Molecular Microbiology (2015) 95(5), 739–753 ■
doi:10.1111/mmi.12857
First published online 8 December 2014

A two-component, multimeric endolysin encoded by a single gene

Daniela Proença,[1] Christophe Velours,[2] Clara Leandro,[1] Miguel Garcia,[1] Madalena Pimentel[3,4] and Carlos São-José[3,4]*

[1]*Technophage, SA, Av. Professor Egas Moniz, Ed. Egas Moniz, piso 2, 1649-028 Lisboa, Portugal.*
[2]*Centre de Recherche de Gif, Centre National de la Recherche Scientifique, Plateforme de biophysique d'IMAGIF, 91198 Gif-sur-Yvette, France.*
[3]*Centro de Patogénese Molecular, Unidade de Retrovírus e Infecções Associadas (CPM-URIA) and*
[4]*Research Institute for Medicines (iMed.ULisboa), Faculdade de Farmácia da Universidade de Lisboa, Av. Prof. Gama Pinto, 1649-003 Lisboa, Portugal.*

Summary

Bacteriophage endolysins are bacterial cell wall degrading enzymes whose potential to fight bacterial infections has been intensively studied. Endolysins from Gram-positive systems are typically described as monomeric and as having a modular structure consisting of one or two N-terminal catalytic domains (CDs) linked to a C-terminal region responsible for cell wall binding (CWB). We show here that expression of the endolysin gene *lys170* of the enterococcal phage F170/08 results in two products, the expected full length endolysin (Lys170FL) and a C-terminal fragment corresponding to the CWB domain (CWB170). The latter is produced from an in-frame, alternative translation start site. Both polypeptides interact to form the fully active endolysin. Biochemical data strongly support a model where Lys170 is made of one monomer of Lys170FL associated with up to three CWB170 subunits, which are responsible for efficient endolysin binding to its substrate. Bioinformatics analysis indicates that similar secondary translation start signals may be used to produce and add independent CWB170-like subunits to different enzymatic specificities. The particular configuration of endolysin Lys170 uncovers a new mode of increasing the number of CWB motifs associated to CD modules, as an alternative to the tandem repetition typically found in monomeric cell wall hydrolases.

Accepted 6 November, 2014. *For correspondence. E-mail csaojose@ff.ul.pt; Tel. (+351) 217946420; Fax (+351) 217986055.

Introduction

Like all viruses, bacteriophages need to parasite host cells, in this case bacteria, in order to multiply. After replication, newly formed virus particles need to escape from infected cells to disseminate. To accomplish this, double-stranded DNA phages have evolved protein systems that compromise the integrity of the bacterial cell envelope in order to cause host cell lysis. These systems may comprehend a set of functions targeting the different cell envelope barriers, but the hallmark of lysis functions is the well-known holin–endolysin dyad (Catalão *et al.*, 2013; Young, 2014). Holins are hydrophobic proteins that oligomerize in the cytoplasmic membrane and induce the formation of holes in a tightly scheduled, saltatory manner. Canonical holins produce micron-scale holes large enough to allow the passage of cytoplasm-accumulated endolysins to the cell wall compartment, whereas pinholins form small channels that serve to depolarize the membrane and activate previously secreted endolysins (Nascimento *et al.*, 2008; Young, 2013; Savva *et al.*, 2014). Endolysins are enzymes that cleave the peptidoglycan (PG) network of the bacterial cell wall. They have been classified into five major functional types according to the bonds of the PG they cleave: N-acetylmuramidases (lysozymes), endo-β-N-acetylglucosaminidases and lytic transglycosylases cleave bonds of the N-acetylmuramic acid (NAM)/N-acetylglucosamine moiety of the PG, but with different specificities and/or end products; N-acetyl-muramoyl-L-alanine amidases hydrolyze the amide bond between NAM and L-alanine residues in the oligopeptide chains, and endopeptidases attack the peptide bonds within or between these chains (São-José *et al.*, 2003; Loessner, 2005). Within each major group, endolysins have been subdivided into families according to the sequence relatedness of their functional domains (Oliveira *et al.*, 2013).

The vast majority of endolysins produced by phages of Gram-positive bacteria and of mycobacteria seem to display a conserved modular architecture (Diaz *et al.*, 1990; Payne and Hatfull, 2012) of two separated functional regions: an N-terminus carrying one to three cata-

lytic domains (CDs) and a C-terminus segment harboring one or several repeats of cell wall binding (CWB) motifs (Fischetti, 2008; Schmelcher *et al*., 2012). With the exception of the multimeric endolysin PlyC, which is composed of a two CD-containing polypeptide (PlyCA) associated to eight PlyCB subunits with CWB activity, with A and B subunits codified by separate genes (Nelson *et al*., 2006; McGowan *et al*., 2012), all studied endolysins are encoded by a single gene and seem to be monomeric when purified.

Some endolysins have the capacity to degrade the cell wall PG and cause cell lysis when added in the form of recombinant proteins to Gram-positive bacteria. This has stimulated intense research to exploit the potential of endolysins as antibacterial agents (for recent reviews, see Fischetti, 2010; Nelson *et al*., 2012; Schmelcher *et al*., 2012). In one of these studies, we showed that the endolysin Lys170 from the enterococcal phage F170/08 exhibited broad lytic activity against *Enterococcus faecalis* clinical strains (Proença *et al*., 2012). Lys170 is a typical modular endolysin displaying an N-terminal amidase CD linked to a putative C-terminal CWB region. Lys170 is virtually identical to the previously described endolysin of *E. faecalis* phage ϕEF24C (Uchiyama *et al*., 2011), with both enzymes showing a single substitution over their 289 amino acid sequence.

We show now that expression of *lys170* results in the production of a *ca*. 12 kDa small protein, basically corresponding to the predicted Lys170 CWB domain, in addition to the expected full-length polypeptide (Lys170FL, 32.6 kDa). We have studied the interaction between these two endolysin components and their contribution to lytic activity. We present data supporting that fully active Lys170 is a multimeric endolysin resulting from the association of ~12 kDa CWB subunits with Lys170FL. We speculate that such endolysin architecture may be quite common, and we discuss its potential advantages in terms of lytic efficacy.

Results

Expression of endolysin gene lys170 *results in two stable polypeptides*

Production of a C-terminal His_6-tagged version of Lys170 in *Escherichia coli* systematically resulted in the accumulation of two polypeptides, one corresponding to expected full-length protein (Lys170FL, 32.6 kDa) and the other with an apparent mass of about 12 kDa. We have missed this smaller protein in previous works because of too long sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) runs (Proença *et al*., 2012). The two proteins co-purified during metal chelate affinity chromatography (AFC) and were immunodetected with anti-His_6 specific antibodies (not shown), which indicated that the ~12 kDa polypeptide was a C-terminal fragment of Lys170FL; its mass suggested that it would essentially correspond to the predicted Lys170 CWB domain (Proença *et al*., 2012), and therefore it was designated CWB170 (Fig. 1A).

Inspection of *lys170* nucleotide sequence revealed the possibility of an internal and in-frame translational start site located at a position compatible with the production of a ~12 kDa protein (Fig. 1B). Elimination of the putative ribosome binding site (RBS) and methionine start codon (Met_{202}) through site-directed mutagenesis resulted in the production of a single protein (mLys170) with the same apparent mass of Lys170FL (Fig. 1C), in agreement with these sequences being a translation signal in *E. coli*. The polypeptide mLys170 has the same amino acid sequence of Lys170FL, except that the internal start methionine was substituted by a leucine residue (Fig. 1D). N-terminal sequencing of a CWB170 band obtained after SDS-PAGE separation of a sample from the Lys170 AFC peak (Fig. 1A) revealed the sequence MY(?)LY, which basically matches the N-terminal sequence MYCLY(. . .) expected for a protein initiated at the putative secondary start site (cysteine residues could not be determined by the method used). To discard the possibility of CWB170 being generated by some sort of cleavage mechanism, which could have been inhibited by the $M_{202}L$ alteration, we have substituted the Met_{170} codon of *lys170* by a stop codon (TAG). Insertion of this stop codon upstream of the putative starting Met_{202} resulted in the production and accumulation of the expected truncated product ($Lys170_{STOP}$) of 17.8 kDa (Fig. 1E and D). Yet, this premature stop in translation had no obvious impact on the synthesis of CWB170 (Fig. 1E), as it would be expected if the latter polypeptide resulted from processing of Lys170FL. We have thus concluded that CWB170 was produced from independent translation initiation at the internal Met_{202} codon, given rise to a product of 11.6 kDa as deduced from the Lys170 primary sequence.

The Lys170 peak fraction from the AFC step (Fig. 1A) was subjected to size-exclusion chromatography (SEC) in an attempt to separate Lys170FL from CWB170. Strikingly, the two polypeptides essentially co-eluted in a single peak during SEC despite their markedly different predicted masses, 32.6 and 11.6 kDa, respectively (Fig. 2). Based on the elution volume of the peak, we estimated a mass of about 62 kDa (see *Experimental procedures*), assuming for the species composing the peak a homogeneous and globular nature analogous to that of the standard proteins run in the same conditions. Interestingly, when a C-terminally His_6-tagged CWB170 independently produced and purified by AFC was subjected to the same SEC, it resulted in a profile clearly distinguishable from that of Lys170, eluting with an apparent mass of ~37 kDa

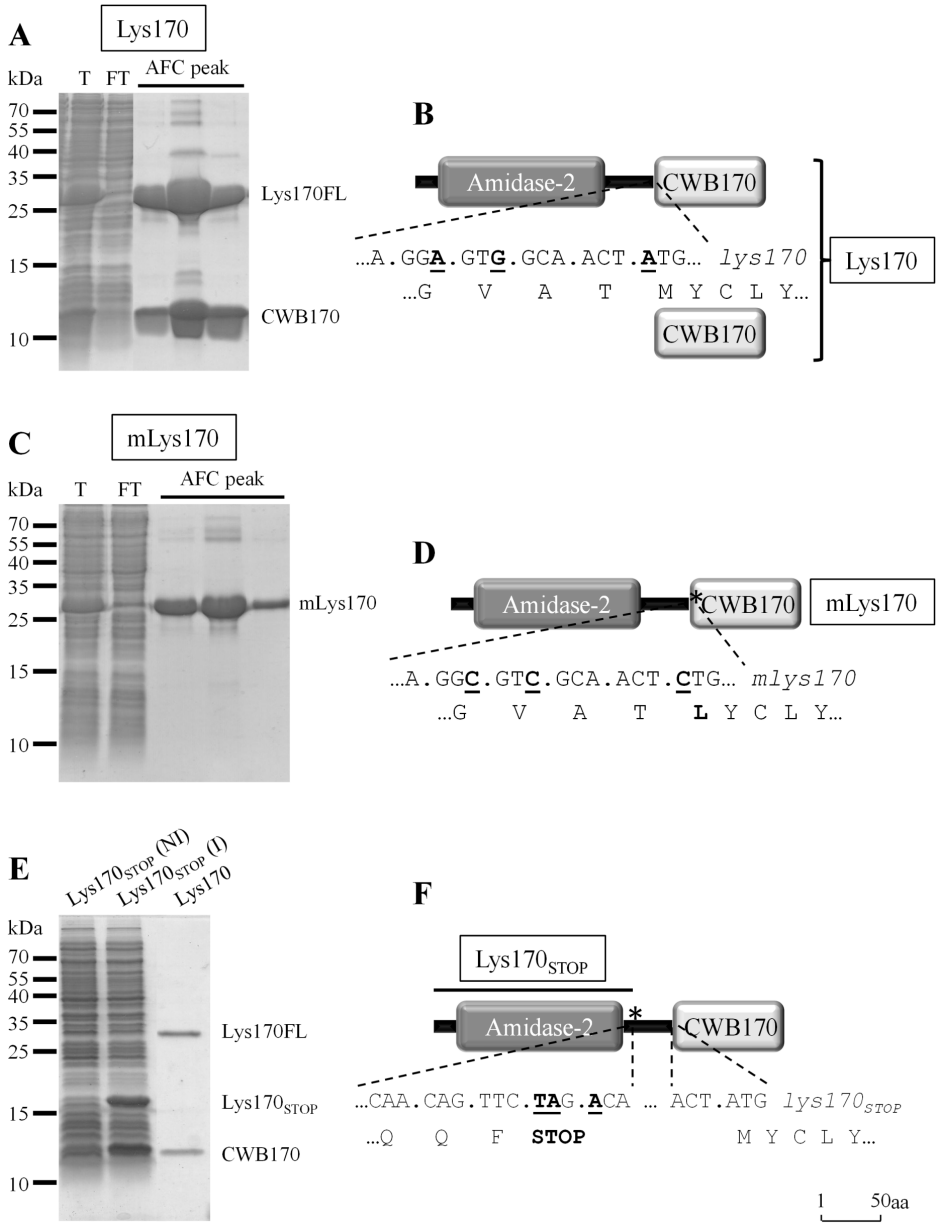

Fig. 1. The two polypeptides of endolysin Lys170.
A. SDS-PAGE analysis of the Lys170 overproduction in *E. coli* and of the peak fractions resulting from the endolysin purification by metal chelate affinity chromatography (AFC). T, total soluble protein extract; FT, affinity column flowthrough. The protein bands corresponding to the full-length endolysin (Lys170FL) and to its C-terminal product CWB170 are indicated.
B. Schematic representation of the predicted Lys170 domain architecture. Details of the putative internal translation start site driving the independent production of CWB170 are shown below the endolysin scheme. The nucleotide positions mutagenized to eliminate the putative ribosome binding site and start codon (Met_{202}) are underlined.
C. SDS-PAGE analysis of the mLys170 overproduction in *E. coli* and of the peak fractions resulting from the protein purification by AFC. T and FT as in panel A. The protein band corresponding to mLys170 is indicated.
D. Domain architecture of mLys170. Below the endolysin scheme, it is shown the nucleotide and amino acid sequences resulting from the site-directed mutagenesis of the putative *lys170* internal translation start site.
E. SDS-PAGE analysis of the Lys170-derived polypeptides produced after substitution of the *lys170* Met_{170} codon by a stop codon. The gel shows total protein extracts produced from noninduced (NI) and induced (I) *E. coli* cultures carrying $lys170_{STOP}$. The lane from the induced culture shows the truncated Lys170 ($Lys170_{STOP}$; 17.8 kDa) and CWB170 polypeptides. A control lane with purified Lys170 shows the positions of Lys170FL and CWB170 proteins.
F. Generation of $Lys170_{STOP}$. Details of the nucleotide and amino acid sequences resulting from the substitution of the *lys170* Met_{170} codon by a stop codon (**TA**G) are shown below the endolysin scheme. Note that the mutagenesis also changed the Ala_{171} codon (GCA) to a Thr codon (**A**CA), which generated an *XbaI* site (TC.**TA**G.**A**) used for screening purposes. The starting Met_{202} codon and N-terminal amino acid sequence of CWB170 is also shown.

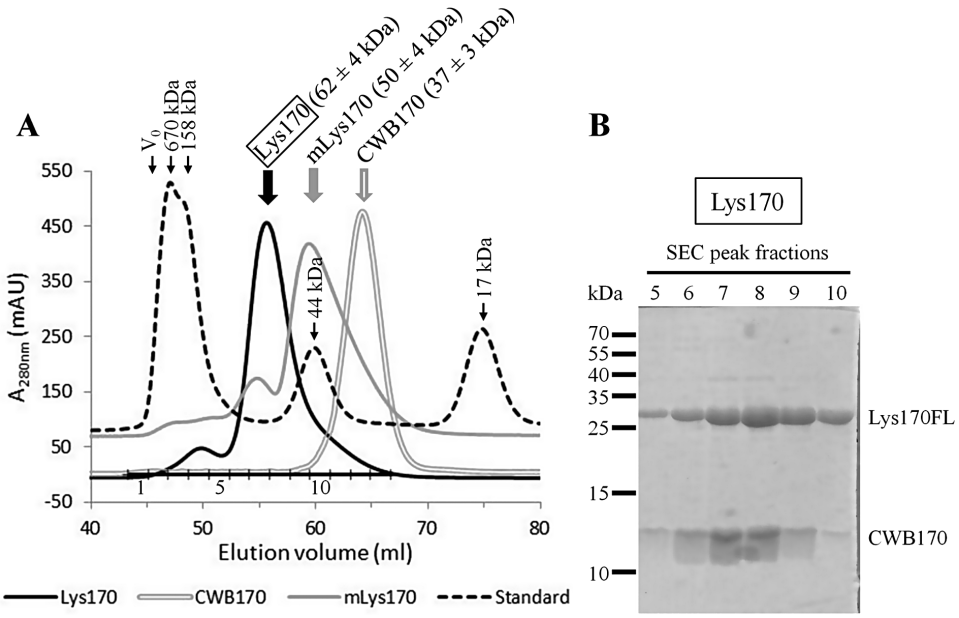

Fig. 2. Size-exclusion chromatography of the Lys170, mLys170 and CWB170 proteins from the affinity purification step.
A. The eluting profile of the proteins was monitored by taking absorbance measurements at 280 nm (A_{280nm}). Representative UV curves were combined in a single graph. The apparent protein masses derived from the experimentally determined partition coefficients (K_{av}, see *Experimental procedures*) are indicated for each protein. The column void volume (V_0), the masses of standard proteins and the fractionation range of Lys170 are also indicated.
B. SDS-PAGE analysis of the SEC peak fractions of Lys170, confirming the co-elution of Lys170FL and CWB170.

(Fig. 2A). mLys170 also peaked at a different elution volume during SEC, in this case with an estimated mass of ~50 kDa (Fig. 2A). Although the masses estimated from simple SEC analysis can be influenced by the proteins Stokes radii, the results suggested that: (i) Lys170 corresponded to a complex of Lys170FL + CWB170 and (ii) CWB170 and mLys170 oligomerized and/or formed elongated structures conferring them apparent masses higher than those expected for monomeric proteins (Erickson, 2009).

Lys170FL and CWB170 polypeptides are required for full endolysin lytic activity in vitro

We showed above that the CWB170 polypeptide seemed to associate with Lys170FL, a result that could hint for a role of the C-terminal fragment in endolysin activity. In fact, when assays of cell suspension turbidity reduction were performed to compare the lytic activity of Lys170 and mLys170 against *E. faecalis* cells, we observed that the latter protein could not elicit any detectable lysis, in clear contrast to the two-component endolysin (data not shown). We have thus reasoned that co-incubation of purified mLys170 and CWB170 might generate active endolysin complexes, resulting in a visible enhancement of lytic activity. To test this hypothesis we have empirically fixed an amount of mLys170 (10 μg, 0.31 nmol) and varied the quantity of the smaller protein to obtain mLys170 : CWB170 molar ratios of 1:0.25, 1:0.5, 1:1, 1:2, 1:4 and 1:6. The different mixtures were spotted on a dense lawn of *E. faecalis* cells (see *Experimental procedures*) and enterolytic activity evaluated based on the presence and relative diameter of the lysis halos developed after overnight incubation. Lys170, mLys170 and CWB170 were also spotted alone and at the maximum concentration used in the different combinations. The results confirmed the much reduced lytic activity of mLys170 when compared with Lys170 and showed that CWB170 was unable to elicit any detectable lysis by itself (Fig. 3, bottom row). However, when mLys170 was pre-incubated with increasing amounts of CWB170, the lytic activity was progressively restored, with the ratio 1:6 producing a lytic effect apparently similar to that of Lys170 (Fig. 3). The results indicated that CWB170 is required for full endolysin activity and again supported an interaction between Lys170FL and CWB170. Somewhat unexpectedly, when we tried to evaluate the lytic effect of the mLys170/CWB170 mixtures in a more quantitative way, by determining lysis kinetics of dense cell suspensions, we could not measure any obvious lysis, even with 1:6 mLys170 : CWB170 molar ratio mixtures. We believe though that this apparently contradictory result can be explained by a low efficiency of production of active complexes upon mLys170/CWB170 co-incubation, which are

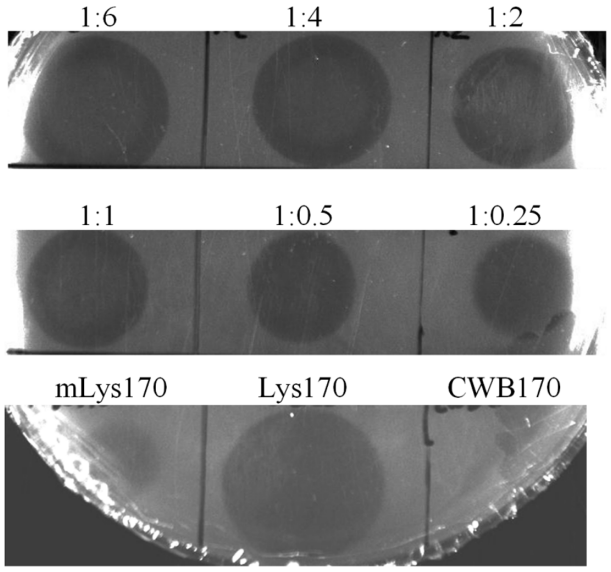

Fig. 3. Impact of CWB170 polypeptide in endolysin activity. Purified mLys170 and CWB170 were co-incubated at the indicated mLys170 : CWB170 molar ratios for 1 h at room temperature. After this period, each protein mixture was spotted on a dense lawn of *E. faecalis* cells. The image shows the lysis halos developed after overnight incubation at 37°C. Lysis halos from individually spotted mLys170 (0.31 nmol), Lys170 (0.31 nmol) and CWB170 (1.86 nmol) are shown in the bottom row.

still sufficient to be detected by the highly sensitive spot test assays (Fig. 3), but not enough to elicit lysis of dense cell suspensions (see *Discussion*).

The two endolysin polypeptides are produced in the phage infection context

As efficient lytic activity of endolysin Lys170 seemed to require the presence of Lys170FL and CWB170, we have anticipated that both polypeptides should be produced in *E. faecalis* cells during infection by phage F170/08. To test this, a mid-log culture of *E. faecalis* strain 926/05 was infected with the phage at an input multiplicity of 2 and samples collected every 10 min for production of total protein extracts (Fig. 4A). These were separated by SDS-PAGE, followed by Coomassie blue-staining (to confirm even loading, Fig. 4B) and Western blot analysis with anti-Lys170 antibodies (Fig. 4C). The results showed that both Lys170FL and CWB170 started to accumulate at $t = 20$ min, reaching their peak 60 min after phage infection. In contrast to what happened with *lys170* expression in *E. coli* (Fig. 1A), in the phage infection context CWB170 seemed to accumulate in great excess when compared with Lys170FL. The results confirmed that the two Lys170 polypeptides produced in *E. coli* are also synthesized during the phage F170/08 infection cycle, excluding the

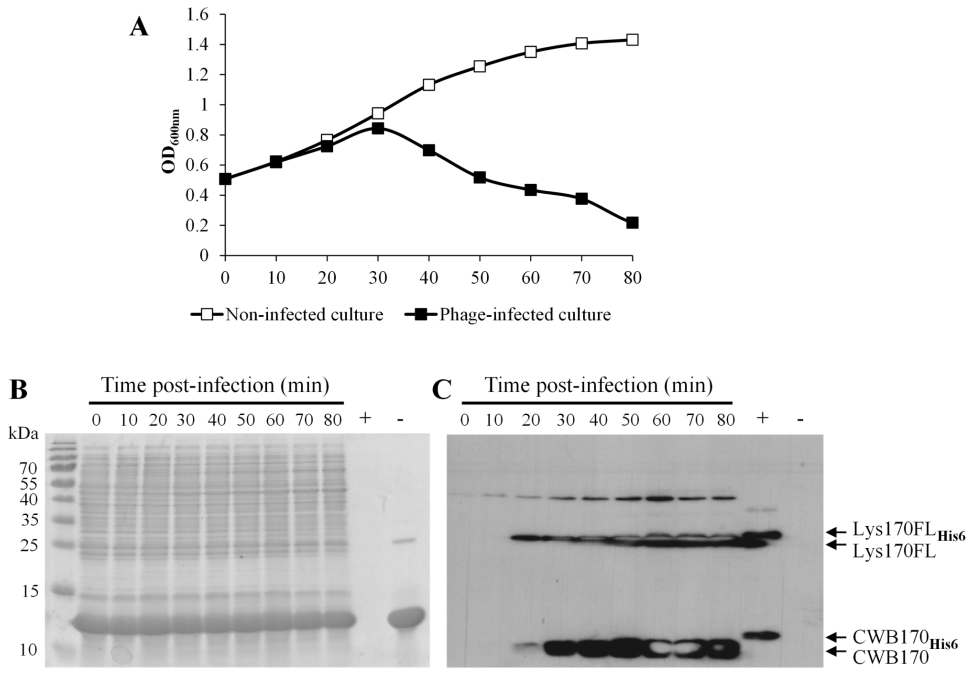

Fig. 4. Time course of Lys170FL and CWB170 synthesis during *E. faecalis* infection by phage F170/08. Total protein extracts were prepared from samples of an infected culture (A) collected every 10 min after infection, during 80 min. Ten micrograms of each extract were separated by SDS-PAGE and analyzed by Coomassie blue-staining (B) and Western blotting with anti-Lys170 antibodies (C). Purified Lys170 (50 ng) was used as positive control (+) and 10 µl of lysis buffer was used as negative control (–). The very intense band observed in B between positions 10 and 15 kDa of the molecular weight marker is from lysozyme, which is present in the lysis buffer at a 2.5 mg ml^{-1} concentration.

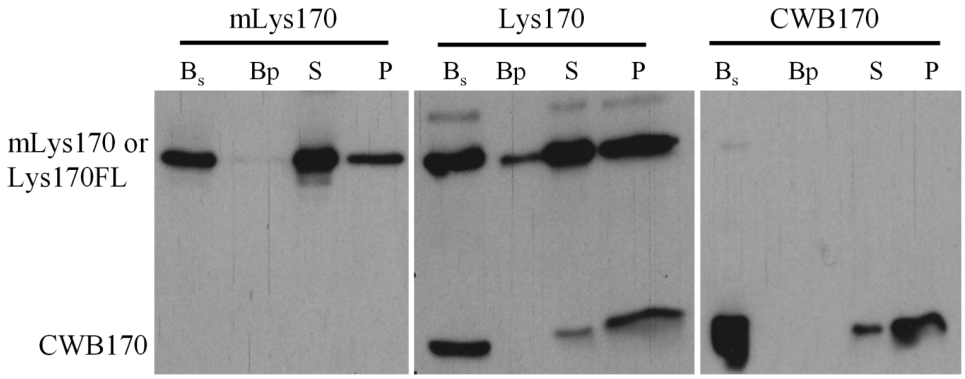

Fig. 5. Binding of purified mLys170, Lys170 and CWB170 to *E. faecalis* cells. Binding reactions were prepared by adding 1 μg of each protein to 100 μl of a concentrated *E. faecalis* cell suspension, followed by 30 min incubation on ice. Each reaction was centrifuged, and the relative distribution of the proteins in the supernatant and pellet fractions analyzed by Western blot using anti-Lys170 antibodies. Negative controls consisting in the addition of the proteins to endolysin buffer only (no target cells) were equally prepared and processed. B_s and B_p, supernatant and pellet fractions of negative controls, respectively; S and P, supernatant and pellet fractions of test assays, respectively.

possibility of an artifact resulting from heterologous expression.

CWB170 promotes endolysin binding to target cells

Based on the typical modular structure of endolysins from Gram-positive systems and also on bioinformatics analysis, we hypothesized that the C-terminal polypeptide CWB170 would mediate the binding of endolysin Lys170 to the cell wall of target bacteria. To study the binding capacity of the different endolysin polypeptides, we have incubated *E. faecalis* cells with purified Lys170, mLys170 and CWB170. After 30 min incubation, the mixtures were centrifuged and we have evaluated how the proteins distributed between the supernatant and cell pellet fractions. In the absence of target cells, the three proteins were essentially present in the supernatant fraction (B_s lanes in Fig. 5), discarding major protein precipitation and consequent sedimentation (B_p lanes in Fig. 5) during the assay. The larger component of Lys170 (Lys170FL) seemed to be more associated with the *E. faecalis* pellet fraction than the mutagenized form mLys170, which was mainly detected in the supernatant fraction (compare fractions S and P for the two polypeptides). The CWB170 portion of Lys170 was only faintly detected in the supernatant, being most of it found in the pellet fraction. The capacity of the CWB170 polypeptide to bind *E. faecalis* was confirmed when the protein was tested alone, as again a much higher fraction of the protein was present in the pellet fraction. The results clearly showed that CWB170 bound to *E. faecalis* cells and strongly suggested that this independent domain was responsible for promoting binding of Lys170FL. The weak binding of mLys170 to the bacterial cell wall, probably due to the absence of independently produced CWB170, may be the cause underlying the decrease of mLys170 activity (see above).

Composition of the Lys170 complex

The elution profiles of Lys170 and CWB170 during SEC (Fig. 2) and the relative lytic activity exhibited by Lys170, mLys170 and mLys170 + CWB170 (Fig. 3) strongly suggested that Lys170FL and the CWB170 polypeptide associated to form the fully active Lys170. The mass estimations of Lys170 (62 kDa) and CWB170 (37 kDa) derived from SEC (Fig. 2A) seemed to be incompatible with Lys170 being a simple 1:1 Lys170FL (32.6 kDa) : CWB170 (11.6 kDa) heterodimer.

To gain insight on the nature of the putative Lys170FL : CWB170 complex, we determined the molar masses of purified Lys170, mLys170 and CWB170 by Size-Exclusion Chromatography – Multi Angle Light Scattering (SEC-MALS), a technique that allows determination of molar mass independently of the protein Stokes radii. SEC-MALS analysis of CWB170 produced a single UV peak (Fig. 6), with the protein eluting in SEC as monodisperse species at 16.1 ml and with a mass of 46 kDa, which indicated that purified CWB170 was a tetramer in solution (theoretical mass of 4CWB170 = 46.4 kDa). Most mLys170 showed its UV peak at 15.1 ml elution volume, with a corresponding molar mass of 34.1 kDa. This mass fits well that expected for a mLys170 monomer (predicted mass = 32.6 kDa) but deviates significantly from that estimated by the simple SEC analysis of Fig. 2 (~50 kDa). Such discrepancies between the masses predicted by conventional SEC and those determined by more accurate methods have been observed previously, and they generally derive from the extended, nonglobular nature of the proteins under analysis (São-José *et al.*, 2006; Ruggiero *et al.*, 2009). Two minor peaks were also detected for mLys170 at around 12.6 and 13.7 ml, corresponding to molar masses of 98.8 and 62.4 kDa, respectively, suggesting that a fraction of mLys170 could form

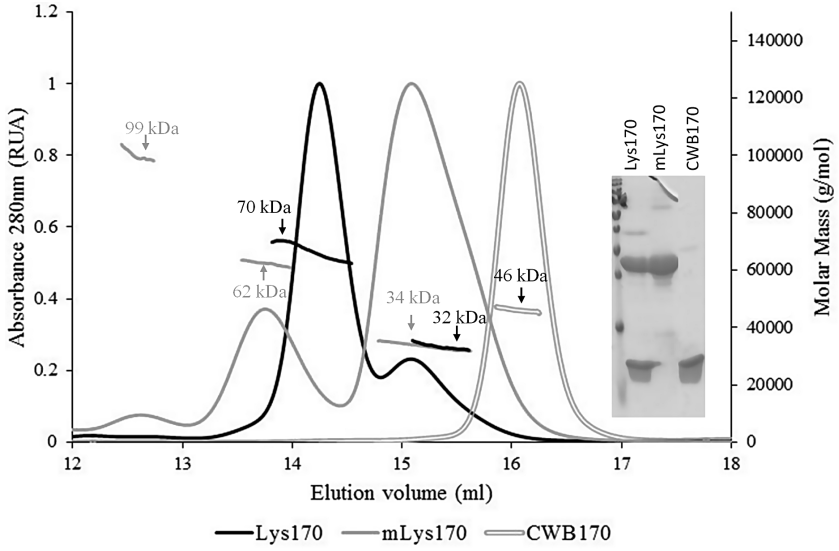

Fig. 6. SEC-MALS analysis of Lys170, mLys170 and CWB170. The relative UV and molar mass (M) curves of the three proteins were combined in a single graph; the refractive index (dRI) and light-scattering (LS) curves were omitted for clarity. The molar masses measured for the detected peaks are indicated. The inset shows an SDS-PAGE loaded with 10 μg samples of each protein preparation used in the SEC-MALS analysis.

homotrimers and homodimers (predicted masses of 97.8 and 65.2 kDa, respectively). Finally, the vast majority of Lys170 (Lys170FL + CWB170) produced an UV peak centered at 14.2 ml elution volume, with monodisperse species detected at 13.9 ml, to which corresponded a molar mass of 70.3 kDa. A minor peak eluted at 15.1 ml, with a measured mass (32 kDa) compatible with monomeric Lys170FL.

Considering that the major Lys170 peak would necessarily contain Lys170FL associated to CWB170, the best fitting model for the measured molar mass of 70.3 kDa was a 1Lys170FL : 3CWB170 complex (67.4 kDa). Interestingly, this 1:3 stoichiometry hypothesis was coherent with the tetrameric state found for isolated CWB170 (see above). Of course the molar mass obtained for the peak could also accommodate Lys170FL dimers (65.2 kDa), but in this case we would expect a hydrodynamic radius (and hence an elution volume) similar to that of mLys170 dimers (62 kDa peak of mLys170, Fig. 6). Thus, the presence of a significant fraction of Lys170FL dimers would be expected to produce a 'shoulder' in the Lys170 UV curve, something that was not observed.

To try sorting out the most likely stoichiometry for the Lys170 complex, we extracted from the SEC-MALS analysis the UV_{280nm} extinction coefficients (ε_p) of the peaks and compared them with the predicted ε_p of the different multimer models, which was computed based on the amino acid sequence of each multimer subunit (see *Experimental procedures*). Table 1 shows the results of this analysis. We observed that the experimental ε_p for the 70 kDa peak of Lys170 (1.953 ml mg^{-1} cm^{-1}) was very close to that predicted for a 1Lys170FL : 3CWB170 complex (1.983 ml mg^{-1} cm^{-1}) and quite distinct from the ε_p expected for a Lys170FL dimer (2Lys170FL, ε_p = 1.379 ml mg^{-1} cm^{-1}). The experimental ε_p for all the other species detected in SEC-MALS (monomeric Lys170FL/mLys170, dimeric mLys170 and tetrameric CWB170) showed a very good match to the corresponding theoretical ε_p (Table 1), supporting the robustness of the analysis.

Table 1. Analysis of the UV_{280nm} extinction coefficient (ε_p) of Lys170 multimer models.

Protein	SEC-MALS M peaks (kDa) Fig. 6	Experimental ε_p, ml mg^{-1} cm^{-1}	Multimer model	Computed molar mass (kDa)	Computed ε_p, ml mg^{-1} cm^{-1}
Lys170	70	**1.953**	**1**Lys170FL : **3**CWB170	67.4	1.**983**
			2Lys170FL	65.2	1.379
	32	1.440	**1**Lys170FL	32.6	1.379
mLys170	62	1.417	**2**mLys170	65.2	1.379
	34	1.410	**1**mLys170	32.6	1.379
CWB170	46	2.457	**4**CWB170	46.4	2.552

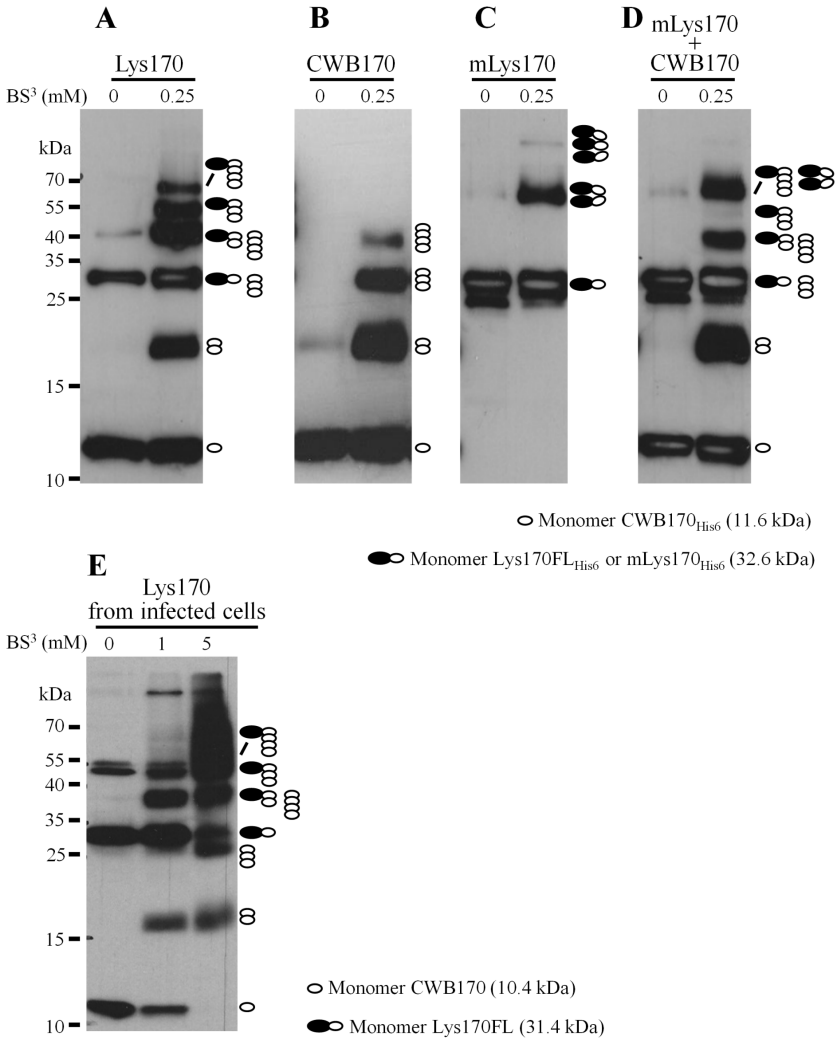

Fig. 7. Cross-linking analysis of endolysin multimers. Purified Lys170 (A), CWB170 (B), mLys170 (C), a pre-incubated mixture of CWB170 with mLys170 (D) and a protein sample from the time-point t = 60 min of the assay shown in Fig. 4 (E) were cross-linked with the indicated concentrations of BS^3 and the reaction products analyzed by Western blot with anti-Lys170 antibodies. When possible, the protein species deduced to compose each band is indicated following the depicted cartoon codes; bands compatible with more than one composition/oligomeric state are marked accordingly.

In summary, from the presented data, we have concluded that the most likely stoichiometry for the 70 kDa Lys170 complex was 1Lys170FL : 3CWB170.

Cross-linking of endolysin multimers

To independently confirm the presence of multimers identified in the SEC-MALS analysis, we used the water-soluble, noncleavable cross-linking agent BS^3, which reacts with primary amines, to treat purified Lys170, CWB170 and mLys170 at micromolar range concentrations (see *Experimental procedures*). Cross-linked and control samples were separated by SDS-PAGE, followed by Western blotting analysis of the resulting products with anti-Lys170 antibodies. Note that the cross-linking reaction conditions were optimized to allow simultaneous detection of the free monomers and of the different multimeric forms of the proteins under analysis in the ensemble of reactions. Several products were detected in the cross-linked sample of Lys170 that were compatible with different homo- and heterooligomeric states (Fig. 7A). In addition to a band expected for the complex 1Lys170FL : 3CWB170 (67.4 kDa), other bands most probably corresponding to different association/dissociation states of the complex were observed: 2CWB170 (23.2 kDa), 3CWB170 (34.8 kDa), 1Lys170FL : 1CWB170 (44.2 kDa) and 1Lys170FL : 2CWB170 (55.8 kDa). Note that the ~55 kDa band in the cross-linking profile of Lys170 can only be explained by a 1Lys170FL : 2CWB170 multimer, further supporting interaction between the two Lys170 subunits. We detected molecular species corresponding to dimers, trimers and tetramers (46.4 kDa) in the cross-linked sample of purified CWB170 (Fig. 7B), whereas in the corresponding sample of mLys170, we detected monomeric (32.6 kDa), dimeric (65.2 kDa) and a faint band of trimeric (97.8 kDa) species (Fig. 7C). In essence, the results were coherent with those obtained in the SEC-MALS analysis, although in the particular assay shown in Fig. 7, the signal

of the CWB170 tetramer was weaker than those obtained in other analogous experiments using higher concentrations of BS^3 (data not shown).

We have also prepared a reaction where pre-incubated mLys170 and CWB170 were cross-linked in the same conditions (Fig. 7D). The cross-linking of the mixture resulted in the appearance of new bands and in the apparent intensification of others when compared with the independent cross-linking of each protein (interpreted as formation of the 44.2 kDa 1Lys170FL : 1CWB170, the 55.8 kDa 1Lys170FL : 2CWB170, and eventually the 67.4 kDa 1Lys170FL : 3CWB170 complexes). The cross-linking pattern resembled that of Lys170, although with altered band intensities. In an attempt to unambiguously identify protein bands resulting from cross-linking of mLys170 and CWB170 subunits, we have generated and purified a version of the latter protein with the hemagglutinin (HA) epitope inserted just upstream of the His_6 tag. Unfortunately, this CWB170-HA protein revealed to be noncompetent to enhance lysis mediated by mLys170 in experiments like that of Fig. 3. In addition, CWB170-HA seemed to be inhibited in its self-association capacity as judged by SEC-MALS analysis (data not shown). We have thus decided not use this protein in cross-linking assays as it seemed affected in its normal biological activity.

To check if the Lys170 cross-linking pattern could be reproduced with endolysin synthesized during phage infection, we used a sample of the $t = 60$ min protein extract of Fig. 4 and cross-linked it with 1 and 5 mM BS^3. The results obtained (Fig. 7E) were consistent with those derived from the cross-linking of purified Lys170 and again supported CWB170 self-association and interaction with Lys170FL. However, caution should be taken when analyzing cross-links of crude protein extracts because of the presence of a high number of other proteins; this may explain the smear observed in presence of 5 mM BS^3 (Fig. 7E).

Discussion

We presented genetic and enzymatic activity studies indicating that maximum lytic activity of endolysin Lys170 requires the full-length enzyme (Lys170FL) and a small polypeptide corresponding to the last 88 amino acid residues of Lys170FL. This C-terminal fragment is produced from an in-frame, secondary translation start site and basically matches the predicted cell wall binding domain of the endolysin (CWB170). In addition, biochemical and biophysical characterization of the endolysin and derived proteins strongly suggests that functional Lys170 corresponds to a Lys170FL : CWB170 heterooligomer, being 1:3 the most likely stoichiometry for the higher order endolysin complex. Lys170 is thus the first two-component multimeric endolysin described to date, whose subunits are produced from a single gene.

The other known multimeric endolysin, PlyC from the streptococcal phage C_1, is composed of a two CD-containing polypeptide (PlyCA) associated to eight PlyCB subunits with CWB activity, but in this case the A and B subunits are encoded by separate genes (Nelson *et al.*, 2006). The ring-like PlyCB assembly, which contains eight potential binding sites for cell wall components, is crucial for PlyC function as the isolated PlyCA module has only residual activity (Nelson *et al.*, 2006; McGowan *et al.*, 2012). The poor lytic activity of the mutant endolysin mLys170, which is unable to produce independent CWB170 subunits, could be at least partially rescued upon incubation with purified tetramers of the latter protein (Fig. 3). This implies that, despite the apparent stability of the CWB170 tetrameric form, the equivalent module in mLys170 can still interact with and displace subunits of the tetramer, something that seems also to be supported by the cross-linking experiments presented in Fig. 7. Yet, the same cross-linking results also suggest that the efficiency of formation of mLys170 : CWB170 complexes during co-incubation of the purified subunits is low. Such indication emerges when comparing the intensity of the bands of cross-linked heterooligomers present in Lys170 and in the mLys170/CWB170 mixtures like, for example, the ~55 kDa cross-linking product corresponding to a 1Lys170FL(or mLys170) : 2CWB170 complex (compare panels A and D of Fig. 7). We could not detect also any obvious new peak (at most, we detected slight peak shifts) during analytical SEC of mLys170/CWB170 mixtures, again suggesting low efficiency of complex formation (data not shown). This should explain why mLys170/CWB170 mixtures could not induce detectable lysis of dense cell suspensions, despite being able to produce the lysis halos of Fig. 3. In conclusion, the production of Lys170FL : CWB170 heterooligomers, that is, Lys170, seems to be much facilitated when the two subunits are produced concomitantly within the same cell.

Our results indicate that the full-length, monomeric endolysin needs to increase its number of CWB motifs for optimal lytic activity, a requirement that is fulfilled by the self-association capacity of the CWB170 subunit (a total of 4 CWB170 motifs present in the 1:3 complex). The assembly of CWB170 modules at the Lys170FL C-terminus seems to increase the endolysin affinity to cells (Fig. 5); it might also contribute to the proper folding/orientation of the enzyme, namely of its N-terminal CD, as suggested for the CWB repeats of the pneumococcal endolysin Cpl-1 (Hermoso *et al.*, 2003). The extended configuration of mLys170 inferred from its 'abnormal' SEC profile (Fig. 6) might be an indication of this.

It has been shown for several PG hydrolases of Gram-positive systems that the number and type of CWB motifs

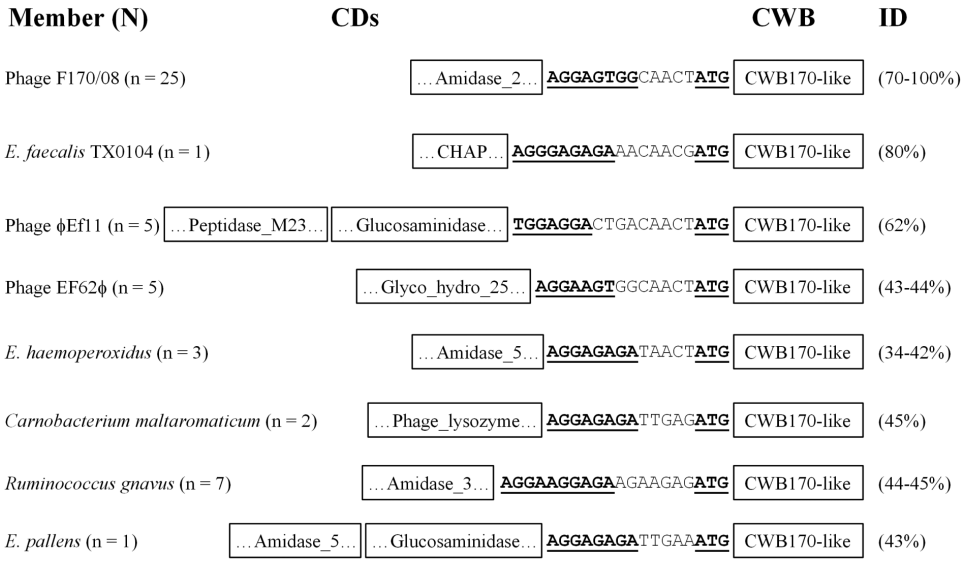

Fig. 8. PG hydrolases with CWB170-like domains. BLASTP searches with CWB170 sequence retrieved several PG hydrolases displaying CDs of different enzymatic specificities. The PG hydrolases are organized according to their domain architecture, and only one member representative of each group is shown. The 'N' value denotes the number of single, nonredundant protein sequences within each group, but each sequence may have been described in many different sequenced genomes (e.g., the represented *E. faecalis* TX0104 sequence is identical to that of 25 different database entries). The GenBank Acc. N0. of the members shown are: *E. faecalis* TX0104, EEI10842; Phage ϕEf11, ACV83371; Phage EF62ϕ, ADX81356; *E. haemoperoxidus*, EOH93425; *Carnobacterium maltaromaticum*, CCO10928; *Ruminococcus gnavus*, EDN76763; *E. pallens*, EOH88591. CDs families are according to Pfam database: Amidase_2, pfam01510; CHAP, pfam05257; Peptidase_M23, pfam01551; Glucosaminidase, pfam0183; Glyco_hydro_25, pfam01183; Amidase_5, pfam05382; Phage_lysozyme, pfam00959; Amidase_3, pfam01520). The in-frame ATG codon defining the beginning of CWB170-like domains and the putative RBS upstream are depicted in bold (also present in many other members of each group, but we have not confirmed if in all of them). ID, sequence identity of the CWB170-like domains to the CWB170 of Lys170, within each group of PG hydrolases (a range or single ID values are presented depending on weather CWB170-like sequences vary or not within a given group, respectively).

can play a critical role on the catalytic regulation and/or affinity of the enzymes to their target cell walls (López and García, 2004; Steen *et al*., 2005; Mesnage *et al*., 2014; Wong *et al*., 2014). This has been particularly studied for the CWB segments of pneumococcal autolysins (bacterial PG hydrolases responsible for the phenomenon of autolysis) and their related phage endolysins, whose tandemly arranged repeats of CWB motifs seem to provide a mechanism to improve cell wall recognition (López and García, 2004; Bustamante *et al*., 2010). For several of these pneumococcal PG hydrolases, the CWB repeats recognize the choline residues of cell wall teichoic acids. The number of choline-binding repeats (ChBRs) of autolysins can vary from 7 to up to 18, but a minimum of 4 ChBRs appears to be required for efficient binding to the cell wall (García *et al*., 1994; López and García, 2004; Moscoso *et al*., 2005). Interestingly, in presence of choline, the C-terminal ChBRs mediate the dimerization of the LytA autolysin, being this dimerization crucial for catalytic activity (Usobiaga *et al*., 1996; Fernández-Tornero *et al*., 2001).

The LytA-like pneumococcal phage amidases Ejl and Pal can exist in solution in a monomer↔dimer equilibrium (and additionally ↔ tetramer in case of Ejl) depending on the choline concentration in the media, with high concentrations favoring the dimeric (and tetrameric Ejl) state (Sáiz *et al*., 2002; Varea *et al*., 2004). The particular ability of Ejl to tetramerize upon substrate binding was proposed as a mechanism to compensate the relatively low affinity of the enzyme for choline (Sáiz *et al*., 2002). Our results do not exclude the possibility of co-existence of the 1Lys170FL : 3CWB170 complex (67.4 kDa) with intermediate molecular species like, for example, 1Lys170FL : 2CWB170 (55.8 kDa) or even 1Lys170FL : 1CWB170 (44.2 kDa). In fact, in the SEC-MALS analysis of Fig. 6, we could observe that elution of the ~70 kDa monodisperse species corresponding to the 1:3 complex was followed by a gradual decrease of the molar mass curve, suggesting the presence of species with lower molar mass. The occurrence of this intermediate species is also compatible with the cross-linking studies (Fig. 7).

We wondered if this particular endolysin CWB domain could be found associated to different CDs, other than the amidase-2 family CD present in Lys170. We performed BLASTP homology searches using the CWB170 primary sequence and, remarkably, the results showed that CWB170-like modules can be found in the C-terminus of PG hydrolases harboring CDs with diverse enzymatic specificities, including lysozymes, glucosaminidases and peptidases, in addition to different families of amidases (Fig. 8). This observation reinforces the independent func-

tional character of this CWB module. Another interesting feature emerging at the DNA level was that all CWB170-like domains analyzed started with a methionine codon preceded by a putative RBS (Fig. 8), suggesting that translation initiation at this internal start sites should be responsible for the independent production of this domain.

While we were preparing this article for publication, Dunne *et al*. (2014) reported on the *E. coli* expression of *Clostridia* endolysins CD27L and CTP1L and the concomitant production of endolysin C-terminal fragments that functioned as a trigger/release factor for these amidases. As for Lys170, these C-terminal fragments essentially correspond to the predicted CWB domains of the endolysins. Moreover, the N-terminal residue of the CD27L fragment was identified as a methionine, whose corresponding codon is preceded by putative RBS (**TGAGGGAGT**TAAA CAG.*ATG*). The CTP1L C-terminal fragment was deduced to start with a valine residue (GTG codon), also preceded by a putative RBS (**AGGGGGA**AGATGAA.*GTG*). Substitution of the initiating Met and Val residues by a proline ceased the production of the C-terminal CWB domains. Strangely, the authors never refer to the hypothesis of CD27L and CTP1L C-terminal fragments being produced from translation initiation at the putative internal start sites. Instead, they propose that the Met and Val residues are critical for an autocleavage event that is responsible for the generation of the C-terminal fragments. The authors also argue that the cleavage mechanism is triggered by one of two possible dimerization modes of the C-terminal fragments and that release of the CWB domain is necessary to activate CTP1L, whereas in the case of CD27L, it might simply facilitate endolysin cell wall penetration.

In the case of endolysin Lys170, we have excluded cleavage as the mechanism generating the independent CWB170 module and proved the functionality of the internal translation site. In addition, in none of the experiments presented in this work, we could obtain evidences for the presence of a 21 kDa N-terminal fragment of Lys170, which would result from a cleavage event at Met_{202}. Such N-terminal endolysin fragment is never shown also in the work of the *Clostridia* endolysins. In contrast to the CD27L endolysin that apparently exhibited fast and continuous cleavage (Dunne *et al*., 2014), independent incubation of Lys170, mLys170 and CWB170 for 3 days at 4°C or at room temperature (RT) did not reveal any obvious alteration in the polypeptides composition, apart from the slight formation of high molecular weight SDS-resistant aggregates at RT (data not shown). The fact that in a great number of PG hydrolases, the putative CWB170-like domain is initiated by an in-frame methionine preceded by a properly spaced RBS (Fig. 8) is also suggestive of the independent production of the C-terminal module by translation initiation at these internal sites. Thus, in our opinion, the autocleavage mechanism proposed to explain the CD27L and CTP1L C-terminal fragments deserves confirmation, namely by studying the effect of inserting a stop codon upstream of the Met_{186} and Val_{195} putative internal starts, respectively.

It is relatively common in phage genomes the existence of fully or partially overlapped genes, some of which encoding lysis proteins that are known to interact. The best studied examples are provided by the *E. coli* phage λ lytic functions. The λ holin *S* gene has a dual translation start that enables the synthesis of the holin and anti-holin functions (Bläsi *et al*., 1989). The last step of λ virus particles release from infected *E. coli* involves the disruption of the host cell outer membrane. This is accomplished by a spanin complex whose components are encoded by genes *Rz* and *Rz1*, with the latter being fully embedded in the + 1 reading frame of *Rz* (Berry *et al*., 2012). Endolysin genes encoding the expected full length and truncated products through alternative translation initiation have been described previously for the staphylococcal and mycobacterial phages 2638A and Ms6, respectively. It was speculated for the 2638A endolysin that interaction between the full length and the truncated product could explain the higher activity observed when the two polypeptides were present (Abaev *et al*., 2013). The two products of the Ms6 endolysin gene were shown to be necessary for the normal timing, progression and completion of host cell lysis during phage infection, but the possible interaction between the two proteins remains elusive (Catalão *et al*., 2011).

In conclusion, we believe that the enterococcal endolysin Lys170 represents the first described example of a class of PG hydrolases, whose optimal lytic activity depends on the assembly of independent CWB subunits at the corresponding module of the full-length monomer. We speculate that this may constitute a new strategy of increasing the number of CWB motifs in these enzymes, as an alternative to the CWB tandem repetition commonly found in monomeric PG hydrolases. Considering the data of Fig. 8 and the discussion above about the *Clostridia* endolysins, we think that this class of lytic enzymes might be more widespread than previously anticipated.

Experimental procedures

Bacteria, plasmids, phage and growth conditions

E. coli strains XL1-Blue and XL1-Blue MRF' (Stratagene), used for plasmid isolation and propagation, were grown at 37°C in Luria Bertani (LB) medium (Sambrook and Russell, 2001). The *E. coli* expression strain CG61 (São-José *et al*., 2000) and its derivatives were grown in LB at 28°C before induction of protein production and at 37°C afterward. Protein production was induced by heat shock at 42°C in a water-bath with shaking for 30 min. When necessary, LB was supplemented with ampicillin (100 μg ml^{-1}), kanamycin (30 μg ml^{-1}) and/or tetracycline 10 μg ml^{-1}. The *E. faecalis* strains 926/05 and 1518/05 (Proença *et al*., 2012) were grown in Trypton Soy Broth (TSB). All culture media components were purchased

from Biokar Diagnostics. The expression vector pIVEX2.3d (Roche Applied Science), used for protein overproduction in *E. coli*, allows the expression of cloned genes under the control of the phage T7 ϕ10 promoter and the production of the corresponding proteins C-terminally fused to a hexahistidine tag. Phage F170/08 was propagated in *E. faecalis* 926/05 as described previously (Proença *et al.*, 2012).

General DNA techniques

Phage F170/08 DNA was extracted from CsCl-purified lysates as described by Vinga *et al.* (2012). DNA polymerase KOD hot start master mix (Novagen) was used for high-fidelity polymerase chain reaction (PCR), whereas screenings by PCR were performed with DNA polymerase NzyTaq green 2 × master mix (NZYTech). Extraction of *E. coli* plasmid DNA and purification of PCR products was performed with the commercial kits QIAprep Spin Miniprep (QIAGEN) and High Pure PCR Product Purification (Roche Applied Science), respectively, following the manufacturers' instructions. Restriction endonucleases and T4 DNA ligase were from Fermentas Molecular Biology Tools (Thermo Scientific). DNA restriction, ligation and conventional agarose gel electrophoresis were carried out essentially as described by Sambrook and Russell (2001). Development of competence and transformation of *E. coli* strains was according to the method of Chung *et al.* (1989). All recombinant plasmids were confirmed by DNA sequencing (Macrogen, Seoul, Korea).

General protein techniques

Sodium dodecyl sulphate polyacrylamide gel electrophoresis was performed as described by Laemmli (1970). Western blotting analysis was carried out basically as described by Renart *et al.* (1979). Endolysin immunodetection was performed using a rabbit anti-Lys170 polyclonal serum (see section *Rabbit immunization with pure Lys170*) as primary antibody, and horseradish peroxidase-conjugated goat anti-rabbit Fc polyclonal antibody (Pierce, Thermo Scientific) as secondary antibody. Antigen/antibody complexes were detected by chemiluminescence using the RapidStep ECL Reagent (Calbiochem). Protein quantification was carried out with the Bradford reagent (Bio-Rad Laboratories) using bovine serum albumin (BSA, Bio-Rad Laboratories) as standard. PageRuler Prestained Protein Ladder (Thermo Scientific) was used as protein marker in SDS-PAGE.

Construction of lys170 *derivatives*

Plasmid pDP2, a pIVEX2.3d derivative carrying gene *lys170*, was described previously (Proença *et al.*, 2012). Gene *lys170* in pDP2 was subjected to site-directed mutagenesis by using the Quick Change II Site directed mutagenesis kit (Stratagene Agilent Technologies), resulting in plasmid pDP3 carrying *mlys170* gene. The introduced nucleotide substitutions eliminated a putative secondary translation start site internal to *lys170* (see text for details). Mutagenic primers were also used to substitute the Met_{170} and Ala_{171} codons of *lys170* by stop and Thr codons, respectively, with the concomitant creation of an *Xbal* restriction site (see text for details). The pIVEX2.3d derivative harboring this mutated gene ($lys170_{STOP}$) was designated pDP5. The DNA segment encoding the endolysin C-terminal region CWB170 was PCR amplified with a primer pair that added *Nde*I and *Xma*I restriction sites to the 5′ and 3′ ends of the coding sequence, respectively. The PCR product was cloned into pIVEX2.3d cut with the referred enzymes, yielding plasmid pDP4.

Production and purification of endolysin polypeptides

Protein overproduction in *E. coli CG61* and subsequent purification by metal chelate affinity chromatography was as previously described (Proença *et al.*, 2012). Peak fractions from the affinity chromatography step were further purified by SEC using a HiLoad 16/600 superdex 75 prep grade column (GE Healthcare Life Sciences), which was equilibrated and run with imidazole-free endolysin buffer (20 mM HEPES-Na, 500 mM NaCl, 1% glycerol and 1 mM DTT, pH 8.0). Fractions containing the purified proteins were pooled, concentrated when necessary and stored at –80°C as small aliquots. Experimentally determined partition coefficients (K_{av}) of proteins were used to estimate Stokes radii and the corresponding relative molecular masses by extrapolation from a plot of Stokes radii of standard proteins versus $(-\log K_{av})^{1/2}$ (Cabré *et al.*, 1989). The column void volume (V_0) was determined with blue dextran 2000 (GE Healthcare Life Sciences). The standard proteins (Bio-Rad Laboratories) were thyroglobulin (molecular mass = 670 kDa; Stokes radius = 8.6 nm), γ-globulin (158 kDa; 4.8 nm), ovalbumin (44 kDa; 2.73 nm), myoglobin (17 kDa; 2.08 nm) and vitamin B12 (1.35 kDa; 0.85 nm) (Cabré *et al.*, 1989; Talmard *et al.*, 2007).

Protein N-terminal sequencing

N-terminal sequencing by the Edman reaction was performed by the Analytical Services Unit, ITQB (Oeiras, Portugal) in a Procise 491 HT Protein Sequencer (Applied Biosystems).

Rabbit immunization with purified Lys170

The service of raising a rabbit polyclonal anti-serum against endolysin Lys170 was purchased to ACIVET, FMV-UTL (Lisbon, Portugal). One New Zealand white rabbit was treated with a total of five subcutaneous injections, where the first one contained 220 μg of purified Lys170 in 1 ml emulsion of Freund's Adjuvant Complete (Sigma-Aldrich) and the remaining four contained each 110 μg of endolysin in 1 ml emulsion of Freund's Adjuvant Incomplete (Sigma-Aldrich). The injections were administered at 2–3 week intervals. The anti-Lys170 reactivity/specificity of the different sera collected from the animal throughout the protocol was analyzed by enzyme-linked immunosorbent assay using HRP-conjugated goat anti-rabbit Fc polyclonal antibody as secondary antibody. Five days after the final boost, the total serum was recovered, aliquoted and stored at –80°C.

Lytic activity of Lys170 and its derivatives

The lytic activity of Lys170, mLys170 and CWB170, alone or in combination, was evaluated by spotting the indicated

protein quantities on a dense lawn of viable target cells, which was prepared as follows. The enterococcal strain 1518/05 was grown overnight at 30°C without aeration, reaching an OD_{600} of approximately 0.8–1.0. Cells were harvested by centrifugation and concentrated 100-fold in fresh TSB. A sample of 300 µl of this bacterial suspension was incorporated in endolysin buffer supplemented with 0.7% agar and poured in a Petri dish. Lysis halos developed during overnight incubation at 37°C. Negative controls were equally prepared by spotting endolysin buffer.

Time course of endolysin production during phage infection

To study the synthesis of Lys170 polypeptides in *E. faecalis* 926/05 during infection by phage F170/08, an exponentially growing culture of the strain was infected with the phage at an input multiplicity of ~2 and incubated at 37°C for 80 min. One milliliter samples were collected every 10 min; cells were pelleted by centrifugation and stored at –80°C. After thawing, cells were resuspended in 40 µl Tris-EDTA buffer (Sambrook and Russell, 2001) supplemented with 2.5 mg ml^{-1} lysozyme, 10 µg ml^{-1} DNase I and 1 × Complete Mini EDTA-free Protease Inhibitor Cocktail (Roche Applied Science) and incubated for 80 min at 37°C for cell lysis. Ten micrograms of total protein from each time point were separated by SDS-PAGE, followed by Western blotting analysis with anti-Lys170 polyclonal serum prepared at 1:10,000 dilution in 1% skim milk in PBS-T (phosphate buffered saline supplemented with 0.02% Tween-20) and HRP-conjugated goat anti-rabbit Fc polyclonal antibody diluted 1:5000 in 3% skim milk in PBS-T.

Binding of endolysin polypeptides to E. faecalis *cells*

Protein samples used in the following experiments were centrifuged (16,000 *g*, 20 min, 4°C) just before their use to ensure elimination of eventual protein aggregates/precipitates. *E. faecalis* strain 1518/05 was grown until OD_{600} 0.5 at 37°C with aeration, pelleted by centrifugation and concentrated 10-fold in endolysin buffer. Samples of 100 µl of this cell concentrate were incubated with 1 µg of endolysin polypeptides (Lys170, mLys170 or CWBD170) for 30 min on ice to minimize cell lysis. The reactions were prepared in microcentrifuge tubes pre-coated with 3% BSA (Sigma-Aldrich) to avoid unspecific protein binding to plastic. Controls were equally prepared with endolysin buffer added instead of target cells. The mixtures were centrifuged (16,000 *g*, 10 min, 4°C), and 10 µl of both supernatant and resuspended pellet fractions (same initial volume) were analyzed by Western blot with anti-Lys170 antibodies as described above.

Size-exclusion chromatography with multi-angle light scattering (SEC-MALS)

Purified Lys170, mLys170 and CWB170 were analyzed using an HPLC-MALS system (Shimadzu), a light scattering detector (mini DAWN TREOS system, Wyatt Technology) and refractive index detector (Optilab T-rEX, Wyatt Technology). A 120 µg sample of each protein was injected in a Superdex 200 10/300 GL Increase column (GE Healthcare Life Sciences) equilibrated in endolysin buffer and run at a flow rate of 0.5 ml min^{-1}. Molar masses of proteins were calculated using ASTRA 6.1 software (Wyatt Technology) using a refractive index increment (dn/dc) value of 0.183 ml g^{-1}.

Protein stoichiometry analysis was performed using the multisignal detection system of SEC-MALS (Nelson *et al.*, 2006). The 'UV extinction from RI peak' method of ASTRA software allows the determination of UV extinction coefficient (ε_p) in units of ml mg^{-1} cm^{-1}. By using the dn/dc value and the UV and RI signals of the protein peaks, experimental ε_p values were calculated and compared with those predicted from the protein amino acid sequences. Theoretical ε_p were computed with ProtParam tool (http://web.expasy.org/protparam/; Gasteiger *et al.*, 2005), which calculates protein extinction coefficients using the Edelhoch method (Edelhoch, 1967), but with the extinction coefficients for Trp and Tyr residues determined according to Pace *et al.* (1995). We have considered the output values assuming reduced Cys residues.

Protein cross-linking experiments

Working solutions of the cross-linking agent bis(sulfosuccinimidyl) suberate (BS^3, Thermo Scientific) were prepared immediately before use to decrease the extent of hydrolysis. BS^3 was first dissolved in ultra-pure water to a final concentration of 20 mM and then diluted to 5 mM in endolysin buffer. Purified Lys170, mLys170 and CWB170 were set to a final concentration of 50 ng $µl^{-1}$ (monomer molar concentrations of 1.5 and 4.3 µM for mLys170 and CWB170, respectively). Protein samples were treated with 250 µM of BS^3 for 30 min at RT, after which the reactions were stopped with 50 mM Tris-HCl at pH 7.5 for 15 min at RT. In the control samples, the cross-linking agent was substituted by endolysin buffer. Two hundred nanograms of each reaction were separated by SDS-PAGE followed by Western blot analysis as described above, except that anti-Lys170 antibodies were diluted 1:40,000. To cross-link Lys170FL and CWBD170 produced during phage F170/08 infection, 5 µg of total protein from time point $t = 60$ min (see section *Time course of endolysin production during phage infection*) were treated with 1 or 5 mM BS^3. Cross-linking conditions were as above and subsequent analysis by Western blot was with a 1:10,000 dilution of anti-Lys170 antibodies.

Bioinformatics analysis

Protein homology searches were carried out with BLASTP (Altschul *et al.*, 1997) using the nonredundant protein sequence database from NCBI (National Center for Biotechnology Information). Protein conserved domains were predicted with NCBI's tool CDD (Marchler-Bauer *et al.*, 2011) and Pfam (http://pfam.xfam.org/). Multiple protein sequence alignments were performed with ClustalW2 (Larkin *et al.*, 2007).

Acknowledgements

D.P. work has been supported through a PhD fellowship (SFRH/BDE/51076/2010) from Fundação para a Ciência e a Tecnologia (FCT, MCTES, Portugal). This work has benefited

from the facilities and expertise of the IMAGIF Structural and Proteomic Biology Unit of the Centre de Recherche de Gif-sur-Yvette (https://www.imagif.cnrs.fr). We thank P. Tavares for the helpful discussions and support.

References

Abaev, I., Foster-Frey, J., Korobova, O., Shishkova, N., Kiseleva, N., Kopylov, P., *et al.* (2013) Staphylococcal phage 2638A endolysin is lytic for *Staphylococcus aureus* and harbors an inter-lytic-domain secondary translational start site. *Appl Microbiol Biotechnol* **97:** 3449–3456.

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25:** 3389–3402.

Berry, J.D., Rajaure, M., Pang, T., and Young, R. (2012) The spanin complex is essential for lambda lysis. *J Bacteriol* **194:** 5667–5674.

Bläsi, U., Nam, K., Hartz, D., Gold, L., and Young, R. (1989) Dual translational initiation sites control function of the λ S gene. *EMBO J* **8:** 3501–3510.

Bustamante, N., Campillo, N.E., García, E., Gallego, C., Pera, B., Diakun, G.P., *et al.* (2010) Cpl-7, a lysozyme encoded by a pneumococcal bacteriophage with a novel cell wall-binding motif. *J Biol Chem* **285:** 33184–33196.

Cabré, F., Canela, E.I., and Canela, M.A. (1989) Accuracy and precision in the determination of Stokes radii and molecular masses of proteins by gel filtration chromatography. *J Chromatogr* **472:** 347–356.

Catalão, M.J., Milho, C., Gil, F., Moniz-Pereira, J., and Pimentel, M. (2011) A second endolysin gene is fully embedded in-frame with the *lysA* gene of mycobacteriophage Ms6. *PLoS ONE* **6:** e20515.

Catalão, M.J., Gil, F., Moniz-Pereira, J., São-José, C., and Pimentel, M. (2013) Diversity in bacterial lysis systems: bacteriophages show the way. *FEMS Microbiol Rev* **37:** 554–571.

Chung, C.T., Niemela, S.L., and Miller, R.H. (1989) One-step preparation of competent *Escherichia coli*: transformation and storage of bacterial cells in the same solution. *Proc Natl Acad Sci USA* **86:** 2172–2175.

Diaz, E., López, R., and García, J.L. (1990) Chimeric phage-bacterial enzymes: a clue to the modular evolution of genes. *Proc Natl Acad Sci USA* **87:** 8125–8129.

Dunne, M., Mertens, H.D., Garefalaki, V., Jeffries, C.M., Thompson, A., Lemke, E.A., *et al.* (2014) The CD27L and CTP1L endolysins targeting *Clostridia* contain a built-in trigger and release factor. *PLoS Pathog* **10:** e1004228.

Edelhoch, H. (1967) Spectroscopic determination of tryptophan and tyrosine in proteins. *Biochemistry* **6:** 1948–1954.

Erickson, H.P. (2009) Size and shape of protein molecules at the nanometer level determined by sedimentation, gel filtration, and electron microscopy. *Biol Proced Online* **11:** 32–51.

Fernández-Tornero, C., López, R., García, E., Giménez-Gallego, G., and Romero, A. (2001) A novel solenoid fold in the cell wall anchoring domain of the pneumococcal virulence factor LytA. *Nat Struct Biol* **8:** 1020–1024.

Fischetti, V.A. (2008) Bacteriophage lysins as effective antibacterials. *Curr Opin Microbiol* **11:** 393–400.

Fischetti, V.A. (2010) Bacteriophage endolysins: a novel anti-infective to control Gram-positive pathogens. *Int J Med Microbiol* **300:** 357–362.

García, J.L., Diaz, E., Romero, A., and García, P. (1994) Carboxy-terminal deletion analysis of the major pneumococcal autolysin. *J Bacteriol* **176:** 4066–4072.

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M.R., Appel, R.D., and Bairoch, A. (2005) Protein identification and analysis tools on the ExPASy server. In *The Proteomics Protocols Handbook*. Walker, J.M. (ed.). New York, NY: Humana Press, pp. 571–607.

Hermoso, J.A., Monterroso, B., Albert, A., Galán, B., Ahrazem, O., García, P., *et al.* (2003) Structural basis for selective recognition of pneumococcal cell wall by modular endolysin from phage Cp-1. *Structure* **11:** 1239–1249.

Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227:** 680–685.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., *et al.* (2007) ClustalW and ClustalX version 2. *Bioinformatics* **23:** 2947–2948.

López, R., and García, E. (2004) Recent trends on the molecular biology of pneumococcal capsules, lytic enzymes, and bacteriophage. *FEMS Microbiol Rev* **28:** 553–580.

Loessner, M.J. (2005) Bacteriophage endolysins – current state of research and applications. *Curr Opin Microbiol* **8:** 480–487.

McGowan, S., Buckle, A.M., Mitchell, M.S., Hoopes, J.T., Gallagher, D.T., Heselpoth, R.D., *et al.* (2012) X-ray crystal structure of the streptococcal specific phage lysin PlyC. *Proc Natl Acad Sci USA* **109:** 12752–12757.

Marchler-Bauer, A., Lu, S., Anderson, J.B., Chitsaz, F., Derbyshire, M.K., DeWeese-Scott, C., *et al.* (2011) CDD: a Conserved Domain Database for the functional annotation of proteins. *Nucleic Acids Res* **39:** D225–D229.

Mesnage, S., Dellarole, M., Baxter, N.J., Rouget, J.B., Dimitrov, J.D., Wang, N., *et al.* (2014) Molecular basis for bacterial peptidoglycan recognition by LysM domains. *Nat Commun* **5:** 4269.

Moscoso, M., Obregón, V., López, R., García, J.L., and García, E. (2005) Allelic variation of polymorphic locus *lytB*, encoding a choline-binding protein, from streptococci of the mitis group. *Appl Environ Microbiol* **71:** 8706–8713.

Nascimento, J.G., Guerreiro-Pereira, M.C., Costa, S.F., São-José, C., and Santos, M.A. (2008) Nisin-triggered activity of Lys44, the secreted endolysin from *Oenococcus oeni* phage fOg44. *J Bacteriol* **190:** 457–461.

Nelson, D., Schuch, R., Chahales, P., Zhu, S., and Fischetti, V.A. (2006) PlyC: a multimeric bacteriophage lysin. *Proc Natl Acad Sci USA* **103:** 10765–10770.

Nelson, D.C., Schmelcher, M., Rodriguez-Rubio, L., Klumpp, J., Pritchard, D.G., Dong, S., and Donovan, D.M. (2012) Endolysins as antimicrobials. *Adv Virus Res* **83:** 299–365.

Oliveira, H., Melo, L.D., Santos, S.B., Nóbrega, F.L., Ferreira, E.C., Cerca, N., *et al.* (2013) Molecular aspects and comparative genomics of bacteriophage endolysins. *J Virol* **87:** 4558–4570.

Pace, C.N., Vajdos, F., Fee, L., Grimsley, G., and Gray, T. (1995) How to measure and predict the molar absorption coefficient of a protein. *Protein Sci* **4:** 2411–2423.

Payne, K.M., and Hatfull, G.F. (2012) Mycobacteriophage endolysins: diverse and modular enzymes with multiple catalytic activities. *PLoS ONE* **7:** e34052.

Proença, D., Fernandes, S., Leandro, C., Silva, F.A., Santos, S., Lopes, F., *et al*. (2012) Phage endolysins with broad antimicrobial activity against *Enterococcus faecalis* clinical strains. *Microb Drug Resist* **18:** 322–332.

Renart, J., Reiser, J., and Stark, G.R. (1979) Transfer of proteins from gels to diazobenzyloxymethyl-paper and detection with antisera: a method for studying antibody specificity and antigen structure. *Proc Natl Acad Sci USA* **76:** 3116–3120.

Ruggiero, A., Tizzano, B., Pedone, E., Pedone, C., Wilmanns, M., and Berisio, R. (2009) Crystal structure of the resuscitation-promoting factor $_{\Delta DUF}$RpfB from *M. tuberculosis*. *J Mol Biol* **385:** 153–162.

Sáiz, J.L., López-Zumel, C., Monterroso, B., Varea, J., Arrondo, J.L., Iloro, I., *et al*. (2002) Characterization of EjI, the cell-wall amidase coded by the pneumococcal bacteriophage Ej-1. *Protein Sci* **11:** 1788–1799.

Sambrook, J., and Russell, D.W. (2001) *Molecular Cloning: A Laboratory Manual*, 3rd edn. New York, NY: Cold Spring Harbor Laboratory Press, Cold Spring Harbor.

Savva, C.G., Dewey, J.S., Moussa, S.H., To, K.H., Holzenburg, A., and Young, R. (2014) Stable micron-scale holes are a general feature of canonical holins. *Mol Microbiol* **91:** 57–65.

São-José, C., Parreira, R., Vieira, G., and Santos, M.A. (2000) The N-terminal region of the *Oenococcus oeni* bacteriophage fOg44 lysin behaves as a bona fide signal peptide in *Escherichia coli* and as a cis-inhibitory element, preventing lytic activity on oenococcal cells. *J Bacteriol* **182:** 5823–5831.

São-José, C., Parreira, R., and Santos, M.A. (2003) Triggering of host cell lysis by double-stranded DNA bacteriophages: fundamental concepts, recent developments and emerging applications. *Recent Res Dev Bacteriol* **1:** 103–130.

São-José, C., Lhuillier, S., Lurz, R., Melki, R., Lepault, J., Santos, M.A., *et al*. (2006) The ectodomain of the viral receptor YueB forms a fiber that triggers ejection of bacteriophage SPP1 DNA. *J Biol Chem* **281:** 11464–11470.

Schmelcher, M., Donovan, D.M., and Loessner, M.J. (2012) Bacteriophage endolysins as novel antimicrobials. *Future Microbiol* **7:** 1147–1171.

Steen, A., Buist, G., Horsburgh, G.J., Venema, G., Kuipers, O.P., Foster, S.J., and Kok, J. (2005) AcmA of *Lactococcus lactis* is an N-acetylglucosaminidase with an optimal number of LysM domains for proper functioning. *FEBS J* **272:** 2854–2868.

Talmard, C., Guilloreau, L., Coppel, Y., Mazarguil, H., and Faller, P. (2007) Amyloid-beta peptide forms monomeric complexes with Cu(II) and Zn(II) prior to aggregation. *Chembiochem* **8:** 163–165.

Uchiyama, J., Takemura, I., Hayashi, I., Matsuzaki, S., Satoh, M., Ujihara, T., *et al*. (2011) Characterization of lytic enzyme open reading frame 9 (ORF9) derived from *Enterococcus faecalis* bacteriophage ϕEF24C. *Appl Environ Microbiol* **77:** 580–585.

Usobiaga, P., Medrano, F.J., Gasset, M., Garciá, J.L., Saiz, J.L., Rivas, G., *et al*. (1996) Structural organization of the major autolysin from *Streptococcus pneumoniae*. *J Biol Chem* **271:** 6832–6838.

Varea, J., Monterroso, B., Sáiz, J.L., López-Zumel, C., García, J.L., Laynez, J., *et al*. (2004) Structural and thermodynamic characterization of Pal, a phage natural chimeric lysin active against pneumococci. *J Biol Chem* **279:** 43697–43707.

Vinga, I., Baptista, C., Auzat, I., Petipas, I., Lurz, R., Tavares, P., *et al*. (2012) Role of bacteriophage SPP1 tail spike protein gp21 on host cell receptor binding and trigger of phage DNA ejection. *Mol Microbiol* **83:** 289–303.

Wong, J.E., Alsarraf, H.M., Kaspersen, J.D., Pedersen, J.S., Stougaard, J., Thirup, S., and Blaise, M. (2014) Cooperative binding of LysM domains determines the carbohydrate affinity of a bacterial endopeptidase protein. *FEBS J* **281:** 1196–1208.

Young, R. (2013) Phage lysis: do we have the hole story yet? *Curr Opin Microbiol* **16:** 790–797.

Young, R. (2014) Phage lysis: three steps, three choices, one outcome. *J Microbiol* **52:** 243–258.