2013-11-29

Crystallization Of The CHAP Domain Of The Endolysin From Staphylococcus Aureus Bacteriophage K

Marta Sanz-Gaitero
Centro Nacional de Biotecnología

Ruth Keary
Cork Institute of Technology

Carmela Garcia-Doval
Centro Nacional de Biotecnología

Aidan Coffey
Cork Institute of Technology

Mark J. van Raaij
Centro Nacional de Biotecnología

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

Crystallization of the CHAP domain of the endolysin from *Staphylococcus aureus* bacteriophage K

Marta Sanz-Gaitero,[a] Ruth Keary,[b] Carmela Garcia-Doval,[a] Aidan Coffey[b] and Mark J. van Raaij[a]*

[a]Departamento de Estructura de Macromoléculas, Centro Nacional de Biotecnología (CNB–CSIC), Calle Darwin 3, 28049 Madrid, Spain, and [b]Department of Biological Sciences, Cork Institute of Technology, Bishopstown, Cork, Ireland

Correspondence e-mail: mjvanraaij@cnb.csic.es

Received 30 September 2013
Accepted 3 November 2013

$CHAP_K$ is the N-terminal cysteine, histidine-dependent amidohydrolase/peptidase domain (CHAP domain) of the *Staphylococcus aureus* bacteriophage K endolysin LysK. It is formed from the first 165 residues of LysK and functions by cleaving specific peptidoglycan peptide bonds, causing bacterial lysis. $CHAP_K$ can lyse *S. aureus* when applied exogenously, making it a good candidate for the treatment of multidrug-resistant *Staphylococcus aureus* infections. Here, the crystallization of $CHAP_K$ and the collection of native and derivative data to high resolution, which allowed structure solution, are reported. The structure may help to elucidate the mechanism of action and in the design of chimeric proteins or mutants with improved antibacterial activity.

1. Introduction

Staphylococci are Gram-positive bacteria that can cause disease in both humans and animals. In recent years, they have become a serious concern owing to their ability to acquire genes associated with antibiotic resistance (Holden *et al.*, 2013). Biofilm formation has been reported to significantly contribute to this problem by causing up to a 1000-fold increase in bacterial antibiotic resistance (Olson *et al.*, 2002). This is an important point in relation to staphylococcal infections, because they are the main cause of nosocomial and medical device-related infections, which are generally biofilm-associated (Otto, 2008).

It is known that bacteriophages, such as the myovirus bacteriophage K (Rees & Fry, 1981; O'Flaherty, Ross *et al.*, 2005), and their endolysins are highly effective at lysing *Staphylococcus aureus*, including methicillin-resistant strains (MRSA; Ralston *et al.*, 1955; Sonstein *et al.*, 1971; O'Flaherty, Coffey *et al.*, 2005; Fenton, Ross *et al.*, 2011). Bacteriophage-encoded endolysins are peptidoglycan hydrolases used by the phages to degrade the bacterial cell wall in the late stages of the phage replication cycle, allowing their progeny to escape from the cell (Loessner, 2005). When applied exogenously to Gram-positive pathogens, endolysins can degrade the peptidoglycan and cause 'lysis from without' or exolysis (Ralston & McIvor, 1967). These enzymes are highly specific owing to the high level of diversity in the peptidoglycan structure in Gram-positive bacteria at both the genera and species level. Also of note is the fact that no bacterial strains resistant to their phage endolysins have yet been found (Loeffler *et al.*, 2001; Schuch *et al.*, 2002).

LysK, the 495 amino-acid endolysin from bacteriophage K, has been shown to kill a wide range of staphylococci, including MRSA (O'Flaherty, Coffey *et al.*, 2005). This enzyme acts synergistically with lysostaphin, an antistaphylococcal bacteriocin synthesized by *S. simulans* (Becker *et al.*, 2008), although it may also be used in isolation. LysK contains three domains: an N-terminal cysteine, histidine-dependent amidohydrolase/peptidase domain (CHAP; residues 35–160), a central amidase domain (amidase 2; residues 197–346) and a C-terminal SH3b cell-wall-binding domain (residues 412–481). The amidase 2 domain cleaves peptidoglycan between *N*-acetylmuramic acid and L-alanine of the stem peptide, while the CHAP domain has been reported to cleave between D-alanine and the first glycine of the pentaglycine cross-bridge (Fig. 1; Becker *et al.*, 2009). $CHAP_K$ is the only domain necessary for exolysis of *S. aureus* cells and a truncated enzyme containing the first 165 amino acids of

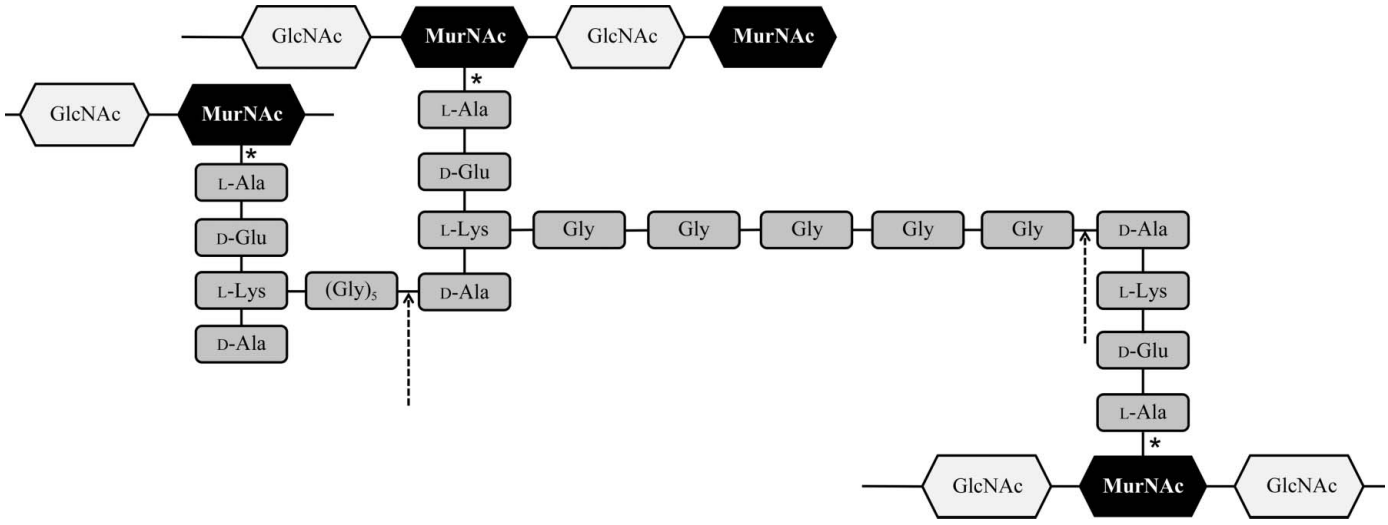

Figure 1
Scheme of the *S. aureus* peptidoglycan structure. Bacteriophage K endolysin (LysK) cleavage sites are indicated. The CHAP domain cleaves at the sites indicated with arrows, while the amidase 2 domain acts on the bonds indicated with asterisks. GlcNAc stands for *N*-acetylglucosamine and MurNAc for *N*-acetylmuramic acid.

LysK showed an activity twofold higher than the native protein (Horgan *et al.*, 2009). The reason for this may be that the SH3b domain is inhibitory when not bound to the cell wall, so removing it may activate the enzyme (Low *et al.*, 2005). $CHAP_K$ has also been demonstrated to prevent and rapidly eliminate staphylococcal biofilms *in vitro* (Fenton *et al.*, 2013).

Here, we report the crystallization of $CHAP_K$, native X-ray diffraction data collected to 1.8 Å resolution and successful structure solution using data collected from a mercury derivative.

2. Methods

Residues 1–165 of LysK endolysin, encoding the entire CHAP domain, were previously cloned into the pQE60 plasmid and transformed into *Escherichia coli* XL1-Blue (Qiagen, Venlo, Netherlands; Horgan *et al.*, 2009). The C-terminal His tag was avoided by including a stop codon; the protein product of this construct is called $CHAP_K$. Expression and purification of $CHAP_K$ by cation-exchange chromatography were performed as described by Fenton, Cooney *et al.* (2011). The purified protein was desalted and concentrated by a buffer exchange with 25 m*M* Tris–HCl pH 7 in an Amicon Ultra centrifugal filter with a molecular-weight cutoff of 10 kDa (Millipore, Billerica, Massachusetts, USA). The concentration of the protein was measured using a Pierce BCA Protein Assay Kit (Thermo Scientific, Rockford, Illinois, USA). The protein was stored in aliquots at 193 K; 4–10 mg of purified protein per litre of bacterial culture was obtained. The integrity and purity of the protein were checked using polyacrylamide gel electrophoresis following Laemmli's protocol (Laemmli, 1970). The concentration of acrylamide in the resolving gel was 12.5%(*w*/*v*). Before performing crystallization trials, the protein was further concentrated using the same centrifugal filters as mentioned previously without a buffer change.

Initial crystallization screening was carried out using sitting-drop vapour diffusion at 294 K, employing different commercial kits, Wizard I and II (Emerald Biosystems, Bainbridge Island, Washington, USA), Morpheus (Molecular Dimensions, Newmarket, Suffolk, England) and JBScreen Basic (Jena Bioscience, Jena, Germany), and a Genesis RSP 150 workstation (Tecan, Männedorf, Switzerland) with MRC 96-well sitting-drop crystallization plates (Molecular Dimensions, Newmarket, Suffolk, England). The reservoir contained 50 µl reservoir solution and drops were formed by mixing 0.2 µl protein solution with 0.2 µl reservoir solution. The conditions that gave the most promising results were scaled up and optimized to fine-tune the precipitant concentration, protein concentration and pH. Sitting-drop vapour-diffusion CompactClover 4-Chamber plates (Jena Bioscience, Jena, Germany) were used for optimization and the reservoir contained 0.15 ml reservoir solution, with drops made up by mixing 2 µl protein solution with 2 µl reservoir solution.

The crystal used for native data collection was from a drop for which the reservoir solution consisted of 22%(*w*/*v*) polyethylene glycol (PEG) 8000, 0.1 *M* 2-(*N*-morpholino)ethanesulfonic acid (MES)–NaOH pH 6.5. For derivatization, a drop for which the reservoir solution consisted of 22%(*w*/*v*) PEG 8000, 0.1 *M* 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES)–NaOH pH 6.5 was used. A few grains of solid methylmercury chloride were transferred into the reservoir and mixed by pipetting the reservoir up and down; 0.5 µl of this modified reservoir was then added to the protein drop and the experiment was re-sealed.

Crystals were cryoprotected by sequential soaking in 5, 10, 15 and 20% glycerol in crystallization solution in the same drop that they were grown in, removing and replacing about 80% of the drop volume at each step (20–30 s per step). For the derivative, this procedure thus effectively involved a quick and probably incomplete back-soak. Crystals were then vitrified by rapid immersion in liquid nitrogen. Data were collected with crystals maintained at 100 K.

iMosflm (Battye *et al.*, 2011) was used for indexing and integration of the native data and *POINTLESS*, *SCALA* and *C-TRUNCATE* from the *CCP*4 suite (Winn *et al.*, 2011) were used for scaling and merging (Evans, 2011). Derivative data were integrated using *xia*2 (Winter *et al.*, 2013), which incorporates *XDS* (Kabsch, 2010) and *AIMLESS* (Evans & Murshudov, 2013). Matthews coefficient determination was performed using *MATTHEWS_COEF* (Matthews, 1968). Phase determination and refinement were performed using *autoSHARP* (Vonrhein *et al.*, 2007). This procedure includes heavy-atom site location by *SHELXC/D* (Sheldrick, 2008), phase refinement with *SHARP* (Bricogne *et al.*, 2003), which found additional sites and rejected sites, solvent flattening and histogram matching with *SOLOMON* (Abrahams & Leslie, 1996), automated model

building with *ARP*/*wARP* (Langer *et al.*, 2008) and refinement with *REFMAC*5 (Murshudov *et al.*, 2011).

3. Results and discussion

The purity of the $CHAP_K$ lysin was verified by denaturing gel electrophoresis, showing a band at approximately 17 kDa (Fig. 2), which is consistent with the expected molecular weight of the protein, 18.6 kDa. Slight amounts of impurities could also be observed. The protein was concentrated to 32 mg ml^{-1} and a wide range of crystallization conditions (288 drops) were screened by vapour diffusion in small-volume sitting drops. Although no crystals were obtained, some drops containing PEG 8000 at pH 6.5 and 10.5 contained crystalline-looking precipitates. After fine-screening of buffer composition, pH and precipitant concentration in larger drops (24 drops in total), large plate-shaped crystals were obtained at pH 6.5 (Fig. 3); these were used for native data collection. Crystals of the same shape were obtained at pH 7.5 and were used for derivatization with methylmercury chloride. The crystals reached maximal dimensions of around 1.0 × 0.5 × 0.05 mm (Fig. 3) in two weeks. Smaller, thinner crystals were obtained in many of the other drops.

The crystals were fragile and needed careful manipulation during harvesting, derivatization and soaking with cryoprotectant solution, which could only be successfully performed in the same drop that the crystals had grown in. Transfer of a crystal from the drop to another solution always led to visible degradation and loss of diffraction properties. Several data sets were collected; well harvested crystals generally diffracted to high resolution, but often showed high mosaicity and/or multiple lattices, making indexing and/or integration impossible. However, we were successful in collecting both a reasonable native and derivative data set from two independent crystals.

The best native crystal diffracted to a resolution of at least 1.8 Å and belonged to the *P*1 triclinic space group. The crystals are most likely to have four molecules of protein per unit cell and a solvent content of 48% (Matthews coefficient of 2.3 Å^3 Da^{-1}; Matthews, 1968). In the highest resolution shell, the merging *R* factor is still reasonable and the signal-to-noise ratio is still relatively high (Table 1). If the detector were to have been positioned closer, useful data to somewhat higher resolution might have been obtained (Evans & Murshudov, 2013). Unfortunately, owing to the low symmetry and the low resulting multiplicity of the data, correlation coefficients between half data sets could not be calculated. With hindsight, we would have collected data corresponding to a total rotation range of 360° instead of 180°, but indexing of the first few images suggested a monoclinic primitive lattice instead of a triclinic space group, on which the data-collection strategy was based. Additional data may well also have been of lower quality, as the crystal visibly suffered from radiation damage after data collection.

When the protein sequence was compared against known structures in the Protein Data Bank, no highly similar sequences were obtained. The top hit was the solution structure of the *S. saprophyticus* CHAP domain solved by nuclear magnetic resonance spectroscopy (PDB entry 2k3a; Rossi *et al.*, 2009), with 26 identical amino acids over a stretch of 94 residues (28%). Structure solution attempts by molecular replacement using this model or less homologous crystallographic models did not give positive results, and therefore derivatization with methylmercury chloride was performed.

The methylmercury-derivative crystals diffracted to 1.7 Å resolution and belonged to the same *P*1 triclinic space group with virtually identical unit-cell parameters. Data were collected at the mercury L_I edge, rather than at the L_{III} edge as intended, owing to misinterpretation of the remote data-collection interface. As a result, the data contained limited anomalous signal; they were also rather incomplete at high resolution owing to spot overlap (Table 1; up to 2.14 Å resolution, 96% of the data were present at 3.3-fold multiplicity). Nevertheless, the derivative crystal structure could be solved after careful data processing and automated phase determination using single-wavelength anomalous dispersion (SAD). Seven potential mercury sites were found; as the $CHAP_K$ protein contains one cysteine residue (Cys54), four were expected. After automatic model building, 619 residues in ten chains were present, clustered into four independent globular protein molecules. If indeed four molecules are present, 94% of the 660 expected residues have been automatically built and docked into sequence. After the refinement included in the automated procedure, values of 23.3 and 28.4% were obtained for R_{work} and R_{free}, respectively. Full refinement of the native and derivative structures, manual model adjustment and structure analysis will be performed and will be reported elsewhere.

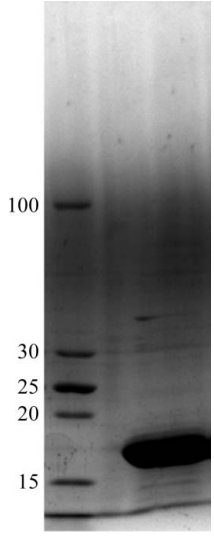

Figure 2
Denaturing gel electrophoresis of purified protein used for crystallization. Low-range molecular-weight markers, with the sizes indicated in kDa, were loaded in the left lane. A sample of the purified protein was loaded in the right lane.

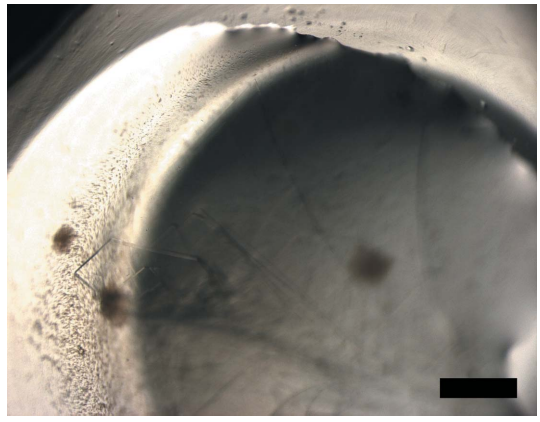

Figure 3
Crystals of the *S. aureus* bacteriophage K endolysin CHAP domain. Crystals grew as clusters of thin, but sizable, plates from which single crystals could be separated by careful manipulation. The size of the scale bar is around 0.5 mm.

Table 1
Crystallographic data collection.

Values in parentheses are for the highest resolution bin. ND: not determined (see text).

	Native	Methylmercury derivative
Synchrotron, beamline	ESRF, ID23-2	DLS, I02
Detector	MarMOSAIC	PILATUS 6M
Crystal-to-detector distance (mm)	170	529
Wavelength (Å)	0.8726	0.8352
No. of images	180	900
Oscillation range (°)	1.0	0.4
Unit-cell parameters (Å, °)	a = 39.2, b = 61.5, c = 73.1, α = 91.5, β = 98.6, γ = 90.0	a = 39.0, b = 61.5, c = 72.8, α = 91.8, β = 98.7, γ = 90.0
Resolution range (Å)	36.53–1.79 (1.88–1.79)	71.92–1.71 (1.76–1.71)
Reflections	62028 (8788)	48498 (572)
Multiplicity	2.0 (1.9)	3.4 (3.2)
Completeness (%)	97.2 (94.3)	64.7 (10.4)
Mean $I/\sigma(I)$	6.3 (2.9)	11.8 (1.8)
$R_{\rm merge}$† (%)	9.1 (25.1)	6.0 (62.0)
CC Imean	ND	0.994 (0.666)
CC anom	ND	0.354 (−0.001)
Wilson B (Å^2)	14.5	14.9
Phasing		
Correlation coefficient (all/weak)		37.79/22.57
Patterson figure of merit		53.52
Correlation coefficient (E)		0.328
$R_{\rm Cullis}$ (anomalous)		0.853
Anomalous phasing power		0.888
Figure of merit		0.305
Solvent flattening (45.6% solvent)		
Final correlation on $\|E\|^2$/contrast		2.2122
Hand score = correlation on $\|E\|^2$/contrast (original/inverted)		0.3245/0.2371

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$ where $I_i(hkl)$ is the intensity of the ith measurement of the same reflection and $\langle I(hkl)\rangle$ is the mean observed intensity for that reflection.

On the basis of the structure, it should be possible to formulate hypotheses about the reaction mechanism, which could then be tested by site-directed mutation and biochemical methods. This knowledge, in turn, might lead to recombinant enzymes with an improved exolysis mechanism or different pathogen specificities.

We thank Jordi Benach (ALBA beamline BL13/XALOC), Max Nanao (ESRF beamline ID23-2), Christian Perrin (ESRF) and James Sandy (DLS beamline I02) for valuable help with synchrotron data collection. We thank ALBA/CELLS (proposal No. 2012010140), the European Synchrotron Radiation Facility (proposal No. MX1477) and the Diamond Light Source (proposal No. MX3808) for access, which contributed to the results presented here. This research was sponsored by grant BFU2011-24843 (MJvR) from the Spanish Ministry of Economy and Competitiveness, a Masters fellowship (MSG) and a predoctoral FPU fellowship (CGD), the latter two from the Spanish Ministry of Education, Culture and Sports. We also acknowledge financial support from TSR-StrandIII:CRS/07/CR03 and FIRM:08RDCIT600 from the Irish Department of Agriculture.

References

Abrahams, J. P. & Leslie, A. G. W. (1996). *Acta Cryst.* D**52**, 30–42.
Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. W. (2011). *Acta Cryst.* D**67**, 271–281.
Becker, S. C., Dong, S., Baker, J. R., Foster-Frey, J., Pritchard, D. G. & Donovan, D. M. (2009). *FEMS Microbiol. Lett.* **294**, 52–60.
Becker, S. C., Foster-Frey, J. & Donovan, D. M. (2008). *FEMS Microbiol. Lett.* **287**, 185–191.
Bricogne, G., Vonrhein, C., Flensburg, C., Schiltz, M. & Paciorek, W. (2003). *Acta Cryst.* D**59**, 2023–2030.
Evans, P. R. (2011). *Acta Cryst.* D**67**, 282–292.
Evans, P. R. & Murshudov, G. N. (2013). *Acta Cryst.* D**69**, 1204–1214.
Fenton, M., Cooney, J. C., Ross, R. P., Sleator, R. D., McAuliffe, O., O'Mahony, J. & Coffey, A. (2011). *Bacteriophage*, **1**, 198–206.
Fenton, M., Keary, R., McAuliffe, O., Ross, R. P., O'Mahony, J. & Coffey, A. (2013). *Int. J. Microbiol.* **2013**, 625341.
Fenton, M., Ross, R. P., McAuliffe, O., O'Mahony, J. & Coffey, A. (2011). *J. Appl. Microbiol.* **111**, 1025–1035.
Holden, M. T. *et al.* (2013). *Genome Res.* **23**, 653–664.
Horgan, M., O'Flynn, G., Garry, J., Cooney, J., Coffey, A., Fitzgerald, G. F., Ross, R. P. & McAuliffe, O. (2009). *Appl. Environ. Microbiol.* **75**, 872–874.
Kabsch, W. (2010). *Acta Cryst.* D**66**, 125–132.
Laemmli, U. K. (1970). *Nature (London)*, **227**, 680–685.
Langer, G., Cohen, S. X., Lamzin, V. S. & Perrakis, A. (2008). *Nature Protoc.* **3**, 1171–1179.
Loeffler, J. M., Nelson, D. & Fischetti, V. A. (2001). *Science*, **294**, 2170–2172.
Loessner, M. J. (2005). *Curr. Opin. Microbiol.* **8**, 480–487.
Low, L. Y., Yang, C., Perego, M., Osterman, A. & Liddington, R. C. (2005). *J. Biol. Chem.* **280**, 35433–35439.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.
O'Flaherty, S., Coffey, A., Meaney, W., Fitzgerald, G. F. & Ross, R. P. (2005). *J. Bacteriol.* **187**, 7161–7164.
O'Flaherty, S., Ross, R. P., Meaney, W., Fitzgerald, G. F., Elbreki, M. F. & Coffey, A. (2005). *Appl. Environ. Microbiol.* **71**, 1836–1842.
Olson, M. E., Ceri, H., Morck, D. W., Buret, A. G. & Read, R. R. (2002). *Can. J. Vet. Res.* **66**, 86–92.
Otto, M. (2008). *Curr. Top. Microbiol. Immunol.* **322**, 207–228.
Ralston, D. J., Baer, B. S., Lieberman, M. & Krueger, A. P. (1955). *Proc. Soc. Exp. Biol. Med.* **89**, 502–507.
Ralston, D. J. & McIvor, M. (1967). *J. Bacteriol.* **88**, 676–681.
Rees, P. J. & Fry, B. A. (1981). *J. Gen. Virol.* **53**, 293–307.
Rossi, P., Aramini, J. M., Xiao, R., Chen, C. X., Nwosu, C., Owens, L. A., Maglaqui, M., Nair, R., Fischer, M., Acton, T. B., Honig, B., Rost, B. & Montelione, G. T. (2009). *Proteins*, **74**, 515–519.
Schuch, R., Nelson, D. & Fischetti, V. A. (2002). *Nature (London)*, **418**, 884–889.
Sheldrick, G. M. (2008). *Acta Cryst.* A**64**, 112–122.
Sonstein, S. A., Hammel, J. M. & Bondi, A. (1971). *J. Bacteriol.* **107**, 499–504.
Vonrhein, C., Blanc, E., Roversi, P. & Bricogne, G. (2007). *Methods Mol. Biol.* **364**, 215–230.
Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.
Winter, G., Lobley, C. M. C. & Prince, S. M. (2013). *Acta Cryst.* D**69**, 1260–1273.