The crystal structure and oligomeric form of *Escherichia coli* L,D-carboxypeptidase A

Karen Meyer,[1] Christine Addy,[1] Satoko Akashi,[1] David I. Roper[2] and Jeremy R. H. Tame[1*]

[1]Graduate School of Medical Life Science
Yokohama City University
Suehiro 1-7-29, Tsurumi, Yokohama 230-0045, Japan

[2]School of Life Sciences
University of Warwick
Gibbet Hill Road
Coventry, CV4 7AL, UK

[*]Corresponding author.
email: jtame@yokohama-cu.ac.jp
tel: +81(0)45 5087228

Abstract

Bacterial peptidoglycan is constructed by cross-linking sugar chains carrying pentapeptide building blocks with two D-alanine residues at the C-terminus. Incorporation into the polymer and subsequent breakdown of peptidoglycan releases a tetrapeptide with a single D-alanine residue. Removal of this residue is necessary for the tripeptide to receive a new D-Ala-D-Ala dipeptide in the synthetic pathway, but proteases are generally unable to work with substrates having residues of unusual chirality close to the scissile bond. Processing of the tetrapeptide is carried out by a dedicated LD-carboxypeptidase, which is of interest as a novel drug target. We describe the high resolution crystal structure of the enzyme from *E. coli*, and demonstrate the dimeric structure is highly conserved.

Keywords

antibiotic ; crystal structure ; recycle pathway ; analytical ultracentrifugation

Introduction

Both Gram positive and Gram negative bacteria possess a cell wall that includes a protective layer of peptidoglycan (PG), built from sugar chains cross-linked via short peptides. Synthesis involves adding fresh units consisting of a disaccharide with an attached pentapeptide that generally terminates with two D-alanine residues. In Gram negative species the peptide has the sequence L-Ala-γ-D-Glu-m-A2pm-D-Ala-D-Ala, where m-A2pm stands for meso-diamopimelic acid [1, 2]. Penicillin-binding proteins (PBPs) [3] are a large group of proteins responsible for the transpeptidation reaction that cross-links the peptides, and also include non-essential enzymes that trim the peptides, and so help to shape the cell and also recycle PG components [4, 5]. The importance of PG to bacteria is underlined by the fact that numerous antibiotics, including the penicillins, function by blocking key enzymes involved in its synthesis. Penicillin and its analogues mimic the D-alanine residues to make suicide substrates [6].

PG is continually reworked as the cell grows and divides [7-9], so that during exponential growth bacterial cells release a considerable quantity of PG break-down products into the growth medium. Some of these cell-wall fragments function as messengers, for example as triggers of immune responses; the mammalian protein NOD1 detects specific PG fragments [10, 11], initiating inflammatory responses. Hydrolases are found that can break each of the glycosidic or amide bonds in PG, and many appear to have overlapping roles [3, 12]. This creates a functionally redundant system that is robust to mutation or chemical inhibition, and allows the cell to re-use much of the material in fresh PG synthesis. Lytic transglycosylases and endopeptidases release anhydromuropeptides that are transported into the cell by the AmpG permease before further degradation through the action of N-acetylglucosaminidase (NagZ), N-acetylmuramyl-L-alanine amidase (AmpD) and L,D-carboxypeptidase (LdcA) [12]. LdcA uniquely cleaves the peptide bond between the third and fourth residues of the tetrapeptide, releasing the C-terminal D-alanine [13].

It is estimated that about half of the PG breakdown products are recaptured in this way, to offset the very considerable metabolic cost of PG synthesis [14]. Growth in rich media affords bacteria the luxury of building PG entirely *de novo*, so that loss of recycling pathways does not impact the speed of cell multiplication. Both Gram positive and negative cells however are found to lyse in stationary phase if recycling is blocked [15, 16]. This discovery has prompted efforts to develop specific inhibitors of LdcA for use as anti-bacterial agents with

an entirely novel mechanism of action, and high-throughput screening conducted some years ago produced a lead candidate [17]. Since then it has become clear that PG recycling plays an important role in triggering the expression of beta-lactamase in pathogenic *Pseudomonas*, and further that strong impairment of PG recycling may severely weaken both fitness and virulence [18, 19]. Given the interest in LdcA as a possible novel drug target, different groups have determined crystal structures of the protein from *Pseudomonas aeruginosa* [20] and *Novosphingobium aromaticuvorans* [21]. These models show the protein is a member of the Peptidase_S66 family (Pfam family PF02016). This family includes the self-immunity protein MccF, whose structure has also been solved by X-ray crystallography [22]. The oligomeric state of *Na*LdcA appears to be dimeric from the crystal structure, but size-exclusion chromotography indicated a monomer [21]. *E. coli* LdcA shows significant sequence differences from the known structures, and includes sequence insertions that might be involved in formation of homo- or hetero-oligomer formation (Supplementary Figure 1). To clarify these issues we have solved the crystal structure of *Ec*LdcA and a mutant at the active site.

Materials and Methods

Cloning

The wild-type *E.coli ldcA* gene was amplified from genomic DNA by PCR, and cloned into pET-28b expression vector between the restriction sites NdeI and BamH1 to allow expression of recombinant protein with a hexa-histidine tag at the N-terminus, cleavable with thrombin. The PCR product before and after restriction enzyme treatment was purified using Nippon Genetics Purification Kit. The insert DNA was ligated into the cut vector using T4 DNA ligase at 16°C for 3 hours before transformation for plasmid preparation. To introduce the site-directed mutation S106A using PCR the following primers were used: GCTCATTTGCGGACATGCGGATTTTACCGCCATTC and GAATGGCGGTAAAATCCGCATGTCCGCAAATGAGC. The expression vector for the mutant protein was produced in the same way as the wild-type.

Expression and purification

The pET-28 plasmid with *ldcA* was transformed into *E. coli* BL-21(DE21) cells and the cells were grown at 37°C with shaking in 1 L LB medium containing kanamycin (20 μ g ml^{-1}). When the OD_{600} of the culture reached 0.6, protein expression was induced by adding IPTG to a final concentration of 0.5 mM and growth was continued overnight at 20°C. The cells were

collected by centrifugation at 4600 *g* at 4°C for 20 min. The pellet was suspended in 10 ml 50 mM Tris-HCl pH 8.5 and 50 mM NaCl per 1 L culture and then lysed by sonication on ice. The lysate was centrifuged at 48000 *g* at 4°C for 30 min. The supernatant was micro-filtered before loading onto a 15 ml nickel Sepharose column equilibrated with 50 mM Tris-HCl pH 8.5, 50 mM NaCl and 10 mM imidazole. After washing, the protein was eluted with 50 mM Tris-HCl pH 8.5, 50 mM NaCl and 250 mM imidazole. After dialysis the protein was loaded onto a HiTrap Q HP column equilibrated with 50 mM Tris-HCl pH 8.5, and eluted with a salt gradient to 1 M NaCl. The pooled fractions containing LdcA were concentrated and loaded onto a gel filtration column equilibrated with 50 mM Tris-HCl pH 8.5 and 50 mM NaCl. The eluted protein was concentrated to 12 mg ml^{-1} using Amicon centrifugal filter units. The same protocol was used to express and purify the mutant LdcA.

Crystallisation and data collection

Purified protein, both native and mutant, was crystallised by vapor diffusion using the hanging-drop method. The reservoir solution (12% (w/v) PEG 3350, 0.2 M sodium thiocyanate, 20 mM TrisHCl pH 8.5) was mixed in a 1:1 ratio with the protein, and the drop then equilibrated against 0.5 mL of reservoir solution at 20°C. A separate crystal form (used for heavy atom soaks) was grown using 85 mM HEPES pH 7.5, 3.66 M sodium chloride and 15% glycerol. Diffraction data used for the final refinements were collected from crystals held at 100K at BL17A of the Photon Factory, Tsukuba using a PILATUS-2M detector. Data processed with XDS [23]. Initial structure solution was performed using PHENIX [24], and data manipulation was performed with the CCP4 suite [25]. Refinement was carried out with REFMAC [26] and COOT [27]. All data collection and refinement statistics are shown in Supplementary Table 1. The structures and X-ray data were deposited in the Protein DataBank and assigned the identifiers PDB: 5Z01 (native) and PDB: 5Z03 (mutant).

Results

Native crystal structure

The cloning, expression and purification of native LdcA were carried out as described in the Materials and Methods section. The protein crystallised in space group *I*222 and $P32_12$. The former crystals, grown in PEG 3350, diffracted to 1.75 Å at beam-line BL17A of the Photon Factory, but the trigonal crystals (grown in sodium chloride) gave weaker data to a

maximum resolution of only 2.54 Å. The trigonal crystals however proved robust to soaking with ethyl-mercury-p-toluenesulphanilide, allowing SAD phasing with a single dataset collected at BL5A using a wavelength of 1.00659 Å. Five mercury sites could be identified and reliably refined using the Autosol [28] procedure of PHENIX [24]. Automatic model building produced a largely connected structure of 248 residues with R-factor/Rfree of 33.26% and 38.67%. PHASER [29] was used to place this search model in the space group and cell of the native data, giving rotation and translation scores (RFZ and TFZ) of 12.5 and 15.7 respectively. Manual improvements were carried out using COOT [27], giving a model with a single continuous chain of 305 residues and 212 water molecules. The data collection and final refinement statistics are given in Supplementary Table 1. The N-terminal histidine tag proved difficult to remove with thrombin and so this step was dropped from the purification protocol, but only a few residues of the tag are visible in the electron density. The Ramachandran plot shows Ser 10 and Ser 106 have unusual backbone angles; the latter residue is the catalytic serine, but even this is not considered an outlier by Molprobity [30]. The ϕ/ψ angles of 68° and 129° are energetically unfavourable, but typical for serine proteases. Some side-chains show unusual rotamers, which can mostly be explained by the fact these are surface residues with no strong side-chain interactions. One exception is Arg 41, which makes strong hydrogen bonds with Asp 56, both residues being very well defined in the electron density.

*Ec*LdcA has a mixed α/β structure with two β sheets, one highly twisted and surface exposed, and the other being sandwiched between α helices (Figure 1). A search for similar models using the monomer model and DALI [31] gave a list of related structures, with the most strongly related (PDB: 3TLB) having a Z-score of 35.9 and 26% sequence identity. This model is of the self-immunity protein MccF, a serine carboxypeptidase that hydrolyzes a natural antimicrobial compound called microcin C7 [22]. The catalytic triads of both enzymes overlay closely, although the active site serine of the MccF model (PDB: 3TLB) has been replaced with alanine so that a substrate peptide can bind without undergoing hydrolysis. Overlaying *Ec*LdcA with MccF using the automated procedure SSM [32] shows that 256 core residues of the two models may be aligned with an rmsd of 1.76 Å. The similarity of the two models is highlighted in Supplementary Figure 2. The principal difference between them is an insertion from residue 174 to 202 in MccF (equivalent to residues 158 and 160 in *Ec*LcdA), which creates a long loop stretching towards the active site. Trp 186 of MccF provides a contact with the substrate, the adenine ring of C7 stacking against it. The peptide substrates of LdcA homologues have no

such feature, and the presence or absence of this loop appears to be related to the substrate specificity. Overall the chain trace of the *Ec*LdcA and MccF models is highly similar, reflecting the high level of sequence identity (29.8%). The *Ec*LdcA model shows greater similarity to this structure than *Pa*LdcA or *Na*LdcA because of more complete secondary structure. The second β sheet is largely absent from the *Na*LdcA model (PDB: 3G23 [21]), and the *Pa*LdcA model (PDB: 1ZRS [20]) terminates in the centre of the final strand, whereas the *Ec*LdcA model described here extends for a further 7 residues. The *Pa*LdcA model in contrast is disordered around the loop preceding helix α5. Notably all the LdcA models and MccF share the same dimer structure (Figure 2). For the *Ec*LdcA model, PISA [33] reports a dimer interface that buries a surface area of roughly 1700 $Å^2$. Overall 48 residues contact the partner chain, the dimer interactions including four salt-bridges (Arg 81-Glu 210, Arg 85-Asp 257) with equivalents in *Na*LdcA but not *Pa*LdcA.

Structure of the S106A mutant

The active site serine (Ser 106) of *Ec*LdcA is found within a highly conserved motif (Supplementary Figure 1). The site-directed mutant of *Ec*LdcA missing the catalytic serine was created in order to help produce crystal structures of a substrate complex. Unexpectedly, loss of the side-chain hydroxyl group altered the crystallization behavior of the protein, and the same crystallization conditions were found to produce crystals in an orthogonal space-group, with some indications of pseudo-translation in the Patterson map. Molecular replacement permitted a clear solution in only $P2_12_12$. The new crystal form diffracts to the same resolution as the native, and contains a dimer in the asymmetric unit, with 283 or 284 ordered residues per subunit. Residues 12-13, 43-48, 235-243 and 304-305 are missing from each subunit, with some other residues contiguous with these gaps also absent from one or other protein chain. Some differences may be due to crystal contacts (residues 240 and 244 making crystal contacts in the native model). The active site itself shows no marked deviations directly due to the mutation, and the missing residues are all at least 10 Å from the catalytic serine. Both Gly 75 and Gly 76 are highly ordered in the native structure but seem to be rather more flexible in the mutant, with Gly 75 undergoing a peptide flip (Supplementary Figure 3). These residues lie opposite the catalytic serine, so that the carbonyl of Gly 75 in the mutant model would clash with a bound substrate. On the opposite side of the substrate pocket, Ile 202 adopts a rather different position in the two models, probably as a result of flexibility in the 235-243 loop of the mutant. In the native model it lies against His 270 of the catalytic triad, and probably holds

it in place throughout catalysis. Movement of this residue provides a possible catalytic control mechanism by partner proteins engaging with the 235-243 loop at the surface. There is no evidence of anything in the active site of the mutant apart from weakly held water molecules, in contrast to the native protein, as described below.

Active site

Comparing the active sites of different models it can be seen that the active site of the LdcA homologues is highly conserved. *Na*LdcA (PDB: 3G23) shows the same peptide flip at Gly 75 of the *Ec*LdcA mutant. Ala 72 and Phe 108 of *Ec*LdcA are found to be swapped in *Na*LdcA. All the LdcA models show one conserved arginine residue from each subunit (equivalent to Arg 211 of *Ec*LdcA) extending towards the active site of the partner chain (Supplementary Figure 4). This feature is also found in MccF. The active site of native *Ec*LdcA is found to hold a significant volume of electron density which cannot easily be reconciled with any solvent molecules present. The maximum peak of difference density reaches 11 σ, although the 2mFo-DFc map reaches only just over 3 σ, and has a ring-like shape (Supplementary Figure 5). For comparison, the map for the mutant is shown in Supplementary Figure 6. The position but not the shape overlaps with the ligand in the MccF structure. The only apparent candidate molecules to fill the active site of the crystal structure are water molecules, but these have been left unmodelled due to the unusual nature of the density.

In order to observe any bound ligand, several attempts were made using cell-wall derived peptides to detect a complex by electro-nanospray ionisation mass spectrometry (ESI). These data indicated a covalent modification of the *Ec*LcdA, with a mass increment of 178 Da over the expected value (Figure 3). However, since the histidine tag was not removed, this covalent adduct can be ascribed to a known gluconoylation reaction of the tag, and not the active site of the protein [34]. The only commercially available cell wall peptide is the pentapeptide of the kind found in Gram positive bacteria, with a lysine residue at the third position. Since the tetrapeptide is the true substrate, and activity on the pentapeptide is very low at best, PBP4 (DacB) from *E. coli* was purified as described previously [35], and used to remove the C-terminal residue. Neither mass spectrometry nor crystal soaking yielded any evidence of binding however, indicating that the third residue may provide important contacts. Antibiotics that operate through inhibition of LdcA may therefore prove to be highly specific for Gram positive or negative species.

Molecular weight

The ESI data indicated that *Ec*LdcA forms a stable dimer under low salt conditions, which is consistent with the large interfacial area seen in both the mutant and wild-type crystal structures. Some uncertainty arose with the publication of the *Na*LdcA crystal structure, since the size-exclusion data of Das and colleagues [21] suggested the protein to be a monomer in solution. To confirm that the protein is indeed dimeric we performed analytical ultracentrifugation sedimentation velocity experiments with *Ec*LcdA in 100 mM salt to reproduce the conditions of the gel filtration experiment (Figure 4). The analysis shows no evidence for any form other than the dimer, and it seems most likely that the published size-exclusion experiment is in error, probably as a result of the distinctly non-spherical shape of the protein.

Discussion

Despite the overall conservation of the fold and sequence, variations clearly exist among the forms of LdcA found in different microbial species, and the oligomeric form of the protein remained uncertain. The crystallographic and other biophysical studies reported here with *E. coli* LdcA provide strong grounds for believing the protein is universally a dimer. The fact that LdcA is not invariably essential, and the cytoplasmic location of the enzyme in most bacteria, detracts from its promise as a drug target, but we hope that the model and other data presented here assist efforts to find clinically useful inhibitors. The recent discovery that gonococcal LdcA may have an outer membrane lipoprotein form suggests that *Neisseria* could benefit from inflammatory responses mediated by the peptidoglycan-sensing protein NOD1, possibly by allowing the bacteria to manipulate apoptosis and autophagy [36]. The flexibility of the enzyme around the active site was not expected, and the different crystallisation behaviour of the native protein and the mutant with alanine in place of the active site serine is unprecedented to the best of our knowledge. This flexibility probably contributes to the fact that structural analysis of a substrate complex of the protein has yet to be reported.

Acknowledgements

We would like to thank the staff at the Photon Factory for assistance with X-ray data collection. This work was performed under the approval of the Photon Factory Program Advisory Committee (Proposal No. 2016G014). JRHT acknowledges financial assistance from OpenEye Scientific. This work did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Figure legends

Figure 1.

(a) A ribbon diagram showing the trace of a single subunit of *Ec*LdcA. Residues conserved among the known LdcA models are shown in red. The catalytic triad at the active site (Ser 106, His 105, Glu 200) is shown as sticks, with carbon atoms shown in yellow, oxygen red and nitrogen blue. (b) The same ribbon diagram coloured from blue (N-terminus) to red (C-terminus).

Figure 2.

(a) A stereo ribbon diagram of native *Ec*LdcA, looking along the dyad axis of the dimer, with the left-hand subunit coloured as a spectrum from N to C-terminus, and the right-hand subunit in light green. (b) The catalytic triad of the left-hand subunit and the conserved Arg 211 of the opposite subunit are shown as sticks. (c) A ribbon diagram of *Na*LdcA, seen in the same orientation as the *Ec*LdcA in (a), with the subunits coloured brown or green. The absence of the second β sheet in both subunits due to disorder is apparent.

Figure 3.

(a) The raw electrospray mass spectrum of denatured *Ec*LdcA in 0.1% trifluroacetic acid/40% acetonitrile, showing the peaks due to the principal species with mass 35600.61 Da and secondary peaks indicating a covalent adduct creating a species with mass 35778.53 Da. (b) Analysis shows only these two species are present in the sample. The larger mass is consistent with spontaneous N-6-phosphogluconoylation of the histidine tag. The mass spectrum was obtained with a SYNAPT G2 HDMS mass spectrometer (Waters).

Figure 4.

A plot of relative species concentration versus molecular weight, determined by sedimentation velocity using a rotor speed of 40,000 rpm. The calculated mass average for native *Ec*LdcA is 69 kDa, with no indication of any species in solution other than the dimer. Data were analysed using SEDFIT [37].

References

[1] W. Vollmer, U. Bertsche, Murein (peptidoglycan) structure, architecture and biosynthesis in *Escherichia coli*, Biochim Biophys Acta, 1778 (2008) 1714-1734.

[2] W. Vollmer, D. Blanot, M.A. de Pedro, Peptidoglycan structure and architecture, FEMS Microbiol Rev, 32 (2008) 149-167.

[3] E. Sauvage, F. Kerff, M. Terrak, J.A. Ayala, P. Charlier, The penicillin-binding proteins: structure and role in peptidoglycan biosynthesis, FEMS Microbiol Rev, 32 (2008) 234-258.

[4] K. Peters, S. Kannan, V.A. Rao, J. Biboy, D. Vollmer, S.W. Erickson, R.J. Lewis, K.D. Young, W. Vollmer, The Redundancy of Peptidoglycan Carboxypeptidases Ensures Robust Cell Shape Maintenance in *Escherichia coli*, MBio, 7 (2016) 00819-00816.

[5] J. van Heijenoort, Peptidoglycan hydrolases of *Escherichia coli*, Microbiol Mol Biol Rev, 75 (2011) 636-663.

[6] D.J. Waxman, J.L. Strominger, Penicillin-binding proteins and the mechanism of action of beta-lactam antibiotics, Annu Rev Biochem, 52 (1983) 825-869.

[7] A.J. Egan, J. Biboy, I. van't Veer, E. Breukink, W. Vollmer, Activities and regulation of peptidoglycan synthases, Philos Trans R Soc Lond B Biol Sci, 370 (2015).

[8] S. Leclercq, A. Derouaux, S. Olatunji, C. Fraipont, A.J. Egan, W. Vollmer, E. Breukink, M. Terrak, Interplay between Penicillin-binding proteins and SEDS proteins promotes bacterial cell wall synthesis, Sci Rep, 7 (2017) 43306.

[9] J.W. Johnson, J.F. Fisher, S. Mobashery, Bacterial cell-wall recycling, Ann N Y Acad Sci, 1277 (2013) 54-75.

[10] M. Chamaillard, M. Hashimoto, Y. Horie, J. Masumoto, S. Qiu, L. Saab, Y. Ogura, A. Kawasaki, K. Fukase, S. Kusumoto, M.A. Valvano, S.J. Foster, T.W. Mak, G. Nunez, N. Inohara, An essential role for NOD1 in host recognition of bacterial peptidoglycan containing diaminopimelic acid, Nat Immunol, 4 (2003) 702-707.

[11] S.E. Girardin, M. Jehanno, D. Mengin-Lecreulx, P.J. Sansonetti, P.M. Alzari, D.J. Philpott, Identification of the critical residues involved in peptidoglycan detection by Nod1, J Biol Chem, 280 (2005) 38648-38656.

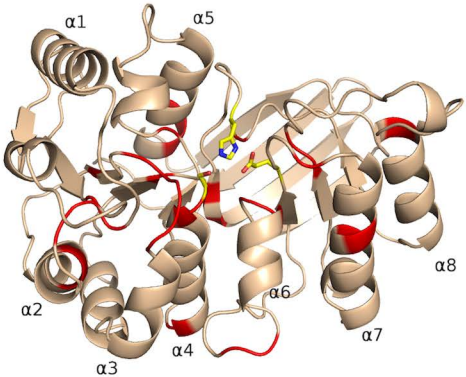
[12] T. Uehara, J.T. Park, Peptidoglycan Recycling, EcoSal Plus, 3 (2008) 5.

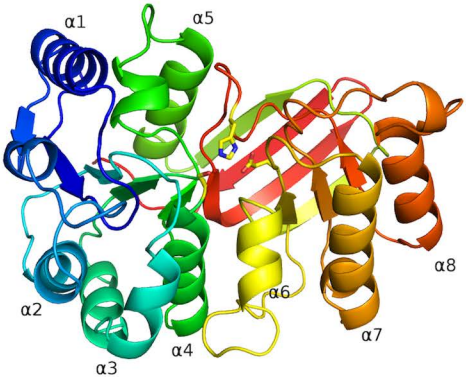
[13] R. Metz, S. Henning, W.P. Hammes, LD-carboxypeptidase activity in Escherichia coli. II. Isolation, purification and characterization of the enzyme from E. coli K 12, Arch Microbiol, 144 (1986) 181-186.

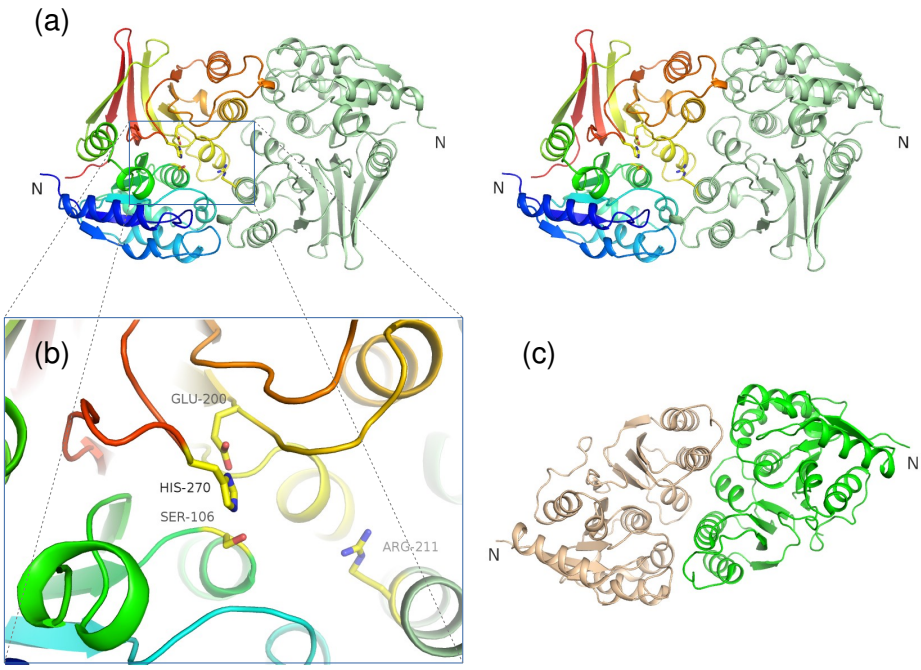
[14] J.T. Park, Why does *Escherichia coli* recycle its cell wall peptides?, Mol Microbiol, 17 (1995) 421-426.
[15] M. Borisova, R. Gaupp, A. Duckworth, A. Schneider, D. Dalugge, M. Muhleck, D. Deubel, S. Unsleber, W. Yu, G. Muth, M. Bischoff, F. Gotz, C. Mayer, Peptidoglycan Recycling in Gram-Positive Bacteria Is Crucial for Survival in Stationary Phase, MBio, 7 (2016) 00923-00916.
[16] M.F. Templin, A. Ursinus, J.V. Holtje, A defect in cell wall recycling triggers autolysis during the stationary growth phase of *Escherichia coli*, Embo J, 18 (1999) 4108-4117.
[17] E.Z. Baum, S.M. Crespo-Carbone, B. Foleno, S. Peng, J.J. Hilliard, D. Abbanat, R. Goldschmidt, K. Bush, Identification of a dithiazoline inhibitor of *Escherichia coli* L,D-carboxypeptidase A, Antimicrob Agents Chemother, 49 (2005) 4500-4507.
[18] M. Perez-Gallego, G. Torrens, J. Castillo-Vera, B. Moya, L. Zamorano, G. Cabot, K. Hultenby, S. Alberti, P. Mellroth, B. Henriques-Normark, S. Normark, A. Oliver, C. Juan, Impact of AmpC Derepression on Fitness and Virulence: the Mechanism or the Pathway?, MBio, 7 (2016) 01783-01716.
[19] G. Torrens, M. Perez-Gallego, B. Moya, M. Munar-Bestard, L. Zamorano, G. Cabot, J. Blazquez, J.A. Ayala, A. Oliver, C. Juan, Targeting the permeability barrier and peptidoglycan recycling pathways to disarm Pseudomonas aeruginosa against the innate immune system, PLoS One, 12 (2017) e0181932.
[20] H.J. Korza, M. Bochtler, *Pseudomonas aeruginosa* LD-carboxypeptidase, a serine peptidase with a Ser-His-Glu triad and a nucleophilic elbow, J Biol Chem, 280 (2005) 40802-40812.
[21] D. Das, M. Herve, M.A. Elsliger, R.U. Kadam, J.C. Grant, H.J. Chiu, M.W. Knuth, H.E. Klock, M.D. Miller, A. Godzik, S.A. Lesley, A.M. Deacon, D. Mengin-Lecreulx, I.A. Wilson, Structure and function of a novel LD-carboxypeptidase a involved in peptidoglycan recycling, J Bacteriol, 195 (2013) 5555-5566.
[22] B. Nocek, A. Tikhonov, G. Babnigg, M. Gu, M. Zhou, K.S. Makarova, G. Vondenhoff, A. Van Aerschot, K. Kwon, W.F. Anderson, K. Severinov, A. Joachimiak, Structural and functional characterization of microcin C resistance peptidase MccF from *Bacillus anthracis*, J Mol Biol, 420 (2012) 366-383.
[23] W. Kabsch, Integration, scaling, space-group assignment and post-refinement, Acta Crystallographica Section D, 66 (2010) 133-144.
[24] P.D. Adams, P.V. Afonine, G. Bunkoczi, V.B. Chen, I.W. Davis, N. Echols, J.J. Headd, L.W. Hung, G.J. Kapral, R.W. Grosse-Kunstleve, A.J. McCoy, N.W. Moriarty, R. Oeffner, R.J.

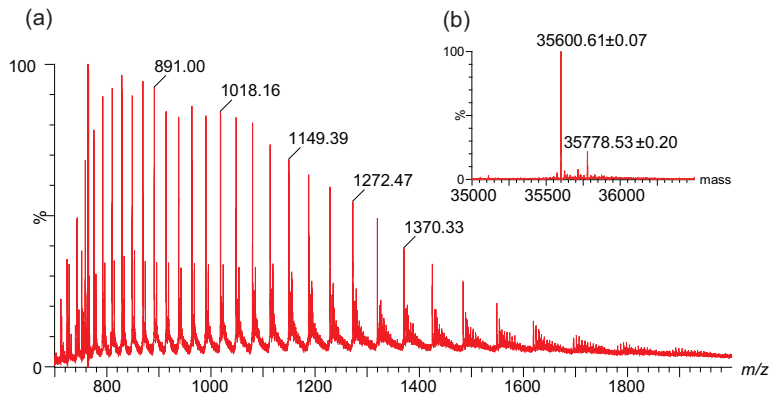
Read, D.C. Richardson, J.S. Richardson, T.C. Terwilliger, P.H. Zwart, PHENIX: a comprehensive Python-based system for macromolecular structure solution, Acta Crystallogr D Biol Crystallogr, 66 (2010) 213-221.
[25] M.D. Winn, C.C. Ballard, K.D. Cowtan, E.J. Dodson, P. Emsley, P.R. Evans, R.M. Keegan, E.B. Krissinel, A.G. Leslie, A. McCoy, S.J. McNicholas, G.N. Murshudov, N.S. Pannu, E.A. Potterton, H.R. Powell, R.J. Read, A. Vagin, K.S. Wilson, Overview of the CCP4 suite and current developments, Acta Crystallogr D Biol Crystallogr, 67 (2011) 235-242.
[26] G.N. Murshudov, P. Skubak, A.A. Lebedev, N.S. Pannu, R.A. Steiner, R.A. Nicholls, M.D. Winn, F. Long, A.A. Vagin, REFMAC5 for the refinement of macromolecular crystal structures, Acta Crystallogr D Biol Crystallogr, 67 (2011) 355-367.
[27] P. Emsley, B. Lohkamp, W.G. Scott, K. Cowtan, Features and development of Coot, Acta Crystallogr D Biol Crystallogr, 66 (2010) 486-501.
[28] T.C. Terwilliger, P.D. Adams, R.J. Read, A.J. McCoy, N.W. Moriarty, R.W. Grosse-Kunstleve, P.V. Afonine, P.H. Zwart, L.W. Hung, Decision-making in structure solution using Bayesian estimates of map quality: the PHENIX AutoSol wizard, Acta Crystallogr D Biol Crystallogr, 65 (2009) 582-601.
[29] A.J. McCoy, R.W. Grosse-Kunstleve, P.D. Adams, M.D. Winn, L.C. Storoni, R.J. Read, Phaser crystallographic software, J Appl Crystallogr, 40 (2007) 658-674.
[30] C.J. Williams, J.J. Headd, N.W. Moriarty, M.G. Prisant, L.L. Videau, L.N. Deis, V. Verma, D.A. Keedy, B.J. Hintze, V.B. Chen, S. Jain, S.M. Lewis, W.B. Arendall, 3rd, J. Snoeyink, P.D. Adams, S.C. Lovell, J.S. Richardson, D.C. Richardson, MolProbity: More and better reference data for improved all-atom structure validation, Protein Sci, 25 (2017).
[31] L. Holm, L.M. Laakso, Dali server update, Nucleic Acids Res, 44 (2016) W351-355.
[32] E. Krissinel, K. Henrick, Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions, Acta Crystallogr D Biol Crystallogr, 60 (2004) 2256-2268.
[33] E. Krissinel, Stock-based detection of protein oligomeric states in jsPISA, Nucleic Acids Res, 43 (2015) W314-319.
[34] K.F. Geoghegan, H.B. Dixon, P.J. Rosner, L.R. Hoth, A.J. Lanzetti, K.A. Borzilleri, E.S. Marr, L.H. Pezzullo, L.B. Martin, P.K. LeMotte, A.S. McColl, A.V. Kamath, J.G. Stroh, Spontaneous alpha-N-6-phosphogluconoylation of a "His tag" in *Escherichia coli*: the cause of extra mass of 258 or 178 Da in fusion proteins, Anal Biochem, 267 (1999) 169-184.

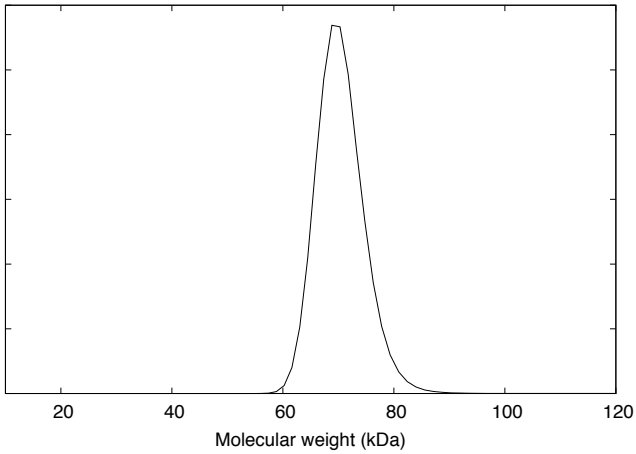
[35] H. Kishida, S. Unzai, D.I. Roper, A. Lloyd, S.Y. Park, J.R.H. Tame, Crystal structure of penicillin binding protein 4 (dacB) from *Escherichia coli*, both in the native form and covalently linked to various antibiotics, Biochemistry, 45 (2006) 783-792.
[36] J.D. Lenz, K.T. Hackett, J.P. Dillard, A Single Dual-Function Enzyme Controls the Production of Inflammatory NOD Agonist Peptidoglycan Fragments by *Neisseria gonorrhoeae*, MBio, 8 (2017) 01464-01417.
[37] P.H. Brown, A. Balbo, P. Schuck, On the analysis of sedimentation velocity in the study of protein complexes, Eur Biophys J, 38 (2009) 1079-1099.


α1
α5
α8
α2
α6
α7
α3
α4


α1
α5
α8
α2
α6
α7
α3
α4


(a)
N
N
N
N
(b)
GLU-200
HIS-270
SER-106
ARG-211
(c)
N
N


(a)
100
%
0
891.00
1018.16
1149.39
1272.47
1370.33
800
1000
1200
1400
1600
1800
m/z
(b)
100
%
0
35600.61±0.07
35778.53±0.20
35000
35500
36000
mass


20
40
60
80
100
120
Molecular weight (kDa)

The crystal structure of *Escherichia coli* L,D-carboxypeptidase A

Karen Meyer, Christine Addy, Satoko Akashi, David I. Roper and Jeremy R. H. Tame

Supplementary material

Figure S1.

A sequence alignment of *E. coli* LdcA with the proteins from *Novosphingobium aromaticuvorans* and *Pseudomonas aeruginosa.* Conserved residues are highlighted in white-on-red. Corresponding structural elements of the *E. coli* model are shown as helices and arrows over the sequence.

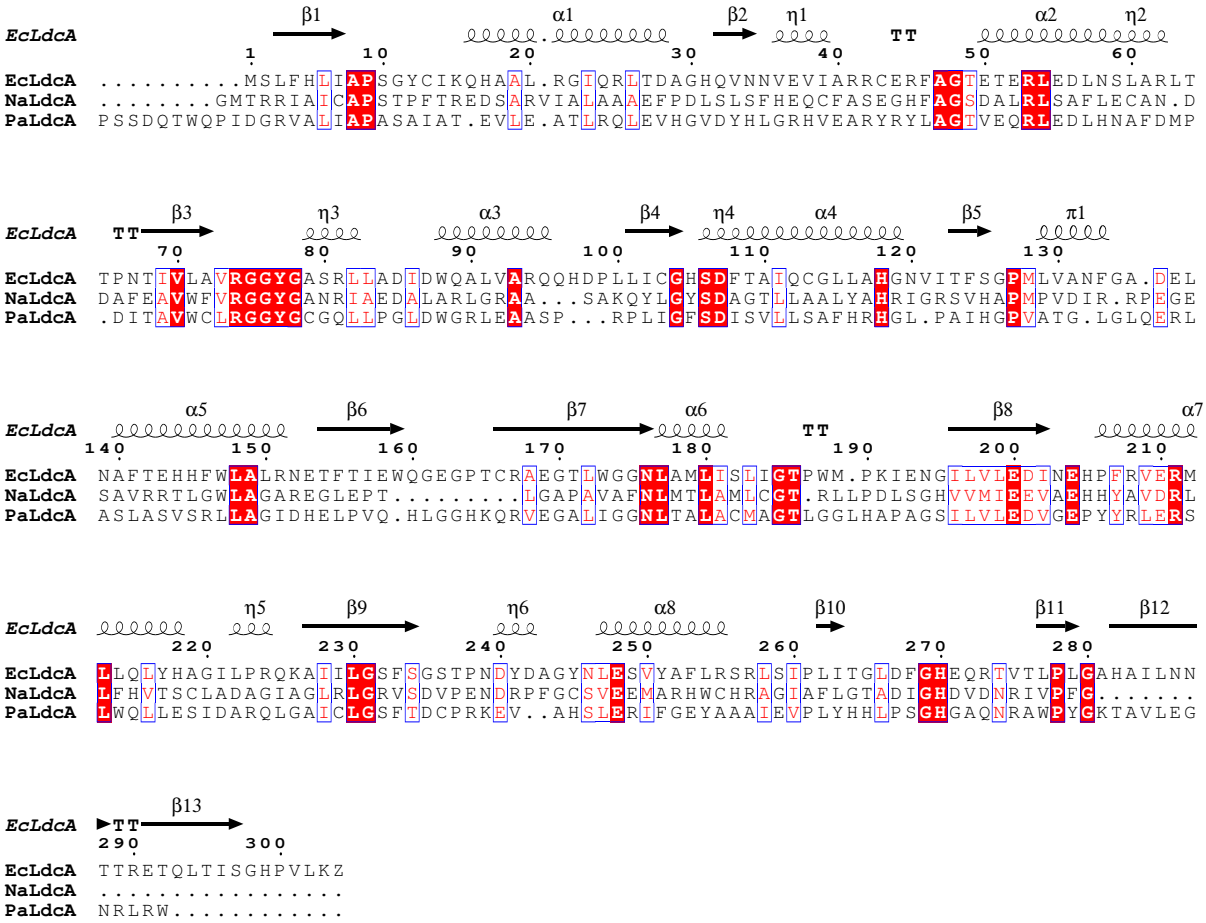


Figure S2.

A least-squares overlay of ribbons representing one subunit of *Ec*LdcA (beige) and MccF (purple), shown in stereo. The extended loop of MccF carrying Trp 186 is shown, and the tryptophan side-chain is shown as sticks. The catalytic triad of *Ec*LdcA is shown as sticks with carbon atoms coloured yellow, oxygen red and nitrogen blue.

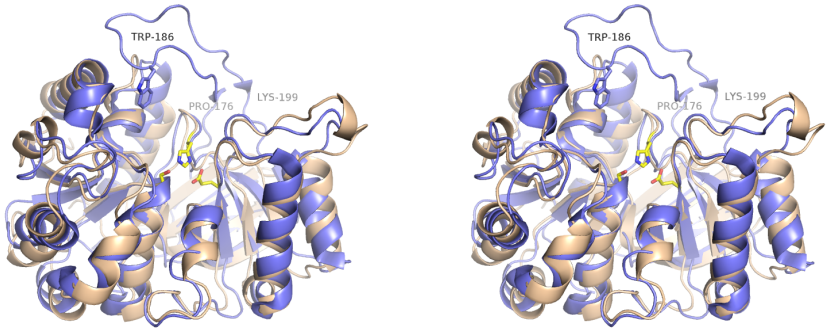


Figure S3.

A stereo figure showing the active site of native *Ec*LdcA (carbon atoms in yellow), superposed with the S106A mutant (with carbon atoms orange). Loss of the active site serine side-chain causes marked main-chain shifts on either side of the active site, notably Gly 75.

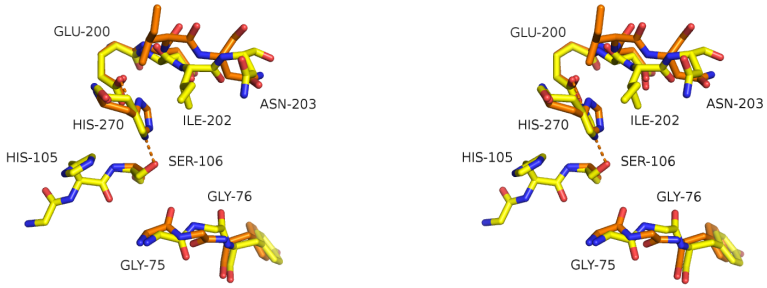


Figure S4.

A stereo figure showing the active site of native *Ec*LdcA (carbon atoms in yellow for one subunit, and in pink for the other). Arg 211 approaches within 8 Ångstroms of the active site serine of the partner subunit. This conserved residue is close enough to bind substrate, suggesting that the dimer is the active form of the enzyme, and the arginine plays some role in substrate interaction.

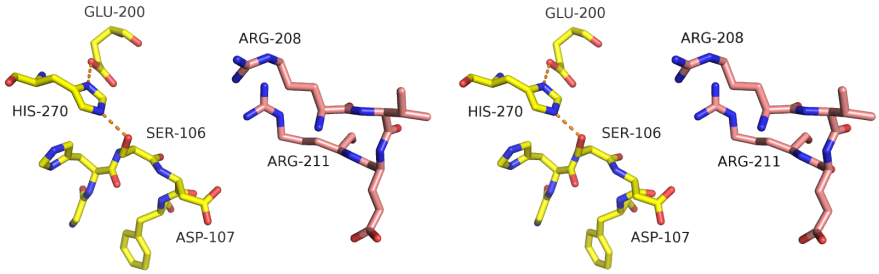


Figure S5.

The 2mFo-DFc electron density map over the catalytic triad of the active site in native *Ec*LdcA, contoured at 1 σ, shown in stereo. Density extending from the active site serine side-chain is found, but mass spectrometry indicates no covalently attached species. The active site density was left unmodelled.

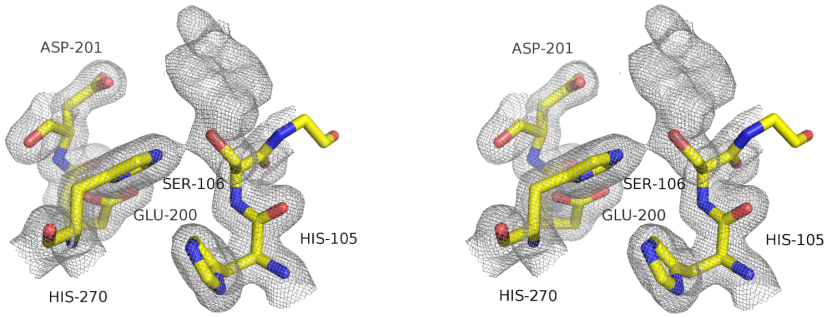
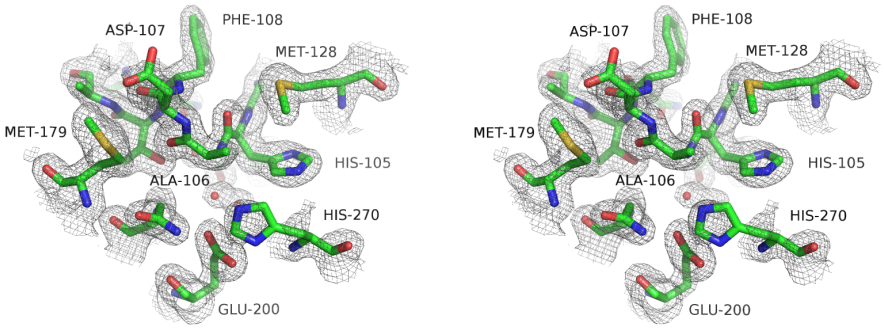


Figure S6.

The 2mFo-DFc electron density map over the catalytic triad of the active site in mutant S106A *Ec*LdcA, contoured at 1 σ, shown in stereo.



Supplementary Table 1. Data collection and structure refinement

Data Set	Native PDB 5Z01	S106A mutant PDB 5Z03
Resolution range (Å)	46.9 – 1.75	46.7 – 1.75
Space group	*I*222	$P2_12_12$
Unit cell dimensions (Å)	*a*=77.9 *b*=89.0, *c*=93.8	*a*=68.8 *b*=93.4, *c*=86.8
Reflections (Measured/Unique)	300,303 / 33,235	363,628 / 56,189
Completeness (%)	99.9 / 99.9	99.1 / 85.0
Mean $\langle I \rangle$ / $\langle \sigma(I) \rangle$	19.4 / 2.2	11.2 / 2.0
Multiplicity	9.0 / 9.3	6.5 / 5.5
*R*merge (%)[a]	7.3 / 101.6	9.7 / 83.2
B factor from Wilson plot (Å^2)	23.5	19.9
cc1/2 (%)	99.9 / 88.9	99.9 / 49.1
Refinement Statistics		
Resolution range (Å)	46.9 – 1.75	46.7 – 1.75
R-factor[b] / free *R*-factor (%)	17.9 / 22.5	22.9 / 27.0
Rms deviations from ideals		
bond lengths (Å) / bond angles (°)	0.02 / 1.80	0.019 / 1.90
No. of water atoms	212	56
B-values (all atoms, Å^2)		
protein/water	30.0 / 34.5	33.4 / 27.7
Ramachandran plot		
residues in most favorable regions	293	533
residues in allowed regions	10	35
residues in outlier regions	0	1

Values in outer shell are for the highest shell with resolution limits of 1.79-1.75 Å.

[a] *R*merge = $\Sigma \mid I_i - \langle I \rangle \mid / \Sigma \mid I_i \mid$, where I_i is the intensity of an observation and $\langle I \rangle$ is the mean value for that reflection and the summations are over all reflections. Free R-factor was calculated with 5% of the data.

[b] *R*-factor = $\Sigma_h \| Fo(h) \mid - \mid Fc(h) \| / \Sigma_h Fo(h)$, where *F*o and *F*c are the observed and calculated structure factor amplitudes, respectively.