Acta Crystallographica Section D
Biological Crystallography
ISSN 1399-0047

Visualization of a substrate-induced productive conformation of the catalytic triad of the *Neisseria meningitidis* peptidoglycan *O*-acetylesterase reveals mechanistic conservation in SGNH esterase family members

Allison H. Williams,[a,b]* Frédéric J. Veyrier,[a,b,c] Mathilde Bonis,[a,b] Yann Michaud,[a,b] Bertrand Raynal,[d] Muhamed-Kheir Taha,[c] Stephen W. White,[e] Ahmed Haouz[f] and Ivo G. Boneca[a,b]*

[a]Institut Pasteur, Unité Biologie et génétique de la paroi bactérienne, Dept. Microbiologie, 28 Rue du Dr. Roux, 75015 Paris, France, [b]INSERM, Groupe Avenir, 75015 Paris, France, [c]Institut Pasteur, Unité Infection Bactériennes Invasives, Dept. Infection et Epidémiologie, 28 Rue du Dr. Roux, 75015 Paris, France, [d]Institut Pasteur, CNRS–UMR3528, Plateforme de Biophysique, 25 Rue du Dr. Roux, 75724 Paris, France, [e]Department of Structural Biology, St. Jude Children's Research Hospital, Memphis, TN 38105, USA, and [f]Institut Pasteur, CNRS–UMR3528, Plateforme de Cristallographie, 25 Rue du Dr. Roux, 75724 Paris, France

Correspondence e-mail: allison.williams@pasteur.fr, bonecai@pasteur.fr

Received 28 February 2014
Accepted 21 July 2014

PDB references: peptidoglycan *O*-acetylesterase, 4k3u; 4k9s; 4k7j; 4k40

Peptidoglycan *O*-acetylesterase (Ape1), which is required for host survival in *Neisseria* sp., belongs to the diverse SGNH hydrolase superfamily, which includes important viral and bacterial virulence factors. Here, multi-domain crystal structures of Ape1 with an SGNH catalytic domain and a newly identified putative peptidoglycan-detection module are reported. Enzyme catalysis was performed in Ape1 crystals and key catalytic intermediates along the SGNH esterase hydrolysis reaction pathway were visualized, revealing a substrate-induced productive conformation of the catalytic triad, a mechanistic detail that has not previously been observed. This substrate-induced productive conformation of the catalytic triad shifts the established dogma on these enzymes, generating valuable insight into the structure-based design of drugs targeting the SGNH esterase superfamily.

1. Introduction

SGNH esterases are ubiquitous enzymes that are involved in a wide range of essential biological processes, for example in viral and bacterial pathogenesis (Leščić Asler *et al.*, 2010). Their fundamental activity is to catalyze the hydrolysis of an acetyl ester, and they have a broad range of substrate specificities and enzyme–substrate promiscuity (Leščić Asler *et al.*, 2010). Their active sites all contain four absolutely conserved residues, Ser, Gly, Asn and His (hence the SGNH acronym), and the catalytic mechanism is predicted to proceed through a nucleophilic Ser which is activated by a catalytic triad that includes the conserved His and an accompanying conserved acidic residue (Leščić Asler *et al.*, 2010; Lo *et al.*, 2003). The superfamily shares a common α/β-hydrolase fold that is frequently observed in structures with catalytic triads such as lipases and thioesterases (Leščić Asler *et al.*, 2010). Although this family of enzymes was identified more than 15 years ago, a full structural understanding of their broad substrate specificities, enzyme promiscuity and catalytic mechanism remains to be elucidated (Leščić Asler *et al.*, 2010).

Recently, a newly described SGNH esterase was reported, peptidoglycan (PGN) *O*-acetylesterase (Ape1), the primary function of which is to catalyze the de-*O*-acetylation of PGN (Weadge *et al.*, 2005). PGN *O*-acetylation in Gram-negative bacteria occurs by the action of PatA and PatB, which facilitate the transfer of a acetate modification to PGN (Fig. 1*a*; Weadge *et al.*, 2005). Ape1 is present in Gram-negative bacteria that *O*-acetylate their PGN, and it catalyzes the hydrolysis of the *O*-acetyl modification specifically at the sixth

carbon position of the muramoyl residue (Fig. 1*b*). No structural data exist for the Ape1 family of enzymes or indeed for any enzyme that is involved in PGN acetylation (Weadge *et al.*, 2005), a process that allows pathogenic bacteria to subvert the host innate immune response (Diacovich & Gorvel, 2010; Aubry *et al.*, 2011; Fig. 1*a*). However, this modification also precludes the normal metabolism and maturation of the peptidoglycan by the endogenous hydrolases of the hexosaminidase superfamily such as lytic transglycosylases (Chan *et al.*, 2012; Bera *et al.*, 2005). Thus, the fine-tuning of this modification is essential for bacterial metabolism and survival. Ape1 performs this fine-tuning in *Neisseria* sp. (Veyrier *et al.*, 2013; Dillard & Hackett, 2005) and is essential for virulence, similar to the viral haemagglutinin-esterase, an SGNH esterase that plays an important role in the pathogenicity of some viruses such as influenza C virus, coronavirus and H5 avian and H9 swine influenza viruses (Rosenthal *et al.*, 1998; Zeng *et al.*, 2008; Desforges *et al.*, 2013; Ha *et al.*, 2002).

Given the availability of substrate analogues, we considered that Ape1 would be an ideal candidate to structurally dissect the active-site topology of this diverse superfamily of enzymes. The catalytic mechanism of the SGNH esterases has long been suggested to be unique because of the hydrogen-bonding pattern and geometric arrangement of the active-site residues (Leščić Asler *et al.*, 2010). Three important outstanding issues to be resolved are the geometry and activation mechanism of the catalytic triad, the precise architecture of the oxyanion hole and the structural basis of the substrate promiscuity. To address these questions, we first determined the crystal structure of Ape1 and then trapped and structurally characterized two intermediates along its catalytic pathway: the bound substrate before catalysis and the covalent acyl-enzyme intermediate.

2. Methods

2.1. X-ray crystallography

Preliminary crystallization screening was carried out by the sitting-drop vapour-diffusion method with a Mosquito (TTP Labtech) automated crystallization system. After extensive screening and optimization, native and selenomethionine-labelled crystals were obtained and cryoprotected using similar conditions with minor variations. All crystals were grown at 18°C using the hanging-drop vapour-diffusion method. Crystals consistently appeared within 1–2 d. The initial crystals of Ape1 were extremely fragile and formed as needle-like crystals arranged in clusters with many nucleation points. To make the transition from needle to rod-like crystals, we added TCEP in combination with a fine gradient of PEG (2–5% increments). Single crystals appeared over a range that was typically between 10 and 40% PEG 10 000 with 5–15 m*M* TCEP and 100 m*M* HEPES pH 7.5. Ape1 solution (10–20 mg ml^{-1}) and well solution were mixed in a 1:1(*v*:*v*) ratio.

Selenomethionine-labelled crystals of Ape1 were grown at 18°C by streak-seeding drops of 10–20 mg ml^{-1} Ape1 in 25 m*M* Tris pH 7.6, 150 m*M* NaCl, 1 m*M* β-mercaptoethanol (BME) equilibrated in a 1:1(*v*:*v*) ratio against a well solution consisting of 100 m*M* HEPES pH 7.5 10–40% PEG 10 000. To generate reaction intermediates, the native crystals of Ape1 were soaked with a tenfold molar excess of muramyl dipeptide (MDP) for 5 and 30 min just before flash-cooling. Crystals were cryoprotected in a mixture of 50% Paratone and 50% paraffin oil and flash-cooled in liquid nitrogen.

Single-wavelength anomalous dispersion diffraction data were collected on X06DA at SLS and native data sets was collected on PROXIMA1 at SOLEIL. SAD data sets were collected at the selenium absorption peak. The highest resolution selenomethionine-labelled crystals diffracted to 2.3 Å resolution with anomalous signal to 2.8 Å resolution (Table 1). Initial phases were found using *SHELXC*/*D*/*E* and the initial model was built using *phenix.autobuild* with 22 selenium sites and a figure of merit of 0.59 (Adams *et al.*, 2010; Sheldrick, 2008).

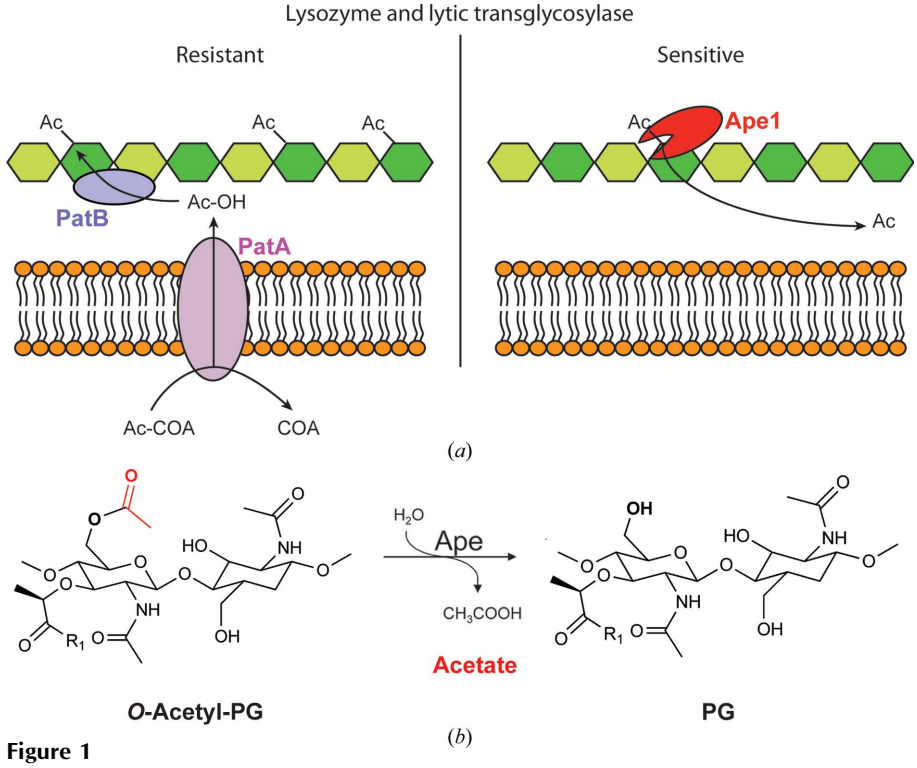


Figure 1
Peptidoglycan is an essential polymeric component of Gram-negative and Gram-positive bacteria. (*a*) *O*-Acetylation is a PGN modification found in many Gram-negative and Gram-positive bacteria carried out by the concomitant action of the enzymes PatA and PatB. This modification is resistant to the action of lysozyme and lytic transglycosylases (LTs). After removal of this modification by Ape1 the bacterium is sensitive to the action of lysosyme and LTs. (*b*) Schematic of the reaction catalyzed by Ape1. Ape1 removes the acetyl modification at the C-6 hydroxyl moiety of the MurNAc residue.

The refined SAD structure was used to determine the phases for the native data set. Data processing for all structures was performed using *XDS* (Kabsch, 2010). Phasing by SAD or molecular replacement was performed using *PHENIX* (Adams *et al.*, 2010). Model building was performed using *Coot* (Emsley & Cowtan, 2004) and restrained refinement was carried out using a combination of the *PHENIX* and *CCP*4 software suites (Winn *et al.*, 2011; Adams *et al.*, 2010). *MolProbity* was used during building and refinement for iterative structure improvements (Chen *et al.*, 2010).

Table 1
Data-collection and refinement statistics.

Values in parentheses are for the outer shell.

	Ape1 Se	Ape1 native	5 min	30 min
Ligand added	None	None	MDP	MDP
Data collection				
Wavelength (Å)	0.9795	1.1	1.1	1.1
Resolution range (Å)	40.41–2.33 (2.42–2.33)	49.00–2.63 (2.73–2.63)	48.64–1.97 (2.04–1.97)	48.54–2.16 (2.23–2.16)
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Unit-cell parameters				
a (Å)	72.19	71.87	70.14	70.12
b (Å)	79.07	79.49	79.98	79.98
c (Å)	123.91	124.38	122.55	122.11
Total reflections	60744	43089	92825	74454
Unique reflections	30271	21361	49265	37615
Multiplicity	2.0 (2.0)	2.0 (2.0)	1.9 (1.9)	2.0 (2.0)
Completeness (%)	98.13 (88.48)	98.30 (94.29)	99.34 (97.72)	99.82 (98.24)
Mean $I/\sigma(I)$	10.34 (2.21)	8.20 (2.05)	8.35 (1.96)	9.91 (2.61)
Wilson *B* factor (Å^2)	24.7	42.4	27.8	29.6
$R_{\rm merge}$†	0.0979 (0.369)	0.0366 (0.176)	0.059 (0.484)	0.0532 (0.287)
Refinement				
R factor‡	0.173 (0.228)	0.165 (0.255)	0.167 (0.246)	0.164 (0.205)
$R_{\rm free}$‡	0.264 (0.346)	0.238 (0.352)	0.198 (0.268)	0.176 (0.272)
No. of atoms	5721	5386	5889	5794
No. of waters	396	85	460	412
No. of protein residues	698	698	700	699
R.m.s.d., bonds (Å)	0.008	0.008	0.007	0.007
R.m.s.d., angles (°)	1.24	1.21	1.15	1.07
Ramachandran plot				
Favoured (%)	96	96	97	96
Allowed (%)	3	3.42	2.72	4
Outliers (%)	1.0	0.58	0.28	0
B factors (Å^2)				
Average	16.20	12.50	24.40	16.40
Macromolecules	16.10	12.60	23.80	16.10
Ligand	—	—	39.50	32.20
Solvent	17.70	6.70	28.10	18.60
All-atom clash score	7.5	5.81	3.80	3.0

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$. ‡ *R* factor $= \sum_{hkl} ||F_{\rm obs}| - |F_{\rm calc}|| / \sum_{hkl} |F_{\rm obs}|$. $R_{\rm free}$ was computed identically except that all reflections belonged to a test set consisting of a 10% random selection of the data.

2.2. Protein expression and purification

All constructs were created using standard molecular-biology techniques. All constructs used in this study were GST fusions from pGEX-4T1 (GE Lifesciences). The native Ape1 protein was expressed in *Escherichia coli* BL21(DE3)pLysS competent cells (Novagen). Cells were induced with 0.6 m*M* IPTG at an OD of 0.7–0.8 and harvested after 4 h of induction at 18°C. Selenomethionine-labelled protein was expressed in *E. coli* B834 (DE3) pLysS cells and then grown in autoinduction medium as described by Studier (2005) and others. After glutathione affinity chromatography and thrombin cleavage, the proteins were purified to homogeneity by ion-exchange (HiTrap Sepharose Fast Flow, GE Healthcare) and size-exclusion chromatography (Superdex 200, GE Healthcare) in 25 m*M* Tris pH 7.6, 150 m*M* NaCl, 1 m*M* BME. After the gel-filtration columns, the proteins were generally used immediately for crystallization. Alternatively, the proteins were flash-cooled in liquid nitrogen and stored at −80°C.

2.3. Analytical ultracentrifugation

Sedimentation-velocity experiments were carried out at 20°C in an XL-I analytical ultracentrifuge (Beckman-Coulter) equipped with double UV and Rayleigh interference detection. Samples were prepared in 25 m*M* Tris pH 7.6, 150 m*M* NaCl, 1 m*M* BME and were spun at 42 000 rev min^{-1} using an An-60 Ti rotor and 12 mm double-sector epoxy centrepieces. The partial specific volume of Ape1 (0.729 ml g^{-1}) as well as the buffer viscosity (η = 1.028 cP) and density (ρ = 1.0092 g ml^{-1}) were estimated using *SEDNTERP* 1.09. Absorbance and interference profiles were recorded every 4 min. Sedimentation coefficient distributions $c(s)$ were determined using *SEDFIT* (Schuck *et al.*, 2002; Schuck, 2000).

2.4. PDB codes

Coordinates and structural data have been deposited in the Protein Data Bank with accession codes 4k3u, 4k40, 4k7j and 4k9s.

3. Results and discussion

3.1. The overall structure and topology of Ape1

The crystal structure of Ape1 was determined using the single anomalous dispersion (SAD) approach with crystals of selenomethionine-derivatized protein (Table 1). Ape1 crystallized in space group $P2_12_12_1$ with two molecules in the asymmetric unit (termed *A* and *B*; see Supplementary Fig. S1[1] for an expanded view) and the structure reveals a two-domain architecture (Figs. 2*a* and 2*b*). The C-lobe is split and comprises residues 45–99 and 227–393, and the intervening N-lobe comprises residues 104–221 (Figs. 2*a* and 2*b* and Supplementary Fig. S2). Two short linkers connect these

[1] Supporting information has been deposited in the IUCr electronic archive (Reference: CB5053).

domains, residues 100–103 and 222–226, and this suggests a tightly packed structure with a limited degree of N/C-lobe domain rearrangement or interaction (Figs. 2*a* and 2*b*). This was further explored using the *ElNémo* normal-mode web server (Suhre & Sanejouand, 2004; http://www.igs.cnrs-mrs.fr/elnemo/). A normal-mode analysis in the lowest energy conformations revealed negligible rearrangement of the N- or C-lobes (Supplementary Fig. S2), although there were some predicted intermolecular movements within the lobes independently (Supplementary Fig. S2).

The C-lobe consists of a central five-stranded parallel β-sheet (β1, β2, β15, β16 and β17) flanked on either side by 11 α-helices (α1–α11) (Figs. 2*a* and 2*b*) and the domain adopts the α/β-hydrolase fold that has previously been observed in other structurally characterized members of the SGNH hydrolase superfamily (Fig. 2*c*) (Desforges *et al.*, 2013).

The fold of the N-lobe is a canonical β-sandwich that consists of two antiparallel β-sheets; β4, β7, β9, β10 and β12 are in sheet 1 and β3 β5, β6, β8, β11 and β13 are in sheet 2 (Figs. 2*a* and 2*b* and Supplementary Fig. S3). Based on sequence analyses, the N-lobe was not predicted to resemble any structure currently in the PDB, but *DALI* and *PDBeFold* revealed that the fold matched that of the carbohydrate-binding module of family 35 (CBM35; Fig. 2*c*). These non-catalytic modules (CBMs) are frequently found in enzymes that need to be targeted to carbohydrate-associated substrates, and are commonly appended to plant cell-wall-degrading enzymes (Montanier *et al.*, 2009). The interaction of the N-lobe with a specific carbohydrate moiety would therefore possibly explain how Ape1 is able to select PGN when presented with an array of possible substrates *in vivo* (Montanier *et al.*, 2009). A more detailed analysis of the structure and substrate specificity of the N-lobe will be presented elsewhere.

3.2. The active site of Ape1

The active site is within the C-lobe (Figs. 2*d* and 3*a*) and contains the four absolutely conserved residues Ser80, Gly236, Asn268 and His369 in their anticipated locations consistent with other SGNH hydrolase family members (Figs. 3*a*–3*d*). Other highly conserved residues located in the active-site

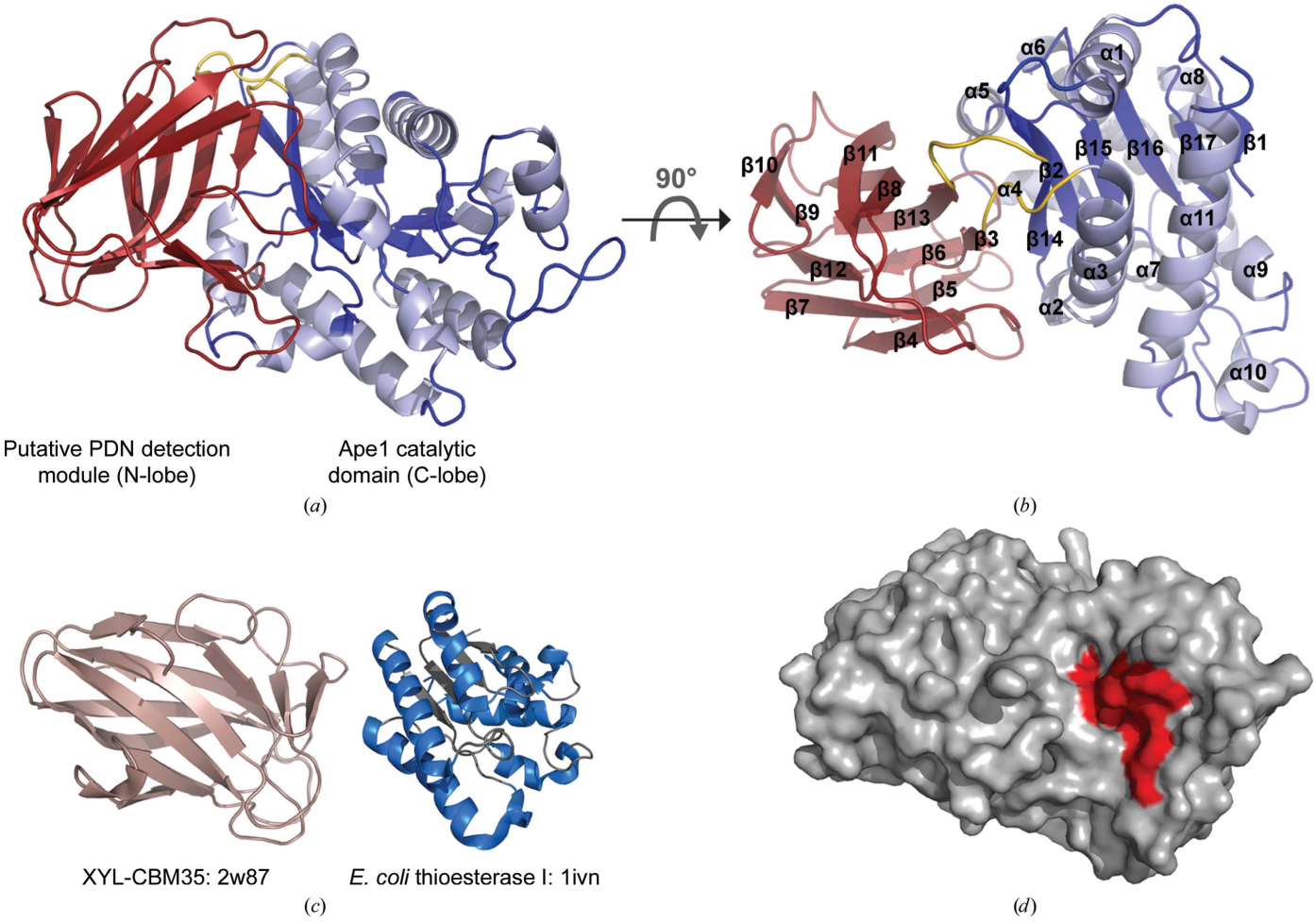


Figure 2
Overall fold and topology of Ape1. (*a*, *b*) Ape1 is represented by a ribbon diagram. The domains are joined by short linker segments coloured yellow. The catalytic domain is coloured light blue and the putative peptidoglycan-binding domain is coloured brick red. In (*b*) Ape1 is rotated 90°. (*c*) Structural comparison of the Ape1 N-lobe with XYL-CBM35 and of the Ape1 C-lobe with *E. coli* thioesterase, a member of the SGNH hydrolase family. PDB codes are included in the labels. (*d*) Surface model of Ape1. Highly conserved residues in the active site of Ape1 are coloured red.

region of Ape1 (Asp79, His81, Val368, Thr267 and Asp366) are highlighted in Fig. 4 (0 min). Previous mutagenesis and kinetic studies and sequence alignments have postulated the important role of the aforementioned residues in catalysis, substrate binding or support of the catalytic residues (Pfeffer *et al.*, 2013).

In agreement with predictions from previous mutagenesis and kinetic studies, Ser80 and His369 together with Asp366 make up the catalytic triad (Pfeffer *et al.*, 2013), and the amide and ND2 atoms of Gly236 and Asn268, respectively, create the oxyanion hole (Pfeffer *et al.*, 2013). In addition, the space afforded by Gly236 together with the OD1 atom of Asn268 are well placed to mediate substrate binding (Fig. 4, 0 min). Within the catalytic triad, Ser80 is flanked by Asp79 and His81, which appear to provide structural support to the catalytic residue (Fig. 4, 0 min) and establish a barricade for the active-site groove, but this is inconsistent with previous studies that suggest that these residues position or facilitate substrate interaction or binding (Pfeffer *et al.*, 2013). Meanwhile, the backbone amide N atom of Val368 appears to orient Asp366 with respect to His369 *via* a hydrogen-bonding interaction and Thr267 structurally supports the catalytic residues.

A key feature of the Ape1 active site in the absence of substrate relates to the geometry of the catalytic triad, in which we do not observe a canonical proton-relay system in which the three catalytic residues are aligned (Fig. 4, 0 and 30 min). Instead, whereas His369 and Asp366 are correctly aligned, the side chain of Ser80 points toward the putative oxyanion hole; it is hydrogen-bonded to the backbone amide N atom of Gly236 and is not appropriately aligned for catalysis (Fig. 4).

3.3. Trapped intermediates in the reaction cycle of Ape1

To explore the catalytic mechanism of Ape1 and to understand the unusual conformation of Ser80 when compared with the canonical orientation observed to date (Figs. 3*a*–3*d*; Leščić Asler *et al.*, 2010), native crystals were soaked with substrate and then flash-cooled in liquid nitrogen with the goal of trapping key catalytic intermediates. The acetylated PGN substrate is insoluble and the available acetylated substrate

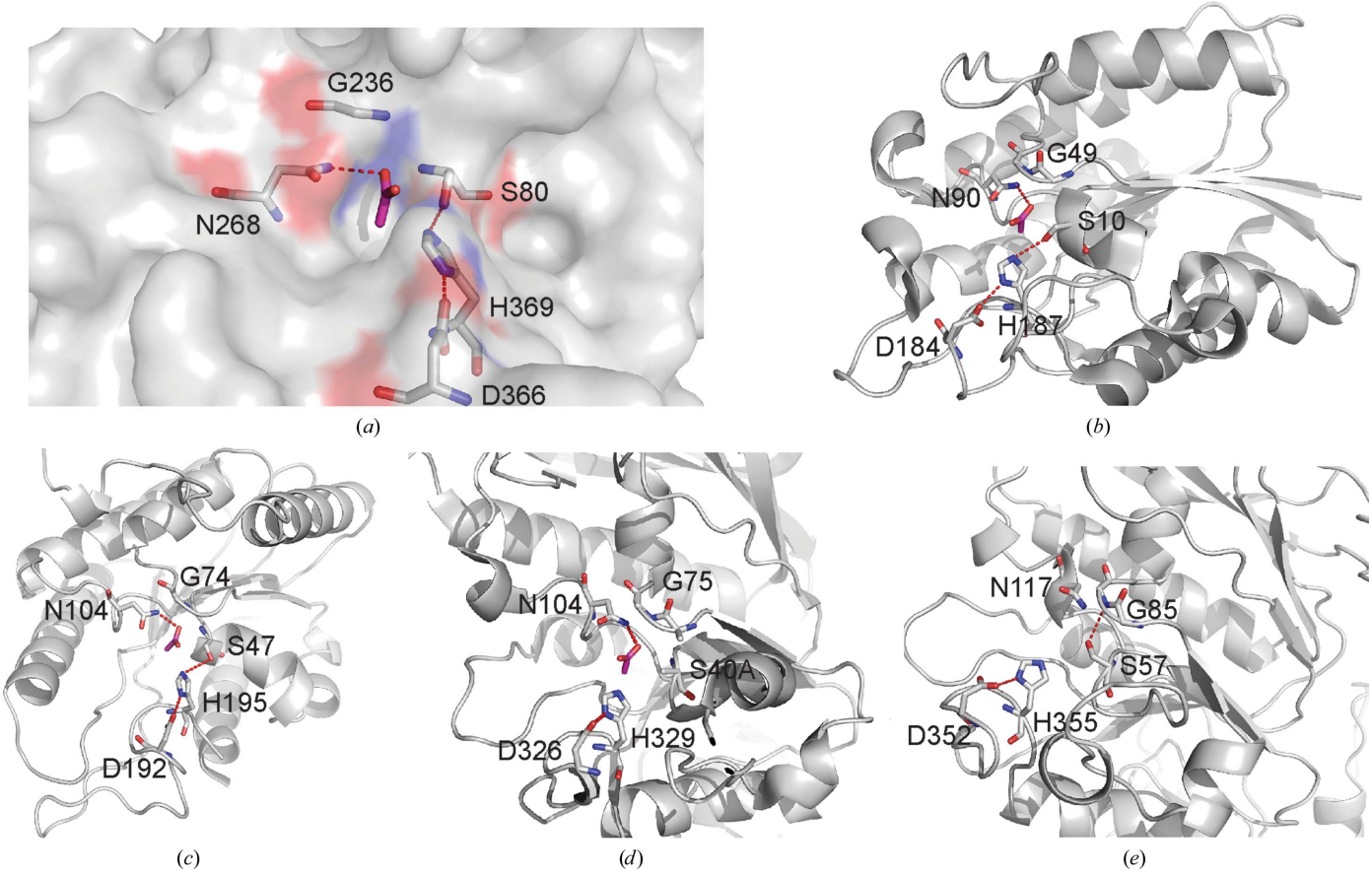


Figure 3
Active-site conservation in the SGNH hydrolase superfamily. (*a*) Surface model of Ape1 displaying the active groove for acetate. According to the active-site arrangement, this groove is absolutely conserved in all members of the SGNH hydrolase superfamily. (*b*) The structure of a putative arylesterase from *Agrobacterium tumefaciens* (PDB entry 3dci; Midwest Center for Structural Genomics, unpublished work). (*c*) Platelet-activating factor with bound acetate in the active-site groove (PDB entry 1wab; Ho *et al.*, 1997). (*d*) Structure of coronavirus haemagglutinin-esterase (PDB entry 3cl5; Zeng *et al.*, 2008). The catalytic serine has to move to create space for the acetate-containing substrate. The entrance of the acetate into the active site ensures the productive conformation of the catalytic triad. A mutation of the catalytic serine to alanine in the coronavirus haemagglutinin-esterase shows a trapped acetate in the active site. (*e*) Close-up of the active site of coronavirus haemagglutinin-esterase.

analogues rapidly turn over to products. We took advantage of the well known fact that some enzymes are able to catalyze both reverse and forward reactions, in some cases rather efficiently. The canonical serine protease, which is the best documented example of an enzyme that proceeds *via* a catalytic triad, can reconstruct peptide hydrolysis (Zakharova *et al.*, 2009). Ape1 catalyzes the removal of the *O*-acetyl group from the muramyl residue of the polymeric PGN, thus releasing an unmodified muramyl residue and free acetate (Fig. 1*b*). We rationalize that a valid product mimic of the reverse reaction would be muramyl dipeptide (MDP), a monosaccharide substructure of the polymeric PGN, along with acetate. Serendipitously, the crystals already contained trace amounts of acetate, a byproduct that is frequently found in the active site of SGNH esterases (Figs. 3*a*–3*d*) and is possibly introduced by the crystallization precipitant. To reconstruct acetyl hydrolysis in the crystalline state, a tenfold excess of MDP was added to crystals that already contained trace amounts of acetate before flash-cooling them in liquid nitrogen at time points of 5 and 30 min (Fig. 4, 5 and 30 min).

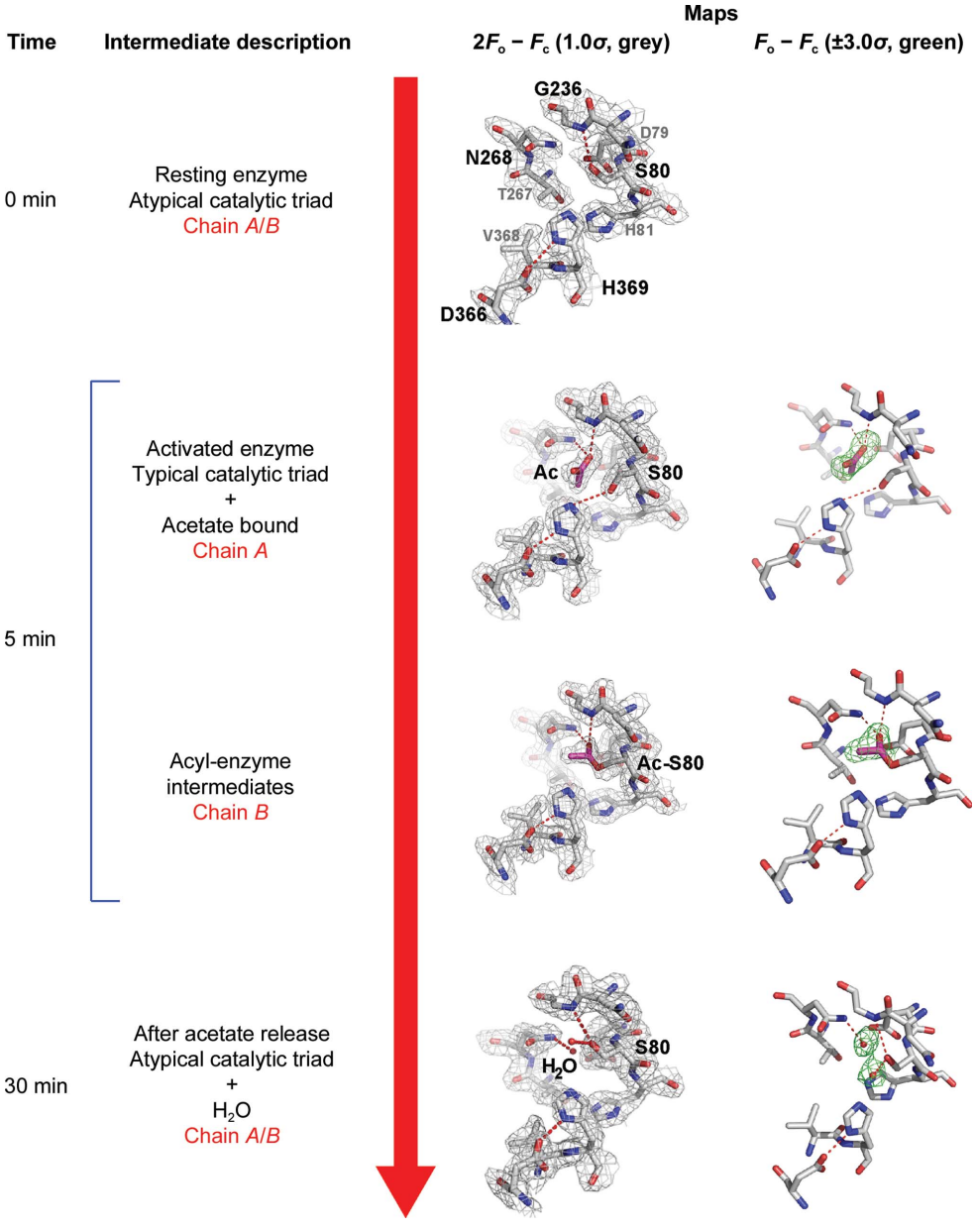


Figure 4
Snapshot of *O*-esterase catalysis *in crystallo*. Trapping of a covalent acyl-enzyme intermediate. All $2F_o - F_c$ and $F_o - F_c$ electron-density maps surrounding Ape1 residues or acetate are contoured at 1.2σ for $2F_o - F_c$ maps and at ±3.0σ for $F_o - F_c$ maps. 0 min: the active site of native Ape1. 5 min, chain *A*: trapped acetate group coloured purple. 5 min, chain *B*: trapped acyl-enzyme intermediate coluored purple. 30 min, chain *A*/*B*: trapped water molecules in the active site of Ape1. See 0 min for complete labelling of the active-site residues.

We were able to capture two key steps in the catalytic mechanism that were well defined in our electron-density maps at the 5 min time point: bound acetate in chain *A* and the covalent acyl-enzyme intermediate in chain *B* (Fig. 4, 5 min; Supplementary Fig. S4). Acetate was not visible in the active site of the apo structure and we are confident that its appearance in these structures is indeed the result of catalysis. Moreover, Ape1 is highly specific for the presence of the *O*-acetyl group and solely catalyzes the *O*-acetyl modification of the polymeric PGN; it does not appear to utilize *N*-acetyl-containing substructures of PGN such as MDP, MurNAc or GlcNAc (Weadge & Clarke, 2006; Veyrier *et al.*, 2013).

Assuming that the bound acetate represents the enzyme–substrate complex prior to catalysis (Fig. 4, 5 min, chain *A*), the catalytic Ser80 side chain has rotated out of the oxyanion hole and now appears to be poised to be activated and to perform a nucleophilic attack on the ester if the acetate is extended to include MurNAc or MDP (Fig. 4, 5 min, chain *A*). The acetate ion is buried in a groove adjacent to Ser80 with one of its O atoms hydrogen-bonded to three N atoms: the backbone amides of Ser80 and Gly235 and ND2 of Asn268 (Fig. 4, 5 min, chain *A*). These three hydrogen bonds were previously noted in the *E. coli* thioesterase (TAP) structure (Lo *et al.*, 2003) and prompted the suggestion that the oxyanion hole unusually comprises three parti-

cipating N atoms. The shape of this groove complements the bound acetate, and this may explain the preference that Ape1 shows for *O*-acetyl-containing substrates such as *p*-nitrophenyl acetate (Fig. 4, 5 min; Weadge & Clarke, 2006; Pfeffer *et al.*, 2013). The acetate-binding groove is the only portion of the active-site region that is buried, suggesting that the rest of the PGN substrate is exposed (Fig. 3 and Fig. 4, 5 min). However, a bound HEPES molecule adjacent to the acetate in the 5 min soaked structure may mimic the PGN substrate, and this suggests that Phe271 and Lys310 may contribute to the substrate-binding site (Supplementary Figs. S5 and S6).

The trapped acyl-enzyme intermediate observed in chain *B* of the 5 min structure demonstrated that catalysis occurred in the crystalline state. The catalytic triad is now observed in the productive conformation, in which the O^{γ} atom of serine points in the direction of His369 and Asp366 (Fig. 4, 5 min, chain *B*). Most significantly, however, the *O*-acetylserine group now points directly towards the putative oxyanion hole, where the O atom does indeed form hydrogen bonds to three N atoms of Ser80, Gly236 and Asn268, as described above. This provided insight into the selectivity of Ape1 for *O*-linked acetyl substrates (Weadge & Clarke, 2006; Veyrier *et al.*, 2013). *N*-linked acetyl substrates such as GlcNAc, MurNAc or MDP would not have the same hydrogen-bonding capabilities in this environment.

In the 5 min structure, chains *A* and *B* display two different steps in the catalytic cycle. A superposition of chains *A* and *B* revealed only minor perturbations in the active-site residues. Globally, based on comparison of the two chains, the largest change was distal to the catalytic groove. We observed that the loop between residues 159 and 164 is in different positions. The high temperature factors and overall weak occupancy of these residues suggests some degree of flexibility in this loop. The different stages in the catalytic cycle that we postulate are possibly because of crystal packing and accessibility.

3.4. Evidence of enzymatic deacetylation in the crystalline state

In order to explore the deacetylation of the acyl-enzyme complex (Fig. 4, 5 min, chain *B*) in the crystalline state, we examined longer time points. In the 30 min structure we note the presence of two water molecules in chains *A* and *B* (Fig. 4, 30 min). One water molecule occupies the space evacuated by the *O*-acetyl residue, while the other is perfectly poised to catalyze deacetylation of the acyl-enzyme complex (Fig. 4, 30 min) and convert Ape1 to its original free form. The regenerated form of the enzyme displays the catalytic triad once again in the nonproductive conformation, suggesting that this is potentially the resting state of the enzyme.

3.5. Analysis of the active site of SGNH hydrolase family members

An exploration of the crystal structures of the SGNH hydrolase superfamily reveals absolute conservation of the architecture of the active-site residues Ser, His, Gly and Asn when compared with Ape1 (Rosenthal *et al.*, 1998; Zeng *et al.*, 2008; Desforges *et al.*, 2013; Ha *et al.*, 2002; Fig. 3). The active-site catalytic serine (Ser57) of haemagglutinin-esterase fusion glycoprotein (HEF), a protein embedded in the viral envelope of influenza C virus and other viruses such as coronavirus and H5 avian and H9 swine influenza viruses (Rosenthal *et al.*, 1998; Zeng *et al.*, 2008; Desforges *et al.*, 2013; Ha *et al.*, 2002), shares a parallel orientation to that found in the native structure of Ape1 (Fig. 3*e*). These structures support possible mechanistic conservation in the SGNH hydrolase superfamily.

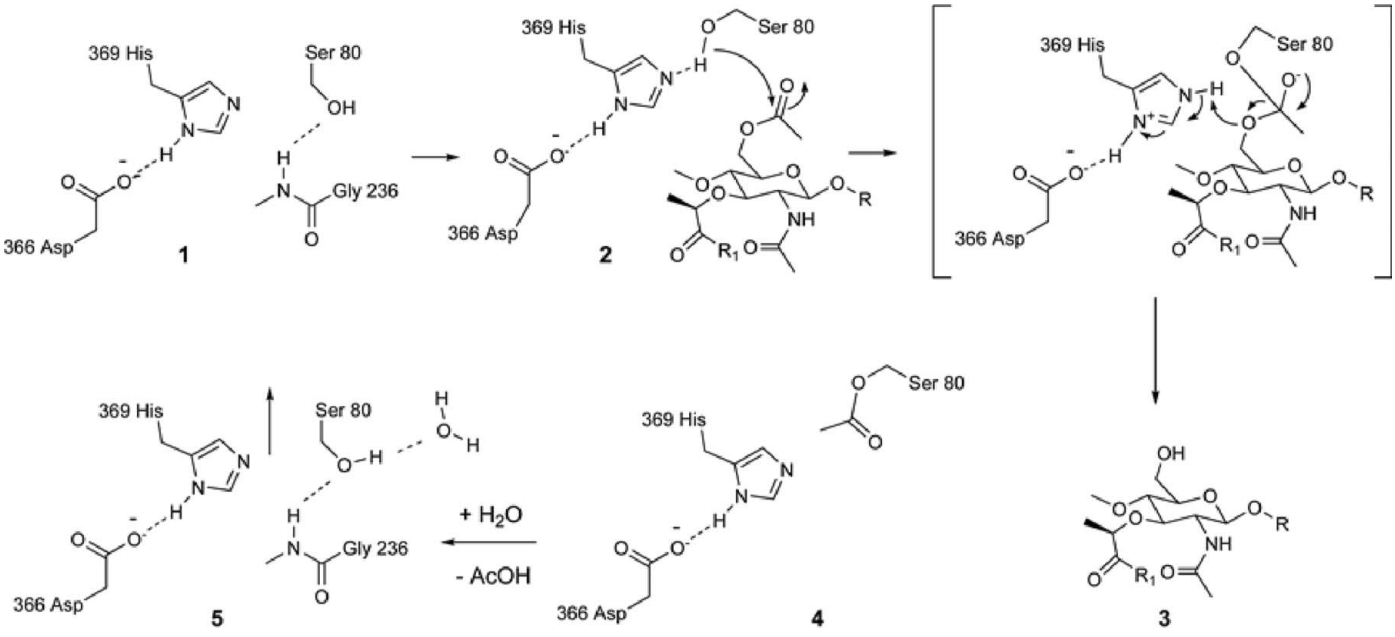


Figure 5
Steps 1 to 5 illustrate the different steps in the catalytic cycle.

3.6. The catalytic mechanism of Ape1

Based on our findings, we propose the following catalytic model for Ape1. In the absence of substrate Ser80 points towards the oxyanion hole, where it is hydrogen-bonded to Gly236, and the enzyme appears to be in the 'parked' position (Fig. 5, step 1). In the presence of substrate, the O^{γ} atom of Ser80 rotates to accommodate the acetate moiety and aligns with His369 to form the catalytic triad (Ser80–His369–Asp366). The now activated Ser80 then attacks the carbonyl C atom of an ester-linked acetate group on the muramoyl residue (Fig. 5, step 2), leading to the formation of a tetrahedral species with the assistance of the now available oxyanion hole (Fig. 5, step 3). The intermediate then collapses and leads to the formation of the covalent acyl-enzyme intermediate (Fig. 5, step 4), with concomitant release of the product. Finally, a water molecule enters and the acetate is hydrolyzed from the catalytic Ser80 (Fig. 5, step 5), thereby returning Ser80 to its protonated state in which the active site is regenerated for another catalytic cycle (Fig. 5, step 5). As regards substrate promiscuity, we suggest that the rotational flexibility of the catalytic serine combined with the generally open substrate-binding groove but the snug acetate-binding pocket may enable Ape1 to bind and process a range of distinct but structurally related *O*-acetylated substrates as large as the polymeric PGN and as small as *p*-nitrophenyl acetate (Weadge & Clarke, 2006).

This work demonstrates the elegant catalytic mechanism used by SGNH esterases to break a covalent bond. In contrast to other serine protease-like mechanisms studied thus far, we observed rotation of the side chain of Ser80 into its catalytically competent position in response to the binding of substrate (Fig. 4, 0 and 30 min). Gly236 plays a crucial role in supporting the formation and hydrolysis of the acyl-enzyme intermediate, but also appears to provide a gate that potentially closes the active site in the absence of substrate (Fig. 4, 5 min, chain *A*). Asn268 plays a key role in stabilizing the substrate and the acyl-enzyme intermediate and possibly the unobserved tetrahedral intermediate. His369 influences the nucleophilic catalysis of Ser80 by maintaining the proton relay of the catalytic triad in the presence of substrate. These snapshots provide a detailed view of the enzyme-catalyzed acetyl transfer of a SGNH esterase. They illustrate for the first time a substrate-induced catalytic triad, an enzyme mechanistic model that is potentially conserved within the diverse SGNH esterase superfamily (Fig. 3). These structures also offer additional insights into the catalytic roles of the absolutely conserved SGNH residues in this diverse superfamily (Figs. 3 and 4) and reveal important intermediates in the catalytic mechanism of Ape1 (Fig. 4).

In the case of the HEF protein, its function is related to the pathogenicity of viruses (Rosenthal *et al.*, 1998; Zeng *et al.*, 2008; Desforges *et al.*, 2013; Ha *et al.*, 2002). It helps in the adhesion of the virus to the intestinal mucus layer (Rosenthal *et al.*, 1998; Zeng *et al.*, 2008; Desforges *et al.*, 2013; Ha *et al.*, 2002). Meanwhile, Ape1 has been linked to the persistence of virulence in a mouse model (Veyrier *et al.*, 2013; Dillard & Hackett, 2005). The substrate-induced catalytic triad may have important implications in bacterial and viral pathogenicity and could define a mechanistic model in which activation of these esterases occurs when the bacterium or virus is presented to the host. These structures pave the way for the design of relevant inhibitors and future antibiotics and antiviral drugs to further explore the biological significance of this family of enzymes both *in vivo* and *in vitro*.

We would like to thank the beamline staff (PROXIMA1 at SOLEIL and X06DA at SLS) for their helpful assistance. We are grateful to Marco Bellinzoni, Frederick Saul and Patrick Weber for advice and assistance. We thank Dr Danny Huang for helpful general comments and criticisms. We thank colleagues across many disciplines from Duke University Medical Center for their helpful criticism on the presentation. AHW was supported by an EMBO long-term fellowship (ALTF 732-2010) and an Institut Carnot–Pasteur Maladies Infectieuses fellowship. FJV was supported by a Marie Curie Excellence grant and a Pasteur–Roux fellowship. This work was supported by an ERC starting grant (PGNfrom-SHAPEtoVIR 202283) to IGB. AHW and IGB designed the research. AHW and YM purified and biochemically characterized the protein. FJV, MKT and MB contributed valuable reagents. BR conducted biophysical experiments. AHW and AH conducted all structural experiments. AH collected the X-ray data. AHW solved and refined all structures. AHW, SWW and IGB analyzed all structural data. AHW and SWW wrote the manuscript. AH, IGB and SWW edited the manuscript. The authors declare that they have no conflicts of interest.

References

Adams, P. D. *et al.* (2010). *Acta Cryst.* D**66**, 213–221.
Aubry, C., Goulard, C., Nahori, M.-A., Cayet, N., Decalf, J., Sachse, M., Boneca, I. G., Cossart, P. & Dussurget, O. (2011). *J. Infect. Dis.* **204**, 731–740.
Bera, A., Herbert, S., Jakob, A., Vollmer, W. & Götz, F. (2005). *Mol. Microbiol.* **55**, 778–787.
Chan, Y. A., Hackett, K. T. & Dillard, J. P. (2012). *Microb. Drug Resist.* **18**, 271–279.
Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* D**66**, 12–21.
Desforges, M., Desjardins, J., Zhang, C. & Talbot, P. J. (2013). *J. Virol.* **87**, 3097–3107.
Diacovich, L. & Gorvel, J.-P. (2010). *Nature Rev. Microbiol.* **8**, 117–128.
Dillard, J. P. & Hackett, K. T. (2005). *Infect. Immun.* **73**, 5697–5705.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Ha, Y., Stevens, D. J., Skehel, J. J. & Wiley, D. C. (2002). *EMBO J.* **21**, 865–875.
Ho, Y. S., Swenson, L., Derewenda, U., Serre, L., Wei, Y., Dauter, Z., Hattori, M., Adachi, T., Aoki, J., Arai, H., Inoue, K. & Derewenda, Z. S. (1997). *Nature (London)*, **385**, 89–93.
Kabsch, W. (2010). *Acta Cryst.* D**66**, 125–132.
Leščić Asler, I., Ivić, N., Kovačić, F., Schell, S., Knorr, J., Krauss, U., Wilhelm, S., Kojić-Prodić, B. & Jaeger, K. E. (2010). *Chembiochem*, **11**, 2158–2167.
Lo, Y.-C., Lin, S.-C., Shaw, J.-F. & Liaw, Y.-C. (2003). *J. Mol. Biol.* **330**, 539–551.
Montanier, C. *et al.* (2009). *Proc. Natl Acad. Sci. USA*, **106**, 3065–3070.

Pfeffer, J. M., Weadge, J. T. & Clarke, A. J. (2013). *J. Biol. Chem.* **288**, 2605–2613.
Rosenthal, P. B., Zhang, X., Formanowski, F., Fitz, W., Wong, C.-H., Meier-Ewert, H., Skehel, J. J. & Wiley, D. C. (1998). *Nature (London)*, **396**, 92–96.
Schuck, P. (2000). *Biophys. J.* **78**, 1606–1619.
Schuck, P., Perugini, M. A., Gonzales, N. R., Howlett, G. J. & Schubert, D. (2002). *Biophys. J.* **82**, 1096–1111.
Sheldrick, G. M. (2008). *Acta Cryst.* A**64**, 112–122.
Studier, F. W. (2005). *Protein Expr. Purif.* **41**, 207–234.
Suhre, K. & Sanejouand, Y.-H. (2004). *Nucleic Acids Res.* **32**, W610–W614.
Veyrier, F. J., Williams, A. H., Mesnage, S., Schmitt, C., Taha, M. K. & Boneca, I. G. (2013). *Mol. Microbiol.* **87**, 1100–1112.
Weadge, J. T. & Clarke, A. J. (2006). *Biochemistry*, **45**, 839–851.
Weadge, J. T., Pfeffer, J. M. & Clarke, A. J. (2005). *BMC Microbiol.* **5**, 49.
Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.
Zakharova, E., Horvath, M. P. & Goldenberg, D. P. (2009). *Proc. Natl Acad. Sci. USA*, **106**, 11034–11039.
Zeng, Q., Langereis, M. A., van Vliet, A. L., Huizinga, E. G. & de Groot, R. J. (2008). *Proc. Natl Acad. Sci. USA*, **105**, 9065–9069.