mM NaN<sub>3</sub>) in steps of 0.5 column volumes. The beads were regenerated using 10–20 column volumes of buffer *R* [100 mM Tris-HCl pH 8.0, 150 mM NaCl, 1 mM EDTA, 1 mM 2-(4-hydroxyphenyl-



**Figure 1**  
Domain organization of Epf.

azo]benzoic acid], followed by four column volumes of 100 mM Tris base and buffer *W*.

EpfN was purified to homogeneity using size-exclusion chromatography. Fractions containing EpfN were pooled and concentrated using a Vivaspin 20 concentrator (Sartorius Stedim). The concentrate was applied onto a Superdex 200 10/300 GL column (GE Healthcare) pre-equilibrated in buffer *C* (10 mM Tris-HCl pH 7.4, 137 mM NaCl, 2.7 mM KCl, 0.3 mM NaN<sub>3</sub>) and mounted on an ÄKTAprime FPLC system at room temperature. EpfN eluted as a single symmetrical peak at a retention volume of 15.5 ml, consistent with a monomeric species. Purified EpfN was 98% pure as judged by SDS-PAGE. The purified protein could be stored at 277 K for several weeks without any sign of degradation.

### 2.3. Crystallization and *in situ* proteolysis of EpfN

For crystallization screening, EpfN in buffer *C* was concentrated to 19–20 mg ml<sup>-1</sup> as measured by absorption at 280 nm assuming an absorption coefficient of 33 920 M<sup>-1</sup> cm<sup>-1</sup> (calculated using *ProtParam*; Gasteiger *et al.*, 2005). The crystallization properties of EpfN were assayed in sitting-drop vapour-diffusion experiments (96-well Intelli-Plates; Hampton Research) using a local 480-condition screen (Moreland *et al.*, 2005) and a Cartesian Honeybee nanolitre dispensing robot (Genomic Solutions). In each well, a protein drop comprising 100 nl EpfN solution mixed with 100 nl reservoir solution was equilibrated against 85 µl reservoir solution. These screening experiments were then incubated at 291 K.

Crystallization conditions for EpfN were also screened in the presence of the protease chymotrypsin using *in situ* proteolysis (Dong *et al.*, 2007). To this end, a 1 mg ml<sup>-1</sup> solution of chymotrypsin (type VII, TLCK-treated; Sigma-Aldrich) dissolved in buffer *C* was added to EpfN at 20 mg ml<sup>-1</sup> to give a final protease:EpfN ratio of 1:1000 (*w:w*). The EpfN–chymotrypsin mixture was then assayed as above using our local 480-condition screen. Subsequent fine screens were performed in hanging-drop vapour-diffusion experiments (24-well VDX plates; Hampton Research) with 20–25% (*w/v*) PEG 3350 and 0.1–0.4 M potassium acetate or potassium chloride as the

reservoir solution at 500 µl per well. Reservoir solution (1 µl) was added to 1 µl of the EpfN–chymotrypsin mixture on a siliconized cover slip (Hampton Research) and incubated at 291 K.

### 2.4. Data collection and processing

Crystals of EpfN grown in the presence of chymotrypsin were transferred to cryoprotectant using a mounting loop and flash-cooled in liquid nitrogen. For the crystals grown in potassium acetate, the cryoprotectant consisted of buffer *C*, 30% (*v/v*) glycerol, 17.5% (*w/v*) PEG 3350, 140 mM potassium acetate. Crystals from the potassium chloride condition were protected using a mixture of 70% (*v/v*) Paratone-N oil and 30% (*v/v*) paraffin oil.

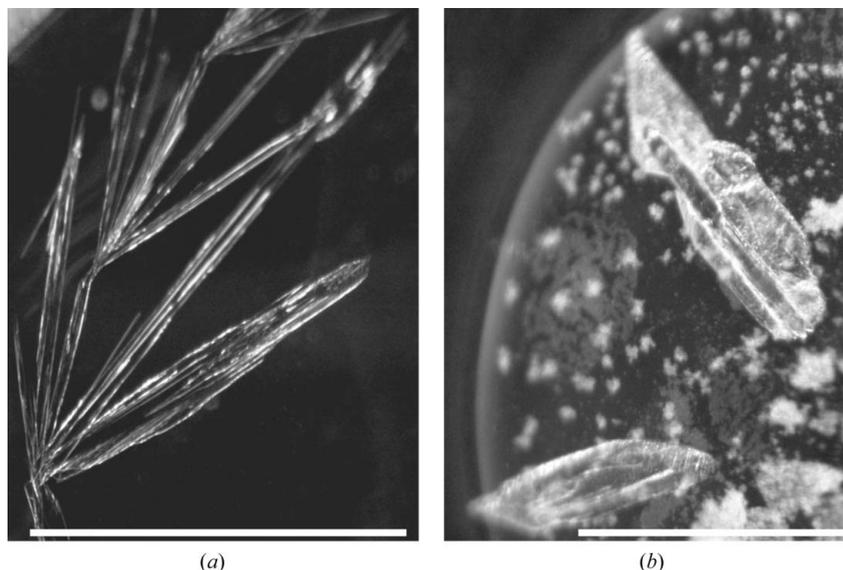
X-ray diffraction data sets for EpfN crystals were collected on beamline MX2 of the Australian Synchrotron (AS), Victoria, Australia (McPhillips *et al.*, 2002) under cooling by a stream of nitrogen at 100 K. The microbeam (37 × 32 µm) available at MX2 was used to screen local diffraction properties, which varied significantly within an EpfN crystal. The crystal region used for data collection was selected after visual inspection of the test images.

Reflections were integrated using *XDS* (Kabsch, 2010), reindexed using *POINTLESS* (Evans, 2011) and scaled using *SCALA* (Evans, 2011) from the *CCP4* software suite (Winn *et al.*, 2011). Data-collection statistics are summarized in Table 2.

### 2.5. Mass spectrometry

Protein molecular mass was determined by electrospray ionization-time of flight (ESI-TOF) mass spectrometry. Samples dissolved in 50% (*v/v*) acetonitrile and 0.1% (*v/v*) formic acid were analysed in positive-ionization mode on a Q-STAR XL Hybrid MS/MS system (Applied Biosystems), with data acquired in the *m/z* range 500–1600. Raw data were processed using the Bayesian Protein Reconstruct tool from the Bioanalyst extensions within *Analyst QS* 1.1 (Applied Biosystems) to give deconvoluted molecular weights.

Peptide identification was performed by reversed-phase liquid chromatography coupled with tandem mass spectrometry (LC-MS/MS). Peptides from the trypsin digest of a protein gel band were



**Figure 2**

Crystals of EpfN grown in the presence of chymotrypsin. The growth conditions were (a) 22% (*w/v*) PEG 3350, 400 mM potassium acetate (crystal form A) and (b) 23% (*w/v*) PEG 3350, 200 mM potassium chloride (crystal form B). The white bars correspond to 1 mm in each picture.

**Table 2**

Crystal and data-collection statistics.

Values in parentheses are for the outermost resolution shell.

Crystal	Form A	Form B
Crystal data		
Growth condition	22% (w/v) PEG 3350, 400 mM potassium acetate	23% (w/v) PEG 3350, 200 mM potassium chloride
Space group	$P2_12_12_1$	$P2_1$
Unit-cell parameters (Å, °)	$a = 60.27, b = 117.78, c = 85.62, \alpha = 90.00, \beta = 90.00, \gamma = 90.00$	$a = 71.08, b = 117.13, c = 73.04, \alpha = 90.00, \beta = 106.46, \gamma = 90.00$
Matthews coefficient (Å <sup>3</sup> Da <sup>-1</sup> )	2.22	2.13
Molecules per asymmetric unit	2	4
Solvent content (%)	44.5	42.2
Data collection		
X-ray source	03ID1 [AS MX2]	03ID1 [AS MX2]
Wavelength (Å)	0.97942	0.97942
Detector	ADSC Quantum 315r	ADSC Quantum 315r
No. of images	540	360
Oscillation angle (°)	1.0	1.0
Exposure time (s)	1.0	1.0
Resolution range (Å)	19.88–1.58 (1.66–1.58)	19.73–2.00 (2.11–2.00)
Total No. of observations	1699412 (137780)	283329 (41558)
Unique reflections	83319 (11097)	76117 (11229)
Multiplicity	20.4 (12.4)	3.7 (3.7)
Completeness (%)	98.7 (91.6)	98.4 (99.2)
Mean $I/\sigma(I)$	35.1 (5.6)	6.4 (2.2)
$I/\sigma(I)$	10 (1.7)	5.6 (1.7)
$R_{\text{merge}}$ (%)	7.0 (45.7)	11.3 (44.1)
$R_{\text{p.i.m.}}$ (%)	1.6 (12.9)	7.6 (30.8)

analysed on the Q-STAR XL Hybrid MS/MS system and identified using the *MASCOT* search engine v.2.0.05 (Matrix Science).

### 3. Results and discussion

In this study, we sought to express and crystallize domains of the adhesin Epf from *S. pyogenes*. Epf has an N-terminal cleavable signal peptide followed by an N-terminal domain (residues 45–386), 18 repeats of a domain of unknown function (DUF1542; PFAM accession No. PF07564) that encompass residues 387–1841 and finally a C-terminal cell-wall sorting signal and hydrophobic helix (Fig. 1). The N-terminal domain displays no significant sequence similarity to any described protein. Here, we refer to this domain as EpfN.

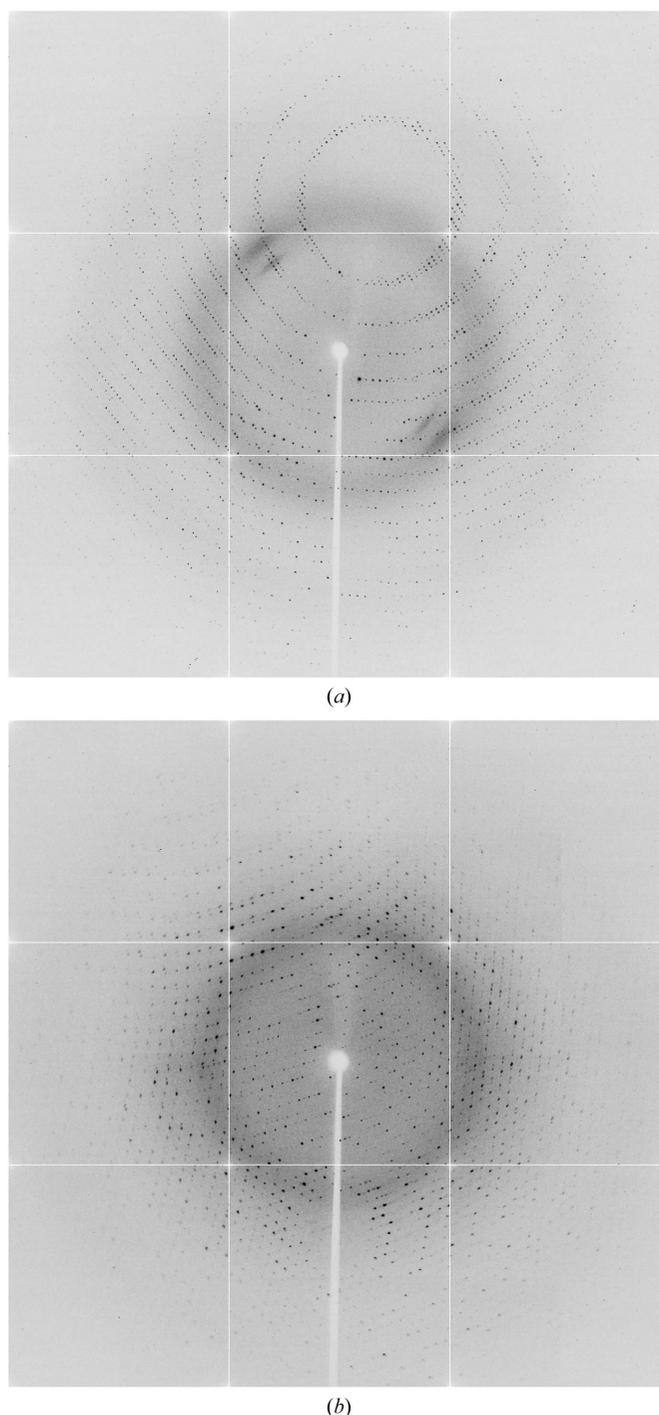
EpfN and constructs with a varying number of DUF1542 repeats were cloned and tested for expression. Whereas the C-terminal DUF1542 repeat domains were found to be either poorly expressed or unstable, the N-terminal domain of Epf (EpfN) was well expressed and stable. Thus, we concentrated our analysis on EpfN.

The final purified EpfN construct corresponded to amino-acid residues 45–386 of Epf (NCBI accession No. ZP\_00366506) with an N-terminal extension (GDHGPEF) and a C-terminal extension (VDLQGDHGLSAWSHPQFEK), the latter including a Strep-Tag derived from the expression vector. The identity of purified EpfN was confirmed by LC MS/MS and ESI-TOF mass spectrometry (observed mass 41 253.4 Da, expected mass 41 253.4 Da).

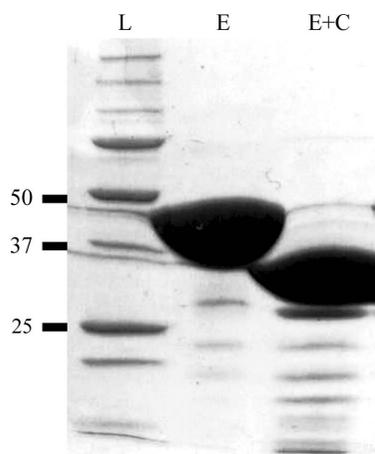
When EpfN was screened for crystallization, crystals of hexagonal shape with a diameter of 100–300 μm were observed in one condition [40% (w/v) PEG 400, 200 mM Li<sub>2</sub>SO<sub>4</sub>, 100 mM Tris-HCl pH 8.5] after 1–2 d. However, these crystals were of poor diffraction quality (maximum resolution 10 Å) and could not be optimized by screening around the original condition or by the use of additives or seeding.

EpfN was also screened in the presence of trace amounts of chymotrypsin. After four weeks, crystals were observed in two related conditions: rod-like crystals in 20% (w/v) PEG 3350, 200 mM potassium acetate (form A) and plate-like crystals in 20% (w/v) PEG 3350,

200 mM potassium chloride (form B). Both crystal forms (with a maximal length of ~1 mm) could be reproduced in the presence of chymotrypsin in hanging-drop vapour-diffusion experiments with 20–25% (w/v) PEG 3350 and 0.1–0.4 M potassium acetate or potassium chloride, respectively (Fig. 2). Complete and redundant X-ray diffraction data of good quality were collected at the Australian Synchrotron from crystals grown under both conditions (Fig. 3, Table 2), with the resolution extending to 1.6 Å for form A and 2.0 Å for form B.



**Figure 3** Diffraction images of crystals of EpfN grown in the presence of chymotrypsin. (a) Form A, (b) form B. The resolution limit at the edge of the image is 1.66 Å in (a) and 2.0 Å in (b).



**Figure 4**

Effect of chymotrypsin on EpfN. SDS-PAGE of untreated EpfN (lane E) and EpfN + chymotrypsin (lane E+C) in a ratio of 1000:1(w:w) after three weeks at 291 K. The protein standards (lane L) are the Precision Plus Protein Standards (Bio-Rad) and are labelled in kDa.

The crystal morphology and symmetry varied between the crystals grown from the two similar conditions. Form *A* crystals were rod-shaped and belonged to the orthorhombic space group  $P2_12_12_1$ , whereas form *B* crystals were plate-like and belonged to the monoclinic space group  $P2_1$ , as determined from systematic absences (Table 2). The crystals also differed in the predicted number of molecules per asymmetric unit, with two for form *A* and four for form *B*, despite similar solvent contents as judged by the Matthews coefficient ( $2.22$  and  $2.13 \text{ \AA}^3 \text{ Da}^{-1}$ , respectively).

To investigate the effect of chymotrypsin on EpfN, a sample of the EpfN–chymotrypsin mixture used for crystallization experiments was incubated at 291 K for three weeks and analysed using SDS-PAGE and mass spectrometry. In SDS-PAGE, only a slight mass shift of EpfN was observed after chymotrypsin treatment (Fig. 4). Using ESI-TOF mass spectrometry, a principal deconvoluted peak with an observed molecular weight of  $34\,277.2 \text{ Da}$  was detected (Supplementary Fig. S1<sup>1</sup>), which corresponds to a loss of  $6976.2 \text{ Da}$  compared with untreated EpfN. The remaining EpfN polypeptide could thus correspond to Epf amino-acid residues 52–357, with an expected molecular weight of  $34\,277.5 \text{ Da}$ . The presence of these residues in the remaining EpfN polypeptide was confirmed by performing LC-MS/MS on a tryptic digest of this species. This gave almost complete coverage for peptides in the residue range 52–357, whereas no peptides were observed outside this range (Table S1<sup>1</sup>).

Recently, *in situ* proteolysis has emerged as a powerful method in protein crystallography (Dong *et al.*, 2007; Wernimont & Edwards, 2009). The growth of high-quality EpfN crystals in the presence of

trace amounts of chymotrypsin is a further example of how a protease can improve crystallization properties. In the case of EpfN, chymotrypsin may have trimmed the EpfN peptide to the actual boundaries of the EpfN domain, which were difficult to predict owing to a lack of structural data on EpfN or its homologues.

Attempts at experimental phasing to solve the structure of EpfN are in progress. Molecular replacement could not be applied, as no protein structure with any significant sequence identity to EpfN could be found in the PDB.

CL was supported by a New Zealand International Doctoral Research Scholarship (Education New Zealand). The work of CL, MJM and ENB was supported by the Health Research Council of New Zealand and the Tertiary Education Commission through funding of the Maurice Wilkins Centre for Molecular Biodiscovery. Access to the Australian Synchrotron was supported by the New Zealand Synchrotron Program. The work of NS and BK was supported by grants from the German Federal Ministry of Education and Research (BMBF) in the framework of the ERANet ‘Patho-GenoMics’ I and II program.

## References

- Clarke, S. R., Harris, L. G., Richards, R. G. & Foster, S. J. (2002). *Infect. Immun.* **70**, 6680–6687.
- Cunningham, M. W. (2000). *Clin. Microbiol. Rev.* **13**, 470–511.
- Dong, A. *et al.* (2007). *Nature Methods*, **4**, 1019–1021.
- Evans, P. R. (2011). *Acta Cryst.* **D67**, 282–292.
- Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D. & Bairoch, A. (2005). *The Proteomics Protocols Handbook*, edited by J. M. Walker, pp. 571–607. Totowa: Humana Press.
- Joh, D., Wann, E. R., Kreikemeyer, B., Speziale, P. & Höök, M. (1999). *Matrix Biol.* **18**, 211–223.
- Kabsch, W. (2010). *Acta Cryst.* **D66**, 125–132.
- Kreikemeyer, B., Nakata, M., Köller, T., Hildisch, H., Kourakos, V., Standar, K., Kawabata, S., Glocker, M. O. & Podbielski, A. (2007). *Infect. Immun.* **75**, 5698–5710.
- Kwinn, L. A., Khosravi, A., Aziz, R. K., Timmer, A. M., Doran, K. S., Kotb, M. & Nizet, V. (2007). *J. Bacteriol.* **189**, 1322–1329.
- Letunic, I., Doerks, T. & Bork, P. (2009). *Nucleic Acids Res.* **37**, D229–D232.
- McPhillips, T. M., McPhillips, S. E., Chiu, H.-J., Cohen, A. E., Deacon, A. M., Ellis, P. J., Garman, E., Gonzalez, A., Sauter, N. K., Phizackerley, R. P., Soltis, S. M. & Kuhn, P. (2002). *J. Synchrotron Rad.* **9**, 401–406.
- Moreland, N., Ashton, R., Baker, H. M., Ivanovic, I., Patterson, S., Arcus, V. L., Baker, E. N. & Lott, J. S. (2005). *Acta Cryst.* **D61**, 1378–1385.
- Patti, J. M., Jonsson, H., Guss, B., Switalski, L. M., Wiberg, K., Lindberg, M. & Höök, M. (1992). *J. Biol. Chem.* **267**, 4766–4772.
- Smith, H. E., Reek, F. H., Vecht, U., Gielkens, A. L. & Smits, M. A. (1993). *Infect. Immun.* **61**, 3318–3326.
- Wernimont, A. & Edwards, A. (2009). *PLoS One*, **4**, e5094.
- Winn, M. D. *et al.* (2011). *Acta Cryst.* **D67**, 235–242.

<sup>1</sup> Supplementary material has been deposited in the IUCr electronic archive (Reference: PU5373).