research papers

Structure and mechanism of a type III secretion protease, NleC

Wenqi Li,[a,b]‡ Yexing Liu,[b,c]‡ Xinlei Sheng,[b,c]‡ Ping Yin,[b,c] Feizhuo Hu,[b,c] Ying Liu,[b,c] Chen Chen,[b,c] Quanxiu Li,[b,c] Chuangye Yan[b,c] and Jiawei Wang[a,b]*

[a]State Key Laboratory of Biomembrane and Membrane Biotechnology, School of Life Sciences, Tsinghua University, Beijing 100084, People's Republic of China, [b]Center for Structural Biology, School of Life Sciences, Tsinghua University, Beijing 100084, People's Republic of China, and [c]Ministry of Education Key Laboratory of Protein Science, School of Life Sciences, Tsinghua University, Beijing 100084, People's Republic of China

‡ These authors contributed equally to this work.

Correspondence e-mail: jwwang@tsinghua.edu.cn

NleC is one of the virulence factors that is injected into infected host cells by enteropathogenic and enterohaemorrhagic *Escherichia coli* (EPEC and EHEC) *via* a needle-like protein complex called the type III secretion system (T3SS). The cytosolic NleC specifically cleaves the p65 subunit of NF-κB in the p65–p50 heterodimeric complex just after the Cys38 site in its N-terminal domain. The degradation of the remainder of the p65 C-terminal domain by the proteasome disrupts the NF-κB signalling pathway, thus dampening the host inflammatory response. Here, the crystal structure of NleC is reported at 1.55 Å resolution. In conjunction with biochemical analyses, the structure reveals that NleC is a member of the zincin zinc protease family and that the configuration of the NleC active site resembles that of the metzincin clan of metallopeptidases but without the canonical Met turn of astacin. The extended zinc-binding motif of NleC (HE*XX*H*XX*T*XXX*D) includes three metal ligands. The fifth zinc ligand, a conserved tyrosine (a bound water molecule is the fourth ligand), lies 45 residues downstream of the zincin motif. Furthermore, the electrostatic potential complementarity between NleC and p65 also contributes to the cleavage activity of the protease. These results not only provide important insights into the mechanism of how NleC recognizes its substrates, but also shed light on the design of new antibiotics for the food-borne diseases arising from EPEC and EHEC.

1. Introduction

Enteropathogenic *Escherichia coli* (EPEC) and enterohaemorrhagic *E. coli* (EHEC) are non-invasive 'attaching and effacing' (A/E) pathogens that are transmitted in contaminated food (Moon *et al.*, 1983; Jerse *et al.*, 1990). EPEC and EHEC cause food-borne diseases worldwide and colonize the cells of the epithelial lining of the intestinal tract (Clarke, 2001). In response, the affected cells initiate an immune response by secreting cytokines that attract immune cells. To prevent their early elimination by the host, these types of intestinal bacteria have evolved an organelle, best known as the type III secretion system (T3SS), to deliver bacterial effector proteins directly from the bacterial cytoplasm into the host cytoplasm to disrupt the NF-κB signalling pathway, which is an essential effector of the host-cell immune response (Garmendia *et al.*, 2005).

Most of the effector proteins in A/E pathogens are encoded at the locus of enterocyte effacement (LEE) and suppress the host inflammatory response in a T3SS-dependent way

(Garmendia *et al.*, 2005). In addition to the LEE effector proteins, a variety of non-LEE effector proteins, *e.g.* NleE (Nadler *et al.*, 2010), NleB (Newton *et al.*, 2010), NleH (Gao *et al.*, 2009) *etc.*, can also reduce inflammatory responses. All of these effector proteins use different mechanisms to modify the host immune response. NleE acts by either directly or indirectly targeting the activation of the IKK complex and IκB degradation (Nadler *et al.*, 2010), while NleB interferes with the upstream components of the TNF-α signalling pathway (Newton *et al.*, 2010). However, all of the effectors identified above cannot completely account for the full immunosuppressive potential of these pathogens (Gao *et al.*, 2009; Nadler *et al.*, 2010; Newton *et al.*, 2010). Yen and coworkers not only identified NleC as yet another type III secretion effector (Yen *et al.*, 2010), but also demonstrated that NleC and NleE are necessary and function cooperatively to inhibit NF-κB activation in a T3SS-deficient mutant (Pearson *et al.*, 2011; Sham *et al.*, 2011). Baruch and coworkers showed that NleC and NleD cooperate and complement other EPEC effectors in accomplishing maximal inhibition of IL-8 secretion (Baruch *et al.*, 2011).

NleC, a zinc-dependent endopeptidase, specifically clips and inactives the p65 subunit of NF-κB in the cytoplasm, and the remaining p65 fragments are digested by the proteasome (Yen *et al.*, 2010; Baruch *et al.*, 2011). Disruption of the p65–p50 complex by NleC prevents the p65–p50 NF-κB from entering the nucleus to induce the transcription of various inflammatory genes, including IL-8. In addition to p65, the host acetyltransferase p300 has been shown to be an additional target of NleC (Shames *et al.*, 2011) and to also be clipped by the metalloprotease domain of NleC. An *in vitro* assay by Yen and coworkers showed that NleC could not digest p50 (Yen *et al.*, 2010). However, Pearson and coworkers suggested that the delivery of NleC by the T3SS of EPEC induced degradation of p65 in infected cells as well as other NF-κB components, *e.g.* c-Rel and p50 (Pearson *et al.*, 2011). It remains controversial whether the cleavage site of p65 by NleC is between residues 10 and 11 (Yen *et al.*, 2010) or residues 38 and 39 (Baruch *et al.*, 2011).

2. Methods

2.1. Protein preparation

Full-length NleC and p65 were subcloned from the *E. coli* (strain O157:H7) cDNA library using a standard PCR protocol. Two-step PCR was employed to generate mutants. All mutants were verified by sequencing the plasmids, which were then transformed and expressed in *E. coli* strain BL21 (DE3) and the proteins were purified using the following protocol. The *E. coli* cells were grown in 1 l Luria–Bertani medium in 2 l flasks. Protein expression was induced with 0.2 m*M* isopropyl β-D-1-thiogalactopyranoside (IPTG) when the cell density reached an OD$_{600}$ of 1.0. After growth at 289 K for 16 h, the cells were shaken at 3000 rev min^{-1} for 10 min and resuspended in lysis buffer consisting of 25 m*M* Tris–HCl pH 8.0, 150 m*M* NaCl. The harvested cells were disrupted by

Table 1
Statistics of data collection and refinement.

Values in parentheses are for the highest resolution shell.

Data	Se-SAD	Native
Integration package	*HKL*-2000	*HKL*-2000
Space group	$P2_1$	$P2_12_12_1$
Unit-cell parameters (Å, °)	$a = 59.51$,	$a = 45.09$,
	$b = 88.66$,	$b = 67.55$,
	$c = 110.65$,	$c = 81.05$,
	$\alpha = \gamma = 90$,	$\alpha = \beta = \gamma = 90$
	$\beta = 92.89$	
No. of molecules in asymmetric unit	4	1
Wavelength (Å)	0.9793	0.9793
Resolution (Å)	40–2.30	40–1.55
R_{merge} (%)†	8.4 (41.1)	6.7 (43.6)
$\langle I/\sigma(I)\rangle$	10.7 (2.1)	19.2 (2.8)
Completeness (%)	98.1 (98.3)	99.8 (100.0)
Total No. of reflections	184220	259912
Unique reflections	98250	69523
Multiplicity	1.9 (1.8)	3.7 (3.6)
Wilson B factor (Å^2)	32.0	16.4
R factor (%)‡	25.27	17.05
R_{free} (%)‡	28.02	19.42
Average B value (Å^2)		
Overall	43.97	23.54
Main chain	44.04	20.20
Side chain	43.58	24.08
Water	48.03	32.58
Other entities		15.24
R.m.s. deviations from ideal values		
Bonds (Å)	0.016	0.007
Angles (°)	1.329	1.184
Ramachandran plot statistics (%)		
Most favourable	90.8	90.8
Additionally allowed	9.0	8.7
Generously allowed	0.1	0.4
Disallowed	0.0	0.0

† $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle|/\sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the mean intensity of i observations of symmetry-related reflections *hkl*. ‡ $R = \sum_{hkl} ||F_{obs}| - |F_{calc}||/\sum_{hkl} |F_{obs}|$, where F_{calc} is the calculated protein structure factor from the atomic model (R_{free} was calculated with a randomly selected 5% of the reflections).

sonication at a power of 800 W at 60% amplitude for 6 min. Cell debris was removed by low-speed centrifugation at 13 000 rev min^{-1} for 1 h at 277 K. Proteins were concentrated and purified using nickel–nitrilotriacetate affinity resin (Ni–NTA column, Qiagen), anion-exchange chromatography (Source 15Q, GE Healthcare) and size-exclusion chromatography (Superdex 200, GE Healthcare), successively. After removal of the hexahistidine tag with the enzyme drICE, the proteins were concentrated to approximately 10 mg ml^{-1} before use in crystallization.

2.2. Crystallization

Crystals were grown at 291 K by the hanging-drop vapour-diffusion method. Extensive crystallization trials were conducted for distinctive NleC mutants using various well conditions. To improve the diffraction resolution, various approaches were employed, including construct modification, crystal optimization, crystal-growth condition screening, various small-molecule additives and post-crystallization manipulation. Finally, the native crystals belonging to space group $P2_12_12_1$ diffracted X-rays to about 1.55 Å resolution

at Shanghai Synchrotron Radiation Facility (SSRF). Crystals were obtained using a well buffer consisting of 0.22 M potassium fluoride, 23%(w/v) polyethylene glycol 3000; selenomethionine-substituted NleC crystals belonging to space group $P2_1$ were obtained under similar conditions. Crystals appeared in 2 d and grew to form needle-shaped crystals in 5 d.

2.3. Data collection, structure determination and refinement

Anomalous and native X-ray data were collected on SPring-8 beamline BL41XU and SSRF beamline BL17U, respectively. The data were integrated and scaled using *HKL*-2000 (Otwinowski & Minor, 1997). Further processing was carried out using programs from the *CCP*4 suite (Winn *et al.*, 2011). The selenium positions were determined using *SHELXD* (Sheldrick, 2008). The identified selenium sites were then refined and initial phases were calculated using *Phaser* (McCoy *et al.*, 2007; Read & McCoy, 2011). Density modification with *DM* (Cowtan, 1994) gave a map of sufficient quality for model building. An initial model was built into the experimental map using *Coot* (Emsley & Cowtan, 2004; Emsley *et al.*, 2010). The sequence docking was aided by the selenium sites in the anomalous difference Fourier map. The structure was refined with *phenix.refine* (Adams *et al.*, 2002, 2010). Model validation was performed with *PROCHECK* (Laskowski *et al.*, 1993) and the *WHATCHECK* routine of *WHAT IF* (Vriend, 1990). All structure figures were prepared with *PyMOL* (DeLano, 2002). The electrostatics of the structures were also calculated in *PyMOL* using a quasi-Coulombic-shaped convolution function.

2.4. Protease assay

In order to compare the protease activity of the NleC mutants, NleC was diluted in 25 mM Tris–HCl pH 8.0, 150 mM NaCl. 1 mg ml^{-1} p65 was incubated with NleC to measure the protease activity of NleC. Each reaction took place for 30 min at a 1:500 molar ratio of NleC to p65 at 293 K. EDTA was added to the mixture to eliminate zinc ions, ending the reaction, after 30 min. SDS–PAGE was used to estimate the amount of degraded p65 semi-quantitatively.
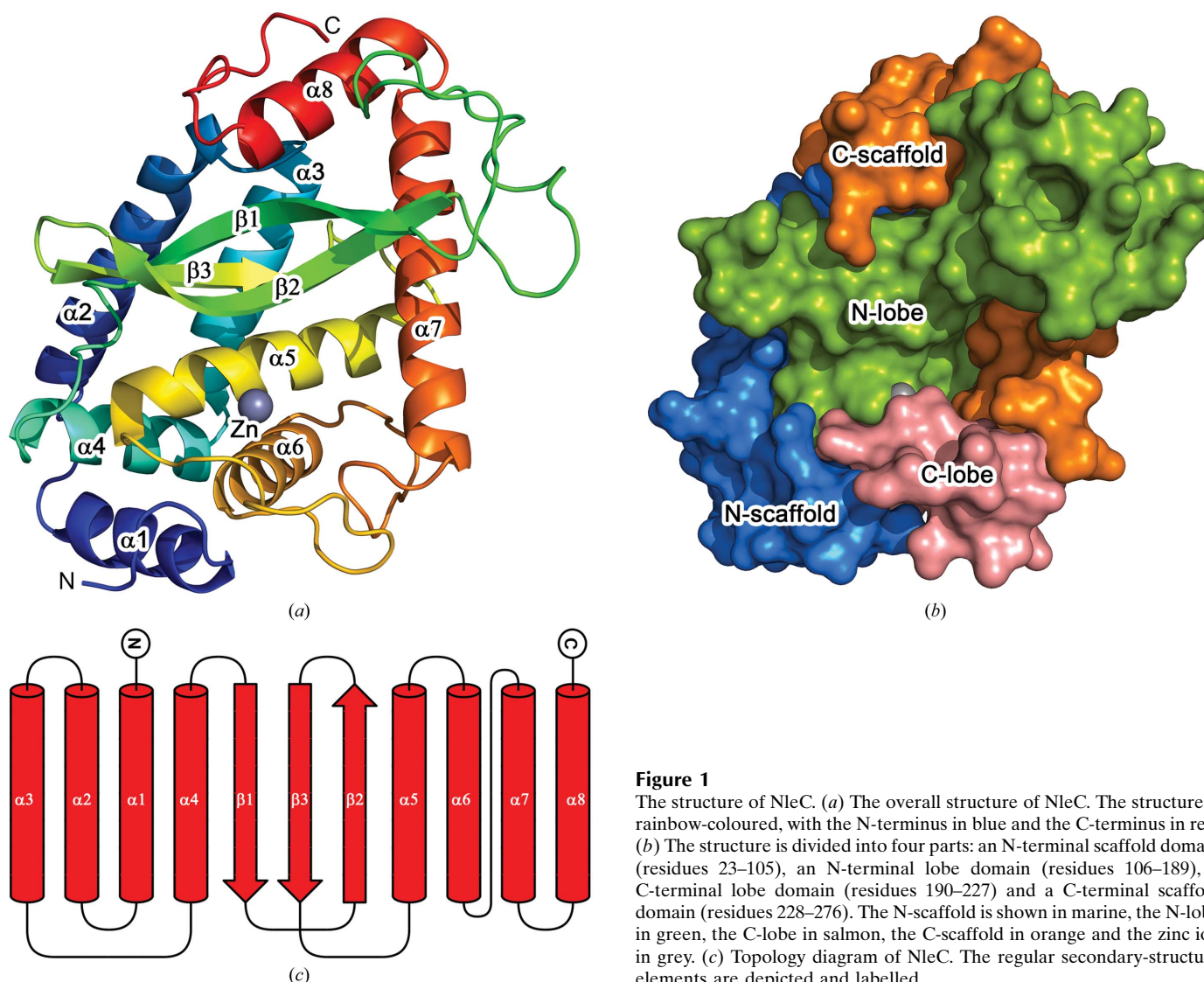
Figure 1
The structure of NleC. (*a*) The overall structure of NleC. The structure is rainbow-coloured, with the N-terminus in blue and the C-terminus in red. (*b*) The structure is divided into four parts: an N-terminal scaffold domain (residues 23–105), an N-terminal lobe domain (residues 106–189), a C-terminal lobe domain (residues 190–227) and a C-terminal scaffold domain (residues 228–276). The N-scaffold is shown in marine, the N-lobe in green, the C-lobe in salmon, the C-scaffold in orange and the zinc ion in grey. (*c*) Topology diagram of NleC. The regular secondary-structure elements are depicted and labelled.
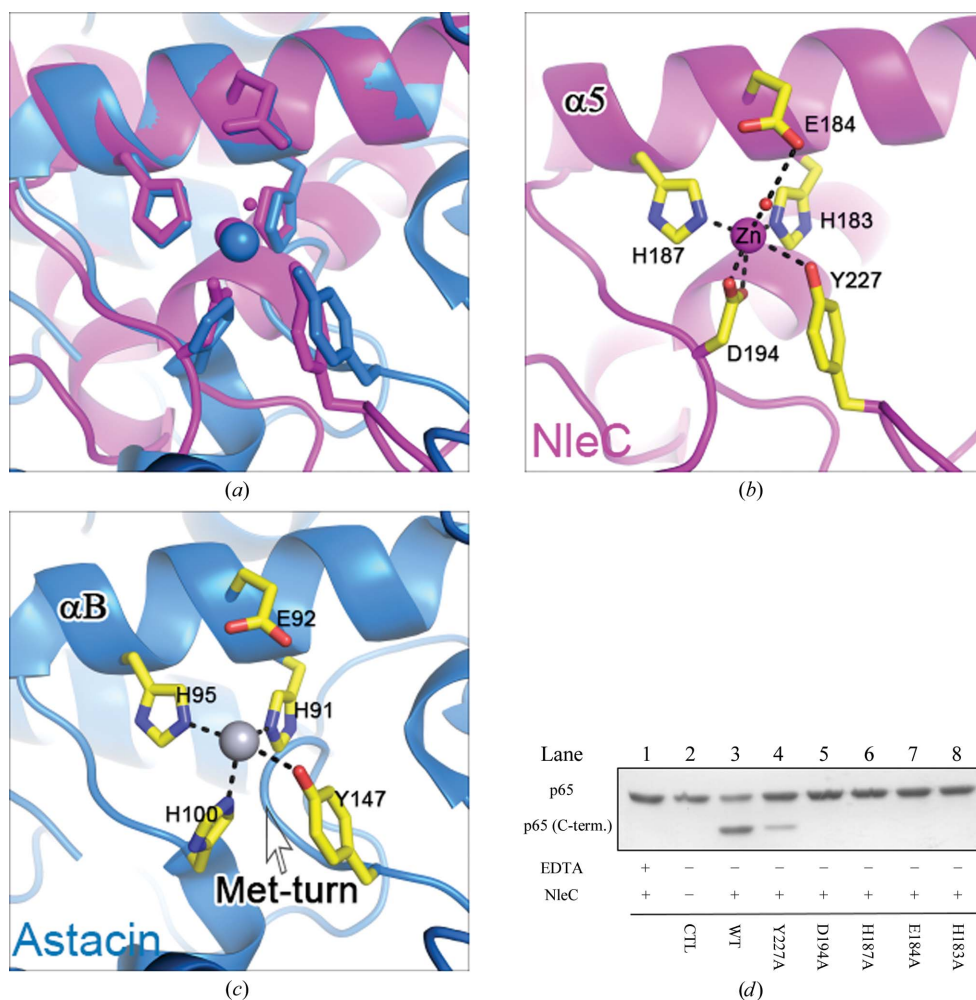
Figure 2
Superposition of the active sites of the catalytic domains of astacin and NleC. (*a*) The superposed active sites of astacin and NleC. The astacin catalytic domain is shown in marine and that of NleC in magenta. (*b*) The active site of NleC. (*c*) The active site of the catalytic domain of astacin. (*d*) NleC shown to be a zincin protease. The mutation of any of the zinc-binding ligands disrupts the protease activity of NleC (lanes 5–8), with the exception of the fifth ligand Tyr227. The Y227A mutant crippled the protease activity (lane 4).

3. Results

3.1. Structure of NleC

The NleC protein (residue 1–330) was expressed as a His-tag fusion protein and the His-tag moiety was removed by drICE cleavage in preparation for crystallization. The full-length NleC protein defied all crystallization attempts, likely owing to the flexibility of its N-terminus and C-terminus. We attempted to shorten this sequence without affecting the NleC protease activity (Supplementary Fig. S1[1]). Removal of the N-terminus (residues 1–26) had no apparent impact on the protease activity and neither did removal of the C-terminus (residues 294–330). Crystals were finally generated of the NleC fragment 23–293. We determined the structure by a SAD experiment using a crystal of selenomethionine-substituted protein. The final atomic model was refined at 1.55 Å resolution (Table 1).

The structure of NleC consists of eight α-helices and three β-strands (Fig. 1*a*). Overall, NleC shows a compact ellipsoid-like shape with maximal dimensions of approximately 39.6 × 46.9 × 26.1 Å. The whole structure was artificially divided into four subdomains of approximately 80 residues: an N-terminal scaffold domain (residues 23–105), an N-terminal lobe domain (residues 106–189), a C-terminal lobe domain (residues 190–227) and a C-terminal scaffold domain (residues 228–276). The Zn atom lies in a deep and narrow active-site cleft between the N-lobe and the C-lobe. The C-shaped N-scaffold domain and the hoop-shaped C-scaffold domain glue the upper N-lobe and the lower C-lobe together. The crystal structure of α1-truncated NleC shows almost no visible density for the C-lobe, except for helix α6 (data not shown).

The N-lobe is a strongly twisted three-stranded β-sheet (β1–β3); its lowermost strand (β2) creates the 'upper rim' of the active-site cleft. The 'active-site helix' (α5) stretches nearly parallel to the strands of the sheet. Helix α5 includes the first three residues of the long zinc-binding consensus sequence $H^{183}EXXHXXTXXXD^{194}$ (single-letter amino-acid code; *X* stands for any residue). Gly191 is the endpoint of both helix α5 and the N-lobe. This glycine allows a sharp turn in the trajectory of the polypeptide so that the peptide chain is able to enter the C-lobe. The C-lobe contains the third zinc-binding residue Asp194. Finally, a tyrosine in the loop between helices α6 and α7 is also engaged in zinc binding. The N-scaffold and C-scaffold domains only contain helices. The helix α7 in the C-scaffold creates the 'right rim' of the active-site cleft.

An exhaustive search of the Protein Data Bank using *DALI* (Holm & Rosenström, 2010) failed to identify any entry that has a closely similar overall folding to NleC. The entry with the highest *Z*-score (a measure of structural similarity) of 6.4 is tricorn protease interactive factor F3, which has a root-mean-square deviation of 3.4 Å for 118 aligned amino acids. This analysis suggests that the structure of NleC may represent a previously uncharacterized fold.
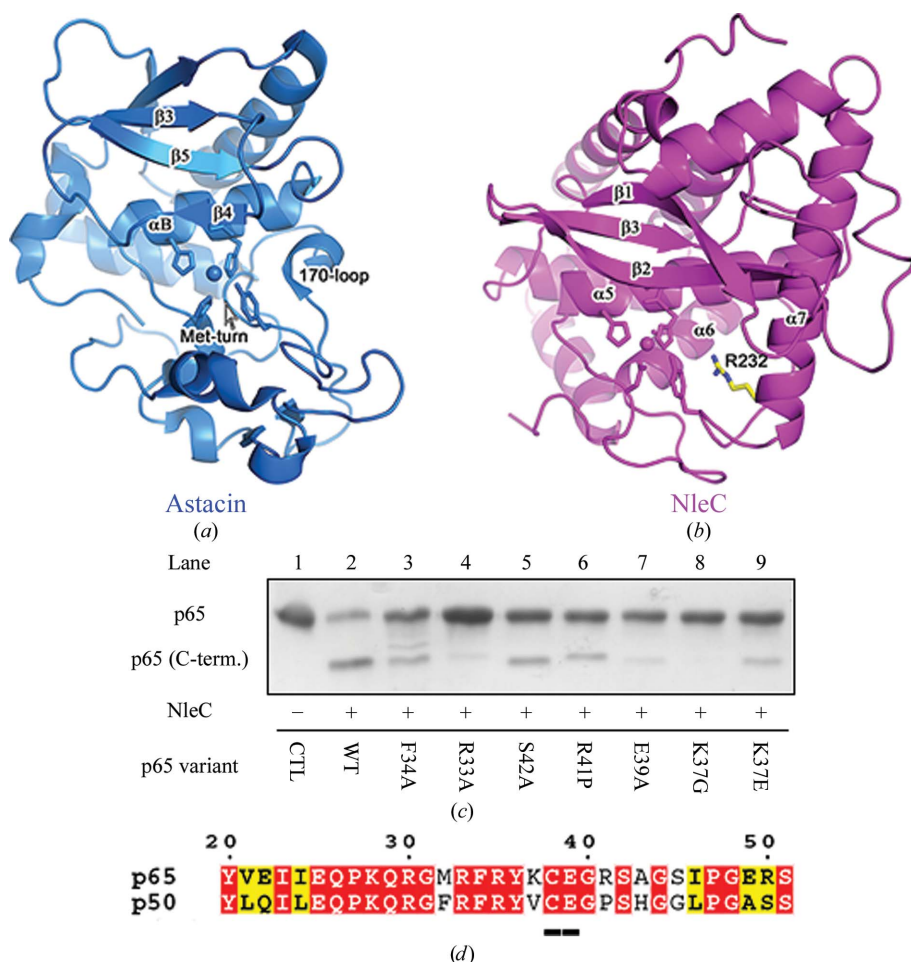
3.2. Active-site cleft and zinc-binding site

The consensus sequence HE*XX*H in NleC is highly conserved across species (Supplementary Fig. S2). A bioinformatics search identified NleC as a zinc protease (Yen *et al.*, 2010). According to the structural analysis, there are five ligands binding to the Zn atom. This binding configuration resembles the metzincin clan of metallopeptidases, except for the conserved methionine in the so-called 'Met-turn' located below the catalytic zinc site, for example in the astacin family and the serratia family (Hooper, 1994). When the active-site residues of astacin, a zinc endopeptidase from the crayfish *Astacus astacus* L. (PDB entry 1iac; Gomis-Rüth *et al.*, 1993), were superposed onto the zinc-binding residues in the structure of NleC, the two active-site residues from the two different folded structures superposed consistently. These observations suggest that astacin and NleC might use a similar mechanism to cleave the substrate peptide.

The addition of a metal-ion chelating agent such as EDTA abrogated the function of the NleC (Fig. 2*d*, lane 1), which

suggests that NleC is a zinc metallopeptidase (Yen *et al.*, 2010). An anomalous difference Fourier map from data collected at the absorption edge of zinc confirmed the existence and the exact position of the bound Zn atom (Supplementary Fig. S3). Mis-sense mutants of the zinc-binding residues all diminished the protease activity of NleC, except for the Tyr227 mutant, which only crippled the peptidase activity. This observation for the Tyr227 mutant might be attributed to the 'tyrosine switch' mechanism found in serralysins and, most likely, pappalysin family members within the metzincins (Gomis-Rüth, 2003).

The catalytic zinc ion resides at the bottom of the active-site cleft. The active-site helix $\alpha 5$ provides two histidine ligands of the metal (His183 and His187) separated by a single helical turn that allows a concerted approach to the metal together with the third zinc-binding aspartate, Asp197. In NleC, the $N^{\varepsilon 2}$ atoms of the two histidines and the $O^{\delta 1}$ and $O^{\delta 2}$ atoms of Asp197 coordinate the metal together with a catalytic solvent molecule anchored to the general acid/base Glu184. The overall metal coordination is tetragonal bipyramidal, with His187 $N^{\varepsilon 2}$ and Tyr227 OH at apical, somewhat larger distances (2.9 Å) and the remaining three ligands in a plane with the metal at distances of ~ 2 Å (Fig. 2*b*).

3.3. N-lobe and C-lobe conformation and the cleavage site of p65

In addition to the resemblance of the detailed active-site configuration between NleC and astacins, the secondary structures in the N-lobe domain are also similar. However, there is no Met-turn beneath the zinc-binding site in the NleC structure (Fig. 3*a*). A strictly conserved methionine (Met147) in astacins forms a Met-turn (methionine-containing turn) conformation which is proposed to act as a plug that inserts laterally into a core structure created by the protein segment engaged in zinc binding, thus contributing to the structural integrity that is indispensable for function (Gomis-Rüth *et al.*, 2012). Nonetheless, there is no such methionine residing beneath the zinc-binding site in the NleC structure; instead, helix $\alpha 6$ is present (Fig. 3*b*).

To determine the exact cleavage point of p65 by NleC, the cleaved p65 was subjected to N-terminal sequencing. The obtained sequence, EGRSA, indicates that NleC cleaves p65 within its conserved DNA-binding domain (Chen *et al.*, 1998) after residue Cys38 (Baruch *et al.*, 2011). Mass-spectrometric results and pull-down experiments (data not shown) also suggest the Cys38 cleavage

Figure 3
Overall structure of the astacin catalytic domain and NleC in comparable orientations. (*a*) Cartoon representation of astacin depicting its secondary structure in the N-lobe subdomain together with the Met-turn and the 170-loop. (*b*) Cartoon representation of NleC; the secondary structures corresponding to those in astacin are labelled. The conserved residue Arg232 is shown in stick representation. (*c*) Activity of mutants of p65 around the scissile bond. (*d*) Sequence alignment between p65 and p50 around the cleavage site.

site (Supplementary Fig. S4). The most important structural element for primed subsites in astacins is the 170-loop. Helix $\alpha 7$ of NleC, corresponding to the 170-loop of astacin, has a conserved arginine (Arg232) that shapes the bottom of a deep

S1′ pocket (Fig. 3b) and is engaged in a salt bridge with Glu39 in p65. This explains why NleC prefers the negatively charged side chain of a glutamate residue in S1′. The mutant E39A disrupts the protease activity of NleC (Fig. 3c, lane 7) owing to the disappearance of the salt-bridge interaction between Arg232 of NleC and Glu39 of p65. The position of NleC helix $\alpha 7$ blocks the primed subsite, corresponding to the 170-loop in astacin. This spatial restriction by helix $\alpha 7$ requires the conserved Gly40 of p65, which makes a sharp turn to avoid clashing with helix $\alpha 7$. The helix $\alpha 7$ is critical for NleC in the process of p65/p50 cellular elimination (Mühlen *et al.*, 2011).

3.4. Complementarity between the electrostatic surface potentials of NleC and p65

The electrostatic surface potential of NleC reveals a strongly negative face cradling the catalytic zinc, which may mediate the recognition of the positively charged NF-κB sub-units that are substrates of NleC (Figs. 4a and 4c). The mutation of negative residues on the supposed interaction interface of NleC and p65 disrupts the NleC protease activity (Fig. 4e), as indicated in Fig. 4(a).

The active-site cleft of astacin with the bound reaction-intermediate analogue BOC-Pro-Lys-PheΨ(PO$_2$)-CH$_2$-C(CH$_3$)CO-Pro-OCH$_3$ (PDB entry 1qji; Grams *et al.*, 1996) was superposed onto the NleC structure. The bound reaction-intermediate analogue from astacin fits the active-site cleft of NleC properly, with the peptide direction from the N-terminus to the C-terminus (Fig. 4c). While the DNA-binding loop of p65 (Fig. 4d) was trans-located into the NleC active-site cleft, with Cys38 occupying the S1 subsite and Gly39 occupying the S1′ subsite, the corresponding positively charged surface of p65 would touch the negatively charged surface of NleC. Any
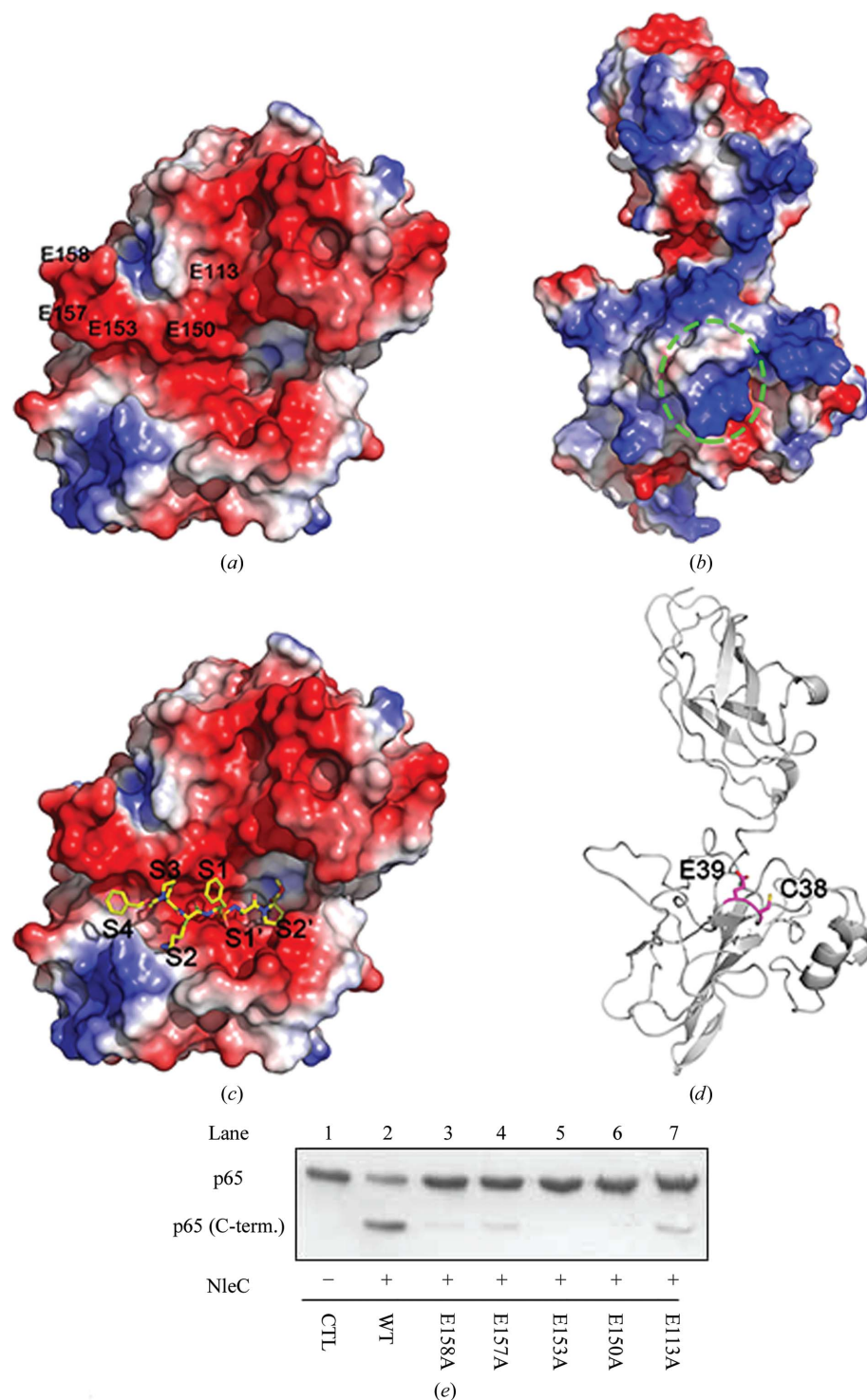
Figure 4
Electrostatic surface potentials of NleC and p65. (*a*) Surface potential of NleC. The mutated negatively charged glutamate residues are labelled. (*b*) Surface potential of p65. The DNA-binding loop cleaved by NleC is indicated by a green dashed circle. (*c*) Surface potential of NleC with the superposed reaction-intermediate analogue from astacin (PDB entry 1qji). (*d*) Cartoon representation of p65. The scissile bond between Cys38 and Glu39 is depicted in magenta. (*e*) Mutation of the negatively charged residues on the surface of NleC abolished the metalloprotease activity.
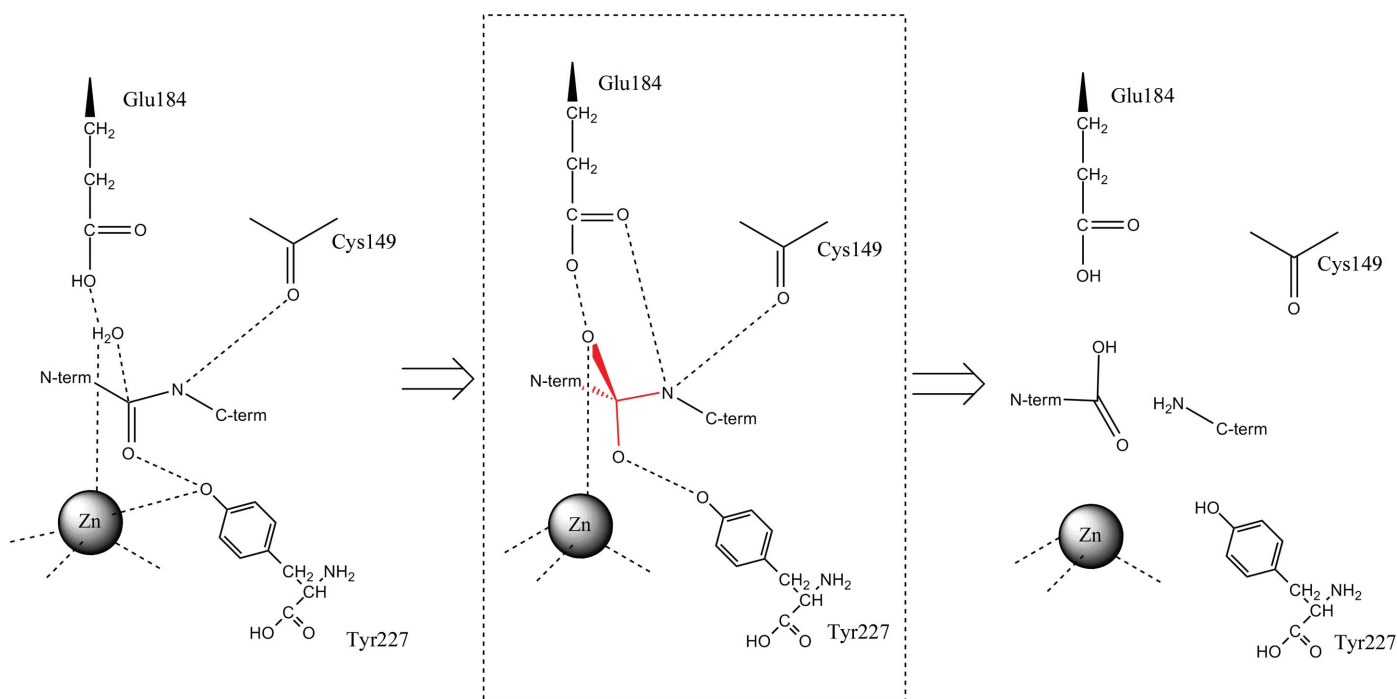
Figure 5
Scheme of the proposed catalytic mechanism of NleC. Substrate hydrolysis by NleC proceeds *via* a high-energy reaction intermediate such as that captured in the structure with PDB entry 1qji. This reaction intermediate is centred on the tetrahedral carboxyl C atom of Cys38 of p65 bound to two *gem*-diolate O atoms, the scissile-bond N atom of Glu39 and the C$^\alpha$ atom of Cys38. The tetrahedral C atom itself is derived from the original scissile-bond carbonyl group by the nucleophilic attack of a solvent molecule polarized by the general acid/base glutamate Glu184. The *para* O atom of Tyr227 is supposed to stabilize one of the *gem*-diolate O atoms of the tetrahedral C atom.

perturbation of the charge complementarity of the surfaces will disrupt the protease and substrate recognition and thus the protease activity (Fig. 4*e*).

4. Discussion

In this study, we report the structure of NleC from *E. coli* at a resolution of 1.55 Å. The configuration of the NleC active site resembles that of the metzincin clan of metallopeptidases, but without the canonical Met-turn of astacin. However, the astacin family of metalloendopeptidases encompasses proteins found in *Hydra* to humans in both mature and in developing systems (Bond & Beynon, 1995). These proteins are synthesized with N-terminal signal and pro-enzyme sequences, and many contain multiple domains at the C-terminal end of the protease domain. The activation of astacin requires proteolysis of the N-terminal propeptide to unveil the active site. These differences between the bacterial effector protein NleC and astacin show that pathogenic *E. coli* has evolved to mimic the proteolysis mechanism of eukaryotes to escape the innate immune system.

We speculate that substrate hydrolysis by NleC proceeds *via* a high-energy reaction intermediate such as that captured in the structure of PDB entry 1qji (Fig. 5). This reaction intermediate is centred on the tetrahedral carboxyl C atom of Cys38 of p65 bound to two *gem*-diolate O atoms, the scissile-bond N atom of Glu39 and the C$^\alpha$ atom of Cys38 (Fig. 5). The tetrahedral C atom itself is derived from the original scissile-

bond carbonyl group by the nucleophilic attack of a solvent molecule polarized by the general acid/base glutamate Glu184 (Gomis-Rüth, 2008; Matthews, 1988; Bayés *et al.*, 2007; Fig. 2*b*). In this intermediate state, the OH atom of Tyr227 is supposed to stabilize one of the *gem*-diolate O atoms of the tetrahedral C atom. This explains why the Y227A mutant only crippled the protease activity of NleC (Fig. 2*d*, lane 4).

These results not only provide important insights into the mechanism of how NleC recognizes its substrates, but also shed light on the design of new antibiotics for the food-borne diseases arising from EPEC and EHEC.

References

Adams, P. D. *et al.* (2010). *Acta Cryst.* D**66**, 213–221.
Adams, P. D., Grosse-Kunstleve, R. W., Hung, L.-W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W., Read, R. J., Sacchettini, J. C., Sauter, N. K. & Terwilliger, T. C. (2002). *Acta Cryst.* D**58**, 1948–1954.

Baruch, K., Gur-Arie, L., Nadler, C., Koby, S., Yerushalmi, G., Ben-Neriah, Y., Yogev, O., Shaulian, E., Guttman, C., Zarivach, R. & Rosenshine, I. (2011). *EMBO J.* **30**, 221–231.

Bayés, A., Fernández, D., Solà, M., Marrero, A., García-Piqué, S., Avilés, F. X., Vendrell, J. & Gomis-Rüth, F. X. (2007). *Biochemistry*, **46**, 6921–6930.

Bond, J. S. & Beynon, R. J. (1995). *Protein Sci.* **4**, 1247–1261.

Chen, F. E., Huang, D.-B., Chen, Y.-Q. & Ghosh, G. (1998). *Nature (London)*, **391**, 410–413.

Clarke, S. C. (2001). *Diagn. Microbiol. Infect. Dis.* **41**, 93–98.

Cowtan, K. (1994). *Jnt CCP4/ESF–EACBM Newsl. Protein Crystallogr.* **31**, 34–38.

DeLano, W. L. (2002). *PyMOL.* http://www.pymol.org.

Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.

Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.

Gao, X., Wan, F., Mateo, K., Callegari, E., Wang, D., Deng, W., Puente, J., Li, F., Chaussee, M. S., Finlay, B. B., Lenardo, M. J. & Hardwidge, P. R. (2009). *PLoS Pathog.* **5**, e1000708.

Garmendia, J., Frankel, G. & Crepin, V. F. (2005). *Infect. Immun.* **73**, 2573–2585.

Gomis-Rüth, F. X. (2003). *Mol. Biotechnol.* **24**, 157–202.

Gomis-Rüth, F. X. (2008). *Crit. Rev. Biochem. Mol. Biol.* **43**, 319–345.

Gomis-Rüth, F. X., Stöcker, W., Huber, R., Zwilling, R. & Bode, W. (1993). *J. Mol. Biol.* **229**, 945–968.

Gomis-Rüth, F. X., Trillo-Muyo, S. & Stöcker, W. (2012). *Biol. Chem.* **393**, 1027–1041.

Grams, F., Dive, V., Yiotakis, A., Yiallouros, I., Vassiliou, S., Zwilling, R., Bode, W. & Stöcker, W. (1996). *Nature Struct. Mol. Biol.* **3**, 671–675.

Holm, L. & Rosenström, P. (2010). *Nucleic Acids Res.* **38**, W545–W549.

Hooper, N. M. (1994). *FEBS Lett.* **354**, 1–6.

Jerse, A. E., Yu, J., Tall, B. D. & Kaper, J. B. (1990). *Proc. Natl Acad. Sci. USA*, **87**, 7839–7843.

Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.

Matthews, B. W. (1988). *Acc. Chem. Res.* **21**, 333–340.

McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.

Moon, H. W., Whipp, S. C., Argenzio, R. A., Levine, M. M. & Giannella, R. A. (1983). *Infect. Immun.* **41**, 1340–1351.

Mühlen, S., Ruchaud-Sparagano, M.-H. & Kenny, B. (2011). *J. Biol. Chem.* **286**, 5100–5107.

Nadler, C., Baruch, K., Kobi, S., Mills, E., Haviv, G., Farago, M., Alkalay, I., Bartfeld, S., Meyer, T. F., Ben-Neriah, Y. & Rosenshine, I. (2010). *PLoS Pathog.* **6**, e1000743.

Newton, H. J. *et al.* (2010). *PLoS Pathog.* **6**, e1000898.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Pearson, J. S., Riedmaier, P., Marchès, O., Frankel, G. & Hartland, E. L. (2011). *Mol. Microbiol.* **80**, 219–230.

Read, R. J. & McCoy, A. J. (2011). *Acta Cryst.* D**67**, 338–344.

Sheldrick, G. M. (2008). *Acta Cryst.* A**64**, 112–122.

Sham, H. P., Shames, S. R., Croxen, M. A., Ma, C., Chan, J. M., Khan, M. A., Wickham, M. E., Deng, W., Finlay, B. B. & Vallance, B. A. (2011). *Infect. Immun.* **79**, 3552–3562.

Shames, S. R., Bhavsar, A. P., Croxen, M. A., Law, R. J., Mak, S. H. C., Deng, W., Li, Y., Bidshari, R., de Hoog, C. L., Foster, L. J. & Finlay, B. B. (2011). *Cell. Microbiol.* **13**, 1542–1557.

Thompson, J. D., Gibson, T. & Higgins, D. G. (2002). *Curr. Protoc. Bioinformatics*, Unit 2.3. doi:10.1002/0471250953.bi0203s00.

Vriend, G. (1990). *J. Mol. Graph.* **8**, 52–56.

Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.

Yen, H., Ooka, T., Iguchi, A., Hayashi, T., Sugimoto, N. & Tobe, T. (2010). *PLoS Pathog.* **6**, e1001231.