Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structural and mechanistic studies of the *orf12* gene product from the clavulanic acid biosynthesis pathway

Karin Valegård,[a] Aman Iqbal,[b] Nadia J. Kershaw,[b] David Ivison,[b] Catherine Généreux,[c] Alain Dubus,[c] Cecilia Blikstad,[d] Marina Demetriades,[b] Richard J. Hopkinson,[b] Adrian J. Lloyd,[e] David I. Roper,[e] Christopher J. Schofield,[b]* Inger Andersson[f] and Michael A. McDonough[b]*

[a]Department of Molecular Biology, Swedish University of Agricultural Sciences, Box 590, S-751 24 Uppsala, Sweden, [b]Department of Chemistry, University of Oxford, Mansfield Road, Oxford OX1 3TA, England, [c]Centre for Protein Engineering, Liége University, B6, Sart-Tilman, B-4000 Liége, Belgium, [d]Department of Chemistry, Uppsala University, Box 576, S-751 23 Uppsala, Sweden, [e]Department of Biological Sciences, University of Warwick, Gibbet Hill Road, Coventry CV4 7AL, England, and [f]Department of Cell and Molecular Biology, Uppsala University, Box 596, S-751 24 Uppsala, Sweden

Correspondence e-mail: christopher.schofield@chem.ox.ac.uk, michael.mcdonough@chem.ox.ac.uk

Received 22 February 2013
Accepted 23 April 2013

PDB References: ORF12, native, 2xep; native, CA complex, 2xf3; Ser173Ala, 2xft; Ser173Ala, CA complex, 2xfs; Ser378Ala, 2xgn; Ser378Ala, CA complex, 2xh9

Structural and biochemical studies of the *orf12* gene product (ORF12) from the clavulanic acid (CA) biosynthesis gene cluster are described. Sequence and crystallographic analyses reveal two domains: a C-terminal penicillin-binding protein (PBP)/β-lactamase-type fold with highest structural similarity to the class A β-lactamases fused to an N-terminal domain with a fold similar to steroid isomerases and polyketide cyclases. The C-terminal domain of ORF12 did not show β-lactamase or PBP activity for the substrates tested, but did show low-level esterase activity towards 3′-*O*-acetyl cephalosporins and a thioester substrate. Mutagenesis studies imply that Ser173, which is present in a conserved S*XX*K motif, acts as a nucleophile in catalysis, consistent with studies of related esterases, β-lactamases and D-Ala carboxypeptidases. Structures of wild-type ORF12 and of catalytic residue variants were obtained in complex with and in the absence of clavulanic acid. The role of ORF12 in clavulanic acid biosynthesis is unknown, but it may be involved in the epimerization of (3*S*,5*S*)-clavaminic acid to (3*R*,5*R*)-clavulanic acid.

1. Introduction

Clavulanic acid {CA; IUPAC name (2*R*,3*Z*,5*R*)-3-(2-hydroxyethylidene)-7-oxo-4-oxa-1-azabicyclo[3.2.0]heptane-2-carboxylic acid} is a bicyclic β-lactam and a potent inhibitor of the largest class of serine β-lactamases: class A (Fig. 1). Because the antibiotic activity of CA alone is insufficient for clinical use, it is used in combination with a more potent penicillin antibiotic such as amoxicillin. The density of functionalization in CA renders its production *via* chemical synthesis difficult. Consequently, CA is biosynthetically produced by fermentation of *Streptomyces clavuligerus*. CA is one of a number of clavams produced by *S. clavuligerus*, which concomitantly produces members of the penicillin and cephalosporin antibiotic families. Like all isolated naturally occurring penicillins and cephalosporins, CA possesses the 3*R*,5*R* stereochemistry (4*R*,6*R* in cephalosporins) that is required for both (effective) antibacterial activity *via* inhibition of cell-wall transpeptidases and inhibition of β-lactamases. However, the 3*R*,5*R* stereochemistry of CA and its immediate biosynthetic precursor (3*R*,5*R*)-clavaldehyde is unusual among the clavam family because almost all of the other identified intermediates in clavam biosynthetic pathways possess the 3*S*,5*S* stereochemistry (Baggaley *et al.*, 1997; Egan *et al.*, 1997; Kershaw *et al.*, 2005).

Three gene clusters have been identified in *S. clavuligerus* that contribute to clavam biosynthesis (Jensen *et al.*, 2000; Tahlan *et al.*, 2004). The best characterized of these (subsequently termed the 'primary' cluster) is currently thought to be comprised of ~18 genes that are directly involved in

biosynthesis, transport and regulation during CA biosynthesis (Fig. 1; Tahlan *et al.*, 2004). This cluster is directly adjacent to a cluster that encodes for the biosynthesis of penicillins and cephalosporins. Another differently regulated clavam gene cluster has also been identified in *S. clavuligerus* that contains a minimum of four genes homologous to open reading frames *orf2*, *orf3*, *orf4* and *orf6* of the better characterized cluster (Jensen *et al.*, 2000).

The functions of the products of *orf10*, *orf11*, *orf12*, *orf14* and *orf15* in the primary gene cluster are unknown. Several reports have demonstrated that the *orf12* gene of the primary gene cluster is essential for CA production (de la Fuente *et al.*, 2002; Jensen *et al.*, 2004; Li *et al.*, 2000). Here, we report structural and biochemical studies of the ORF12 protein. Sequence and crystallographic analyses reveal ORF12 to be a member of the PBP/β-lactamase structural family with an additional N-terminal domain. However, ORF12 does not appear to possess β-lactamase or PBP activity and the formation of a covalent acyl-enzyme complex with β-lactams was not observed under the conditions tested. Instead, ORF12 was found to have cephalosporin esterase activity, catalysing the hydrolysis of the ester group of 3′-*O*-acetyl cephalosporins, similar to the family VIII esterase EstB (Petersen *et al.*, 2001).

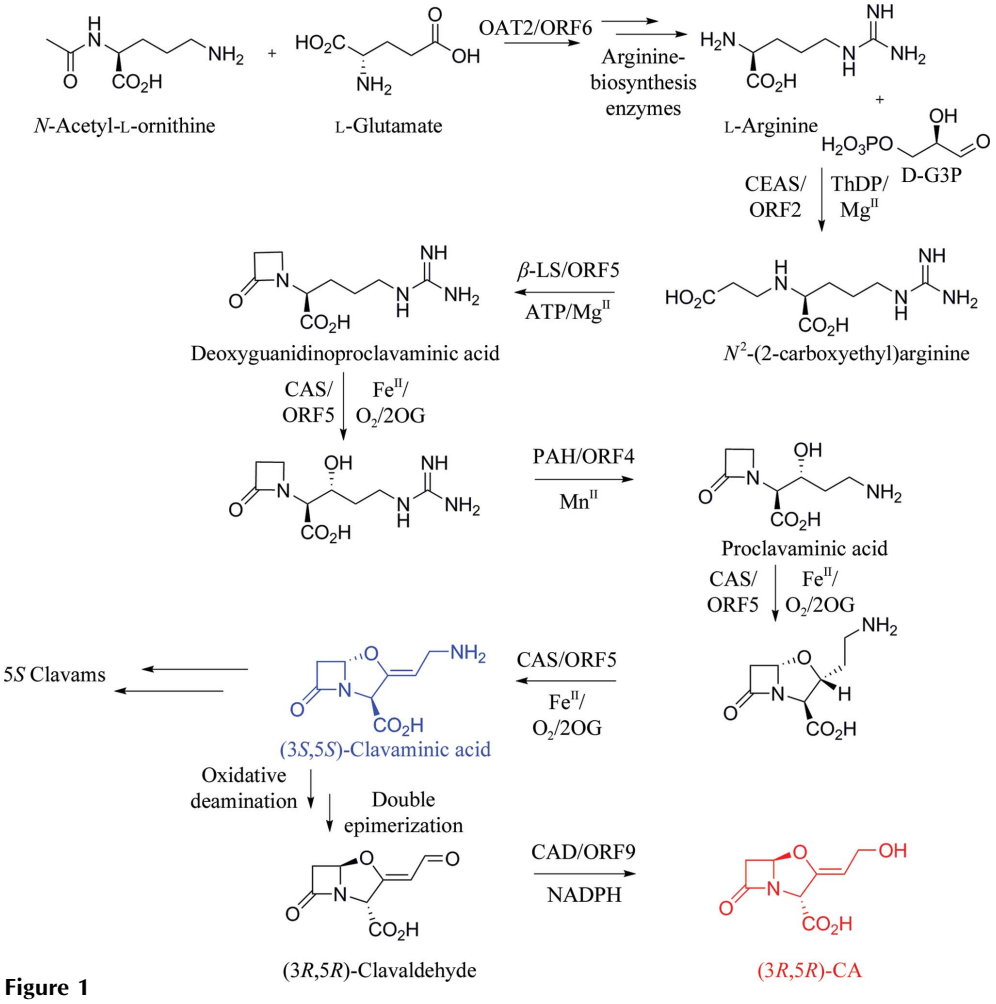

Figure 1
Outline of the clavulanic acid biosynthesis pathway in *S. clavuligerus*, showing identified intermediates and the enzymes involved. The proposed branch-point intermediate clavaminic acid is shown in blue and CA is shown in red [CEAS (ORF2), carboxyethyl arginine synthase; β-LS (ORF3), β-lactam synthase; CAS (ORF5), clavaminate synthase; PAH (ORF4), proclavaminate synthase; OAT2 (ORF6), ornithine acetyltransferase; CAD (ORF9), clavulanate dehydrogenase; Kershaw *et al.*, 2005; not all cosubstrates/cofactors are shown].

2. Materials and methods

2.1. Cloning, expression and purification of ORF12

Three potential versions of the *orf12* gene have been proposed with slightly different N-termini, resulting in 430-, 458- and 466-residue proteins all with the same C-terminus (Jensen *et al.*, 2004; Li *et al.*, 2000; Mellado *et al.*, 2002). Each was amplified by PCR (using a vector kindly donated by GlaxoSmithKline, UK). The forward primers used were the 'Mellado sequence', 5′-GGT GGT **CAT ATG** GCC GAG GGC CG-3′ (466 residues, 50.6 kDa), the 'Jensen sequence', 5′-GGT GGT **CAT ATG** ATG AAG AAA GC-3′ (458 residues, 49.8 kDa), and the 'Li sequence', 5′-GGT GGT **CAT ATG** GGC GAC GCG GCC-3′ (430 residues, 47.0 kDa). The reverse primer was the same for all three PCR reactions: 5′-GGT GGT **GGA TCC** TCA TCG CCG GG-3′. The PCR products were directly cloned into the pET24a(+) vector (Novagen) using *Nde*I/*Bam*HI restriction sites. Expression trials for the three constructs only showed soluble expression for the 458-residue 'Jensen sequence' and this was the construct that was used for all experiments described in this work. For production, the plasmids were transformed into *Escherichia coli* BL21 (DE3) cells and grown at 310 K with shaking at 250 rev min^{-1} in 2TY medium containing kanamycin at 50 µg ml^{-1}. When the OD$_{600}$ reached 0.6–0.8, the temperature was lowered to 288 K and protein production was induced by the addition of 1 m*M* isopropyl β-D-1-thiogalactopyranoside (IPTG). Following overnight incubation, the cells were harvested by centrifugation at 6000*g* for 15 min at 277 K. The cells were resuspended in 50 ml 50 m*M* Tris–HCl pH 7.5, 2 m*M* $MgCl_2$, 10 µl (250 U) Benzonase Nuclease (Novagen, Denmark) and one Complete EDTA-free Protease Inhibitor tablet (Roche

Diagnostics GmbH, Mannheim, Germany). Cell lysis was performed using either a One Shot cell disruptor (Constant Cell Disruptor Systems, Daventry, England) or a Vibra Cell sonicator and the lysate was cleared by centrifugation at 20 000*g*. Q Sepharose Fast Flow, HiPrep 26/10 Desalting, Mono Q and Superdex 200 columns (GE Healthcare Biosciences AB, Uppsala, Sweden) were subsequently employed for protein purification. The purified protein was concentrated to approximately 12 mg ml^{-1} (the estimated yield was 55 mg purified protein per litre of culture).

The ORF12 Ser173Ala, Ser234Ala and Ser378Ala variants were produced by standard site-directed mutagenesis methodology using the *orf12*/pET24a(+) vector. Note that the numbering of the ORF12 protein sequence is based on the protein predicted by Jensen *et al.* (2004) (Supplementary Fig. S1[1]). After mutagenesis, the genes were subcloned into the pET28a(+) vector, which encodes proteins with an N-terminal hexahistidine tag.

The N-terminal domain of ORF12 (ORF12$_{1–127}$) was prepared by changing the codon encoding Met128 into a stop codon. The C-terminal domain of ORF12 was prepared by amplifying the gene fragment (residues 128–458) and subcloning it into pET24a(+) vector to give the ORF12$_{128–458}$ pET24a(+) construct.

The Ser234Ala, Ser378Ala and Ser173Ala ORF12 variants were produced in soluble form in *E. coli* BL21 (DE3) cells (as the N-terminally His$_6$-tagged form) and were purified to >90% purity using Ni^{2+}-affinity chromatography. In addition, both the C-terminal domain [using the ORF12$_{128–458}$ pET24a(+) vector] and the N-terminal domain [using the ORF12$_{1–127}$ pET28a(+) vector] of ORF12 could be produced separately in soluble form in *E. coli* BL21 (DE3) cells, but the expression level was lower for ORF12$_{128–458}$ (10% of the total cell lysate proteins compared with 20% for the wild-type ORF12). ORF12$_{1–127}$ was purified by Ni^{2+}-affinity chromatography to >80% purity, while ORF12$_{128–458}$ was purified using the same chromatography protocol as used for wild-type ORF12.

Selenomethionine (SeMet) substituted ORF12 was produced using a metabolic inhibition protocol. Cells were grown at 310 K in 2.5 ml 2TY medium containing kanamycin at 50 µg ml^{-1}. The starter culture was used to inoculate 50 ml M9 minimal medium (50 m*M* Na_2HPO_4, 3 g l^{-1} KH_2PO_4, 0.5 g l^{-1} NaCl, 1 g l^{-1} NH_4Cl) and the cells were grown at 310 K until an OD_{600} of 0.5 was reached. A 25 ml inoculum was added to 500 ml M9 medium supplemented with 2 m*M* $MgSO_4$, 0.1 m*M* $CaCl_2$, 0.4%(*w*/*v*) glucose and 50 µg ml^{-1} kanamycin. The cells were grown at 310 K until an OD_{600} of 0.4–0.6 was reached, whereupon the following L-amino acids were added: Lys, Phe and Tyr (100 mg l^{-1}), Leu, Ile and Thr (50 mg l^{-1}) and the selenomethionine derivative L-selenomethionine (Acros Organics) at a final concentration of 60 mg l^{-1}. Complete (>95%) SeMet incorporation was achieved as observed by ESI-MS. The SeMet-ORF12 protein was purified as described for wild-type ORF12.

2.2. Substrate screening and enzyme assays

2.2.1. HPLC-based substrate screening. Activity assays employed reverse-phase HPLC using a C18 column (3.5 µm, 75 × 4.6 mm) at a flow rate of 1.5 ml min^{-1}. Each assay consisted of 10 m*M* 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) buffer pH 7.0, 10 µ*M* ORF12 and a range of substrate (cephalosporin/clavulanic acid derivative side chains of various forms; see Supplementary Fig. S2 and Table S1) concentrations (150, 300, 600 and 1200 µ*M*) in a final volume of 200 µl at 298 K. Samples (10 µl) were injected onto the HPLC system at regular time points (0, 60, 120, 180, 240, 300, 360 and 420 min). A control reaction consisting of 10 m*M* HEPES pH 7.0, 1200 m*M* substrate and bovine serum albumin was analysed at the start and end of the data collection to test for non-enzyme-catalysed substrate hydrolysis. For HPLC studies, solution *A* was potassium phosphate pH 2.5 and solution *B* was 100% acetonitrile. Isocratic flow of 98% solution *A* for 1 min was followed by a linear gradient to 82% solution *A* over 3 min, a further linear gradient to 20% solution *A* over 1 min, isocratic flow for 2 min and then a gradient back to the initial conditions over 1 min. The absorbance at 260 nm was monitored. Alternative substrates were assayed under similar conditions. Data were analysed using *OriginPro* 8.5.1.

2.2.2. Acetyltransferase assay. To detect acetyltransferase activity, assays were performed using reaction mixtures consisting of 100 m*M* Tris–HCl pH 7.5, 10 m*M* substrate, 30 µ*M* ORF12 at 293 K for between 30 and 180 min. Reaction products were analysed by LC-MS using a Waters 600 Controller Pump with a Waters 2700 Sample Manager in combination with a Micromass ZMD mass spectrometer to perform ESI-MS. Samples were separated using a C18 (150 × 4.60 mm, 5 m) reverse-phase column [LUNA 5 m C18(2) 100A, Phenomenex] with a gradient from 5 to 30% acetonitrile in 100 m*M* ammonium formate pH 4.0 at a flow rate of 1 ml min^{-1} over 15 min followed by a wash with 30% acetonitrile in 100 m*M* ammonium formate pH 4.0 for 10 min before the column was re-equilibrated to 5% acetonitrile in 100 m*M* ammonium formate pH 4.0 over 5 min. *MassLynx* software (Waters) was used for data analysis and processing.

2.2.3. NMR-based deacetylase assay. For the ^{1}H NMR experiments, a reaction mixture consisting of 0.2 m*M* ORF12 (or an ORF12 variant), 3.5 m*M* substrate, 50 m*M* Tris–HCl pH 7.5 and D_2O to a final volume of 200 µl was prepared and scans were recorded every 30 min at 298 K using a Bruker Avance II 500 MHz NMR instrument.

2.2.4. Carboxypeptidase assay. Continuous spectrophotometric assays for PBP enzyme activity which followed the release of D-alanine from the UDP-MurNAc pentapeptide as reported by D-amino-acid oxidase/horseradish peroxidase and Amplex Red were based on those described by Clarke *et al.* (2009). Specifically, assays in a final volume of 0.2 ml consisting of 50 m*M* HEPES pH 7.5, 10 m*M* $MgCl_2$, 50 m*M*

[1] Supplementary material has been deposited in the IUCr electronic archive (Reference: KW5065). Services for accessing this material are described at the back of the journal.

Table 1
Data-collection and refinement statistics.

Values in parentheses are for the outer resolution shell.

	SeMet-ORF12 SAD	Native ORF12	ORF12–CA	ORF12-Ser173Ala	ORF12-Ser173Ala–CA	ORF12-Ser378Ala	ORF12-Ser378Ala–CA
PDB entry	NA	2xep	2xf3	2xft	2xfs	2xgn	2xh9
X-ray source	ID14-3, ESRF	ID14-1, ESRF	ID14-4, ESRF	ID14-2, ESRF	ID14-2, ESRF	ID14-3, ESRF	BM16, ESRF
Resolution (Å)	60.4–2.30 (2.38–2.30)	81.6–1.50 (1.53–1.50)	35.0–1.55 (1.61–1.55)	81.1–1.80 (1.86–1.80)	81.1–1.80 (1.86–1.80)	81.0–1.70 (1.76–1.70)	81.1–1.80 (1.86–1.80)
Space group	$P2_1$	$P2_1$	$P2_1$	$P2_1$	$P2_1$	$P2_1$	$P2_1$
Unit-cell parameters (Å, °)	a = 60.7, b = 97.9, c = 81.9, β = 93.8	a = 61.1, b = 99.1, c = 81.6, β = 94.2	a = 61.6, b = 100.8, c = 81.2, β = 94.0	a = 61.1, b = 100.0, c = 81.3, β = 93.6	a = 61.6, b = 100.6, c = 81.2, β = 93.8	a = 61.3, b = 99.8, c = 81.3, β = 94.0	a = 61.5, b = 100.9, c = 81.4, β = 94.0
Molecules in asymmetric unit	2	2	2	2	2	2	2
No. of observations	358062	878375	770269	382567	381178	449998	335907
No. of unique reflections	42537	151975	139836	90179	89838	106976	89909
Completeness (%)	100.0 (100.0)	98.7 (96.9)	97.5 (99.8)	100.0 (100.0)	98.3 (98.4)	100.0 (100.0)	98.1 (91.3)
Multiplicity	8.4 (7.4)	3.8 (3.5)	5.5 (5.1)	4.2 (3.4)	4.2 (3.7)	4.2 (3.8)	3.7 (3.1)
R_{meas}†	0.116 (0.305)	0.075 (0.283)	0.081 (0.249)	0.061 (0.217)	0.102 (0.431)	0.080 (0.407)	0.096 (0.251)
No. of Se sites	27						
Figure of merit	0.34						
Phasing power	1.003						
Residues in model		*A*7–*A*41, *A*47–*A*59, *A*64–*A*434, *B*8–*B*58, *B*64–*B*434	*A*8–*A*44, *A*49–*A*434, *B*8–*B*435	*A*10–*A*41, *A*47–*A*59, *A*64–*A*434, *B*8–*B*59, *B*64–*B*435	*A*8–*A*43, *A*47–*A*59, *A*64–*A*434, *B*8–*B*59, *B*64–*B*435	*A*8–*A*43, *A*47–*A*59, *A*64–*A*434, *B*8–*B*59, *B*64–*B*435	*A*8–*A*43, *A*47–*A*59, *A*64–*A*434, *B*8–*B*62, *B*64–*B*435
No. of solvent molecules		768	1039	586	601	668	643
B factor (overall) (Å^2)		18.1	17.1	19.6	21.3	17.7	22.1
R_{cryst}‡		0.200	0.187	0.194	0.214	0.191	0.196
R_{free}§		0.217	0.210	0.236	0.261	0.224	0.226
Wilson B factor (Å^2)		14.1	14.3	18.2	19.9	15.8	20.8
R.m.s. deviations from ideal geometry							
Bond lengths (Å)		0.006	0.010	0.014	0.017	0.011	0.014
Bond angles (°)		1.058	1.263	1.395	1.888	1.339	1.765
Ramachandran¶ (%)		92.5, 7.0, 0.3, 0.3	91.9, 7.6, 0.5, 0.0	92.3, 7.1, 0.4, 0.1	90.9, 8.5, 0.6, 0.0	92.8, 6.5, 0.2, 0.5	91.4, 8.0, 0.6, 0.0
Rotamer outliers†† (%)		0.7	0.4	0.9	0.7	0.9	0.7
Clashscore††		3.53	2.13				

† As defined by Diederichs & Karplus (1997). ‡ $R_{work} = \sum_{hkl} ||F_{obs}| - |F_{calc}|| / \sum_{hkl} |F_{obs}|$, where F_{obs} and F_{calc} are the observed and calculated structure-factor amplitudes, respectively. § R_{free} was calculated from a randomly chosen 5% of all unique reflections. ¶ Ramachandran plot statistics (most favoured, additionally allowed, generously allowed, disallowed). †† As output from *MolProbity* (Chen *et al.*, 2010).

KCl, 0.1 m*M* Amplex Red (Molecular Probes Inc.), 0.27 mg ml^{-1} horseradish peroxidase, 0.38 mg ml^{-1} D-amino-acid oxidase (both from Sigma and used without further purification). ORF12 was added to a final concentration of 0.11 mg ml^{-1}. The background absorbance at 555 nm at 310 K was followed and PBP activity was initiated by the addition of the UDP MurNAc pentapeptide (lysine or diaminopimelic acid variant) as detailed in Supplementary Fig. S3. To establish the functionality of the assay, where necessary, 16.4 µg ml^{-1} *Escherichia coli* (DacB) D,D-carboxypeptidase was added as a positive control at the end of the assay (Supplementary Fig. S3).

2.2.5. Fluorescent penicillin-binding assay. To confirm and extend the data obtained from spectrophotometric assays of PBP-catalysed D-alanine release, the ability of ORF12 to bind β-lactams was probed with the fluorescent penicillin V derivative bocillin as described by Zhao *et al.* (1999) (Supplementary Fig. S5). ORF12, *S. pneumoniae* PBP2x (minus the N-terminal 48 amino acids corresponding to the transmembrane helix), *S. pneumoniae* 2B (minus the N-terminal 39 amino acids corresponding to the transmembrane helix), both purified according to Abrahams (2011), and *E. coli* DacB purified according to Clarke *et al.* (2009) (20 µg each) were incubated at 298 K in final volume of 20 µl in 50 m*M* HEPES, 10 m*M* $MgCl_2$ pH 7.6 for 30 min with or without 0.1 m*M* ampicillin (Clarke *et al.*, 2009). All samples were then supplemented with bocillin at a 5:1 molar ratio of protein to bocillin, incubated for a further 60 min and analysed by SDS–PAGE (10% acrylamide).

2.3. Crystallization and X-ray diffraction data collection

Purified protein was crystallized using the sitting-drop vapour-diffusion method at 293 K. Crystals were obtained by streak-seeding from crushed needle-cluster crystals grown in 13–14%(*w*/*v*) polyethylene glycol 4000, 0.1 *M* sodium acetate, 0.1 *M* sodium chloride, 10%(*w*/*v*) glycerol, 0.1 *M* Tris–HCl pH 8.0. Selenomethionine-substituted ORF12, Ser173Ala-ORF12 and Ser378Ala-ORF12 were crystallized in the same conditions. The CA complexes were prepared by soaking crystals at 293 K for 20 min in a solution containing 20 m*M* CA. The crystals were flash-cooled in liquid nitrogen.

2.4. Structure determination and refinement

The structure of ORF12 was determined using single-wavelength anomalous dispersion (SAD) data collected to 2.2 Å resolution at 100 K on beamline ID14-3 at the ESRF, Grenoble, France. The data were indexed with *DENZO* and scaled with *SCALEPACK* (Otwinowski & Minor, 1997). The crystals belonged to space group $P2_1$, with unit-cell parameters a = 61.1, b = 99.1, c = 81.6 Å, β = 94.7° and two monomers in the asymmetric unit. 12 selenium sites were located using *HKL2MAP* (Pape & Schneider, 2004) and were cross-validated and refined using *autoSHARP* (de La Fortelle & Bricogne, 1997), during which 15 new sites were added. Solvent flattening and density modification was performed as implemented in *autoSHARP* and the resulting maps were used for initial chain-tracing in *O* (Jones *et al.*, 1991).

Native data to 1.5 Å resolution were collected on beamline ID14-1 of the ESRF. The native structure was generated using rigid-body refinement in *REFMAC*5 (Murshudov *et al.*, 2011) using the selenomethionine-derivative ORF12 structure as the initial model, followed by minor rebuilding. The structure was initially refined in *PHENIX* (Adams *et al.*, 2010) using simulated annealing. Refinement continued using *REFMAC*5. The $R_{\rm free}$ value was calculated using a randomly selected subset of the data (5%). Solvent molecules were added using *ARP*/*wARP* (Lamzin *et al.*, 2001) and were manually inspected in *O* (Jones *et al.*, 1991). The final rebuilding/refinement round resulted in a model with an R factor of 20.0% and an $R_{\rm free}$ of 21.7%. High-resolution data were collected for the variants and their complexes with CA at 100 K on beamlines ID14-4 (ORF12–CA), ID14-2 (ORF12-Ser173Ala and ORF12-Ser173Ala–CA), ID14-3 (ORF12-Ser178Ala) and BM16 (ORF12-Ser178Ala–CA) at the ESRF. Data for ORF12 (and variants) without ligand and in complex with CA were indexed, processed and scaled with *DENZO* and *SCALEPACK*. Models were generated using rigid-body refinement in *REFMAC*5 using the native protein model, followed by minor rebuilding. The structures were evaluated using *PROCHECK* (Laskowski *et al.*, 1993). A summary of the data-collection and refinement statistics for all structures is given in Table 1.

3. Results and discussion

3.1. Enzymatic activity of ORF12

3.1.1. β-Lactam reactivity. Given that ORF12 is vital for CA biosynthesis (de la Fuente *et al.*, 2002; Jensen *et al.*, 2004; Li *et al.*, 2000) and that CA is an antibiotic/β-lactamase inhibitor, we considered the possibility that ORF12 may also act as a β-lactamase and possibly also as part of a self-resistance mechanism. However, various β-lactams (nitrocefin, clavulanic acid, benzylpenicillin and ampicillin) were not observed to be hydrolysed above background levels when incubated with ORF12 under the conditions tested, as observed by either high-performance liquid chromatography (HPLC)/UV detection or nuclear magnetic resonance (NMR) based assays (Supplementary Fig. S2 and Table S1). We also examined the possibility of CA reacting and remaining bound covalently as the acyl-enzyme complex with ORF12 by ESI-MS (Brown *et al.*, 1996). However, no such interaction was observed.

3.1.2. Lack of penicillin binding. To further investigate whether or not ORF12 binds covalently to β-lactam antibiotics, we ran three positive-control PBPs using the fluorescent β-lactam bocillin FL (Molecular Probes): *S. pneumoniae* PBP2x, *S. pneumoniae* PBP2B (both high-molecular-weight transpeptidases) and *E. coli* DacB (a low-molecular-weight D,D-carboxypeptidase). The assay was performed after incubation in the presence or absence of ampicillin and the exercise was repeated for ORF12. The PBPs and ORF12 were then subjected to SDS–PAGE to separate bound bocillin from the free compound and to allow visualization of β-lactam labelling of the PBP polypeptide. The three bona fide PBPs all showed labelling with bocillin (Supplementary Fig. S3); the labelling was eliminated by treatment with ampicillin, indicating that the fluorescent labelling occurs at the active site. In contrast, ORF12 failed to bind bocillin in this experiment (Supplementary Fig. S3), with no fluorescence being detected.

3.1.3. C-3′-*O*-Acetyl cephalosporin esterase activity. Unexpectedly, however, when ORF12 was incubated with cephalosporin C followed by HPLC analysis, we observed the formation of a product peak that retained its absorption maximum at 260 nm characteristic of β-lactams. The product obtained was shown to have the same retention time as deacetylcephalosporin C, indicating that ORF12 has a low level of cephalosporin deacetylase activity (Fig. 2*a*). Kinetic parameters were obtained for cephalosporin C ($K_{\rm m}$ = 79.6 μ*M* and $k_{\rm cat}$ = 0.00125 min^{-1} by HPLC analysis; Supplementary Fig. S2). Other C-3′-*O*-acetyl cephalosporin esters were also found to be weak substrates for the esterase activity (cephaloglycin, cephalothin and cefotaxime). Although the levels of activity are low, the results reveal that ORF12 accepts cephalosporins with acetyl-ester side chains at the C-3′ position, but not carbamates (Supplementary Table S1). This implies that the biological function of ORF12 may not be as an esterase, but supports the proposal that ORF12 is a hydrolase/nucleophilic enzyme.

ORF12 was also shown to possess acetyltransferase activity by LC-MS analyses. Thus, incubation of deacetylcephalosporin C (373 Da) and cephaloglycin (405 Da) led to the products cephalosporin C (416 Da) and deacetylcephaglycin (364 Da), respectively. Acetate alone was not able to act as an acetyl donor in acetyltransferase catalysis, and acetyl transfer to CA from cephalosporin C was not observed. The observed acetyltransfer activity of ORF12 supports a mechanism in which a covalently bound acetyl-enzyme complex is reversibly formed, possibly in a manner related to that occurring for the PBP transpeptidases.

There is a precedent for the cephalosporin acetylase activity observed with ORF12. Like ORF12, the structurally related enzyme EstB from *Burkholderia gladioli* shows deacetylation activity towards cephalosporins and also shows no detectable β-lactamase activity (Petersen *et al.*, 2001). EstB is also active towards short-chain triglycerides and nitrophenyl esters, suggesting that it is a typical carboxylesterase. However, no esterase activity was observed when *p*-nitrophenylbutyrate

and *p*-nitrophenylacetate were assayed as substrates for ORF12. EstB belongs to esterase family VIII, which contains four broad-specificity carboxylesterases that show sequence homology to class C β-lactamases (differing from class A, which is most similar to ORF12; Supplementary Fig. S4; Wagner *et al.*, 2002). Since some PBPs have been shown to possess esterase activity (Kelly *et al.*, 1986), we then tested ORF12 with a thioester substrate, (*R*)-2-[(2-benzamidopropanoyl)thio]acetic acid (S2d), a mimic of D-Ala-Gly which is used to assay PBP activity (Fig. 2*b*; Adam *et al.*, 1990, 1991; Wilkin *et al.*, 1993). S2d hydrolysis was catalysed by ORF12 as measured by NMR, with an initial rate of 0.037 m*M* min^{-1}.

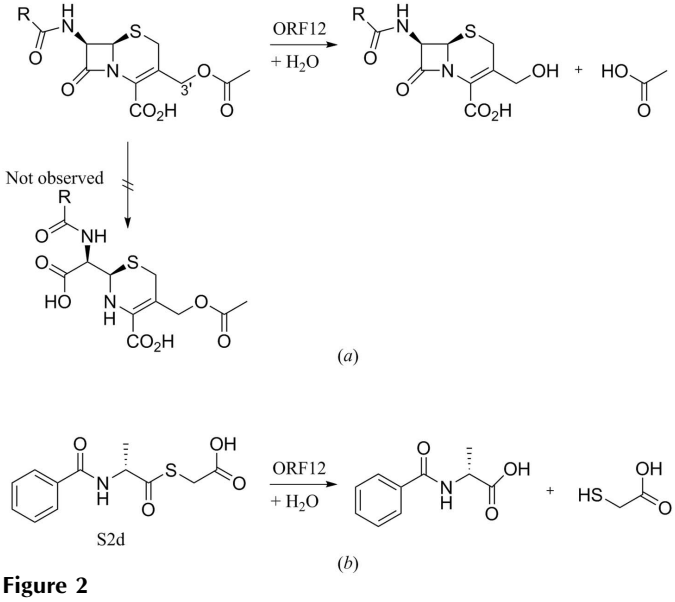

Figure 2
(*a*) ORF12 possesses cephalosporin esterase activity. Cephalosporins with a 3′-*O*-acetyl group are substrates of ORF12 (cephalosporin C, cephaloglycin, cephalothin and cefotaxime). Other tested cephalosporins were not found to be substrates, including those with C3′-carbamides. No ORF12-catalysed β-lactamase activity was observed. (*b*) ORF12 also catalyses the hydrolysis of the thioester S2d. See Supplementary Table S1 for a comparison of the substrates.

3.1.4. The C-terminal domain construct shows esterase activity. Independent ^{1}H NMR activity assays were carried out with separate constructs of the N-terminal domain (ORF12$_{1–127}$) and the C-terminal domain (ORF12$_{128–458}$) of ORF12 using cephalosporin C. The N-terminal domain construct ORF12$_{1–127}$ showed no esterase activity, while the C-terminal domain construct ORF12$_{128–458}$ showed only a slight reduction in activity (V_{max} = 0.019 m*M* min^{-1}) compared with full-length ORF12 (V_{max} = 0.047 m*M* min^{-1}), demonstrating that the esterase activity resides in the C-terminus of ORF12.

3.1.5. Lack of amidase and carboxypeptidase activities. In order to explore whether ORF12 can hydrolyse amides as well as esters, ε-*N*-acetyl-lysine, *N*-acetyl-glycine, *N*-acetyl-glycyl-glycine, *N*-acetyl-ethanolamine, *N*-(2-chloroethyl)acetamide and 4-acetamidoiodobutyric acid were also assessed as substrates by NMR. No hydrolase activity was observed with these substrates, suggesting that ORF12 is not an amidase, or at least not a nonspecific one.

We then tested ORF12 for carboxypeptidase activity using the Amplex Red-based spectroscopic assay (Clarke *et al.*, 2009) with peptidoglycan precursors found in *Streptomyces* as PBP substrates (Supplementary Fig. S5). In this assay, activity is detected by the release of D-alanine from either UDP-MurNAc-L-alanyl-γ-D-glutamyl-*meso*-diaminopimelyl-D-alanyl-D-alanine (UDP-MurNAc pentapeptide; DAP) or UDP-MurNAc-L-alanyl-γ-D-glutamyl-L-lysyl-D-alanyl-D-alanine (UDP-MurNAc pentapeptide; Lys) owing to D,D-carboxypeptidase activity. The D-alanine is oxidized by D-amino-acid oxidase and the resulting hydrogen peroxide is consumed by horseradish peroxidase with the concomitant conversion of Amplex Red (10-acetyl-3,7-dihydroxyphenoxazine) to resorufin with the production of intense absorbance at 555 nm. With ORF12, the assay did not reveal any evidence of D-alanine products after incubation with either potential substrate. In contrast, use of *E. coli* DacB as a positive control demonstrated that the assay was functional.

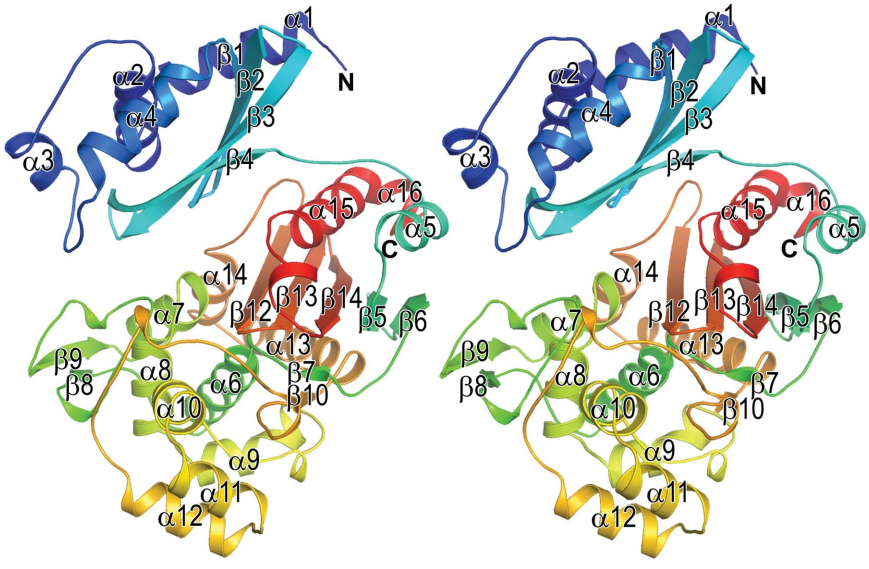

Figure 3
The overall structure of ORF12. Stereoview ribbons representation of ORF12 coloured from blue at the N-terminus to red at the C-terminus, with the secondary-structure elements labelled as designated by the *HERA* subroutine of *PROMOTIF* (Hutchinson & Thornton, 1990, 1996).

3.2. Crystal structure of ORF12

3.2.1. Overall structure. To investigate possible structure–function relationships between ORF12 and esterases/β-lactamases/PBPs, the ORF12 structure was determined to 2.2 Å resolution by the single-wavelength anomalous dispersion (SAD) method using selenomethionine-substituted protein based on the Jensen start codon (458 residues in length, 49.8 kDa; Table

1). The phases were subsequently improved and extended to 1.5 Å resolution using data from the native ORF12 crystal.

The ORF12 crystal structures reveal a monomer consisting of two domains: an N-terminal domain (residues 7–121) and a C-terminal domain (residues 122–435) (Fig. 3). The N-terminal domain consists of one four-stranded antiparallel β-sheet (β1–β4) surrounded by four α-helices (α1–α4). The C-terminal domain consists of a mixed α/β 'β-lactamase-like' fold characterized by a central five-stranded antiparallel β-sheet (β5, β6, β12, β13 and β14) flanked by α-helices α5 and α15 (Fig. 3) and an additional cluster of mixed 3_{10}-helices and α-helices.

3.2.2. Sequence analysis and the N-terminal domain. A sequence-similarity search using NCBI *BLAST* identified two sequences with close homology to ORF12, both of which have N- and C-terminal domains related to those of ORF12 (Supplementary Fig. S6). One was from *Streptomyces flavogriseus* (55% identity) and the other was from *Saccharomonospora viridis* (45% identity). In addition, a large subfamily annotated as lipoprotein LpqF with unknown function from *Mycobacterium* species appears to be distantly related to ORF12 (~35% identity across the majority of these sequences; Supplementary Figs. S6 and S7). The gene encoding the hypothetical LpqF protein is immediately flanked by the TB11.2 gene [a putative copper(II)-binding antibiotic biosynthesis monooxygenase] upstream and by a putative PBP, Rv3594, downstream, suggesting that like ORF12 it may have a biosynthetic role.

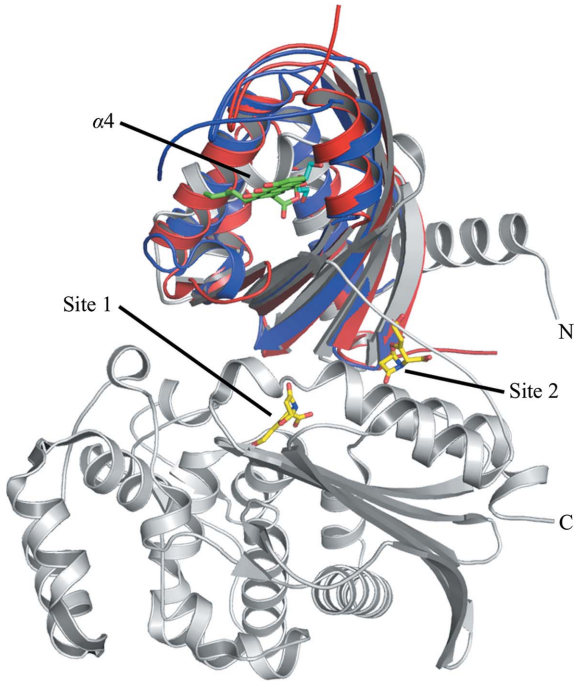

Figure 4
Comparison of the ORF12–CA complex (white ribbons with CA shown as yellow sticks) with similar structures derived from a *DALI* search using the N-terminal domain of ORF12. Red ribbons (PDB entry 3kkg; putative SnoaL-like polyketide cyclase; Joint Center for Structural Genomics, unpublished work) with PEG bound to the active site (cyan sticks). Blue ribbons (PDB entry 2f99; polyketide cyclase AknH; Kallio *et al.*, 2006) with 3-[(1*R*,3*S*)-1,3-dihydroxypentyl]-4,5,9,10-tetrahydroxy-2-anthryl acetate bound in the active site (green sticks). The PKTC active sites are open cavities; ORF12 does not have a similar cavity owing to the position of helix α4.

The N-terminal domain fold of ORF12 has not previously been observed in any of the PBP/β-lactamase or esterase VIII family members and its function is unknown. A similarity search for the N-terminal domain using the *DALI* server (Holm & Rosenström, 2010) indicated structural similarity to an uncharacterized nuclear transport factor 2-like protein from *Mesorhizobium loti* (PDB entry 3fh1; *Z*-score 7.8, r.m.s.d. of 2.4 Å based on 80 of 122 C^{α} atoms, 8% identity; Joint Center for Structural Genomics, unpublished work), a putative ketosteroid isomerase from *Shewanella frigidimarina* (PDB entry 3bb9; *Z*-score 7.8, r.m.s.d. of 2.9 Å based on 91 of 123 C^{α} atoms, 11% identity; Joint Center for Structural Genomics, unpublished work), a putative steroid δ-isomerase from *Pseudomonas aeruginosa* (PDB entry 3mso; *Z*-score 7.6, r.m.s.d. of 2.6 Å based on 88 of 134 C^{α} atoms, 11% identity; Joint Center for Structural Genomics, unpublished work), the polyketide cyclase (PKTC) AknH from *Streptomyces galilaeus* (PDB entry 2f98; *Z*-score 6.8, r.m.s.d. of 2.9 Å based on 87 of 141 C^{α} atoms, 20% identity; Kallio *et al.*, 2006) and the polyketide cyclase SnoaL from *S. nogalater* (PDB entry 1sjw; *Z*-score 6.9, r.m.s.d. of 2.9 Å based on 87 of 142 C^{α} atoms, 17% identity; Sultana *et al.*, 2004). Given that ORF12 may be involved in epimerization during CA biosynthesis, its structural relationship to steroid isomerases is notable. The PKTCs are involved in the biosynthesis of aclacinomycin (converting aklanonic acid methyl ester to aklaviketone; Kantola *et al.*, 2000) and the polypeptide nogalamycin (ring closure *via* aldol condensation to give nogalaviketone; Torkkell *et al.*, 2000); a polar active-site pocket is found between their β-sheet and α-helices. Structures are available of complexes of PKTCs with ligands. However, upon superimposition of the PKTC complexes with the N-terminal domain of ORF12, no apparent polar-lined active-site pocket or conserved catalytic residues analogous to those of either the steroid isomerases or PKTCs were observed in ORF12. Instead, helix α4 of the N-terminal domain of ORF12 occupies the analogous PKTC active-site cavity (Fig. 4).

3.2.3. Structural similarity to β-lactamases/D-Ala carboxypeptidases and a family VIII esterase. The results from the *DALI* search for structural homologues of ORF12 imply that the C-terminal domain fold is related to the β-lactamase/D-Ala carboxypeptidase-like fold, the β-lactamase/D-Ala carboxypeptidase superfamily and the β-lactamase class A family as defined in the Structural Classification of Proteins (SCOP) database (Murzin *et al.*, 1995; Holm *et al.*, 2008). Many reported structures are available for this fold (824 hits using *DALI*). The members of this structural family have been shown to catalyse a variety of different reactions, including (but not limited to) β-lactam hydrolysis (β-lactamases), carboxypeptidation/transpeptidation (PBPs), thioester/ester hydrolysis (EstB) and deamidation (DAA) (Asano *et al.*, 1989; Ghuysen, 1991; Petersen *et al.*, 2001).

The highest structural similarity found using the ORF12 C-terminal domain was to penicillin-binding protein A (PBP-A) from the cyanobacterium *Thermosynechococcus elongatus*

(PDB entry 2jbf; *Z*-score 23.9, r.m.s.d. of 3.0 Å for 242 of 265 residues with 17% identity); although described as being most closely structurally related to the class A β-lactamases, the *T. elongatus* PBP-A enzyme was found not to be active as a β-lactamase but could catalyse the hydrolysis of thioesters with similar kinetic parameters as those of known PBP carboxypeptidase/transpeptidases (Urbach *et al.*, 2008, 2009). As such, PBP-A is predicted to function as a PBP carboxypeptidase/transpeptidase and not as a β-lactamase. Following PBP-A, the top 100 *DALI* hits for the ORF12 structure included class A β-lactamases (TEM and SHV variants; *Z*-score range 23.5–22.9, r.m.s.d. range 2.8–3.3 Å; Jelsch *et al.*, 1992; Kuzin *et al.*, 1999). The family VIII esterase EstB was a low-rated hit at 537th on the list (*Z*-score 14.8, r.m.s.d. of 3.5 Å for 218 out of 377 residues with 13% identity; Wagner *et al.*, 2002).

3.2.4. Active site. Superimposition of the ORF12 active site with those of a representative set of β-lactamase/D-Ala carboxypeptidase superfamily members was carried out (Supplementary Table S2). Striking similarities between the ORF12 active-site architecture and those of the class A β-lactamases (*e.g.* TEM-1 from *E. coli*) and low-molecular-weight class C PBPs (*e.g.* DD-peptidase from *Actinomadura* sp. strain R39; R39) are apparent (Jelsch *et al.*, 1992; Sauvage *et al.*, 2005). The serine β-lactamases are typically characterized by the presence of three active-site sequence motifs: the S*XX*K (containing the nucleophilic serine), SDN and KTG motifs (Supplementary Table S2; Galleni *et al.*, 1995). Residues from these motifs form a Ser1/Ser2(Tyr)/Lys1(His)/Lys2 catalytic tetrad (Pratt & McLeish, 2010) which is also

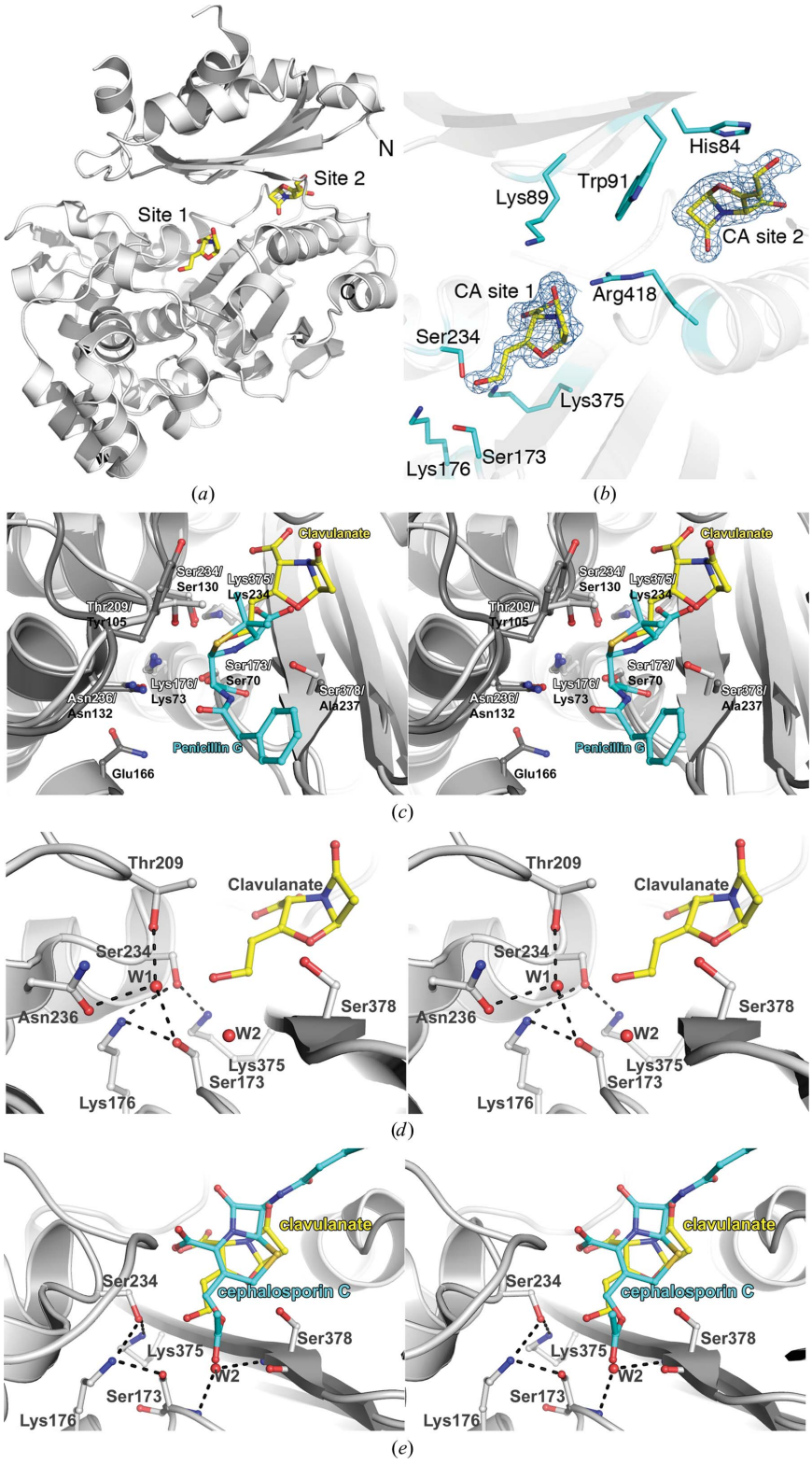

Figure 5
The binding of CA to ORF12. (*a*) ORF12 (white ribbons) has two CA (yellow sticks) binding sites observed in the crystal structures. One CA molecule binds to the active site (site 1) and the second CA molecule binds between the two domains (site 2). (*b*) Close-up showing the $2F_o - F_c$ electron-density map contoured at 1σ around CA in site 1 and site 2. A selected subset of residues lining the binding pockets are depicted as cyan sticks. (*c*) Stereoview of superimposition of the TEM-1 β-lactamase (PDB entry 1fqg; Strynadka *et al.*, 1992; dark grey ribbons and sticks) in complex with penicillin G (cyan sticks) onto the ORF12–CA complex (ORF12 in white ribbons and sticks and CA in yellow sticks). (*d*) Stereoview of the ORF12 (white ribbons) active site (white sticks) with CA (yellow sticks) bound. Water W2 (red sphere) is bound in the oxyanion hole formed by the backbone amides of residues Ser173 and Ser378. Water W1 (red sphere) is believed to be representative of the water involved in the deacylation step of the esterase reaction. Black dashes represent the hydrogen-bonding network in the active site. (*e*) Docking cephalosporin C (cyan sticks) into the ORF12 (white ribbons and sticks) active site based on the ORF12–CA structure (CA in yellow sticks). Note the proposed docked position of the cephalosporin ester carbonyl O atom in the oxyanion hole formed by the backbone amides of residues Ser173 and Ser378. Black dashes represent the hydrogen-bonding network in the active site.

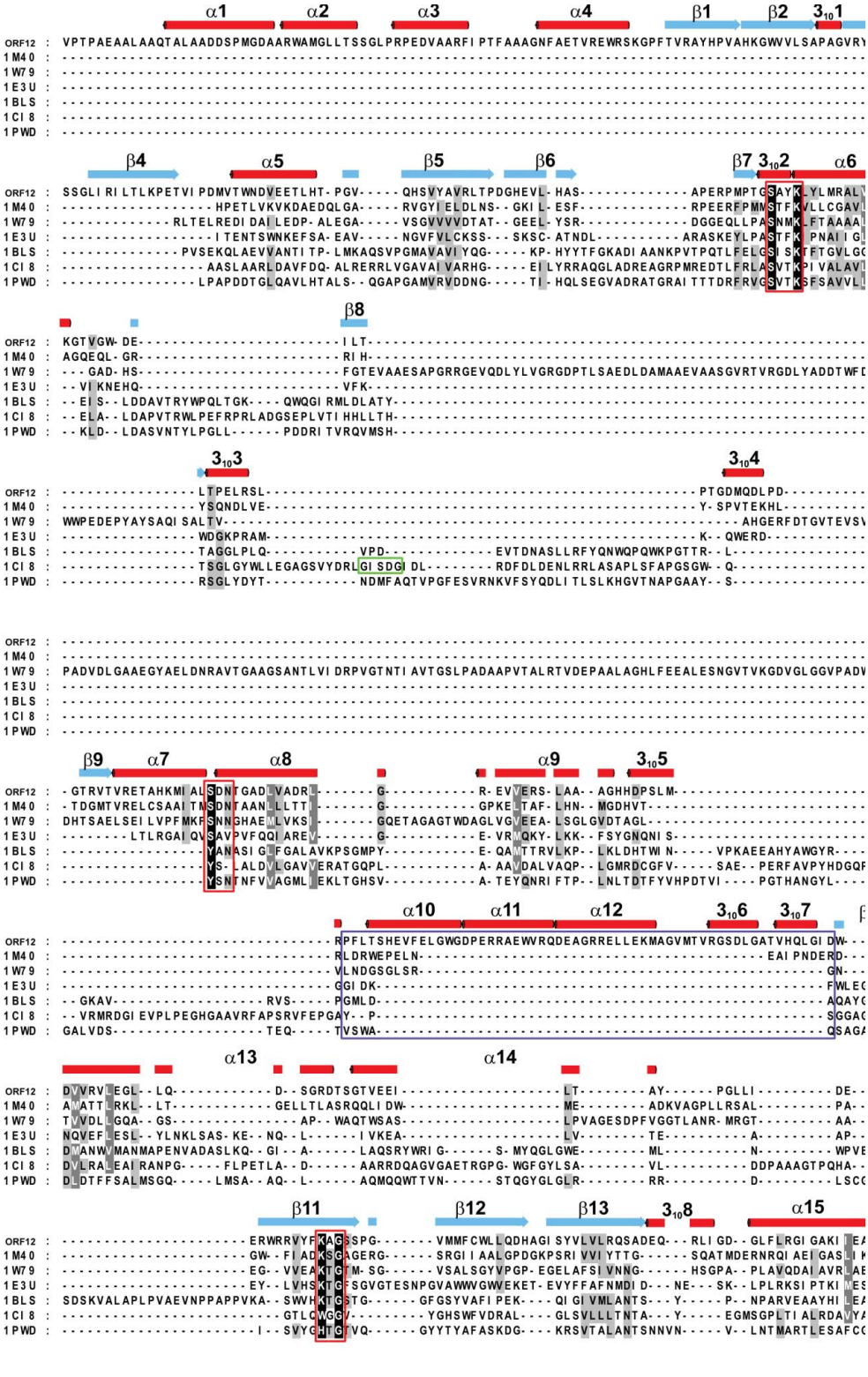

Figure 6
Structure-based sequence alignment of ORF12 with representative members from the β-lactamase/D-Ala carboxypeptidase fold superfamily: PDB entries 1m40 (TEM-1; Minasov *et al.*, 2002), 1w79 (R39; Sauvage *et al.*, 2005), 1e3u (OXA10; Maveyraud *et al.*, 2000), 1bls (P99; Lobkovsky *et al.*, 1994), 1ci8 (EstB; Wagner *et al.*, 2002) and 1pwd (R61; Silvaggi *et al.*, 2005) (see Supplementary Table S1 for details of the catalytic residues of each enzyme). The three motifs S*XX*K, SDN and KTG are indicated by red boxes, the functional Ω loop found only in class A β-lactamases is indicated by a purple box and the proposed G*X*S*X*S esterase motif of EstB is indicated by a green box. The secondary structure of ORF12 is indicated above the sequence with helices in red and strands in light blue.

apparent in ORF12 (Ser173/Ser234/Lys176/Lys375; Figs. 5 and 6). The class A β-lactamases are distinguished from the class C and class D β-lactamases by the presence of an Ω loop, which contains a strictly conserved residue, Glu166, believed to be near essential for orientating/deprotonating a water molecule in both the acylation and deacylation steps of the reaction (Lobkovsky *et al.*, 1994). Unlike the class A β-lactamases, however, ORF12 has a large insert in its Ω-loop region with no apparent Glu166 surrogate (Figs. 6 and 7). Class C and D β-lactamases and PBP D-Ala carboxypeptidases are believed to use the Ser2 (or Tyr2) and/or Lys1 (or His1) residues of the catalytic tetrad, which are also present in ORF12 (Ser234 and Lys176), in roles as the general base in deacylation in place of the Ω-loop Glu166 found in the class A β-lactamases (Figs. 5 and 6; Herzberg & Moult, 1987; Lobkovsky *et al.*, 1994).

The similarity of the ORF12 C-terminal domain to the class A β-lactamases (*e.g.* TEM-1) and the low-molecular-weight class C PBPs (*e.g.* R39) is notable in part because ORF12 has low-level esterase activity (see below). The family VIII esterase EstB from *B. gladioli* has closer structural homology to the class C β-lactamases (*e.g.* P99) and the low-molecular-weight class B PBPs (*e.g.* R61; Wagner *et al.*, 2002; Kelly *et al.*, 1986; Lobkovsky *et al.*, 1994). Other esterase families not belonging to the family VIII esterases and without a β-lactamase-type fold have a conserved G*X*S*X*G motif containing the nucleophilic serine. A G*X*S*X*G motif is present in the N-terminal portion of the EstB sequence containing Ser149 (Fig. 5, green box); however, the Ser149Ala substitution did not eliminate the EstB esterase activity. Instead, Wagner and

coworkers demonstrated that independent substitution of the Ser75 (Ser1) or Tyr181 (Ser2) residues of EstB (equivalent to Ser173 and Ser234 in ORF12, respectively) with alanine residues both result in a complete loss of activity (Wagner *et al.*, 2002). ORF12 also contains a G*X*S*X*G motif including Ser378, the first glycine residue of which is part of the KTG motif found in β-lactamases/D-Ala carboxypeptidases (Galleni *et al.*, 1995; KAG in ORF12; Fig. 6).

To test the potential roles of the different serines in the active site of ORF12, we produced Ser173Ala, Ser234Ala and Ser378Ala variants and assayed them by NMR using wild-type ORF12 for comparison. Wild-type ORF12 catalysed the complete conversion of cephalosporin C to deacetylcephalosporin C in 2 h under the standard conditions (V_{max} = 0.047 m*M* min^{-1}). Consistent with the role of Ser75 as a nucleophile in EstB catalysis, the Ser173Ala variant of ORF12 showed a 100-fold reduction in activity (V_{max} = 0.0004 m*M* min^{-1}). A bovine serum albumin control showed similar kinetics ($V_{max} \simeq$ 0.00037 m*M* min^{-1}), indicating that Ser173Ala was virtually inactive. The Ser234Ala and Ser378Ala variants displayed smaller, but still substantial, reductions in enzyme activity (V_{max} = 0.0039 and 0.001 m*M* min^{-1}, respectively) compared with the wild-type ORF12 and consistent with the loss of activity observed for the EstB Tyr181Ala variant (equivalent to Ser234 in ORF12). Thus, the combined site-directed mutagenesis and structural information suggest that, as for β-lactamases and PBPs, it is the serine (Ser173 in ORF12 or Ser75 in EstB) of the S*XX*K motif (Ser1) that acts as the nucleophilic catalyst. Further detailed kinetic analyses are required to dissect the precise roles of Ser234 and Ser378 in ORF12 catalysis.

Overall, these analyses suggest that the structural fold and active-site architecture of ORF12 shows greater similarity to class A β-lactamases than to the esterase VIII family. It has been proposed that β-lactamases evolved from PBPs (Kelly *et al.*, 1986). A bioinformatic/phylogenetic comparison of the EstB and ORF12 sequences and structures suggests that their esterase activities may have evolved separately from evolutionarily different phylogenetic branches (Supplementary Fig. S8). Although the level of esterase catalysis measured for ORF12 is low and might reflect a 'moonlighting activity', it is also possible that the evolution of the

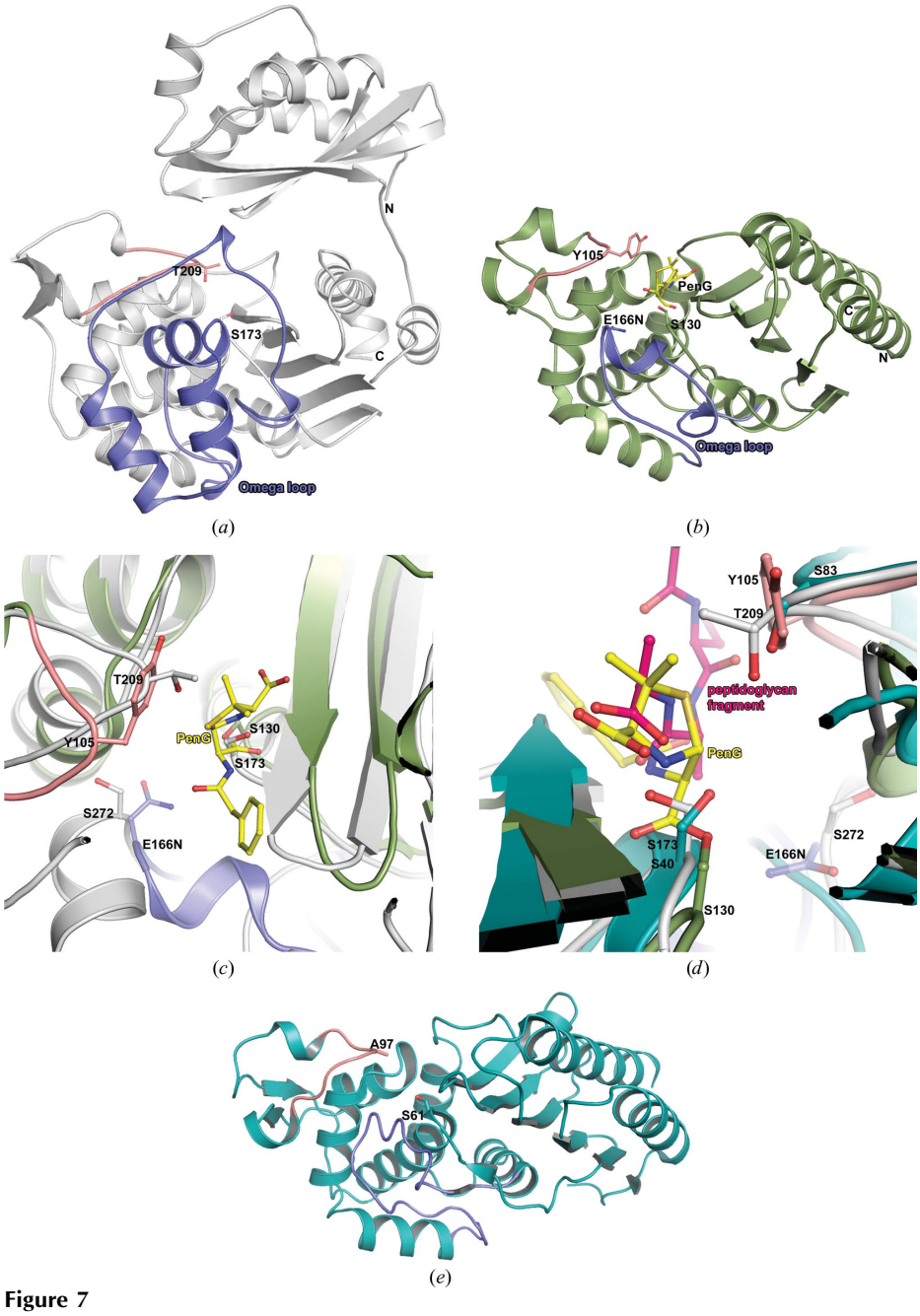

Figure 7
Comparison of TEM1 and PBP-A with ORF12. (*a*) Ribbons representation of ORF12 with its Ω-loop region highlighted in slate blue. Another loop containing Thr209 is positioned to hinder β-lactam binding (salmon). (*b*) A view of the TEM1(E166N)–PenG complex (PDB entry 1fqg; Strynadka *et al.*, 1992) with the Ω loop coloured slate blue and the equivalent Thr209 loop coloured salmon; the orientation is the same as in (*a*). (*c*, *d*) Close-up views of the ORF12 active site (white ribbons) superimposed on the TEM1–PenG complex (green ribbons) showing the position of Ser272, which is closest to E166N (slate blue sticks) of the Ω loop in TEM1, and the position of Thr209, showing its proximity to the covalently bound PenG (yellow sticks) in TEM1. (*e*) Ribbons representation of the PBP-A–PenG complex (cyan; PDB entry 2jbf; Urbach *et al.*, 2009) with the Ω loop highlighted in slate blue and the Thr209 loop in salmon, with the equivalent residue Ala97 shown as sticks. Note that the PBP-A residue Glu158 is in the equivalent position to Glu166 in the Ω loop of TEM1.

esterase activity of both ORF12 and EstB is an example of convergent functional evolution occurring from a related starting scaffold (the β-lactamase/D-Ala peptidase fold) but with a partially different set of catalytically important residues, Ser173/Ser234/Lys176/Lys375 in ORF12 and Ser75/Tyr181/Lys78/Trp348 in EstB, to perform esterase reactions on similar substrates (Supplementary Table S2).

3.2.5. Binding of clavulanic acid. To investigate the possible function of ORF12, we determined the structures of ORF12 in crystals soaked with fragments of varying CA-biosynthesis mimics representing substrates and/or products (including deacetylcephalosporin C, cephalosporin C, deacetoxycephaloglycin, cephamycin, cephaloglycin, *N*-acetyl-glycyl-*N*-acetyl ornithine and *N*-acetyl-glycyl clavaminic acid; Supplementary Table S1). However, the electron density observed in the structures obtained by cocrystallization with these compounds was not sufficient for modelling owing to very low occupancy. Nevertheless, in the case of the crystals soaked with the CA-biosynthetic pathway product clear evidence for CA binding was observed, as described below (Fig. 5).

To obtain ORF12–CA complex structures, crystals of apo ORF12 (wild type or variants) were soaked in a solution containing CA. Data were collected to 1.55 Å resolution for the wild-type ORF12–CA complex (Table 1). The electron-density maps clearly show two CA molecules binding to one ORF12 molecule (Figs. 5*a* and 5*b*). The final model for the wild-type ORF12–CA complex structure was refined to an R factor of 18.7% and an $R_{\rm free}$ of 21.0%. One CA molecule binds to the active site (site 1) and the other molecule binds between the two domains (site 2) (Fig. 5*a*). The binding of CA did not result in any significant domain movements. In binding site 1, CA hydrogen bonds to the side chains of Tyr359 (2.7 Å) from the loop between α14 and β11, Lys89 (2.9 Å) from strand β2 and Arg418 (2.6 Å) from helix α15. CA is positioned in the active-site pocket lined by residues from both the N- and the C-terminal domains (His88, Ser173, Thr209, Ser234, Ser278, Met383, Phe374, Ala376 and Phe385). The CA β-lactam core binds in the pocket with its β-face exposed to solvent and with its α-face packed adjacent to Met383, Phe385 and Ala376. This binding mode is very different from that observed in the complexes of PBPs/β-lactamases with β-lactams (Fig. 5*c*). The C-2 carboxylate of CA is positioned deep in the pocket, making electrostatic interactions with Lys375, Arg418 and Lys89. Lys375 is part of the KTG motif; the equivalent residue of PBPs/β-lactamases has previously been proposed/observed to interact with the analogous carboxylates of the penicillin/cephalosporin core and this interaction may play a role in activating the catalytic tetrad. However, in the ORF12 complex the CA lactam-ring carbonyl O atom is orientated in the opposite direction, positioning the lactam ring away from the nucleophilic serine and the oxyanion hole (~9.0 Å; Figs. 5*c*, 5*d* and 5*e*).

Superimposition of the ORF12–CA complex structure with that of the catalytically inactive mutant (E166N) class A TEM-1 β-lactamase acyl-enzyme complex with penicillin G (PenG; PDB entry 1fqg; Strynadka *et al.*, 1992) and that of PBP6 in complex with a fragment of peptidoglycan substrate (PDB entry 3itb; Chen *et al.*, 2009) provides a possible explanation as to why ORF12 does not show β-lactamase activity (at least with the tested substrates; see below; Fig. 5*c*). In the ORF12–CA complex the carbonyl O atom of the CA β-lactam ring is positioned ~9.0 Å away from the water molecule occupying the analogous 'oxyanion hole' formed by the backbone amides of Ser173 and Ser378 (Fig. 5*e*). In order for β-lactamases to hydrolyse β-lactams, the β-lactam carbonyl O atom must first be positioned in the oxyanion hole (the geometry of this interaction is likely to be essential for productive catalysis), which stabilizes the partial charge formed on the O atom during nucleophilic attack of the carbonyl C atom (Fig. 5*c*). In addition, the side chain of Thr209 of ORF12 is positioned where the β-lactam core typically binds in the PBP/β-lactamase enzymes and where the terminal D-Ala from the peptidoglycan substrate binds in PBP6, thereby potentially sterically hindering the appropriate positioning of the β-lactam carbonyl O atom or the penultimate D-Ala carbonyl O atom of the peptidoglycan substrate in the oxyanion hole of ORF12 (Figs. 5*c* and 7*d*). Therefore, the structure of the ORF12–CA complex helps to rationalize why ORF12 is unable to catalyse β-lactam hydrolysis or peptidoglycan carboxypeptidation/transpeptidation.

In CA-binding site 2 there are no direct hydrogen bonds that stabilize CA binding to ORF12. The CA molecule binds in a mostly hydrophobic cleft between residues His84, Trp91, Leu362, Leu415, Arg418 and Ala422. Two of these residues, Trp91 and Arg418, are conserved in the two closest ORF12 subfamily members and may act as 'gatekeeper' residues for access to the active site (Fig. 5*b*).

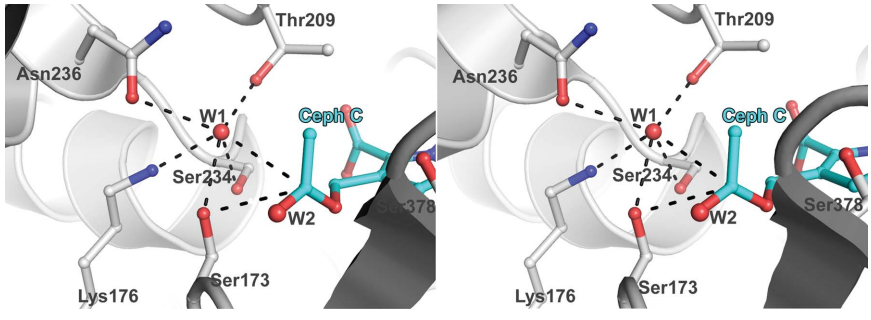

Figure 8
Stereoview image of ORF12 with cephalosporin C (Ceph C) docked into the active site. As with the water molecule observed in class A β-lactamases stabilized by Glu166 of the Ω loop (Lobkovsky *et al.*, 1994; Zawadzke *et al.*, 1996), an analogous water molecule (W1) is observed in the ORF12–CA complex structure at the about the correct position for a role in deacylation (Fig. 5*d*). This water molecule is absent in both the apo wild-type and the CA-bound catalytic mutants of ORF12, supporting a catalytic role, assuming that bound CA represents bound reaction product. Note that the side chain of Thr209 apparently sterically hinders the bicyclic β-lactam core from binding in a mode productive for β-lactamase activity, rationalizing why ORF12 has cephalosporin esterase but not β-lactamase activity.

3.3. Mechanistic considerations

Modelling cephalosporin C into the active site of ORF12 based on the CA complex indicates favourable positioning of the ester carbonyl O atom into the oxyanion hole formed by the backbone amide N atoms of Ser173 and Ser378 (occupied by a water molecule in all ORF12 structures; Figs. 5*e* and 8), and the O atom of the Ser173 side chain is in an apparently stereoelectronically favoured position for nucleophilic attack of the carbonyl C atom of the 3′-*O*-acetyl cephalosporin ester. Stabilization of the tetrahedral intermediate and/or the positioning of a hydrolytic water molecule may be partially enabled by the side chain of Ser378, which would explain the reduction in activity by the Ser378Ala mutant (Fig. 8). Other residues that may be involved in stabilization of the tetrahedral intermediate and/or the positioning of a hydrolytic water molecule include Thr209 and Asn236 (Fig. 8). The subsequent deacylation reaction step may potentially be aided by a hydrogen-bonding network through the side chains of Lys375, Asn236 and/or Ser234, which would explain the partial loss of activity in the Ser234Ala mutant owing to its potential role in proton shuttling between Ser234 and Lys375 (Figs. 5*d* and 5*e*).

4. Conclusions

Overall, our results reveal that ORF12 is functionally similar to the esterase VIII superfamily in that it possesses *O*-acetyl esterase activity towards 3′-*O*-acetyl cephalosporin derivatives, albeit at a low level. The catalytic esterase activity of ORF12 resides in the C-terminal domain. ORF12 did not display β-lactamase activity with the tested β-lactams, suggesting that it is likely that it does not act as a β-lactamase despite showing high structural similarities to class A β-lactamases; it also does not display DD-carboxypeptidase activity as in PBPs. We cannot rule out the possibility that the functionally relevant catalytic activity requires post-translational modification, including (but not limited to) proteolysis (*i.e.* separation of the N- and C-terminal domains). In addition to CA and clavams, *S. clavuligerus* produces a number of other β-lactam metabolites, including penicillin N, cephamycin C and deacetylcephalosporin C. The essential involvement of ORF12 in cephamycin C production can be ruled out on the basis that the gene cluster has been characterized and all of the enzymes that are required for its biosynthesis have been identified (Alexander & Jensen, 1998). Gene-disruption studies reveal that the deletion of *orf12* has no effect on the levels of penicillin, cephalosporin and cephamycin antibiotics produced by *S. clavuligerus*, but that it is essential for CA production, indicating that ORF12 is specifically required for CA biosynthesis but not for cephalosporin biosynthesis (Li *et al.*, 2000). Therefore, we believe it to be likely that the activity of ORF12 towards 3′-*O*-acetyl cephalosporins is not an essential function but may reflect promiscuity based on substrate similarity.

It is possible that ORF12 is involved in the requisite epimerization step during the conversion of clavaminic acid to CA. The structure of the ORF12–CA complex may in part reflect an enzyme–product (analogue) complex (Supplementary Fig. S9). In this regard it is notable that the fold of the N-terminal domain of ORF12, which has not previously been identified, is related to the structures of PKTC and steroid isomerase enzymes. One possibility is that ORF12 reacts with an oxidized clavulanic acid derivative at the side-chain allylic position to form a covalently linked intermediate that tethers the bicyclic β-lactam to ORF12, thus enabling epimerization, perhaps as shown in Supplementary Fig. S9. This could be aided by residues lining the β-lactam core-binding region, including the residues Lys89 and/or Arg418. Further studies on the potential functional residues described here will be required to give a more complete description of the enzymatic mechanism that ORF12 uses in the biosynthesis of CA. Given the observation that ORF12 has the potential to bind to two molecules of CA, we also cannot rule out the possibility that ORF12 plays a role in β-lactam transport and/or in self-resistance.

We thank the ESRF/EMBL, Grenoble for providing beam time and data-collection facilities. This work was supported by grants from the Swedish Natural Science Research Council (VR), the Biotechnology and Biological Sciences Research Council (UK), MRC research grants G500643, G0600801 and G1001023 and Wellcome trust equipment grants 071998 and 068598. AJL is a Science City Interdisciplinary Research Alliance Research Fellow supported by the Birmingham–Warwick Science City Translational Medicine Initiative.

References

Abrahams, K. A. (2011). Thesis. Warwick University, England.
Adam, M., Damblon, C., Jamin, M., Zorzi, W., Dusart, V., Galleni, M., el Kharroubi, A., Piras, G., Spratt, B. G., Keck, W., Coyette, J., Ghuysen, J.-M., Nguyen-Distèche, M. & Frère, J.-M. (1991). *Biochem. J.* **279**, 601–604.
Adam, M., Damblon, C., Plaitin, B., Christiaens, L. & Frère, J.-M. (1990). *Biochem. J.* **270**, 525–529.
Adams, P. D. *et al.* (2010). *Acta Cryst.* D**66**, 213–221.
Alexander, D. C. & Jensen, S. E. (1998). *J. Bacteriol.* **180**, 4068–4079.
Asano, Y., Mori, T., Hanamoto, S., Kato, Y. & Nakazawa, A. (1989). *Biochem. Biophys. Res. Commun.* **162**, 470–474.
Baggaley, K. H., Brown, A. G. & Schofield, C. J. (1997). *Nat. Prod. Rep.* **14**, 309–333.
Bompard-Gilles, C., Remaut, H., Villeret, V., Prangé, T., Fanuel, L., Delmarcelle, M., Joris, B., Frère, J. & Van Beeumen, J. (2000). *Structure*, **8**, 971–980.
Brown, R. P., Aplin, R. T. & Schofield, C. J. (1996). *Biochemistry*, **35**, 12421–12432.
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.
Chen, Y., Zhang, W., Shi, Q., Hesek, D., Lee, M., Mobashery, S. & Shoichet, B. K. (2009). *J. Am. Chem. Soc.* **131**, 14345–14354.
Clarke, T. B., Kawai, F., Park, S.-Y., Tame, J. R. H., Dowson, C. G. & Roper, D. I. (2009). *Biochemistry*, **48**, 2675–2683.
Contreras-Martel, C., Job, V., Di Guilmi, A. M., Vernet, T., Dideberg, O. & Dessen, A. (2006). *J. Mol. Biol.* **355**, 684–696.
Diederichs, K. & Karplus, P. A. (1997). *Nature Struct. Biol.* **4**, 269–275.
Egan, L. A., Busby, R. W., Iwata-Reuyl, D. & Townsend, C. A. (1997). *J. Am. Chem. Soc.* **119**, 2348–2355.

Fonzé, E., Vermeire, M., Nguyen-Distèche, M., Brasseur, R. & Charlier, P. (1999). *J. Biol. Chem.* **274**, 21853–21860.
Fuente, A. de la, Lorenzana, L. M., Martín, J. F. & Liras, P. (2002). *J. Bacteriol.* **184**, 6559–6565.
Galleni, M., Lamotte-Brasseur, J., Raquet, X., Dubus, A., Monnaie, D., Knox, J. R. & Frère, J.-M. (1995). *Biochem. Pharmacol.* **49**, 1171–1178.
Ghuysen, J.-M. (1991). *Annu. Rev. Microbiol.* **45**, 37–67.
Herzberg, O. & Moult, J. (1987). *Science*, **236**, 694–701.
Holm, L., Kääriäinen, S., Rosenström, P. & Schenkel, A. (2008). *Bioinformatics*, **24**, 2780–2781.
Holm, L. & Rosenström, P. (2010). *Nucleic Acids Res.* **38**, W545–W549.
Hutchinson, E. G. & Thornton, J. M. (1990). *Proteins*, **8**, 203–212.
Hutchinson, E. G. & Thornton, J. M. (1996). *Protein Sci.* **5**, 212–220.
Jelsch, C., Lenfant, F., Masson, J. M. & Samama, J.-P. (1992). *FEBS Lett.* **299**, 135–142.
Jensen, S. E., Elder, K. J., Aidoo, K. A. & Paradkar, A. S. (2000). *Antimicrob. Agents Chemother.* **44**, 720–726.
Jensen, S. E., Paradkar, A. S., Mosher, R. H., Anders, C., Beatty, P. H., Brumlik, M. J., Griffin, A. & Barton, B. (2004). *Antimicrob. Agents Chemother.* **48**, 192–202.
Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.
Kallio, P., Sultana, A., Niemi, J., Mäntsälä, P. & Schneider, G. (2006). *J. Mol. Biol.* **357**, 210–220.
Kantola, J., Kunnari, T., Hautala, A., Hakala, J., Ylihonko, K. & Mäntsälä, P. (2000). *Microbiology*, **146**, 155–163.
Kelly, J. A., Dideberg, O., Charlier, P., Wery, J. P., Libert, M., Moews, P. C., Knox, J. R., Duez, C., Fraipont, C., Joris, B., Dusart, J., Frère, J.-M. & Ghuysen, J.-M. (1986). *Science*, **231**, 1429–1431.
Kelly, J. A. & Kuzin, A. P. (1995). *J. Mol. Biol.* **254**, 223–236.
Kershaw, N. J., Caines, M. E., Sleeman, M. C. & Schofield, C. J. (2005). *Chem. Commun.*, pp. 4251–4263.
Koch, A. L. (2000). *Crit. Rev. Microbiol.* **26**, 205–220.
Kuzin, A. P., Nukaga, M., Nukaga, Y., Hujer, A. M., Bonomo, R. A. & Knox, J. R. (1999). *Biochemistry*, **38**, 5720–5727.
La Fortelle, E. de & Bricogne, G. (1997). *Methods Enzymol.* **276**, 472–494.
Lamzin, V. S., Perrakis, A. & Wilson, K. S. (2001). *International Tables for Crystallography*, Vol. *F*, edited by M. G. Rossman & E. Arnold, pp. 720–722. Dordrecht: Kluwer Academic Publishers.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Li, R., Khaleeli, N. & Townsend, C. A. (2000). *J. Bacteriol.* **182**, 4087–4095.
Lim, D. & Strynadka, N. C. (2002). *Nature Struct. Biol.* **9**, 870–876.
Llinás, A., Ahmed, N., Cordaro, M., Laws, A. P., Frère, J.-M., Delmarcelle, M., Silvaggi, N. R., Kelly, J. A. & Page, M. I. (2005). *Biochemistry*, **44**, 7738–7746.
Lobkovsky, E., Billings, E. M., Moews, P. C., Rahil, J., Pratt, R. F. & Knox, J. R. (1994). *Biochemistry*, **33**, 6762–6772.
Massova, I. & Mobashery, S. (1998). *Antimicrob. Agents Chemother.* **42**, 1–17.
Maveyraud, L., Golemi, D., Kotra, L. P., Tranier, S., Vakulenko, S., Mobashery, S. & Samama, J.-P. (2000). *Structure*, **8**, 1289–1298.
Mellado, E., Lorenzana, L. M., Rodriguez-Saiz, M., Diez, B., Liras, P. & Barredo, J. L. (2002). *Microbiology*, **148**, 1427–1438.
Minasov, G., Wang, X. & Shoichet, B. K. (2002). *J. Am. Chem. Soc.* **124**, 5333–5340.
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.
Murzin, A. G., Brenner, S. E., Hubbard, T. & Chothia, C. (1995). *J. Mol. Biol.* **247**, 536–540.
Negoro, S., Ohki, T., Shibata, N., Sasa, K., Hayashi, H., Nakano, H., Yasuhira, K., Kato, D., Takeo, M. & Higuchi, Y. (2007). *J. Mol. Biol.* **370**, 142–156.
Nicola, G., Peddi, S., Stefanova, M., Nicholas, R. A., Gutheil, W. G. & Davies, C. (2005). *Biochemistry* **44**, 8207–8217.
Okazaki, S., Suzuki, A., Komeda, H., Yamaguchi, S., Asano, Y. & Yamane, T. (2007). *J. Mol. Biol.* **368**, 79–91.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Pape, T. & Schneider, T. R. (2004). *J. Appl. Cryst.* **37**, 843–844.
Petersen, E. I., Valinger, G., Sölkner, B., Stubenrauch, G. & Schwab, H. (2001). *J. Biotechnol.* **89**, 11–25.
Pratt, R. F. & McLeish, M. J. (2010). *Biochemistry*, **49**, 9688–9697.
Sauvage, E., Herman, R., Petrella, S., Duez, C., Bouillenne, F., Frère, J.-M. & Charlier, P. (2005). *J. Biol. Chem.* **280**, 31249–31256.
Sauvage, E., Kerff, F., Terrak, M., Ayala, J. A. & Charlier, P. (2008). *FEMS Microbiol. Rev.* **32**, 234–258.
Silvaggi, N. R., Josephine, H. R., Kuzin, A. P., Nagarajan, R., Pratt, R. F. & Kelly, J. A. (2005). *J. Mol. Biol.* **345**, 521–533.
Strynadka, N. C., Adachi, H., Jensen, S. E., Johns, K., Sielecki, A., Betzel, C., Sutoh, K. & James, M. N. G. (1992). *Nature (London)*, **359**, 700–705.
Sultana, A., Kallio, P., Jansson, A., Wang, J.-S., Niemi, J., Mäntsälä, P. & Schneider, G. (2004). *EMBO J.* **23**, 1911–1921.
Tahlan, K., Park, H. U., Wong, A., Beatty, P. H. & Jensen, S. E. (2004). *Antimicrob. Agents Chemother.* **48**, 930–939.
Torkkell, S., Kunnari, T., Palmu, K., Hakala, J., Mäntsälä, P. & Ylihonko, K. (2000). *Antimicrob. Agents Chemother.* **44**, 396–399.
Urbach, C., Evrard, C., Pudzaitis, V., Fastrez, J., Soumillion, P. & Declercq, J.-P. (2009). *J. Mol. Biol.* **386**, 109–120.
Urbach, C., Fastrez, J. & Soumillion, P. (2008). *J. Biol. Chem.* **283**, 32516–32526.
Wagner, U. G., Petersen, E. I., Schwab, H. & Kratky, C. (2002). *Protein Sci.* **11**, 467–478.
Wilkin, J. M., Jamin, M., Damblon, C., Zhao, G.-H., Joris, B., Duez, C. & Frère, J.-M. (1993). *Biochem. J.* **291**, 537–544.
Zawadzke, L. E., Chen, C. C. H., Banerjee, S., Li, Z., Wäsch, S., Kapadia, G., Moult, J. & Herzberg, O. (1996). *Biochemistry*, **35**, 16475–16482.
Zhao, G., Meier, T. I., Kahl, S. D., Gee, K. R. & Blaszczak, L. C. (1999). *Antimicrob. Agents Chemother.* **43**, 1124–1128.