PROTEINS:
Structure, Function, and Bioinformatics

Structural and Functional Analysis of SleL, a Peptidoglycan Lysin Involved in Germination of Bacillus Spores

Journal:	*PROTEINS: Structure, Function, and Bioinformatics*
Manuscript ID:	Prot-00148-2015.R1
Wiley - Manuscript type:	Research Article
Date Submitted by the Author:	19-Jun-2015
Complete List of Authors:	Christie, Graham; University of Cambridge, Department Chemical Engineering and Biotechnology Ustok, Fatma; University of Cambridge, Department of Haematology Chirgadze, Dimitri; University of Cambridge, Department of Biochemistry
Key Words:	Bacillus, Spore, Germination, Peptidoglycan lysin, Crystal structure

Structural and Functional Analysis of SleL, a Peptidoglycan Lysin Involved in Germination of *Bacillus* Spores

Short title: Crystal structures for the *Bacillus* spore lytic enzyme SleL

Fatma Işık Üstok[1, 3], Dimitri Y. Chirgadze[2], and Graham Christie[1*]

[1]Institute of Biotechnology, Department of Chemical Engineering and Biotechnology, University of Cambridge, Cambridge, United Kingdom

[2]Crystallography and Biocomputing Unit, Department of Biochemistry, University of Cambridge, Cambridge, United Kingdom

[3]Current Address: Department of Haematology, Division of Structural Medicine and Thrombosis Research Unit, Cambridge Institute for Medical Research, University of Cambridge, Cambridge, United Kingdom

*Correspondence to: Graham Christie, Institute of Biotechnology, Department of Chemical Engineering and Biotechnology, University of Cambridge, Cambridge, United Kingdom,
Tel (+44) 1223 334-166; E-mail: gc301@cam.ac.uk

Keywords: cortex lytic enzyme, crystal structure

Abstract

A major event in the germination of *Bacillus* spores concerns hydrolysis of the cortical peptidoglycan that surrounds the spore protoplast, the integrity of which is essential for maintenance of dormancy. Cortex degradation is initiated in all species of *Bacillus* spores by the combined activity of two semi-redundant cortex-lytic enzymes, SleB and CwlJ. A third enzyme, SleL, which has *N*-acetylglucosaminidase activity, cleaves peptidoglycan fragments generated by SleB and CwlJ. Here we present crystal structures of *B. cereus* and *B. megaterium* SleL at 1.6 angstroms and 1.7 angstroms, respectively. The structures were determined with a view to identifying the structural basis of differences in catalytic efficiency between the respective enzymes. The catalytic $(\alpha/\beta)_8$-barrel cores of both enzymes are highly conserved from a structural perspective, including the spatial distribution of the catalytic residues. Both enzymes are equipped with two N-terminal peptidoglycan-binding LysM domains, which are also structurally highly conserved. However, the topological arrangement of the respective enzymes second LysM domain is markedly different, and this may account for differences in catalytic rates by impacting upon the position of the active sites with respect to their substrates. A chimeric enzyme comprising the *B. megaterium* SleL catalytic domain plus *B. cereus* SleL LysM domains displayed enzymatic activity comparable to the native *B. cereus* protein, exemplifying the importance of the LysM domains to SleL function. Similarly, the reciprocal construct, comprising the *B. cereus* SleL catalytic domain with *B. megaterium* SleL LysM domains, showed reduced activity compared to native *B. cereus* SleL.

Introduction

Bacterial spores of the genera *Bacillus* and *Clostridium* represent nature's most resilient cells. Formed *via* the cellular differentiation process of sporulation in response to nutrient starvation, they are entirely metabolically dormant, and can persist as survival structures in the environment for perhaps millennia [1]. Spores of a number of species, including *B. cereus* and the closely related *B. anthracis*, are important vectors in the transmission of disease and toxigenesis. However, spores *per se*, including those from pathogenic species, are generally considered to be non-toxic. Instead, spores must first undergo the processes of germination and outgrowth before initiating expression of virulence factors, including toxins [2,3]. The physiological route to spore germination in all *Bacillus* species is initiated by the activation of membrane-bound germinant receptor proteins that are buried deep on the inside of the spore [4]. By traversing the protective outer layers of the spore, germinant molecules (typically amino acids, sugars or nucleosides) interact with their cognate receptors, resulting in the irreversible commitment of the spore to proceed through the series of ion, analyte and water fluxes, and degradative biophysical reactions, that comprise germination [5,6].

A major, but relatively late germination event concerns the enzymatic depolymerisation of the specialised layer of peptidoglycan (PG), or spore cortex, that surrounds the protoplast-containing spore core. The spore cortex serves to maintain the relatively dehydrated status of the spore protoplast during dormancy, presumably by mechanical means [7,8]. Structurally, cortical PG is characterised by an extremely low cross-linking index and by the presence of the muramic-δ-lactam moiety (MδL), where the side-chain of every alternate N-acetyl muramic acid (NAM) sugar is cyclised to form a lactam ring [9]. The MδL moiety serves as a structural feature that permits recognition of only cortical PG, and not the thin innermost layer of germ-cell wall PG that maintains the integrity of the germinating protoplast, by spore PG lysins that are active during germination [10].

Work conducted by several groups has shown that spores of *Bacillus* species generally have a requirement for one of two semi-redundant cortex-lytic enzymes (CLEs), SleB or CwlJ, to initiate degradation of cortical PG during germination [11-13]. SleB, which is present in the spore in a mature form, has been characterised both structurally [14,15] and biochemically [16-18], but its mode of activation has not yet been elucidated [19]. CwlJ, like SleB, is thought to be a lytic transglycosylase [14,19], and although it has yet to be definitively characterised, genetic evidence indicates that it is activated by the efflux of calcium dipicolinate (Ca^{2+}-DPA) from the spore core during germination [20]. A third CLE common to *Bacillus* species, SleL, has been characterised as an N-acetylglucosaminidase, with catalytic activity against cortical fragments generated by SleB and CwlJ activity [21-23]. Despite its characterisation as a cortical fragment lytic enzyme (CFLE), SleL can seemingly initiate cortex hydrolysis in *B. megaterium sleB cwlJ* double mutant spores, as long as YpeB, a non-lytic protein with unknown function, is also present [24]. However, *in vitro* analyses conducted with purified spore PG sacculi indicate that SleL is a CFLE; hence its role in initiating cortical hydrolysis together with YpeB in the absence of SleB and CwlJ *in vivo* is not clear.

What is clear, however, is that orthologous SleL enzymes from different species of *Bacillus* are associated with differing levels of catalytic activity. SleL from *B. cereus* spores, for example, is considerably more efficient in hydrolysing purified cortical fragments *in vitro* than *B. megaterium* SleL [24]. Similarly, analysis of PG fragments released to the germination medium of pathogenic *B. anthracis* and *B. cereus* spores reveals that they contain significant amounts of SleL-derived N-acetylglucosaminidase products [22], whereas germination exudates from non-pathogenic *B. subtilis* and *B. megaterium* spores are characterised by SleL-associated epimerase products [17,25,26]. The latter have not been detected *in vitro*, however, where at least for *B. megaterium* SleL only N-acetylglucosaminidase products are detected [24]. It remains to be determined whether apparent differences in SleL activity between pathogenic

and non-pathogenic spores, which influences the rate at which PG fragments are released into the surrounding germination medium, confers a physiological advantage, such as in the recycling of PG fragments.

In an effort to address some of the questions relating to SleL activity during spore germination, we have solved crystal structures for both *B. cereus* and *B. megaterium* SleL by X-ray crystallography. Comparison of the resultant models, allied to functional analyses conducted with variant and chimeric proteins, has revealed that the most significant structural factor that influences the efficiency of the respective enzymes is associated with the substrate-binding LysM domains as opposed to the catalytic domains.

Materials and Methods

Cloning, protein expression and purification - Genes encoding SleL were amplified by PCR using genomic DNA from *B. cereus* ATCC 10876 (NCBI accession number EEK49693) and *B. megaterium* QM B1551 (NCBI accession number ADE67156). PCR fragments encoding entire open-reading frames were cloned using a combined ligation-independent cloning (LIC) and vector backbone exchange (VBEx) protocol to construct pNZ-SleL derived plasmids for expression in *Lactococcus lactis* DML1 [27], essentially as described previously [19,24]. Intermediate pRE-SleL plasmids were used as templates for site-directed mutagenesis procedures, which were conducted using a QuikChange Lightning Site-Directed Mutagenesis kit (Agilent Technologies, Wokingham, UK). Plasmids with the correct substitutions were identified by DNA sequencing, and then subject to the aforementioned VBEx procedure to create *L. lactis* expression plasmids. Similarly, chimeric SleL proteins comprising SleL domains from *B. cereus* or *B. megaterium* fused to the catalytic domain from the reciprocal species (i.e. Bm-LysM Bc-Cat and Bc-LysM Bm-Cat) were prepared using an overlap PCR technique. Briefly, PCR amplicons encoding the first 100 codons of either SleL orthologue,

which encodes the respective LysM domains, were used as templates alongside additional PCR amplicons encoding the respective catalytic domains. The resultant fusion PCR products were purified, validated by sequencing, and cloned using the LIC/VBEx procedure.

All SleL and variant proteins were expressed in *L. lactis* essentially as described previously [19,24], inducing with nisin (1µg/L) when the optical density (A660 nm) of the cultures reached 0.6, with expression for 6 hours prior to harvesting the cells. Cell pellets were resuspended in 16 ml of ice-cold breakage buffer (20 mM NaH_2PO_4 [pH 7.4], 500 mM NaCl, 1 mM PMSF, and 20 mM imidazole) and then the cells disrupted by four passes through a One-Shot cell disrupter (Constant Systems, Northampton, United Kingdom) operating at 40 x 10^3 lb/in^2. Recombinant proteins containing C-terminal His_{10} tags were purified from clarified cell lysates via Ni^{2+}-nitrilotriacetic acid (NTA) affinity chromatography, followed by tobacco etch virus (TEV S219V) protease cleavage of the His_{10} tag, and gel filtration (Superdex 75, GE Healthcare, Little Chalfont, UK) chromatography. The resultant proteins, all of which contained additional vector-derived MGGGFA and ENLYFQ residues at the respective N- and C-termini, were concentrated to 10 mg/mL by ultrafiltration in 5 mM sodium phosphate (pH 7.0) plus 25 mM NaCl, and stored at -80^{o}C.

Crystallisation - Crystallisation trials were performed using the vapour diffusion sitting-drop technique in 96-well MRC 2-drop crystallization plates (SWISSCI, Wokingham, UK). □□□□nL of the crystallisation screen conditions were mixed with 200 nL of protein solution and set against 70 µL of reservoir using a crystallisation robot (Crystal Phoenix, Art Robbins Instruments, Inc.). A number of crystallisation trials using various crystallisation screening kits were performed, incubated at 19°C and monitored in a Rock Imager 500 (Formulatrix, Inc) automated imaging system. The crystallisation hits were observed in condition A2 of the JCSG crystallisation screen for *B. cereus* SleL and in condition E8 of the PACT

crystallisation screen for *B. megaterium* SleL. Initial crystallisation conditions were then optimised by creating screens with conditions containing varying concentrations of precipitant agents. These optimisation trials were conducted using the vapour diffusion hanging-drop technique in 24-well crystallisation plates (Hampton Research). The final crystallisation condition for the *B. cereus* SleL protein was 0.1 M trisodium citrate buffer, pH 5.5, 20% (w/v) PEG 3,000, whereas for the *B. megaterium* SleL protein the final condition was 0.2 M sodium sulfate, 20% (w/v) PEG 3,350. Optimisation trials for both proteins were conducted at 19^{o}C, with crystals usually appearing after a few days and growing to their maximum size of about (0.4 x 0.1 x 0.1 mm^{3}) in 1 – 2 weeks.

Diffraction Data Collection and Processing - Fully grown crystals of the *B. cereus* SleL and *B. megaterium* SleL proteins were cryo-protected by immersing in a drop containing the crystallisation condition plus 26 % (v/v) ethylene glycol for a few seconds, and then flash-freezing in liquid nitrogen. The X-ray diffraction data collection were performed either at Diamond Light Source, beamline I04-1 (Oxford, UK) for *B. cereus* SleL crystals, or at the European Synchrotron Radiation Facility, beamline ID23-1 (Grenoble, France) for *B. megaterium* SleL crystals. Raw diffraction data were indexed, scaled, and merged using XDS software [28]. Crystals of the *B. cereus* SleL and *B. megaterium* SleL proteins diffracted to a maximum resolution of 1.6Å and 1.7Å and had P1 and $P2_12_12_1$ space groups respectively. The analysis of solvent content of crystals using Matthews Coefficient indicated that there are three molecules of the *B. cereus* SleL protein in the asymmetric unit, whereas the crystals of *B. megaterium* SleL have only one molecule of the protein. The crystallographic data collection statistics are summarised in Table 1.

Crystal Structure Determination, Model Building and Refinement - Both SleL crystal

structures were solved by the Molecular Replacement (MR) method. First, the structure of the *B. cereus* SleL protein was solved using the unpublished but deposited into the Protein Data Bank (PDB) structure of the *B. subtilis* YdhD protein (PDB-ID: 3CZ8) as the MR probe. All MR calculations were performed in PHASER, part of the PHENIX crystallographic software suite [29]. The positions of the three *B. cereus* SleL molecules within the asymmetric part of the unit cell were successfully identified giving the translation factor Z-score of 23.4 and the R-factor of 45.5%. The obtained model was then subjected to several rounds of alternating manual rebuilding performed in molecular graphics software suite COOT [30] and crystallographic refinement calculations with either PHENIX [29] or REFMAC software [31]. The non-crystallographic symmetry restraints were used at the initial stages of the refinement except for the last two rounds of refinement. The R_{cryst} and R_{free} converged to the values of 15.7% and 17.7%, respectively. This structure was then used as the MR probe for the crystal structure solution of *B. megaterium* SleL; the position of one molecule was identified with the translation factor Z-score of 21.6 and the R-factor of 47.8%. The model was also manually rebuilt according to the sequence of *B. megaterium* SleL and refined using the same software described above. The R_{cryst} and R_{free} converged to the values of 16.7% and 18.4%, respectively. The crystallographic statistics and structural validation aspects are shown in Table 1.

Enzyme assays - Enzymatic activity of the various forms of SleL was assessed by incubating purified enzymes with *B. subtilis* spore PG sacculi (purified as described previously [24]), suspended at an optical density (A600 nm) of 0.5 in 5 mM Tris-HCl (pH 7.8) plus 1 mM DTT at 25ºC. Sacculi were digested with recombinant *B. megaterium* SleB (0.5 µM) for 40 min, and then SleL proteins added to the suspensions in 96-well plates at a final concentration of 1 µM. Plates were incubated at 25°C in a PerkinElmer Envision-Xcite

multilabel plate reader, agitated orbitally for 10 s every 2 min, and absorbance measurements recorded at 600 nm over a 2.5 h period. Assays were conducted in triplicate.

PG binding assays - Qualitative analysis of the binding affinity of different SleL proteins for PG was assessed by incubating 0.1 mg of purified recombinant protein in 50 µL reactions containing *B. subtilis* spore sacculi (OD_{600} ~20) resuspended in 5 mM Tris-HCl (pH 7.5). Reaction mixtures were incubated on ice for 30 min, and then centrifuged (15,000 g x 10 min) to separate supernatant (unbound) and pellet (bound) fractions. The pellet fractions were then resuspended in 50 µL of the same buffer. Samples (20 µL) were combined with SDS-PAGE sample buffer, boiled for 10 min, centrifuged, and then analysed on 4-12% (w/v) gradient SDS-PAGE gels stained with colloidal coomassie blue (Life Technologies, Paisley, UK).

Accession codes - The atomic coordinates and structure factors for the reported crystal structures have been deposited with the Protein Data Bank (PDB) under accession code 4S3J for *B. cereus* SleL and 4S3K for *B. megaterium* SleL.

Results

SleL crystal structures - The three dimensional structure of *B. cereus* SleL (hereafter Bc-SleL) was solved by molecular replacement and was refined to 1.6 Å, with an R_{cryst} of 15.7% and R_{free} of 17.7 % (Table 1). Analysed crystals contain three SleL molecules in the asymmetric unit, along with 12 ligand/ion atoms, plus 1336 water atoms. The calculated electron density map allowed largely unambiguous tracing of all three molecules in the asymmetric unit, which were structurally alike i.e. r.m.s.d. values for superposition of molecule A with molecules B and C were 0.74Å and 0.46Å respectively, and 0.52Å for

superposition of molecule B with molecule C. Electron density for the side-chain of K333 was not observed, hence only the C_β of the residue side chain has been modelled. Similarly, vector-derived MGGGF and LYFQ residues were not evident in the density map, presumably owing to structural disorder at the respective N and C-termini. All structural analyses for Bc-SleL were confined to one monomer (molecule B, as the molecule with the lowest overall temperature factor), cartoon and solvent-accessible representations for which are shown in Figure 1. Subsequently, the *B. megaterium* SleL (hereafter Bm-SleL) crystal structure was solved at 1.7Å resolution, also by molecular replacement, and was refined to an R_{cryst} of 16.7% and an R_{free} of 18.4% (Table 1). Bm-SleL crystals contain a single molecule in the asymmetric unit. Excellent electron density was observed for most residues, with the exception of vector-derived MG at the N-terminus, plus the final three residues (TQP) and vector derived ENLYFQ from the C-terminal end of the protein. In a handful of cases electron density was such that only C_β of side chain atoms could be placed with certainty (K259, K261, W262). Additionally, a short apparently unstructured region (W285-S288) lacked any observable electron density (Figure 1). The significance of this region is examined in this work.

The Bc-SleL molecule is roughly kidney shaped, with dimensions of 47 x 70 x 52 Å. The N-terminal region of the protein, encompassing approximately 100 residues, comprises two LysM (lysin motif; pfam01476) peptidoglycan-binding domains (LysM1 [I2 – K50] and LysM2 [G51 - P97]), both of which adopt the βααβ fold that is characteristic of this domain [32]. A long loop that extends down the back of the molecule as oriented in Figure 1(a) connects the second LysM domain to strand β1 of the $(\alpha/\beta)_8$-barrel core that forms the C-terminal catalytic domain of the protein. As with all $(\alpha/\beta)_8$ barrels, the SleL catalytic domain comprises an internal 8-stranded parallel β-sheet, covered by 8 external α-helices [33]. Several of the βα loops that extend from the C-terminal ends of the β-strands contain additional

secondary structural elements that define the geometry of the barrel, which is marked by a large groove that demarcates the face of the barrel into upper and lower lobes. The most notable of these loops, βα-loop 7, forms a structure comprising two antiparallel β-sheets separated by helical and loop regions, which together with the long βα-loop 6, shapes much of the canopy that protrudes from the upper lobe of the barrel as oriented in Figure 1a - c. The lower lobe is formed by βα-loop 2, which contains a 2-strand antiparallel β-sheet, and βα-loop 3, which includes a short α-helix. The base of the barrel, when oriented as in Figure 1a, is formed predominantly by the α3 and α4 helices, with the relatively short α1, α2 and α5 helices forming the sides of the lower lobe. Helices α7 and α8 sit roughly parallel to each other, slightly offset to the barrel axis, forming a platform upon which the rear-mounted LysM domains are positioned.

Bm-SleL has dimensions of 45 x 80 x 54Å, and adopts a very similar fold to the orthologous *B. cereus* protein. Superposition of the two molecules (Figure 1), which share 46% amino acid identity, reveals a high level of structural conservation (r.m.s.d. of 1.2Å for 371 aligned residues, and 0.9Å for 318 aligned main chain atoms of residues restricted to the barrel domains). As with the *B. cereus* SleL molecule, LysM1 (Q2 – I46) is positioned on top of α7, α8 and the mixed helical-β-sheet structure formed by βα-loop 7 when the molecule is oriented as in Figure 1d. However, in contrast with *B. cereus* SleL, the LysM2 domain (P47 – P96) is positioned behind and slightly offset to the right of LysM1, in a markedly elevated position that is essentially in line with the long axis of the molecule.

Structural comparison of the SleL catalytic domain with ChiA - Identification of significant features of the catalytic domain of SleL was achieved principally by structural comparison with other members of the GH-18 family members (pfam00704), in particular *Serratia marcescens* chitinase A (ChiA). Crystallographic analyses of the latter in complex with

oligo-NAG substrates (e.g. PDB accession number 1EHN) have yielded considerable insight to the molecular basis of substrate binding and catalysis in the GH-18 family [34]. Despite sharing limited sequence identity across this region (~17%), superposition of the catalytic $(\alpha/\beta)_8$-barrels of Bc-SleL and Bm-SleL with the ChiA barrel reveals a close structural alignment, with an r.m.s.d. of 2.4Å (over 278 aligned C_α atoms) for Bc-SleL and 2.7Å (over 247 aligned C_α atoms) for Bm-SleL. In contrast to ChiA, which forms a semi-closed tunnel, surface representation of SleL reveals a large groove that traverses the carboxy-terminal end of the β–strands of the barrel, which appears open at both ends (Figure 1c). The centre of the groove, which measures approximately 39Å by 13Å in Bc-SleL, is marked by a cavity, the electrostatic surface potential of which is negative (Figure 2a). Collectively, and by analogy with other $(\alpha/\beta)_8$-barrel enzymes, these sites are proposed to form the SleL substrate binding groove and active site.

The catalytic mechanism of GH-18 family enzymes, which includes chitinases and *N*-acetyl glucosaminidases, is proposed to proceed via a substrate-assisted mechanism. Briefly, as elucidated from analysis of ChiA-oligo-NAG complexes [34], this entails binding of substrates in an extended form in the active site cleft, positioned such that extensive hydrogen bonding and hydrophobic interactions occur at several ligand-binding subsites (from -6 to +2 in ChiA) that accommodate the aldohexose rings of the carbohydrate. Upon binding, the enzyme bends and rotates the substrate in the vicinity of the scissile glycosidic bond, which bridges sugar moieties located at the -1 and +1 subsites. The glycosidic bond is broken upon protonation by a glutamic acid residue that is highly conserved in GH-18 family enzymes, as are aspartate and tyrosine residues that participate directly in the hydrolytic reaction. Multiple sequence alignment reveals that all three residues essential to ChiA activity are conserved in SleL (D217, E219 and Y283 in Bc-SleL, and D216, E218 and Y283 in Bm-SleL) and orthologues of the spore cortex-lytic enzyme YdhD (Figure S1). Structural

comparison, achieved by superposing ChiA with both Bc-SleL and Bm-SleL, reveals that all three residues occupy essentially identical space to the analogous ChiA residues within the active site (Figure 2b). Notably, Bc-SleL D217 and Bm-SleL D216 are pointing towards catalytic E219/218 whereas ChiA D313 is pointing in the opposite direction. However, this residue is known to adopt two conformations in ChiA, a feature that is important to the catalytic mechanism of the enzyme [34]. Substantial loss of activity of Bc-SleL variant proteins bearing D217A, E219Q and Y283F mutations against partially digested spore sacculi, indicates that these residues comprise the catalytic machinery of SleL (Figure 3b).

Having identified positional conservation of probable catalytic residues between SleL and ChiA, our analysis then turned to a comparison of the carbohydrate-binding cleft of the proteins. This was facilitated by the availability of a high resolution ChiA E315Q–$(NAG)_8$ complex, which, in the original study [34], permitted the identification of a large number of typically aromatic and charged residues that contribute to binding of the substrate. The most significant enzyme-substrate contacts observed in the ChiA E315Q-$(NAG)_8$ complex were found to be localised to the +2, +1 and -1 subsites, with weaker interactions extending to the -6 position. However, before considering the significance to carbohydrate binding of residues that are conserved structurally in both substrate-binding clefts, one has to consider that ChiA and SleL have different substrate specificities. Whereas the former hydrolyses $(NAG)_2$ units in a processive manner from the reducing end of chitin, which is a NAG polymer, SleL cleaves PG that contains NAM and MδL moieties in addition to NAG [21]. Consideration of the products generated by SleL activity against spore sacculi partially digested with the lytic transglycosylase SleB – which include tetrasaccharide (MδL-NAG-NAM-NAG) and anhydro-trisaccharide (MδL-NAG-anhydroNAM) derivatives [24] – indicates that NAG must occupy at least the +2 and -1 binding subsites. NAG may also occupy the +4 and -3 positions, although this will depend on the position that products are released from the active

site. Similarly, the MδL moiety must occupy the +1 and potentially -4 subsites, with NAM (or anhydro NAM) at +3 and -2 sites.

Superposition of Bc-SleL with ChiA E315Q-$(NAG)_8$ reveals good structural conservation of residues at the +2 and, in particular, the -1 subsites i.e. those predicted to accommodate NAG moieties (Figure 4). Notably, a residue analogous to Y418, which essentially blocks the binding cleft after the +2 position in ChiA, and against which the reducing end of the sugar docks [34], is absent in both SleL molecules. Instead, I298, which is found at a similar position at the end of α-helix 6 in both SleL proteins, points away from the groove, leaving the binding cleft open at both ends and presumably enabling substrate binding at putative +3 subsites, and beyond. Another key ChiA +2 residue, D391, which is also involved in substrate interactions at the +1 site, is positionally conserved in Bc-SleL (D284). The same position is occupied by E284 in Bm-SleL, whose side-chain points down into the active site of the enzyme. Analysis of the -1 subsite, predicted to accommodate NAG in SleL, and which makes more substrate contacts (eleven) than any other site in ChiA, reveals that virtually all residues are structurally conserved (Figure 4). The -1 position includes all three catalytic residues, and so perhaps it is not surprising that conservation in this region is particularly high.

Of the six residues identified as making contacts at the +1 subsite, which accommodates NAG in ChiA but is predicted to accommodate MδL in SleL, three residues are structurally conserved, including E219, which is catalytic, D284, and M281 (M388 in ChiA). The sulphur atom of the latter is predicted to stabilise reaction intermediates in ChiA catalysis [34] and presumably contributes a similar role to SleL function. Intriguingly, residue W275, whose indole ring lies in the plane of the lower surface of the binding cleft to form a major component of the +1 binding site in ChiA, is replaced by a conserved phenylalanine residue in both SleL proteins (and in YdhD). In contrast to ChiA W275, the aromatic ring of

F180 of Bc-SleL (F179 in Bm-SleL) is perpendicular to the surface of the putative SleL substrate binding cleft and, additionally, is further away from the superposed $(NAG)_8$ substrate. Although speculative, both modifications may be important in facilitating interactions with the MδL moiety. Similarly, whereas ChiA R446 protrudes from the top of the binding cleft into the barrel, where it is in close proximity to the NAG units at the +1 and -1 sites, the analogous SleL residues (F326 in Bc-SleL and Y326 in Bm-SleL) are further removed from the superposed substrate. Instead, Q288 occupies a similar space to ChiA R446 in Bc-SleL and may be involved in mediating MδL binding in the +1 site. Residues in the +1 site immediately adjacent in sequence to the catalytic glutamic acid residue are structurally conserved between ChiA (F316) and Bm-SleL (F219), but not in Bc-SleL (S220), indicating that this position isn't crucial to recognition of MδL.

Analysis of the remaining substrate-binding subsites reveals that the -2 position, predicted to bind NAM in SleL, is poorly conserved between ChiA and SleL. Indeed, none of the three residues that uniquely mediate contacts with NAG in this position in ChiA are conserved structurally in SleL, although both SleL molecules are essentially identical in this region. Similarly, the -3 subsite, which mediates weak contacts with NAG in ChiA, and is predicted to accommodate the same sugar in SleL, is poorly conserved. A number of residues identified as being structurally conserved in the substrate binding cleft, or which occupy prominent positions in the active site, were assessed for their importance to Bc-SleL function by SDM (Figure 3). Activity assays show that in each case examined, the activity of the enzyme is severely impaired (Figure 3b), with non-catalytic T249, V279 and W407 alanine substitutions exerting the greatest deleterious effect upon protein function.

Structural comparison of Bc-SleL and Bm-SleL - A key objective of the current work was to identify structural differences that could account for previously observed differences in Bc-

SleL and Bm-SleL catalytic activity. Bc-SleL, for example, rapidly hydrolyses suspensions of purified and fragmented spore cortical PG, whereas equimolar amounts of Bm-SleL show no apparent activity [24]. Comparison of the Bm-SleL molecule with Bc-SleL reveals three potentially significant structural differences (Figure 5). First, Bm-SleL has a short section of residues (W285-S288), which could not be modelled due to missing electron density in this region. This may represent a disordered part of the molecule or could simply be an artefact of close molecular packing within this region of the crystal lattice. These residues are part of the extended βα-loop 6 that forms part of the top canopy of the substrate-binding cleft (Figure 5). Two residues in this same location in Bc-SleL – W287 and Q288 – are predicted to be surface exposed and in a good position to mediate contacts in the +2 and +1 substrate-binding sites. Hence, putative disorder in this region in Bm-SleL may impact adversely upon substrate binding. Second, one side of the entrance to the Bm-SleL substrate binding cleft may be partially obscured by a prominent asparagine residue (N114) located on βα-loop 1. The same loop is positioned slightly higher up in Bc-SleL, where I115 shapes the entrance to the upper-left side of the groove. The third potentially significant difference between the two structures concerns the topology of the LysM domains, particularly LysM2 (Figure 1f). While both LysM domains adopt a fairly compact position adjacent to each other on the Bc-SleL $(\alpha/\beta)_8$-barrel, LysM2 of Bm-SleL is positioned somewhat distally to the barrel via a long connective loop. Analysis of average B-factors for Bm-SleL reveals a high level of atomic displacement or intrinsic mobility associated with this LysM2 domain relative to the remainder of the molecule (Figure 5), and indeed to the analogous domain in Bc-SleL.

In order to assess the significance of these structural variables to SleL function, a series of variant and chimeric proteins were constructed and tested for activity against partially digested spore sacculi. Analysis of the enzymatic activity of these constructs is revealing (Table 2). As expected, Bm-SleL showed no activity against partially digested PG

substrate, and neither did the variant proteins bearing N114A or the βα-loop 6 E284D, Y287W and S288Q mutations. However, the chimeric Bc-LysM Bm-Cat protein showed a level of activity commensurate with the native Bc-SleL protein. The reciprocal construct, Bm-LysM Bc-Cat, showed reduced activity (40%) compared to native Bc-SleL. Similarly, Bc-SleL bearing βα-loop 6 substitutions (D284E, W287Y, Q288S) also showed reduced activity (67%), indicating collectively that these changes are deleterious to Bc-SleL function. However, the Bc-LysM Bm-Cat protein can tolerate the βα-loop 6 and N114A substitutions, showing 98% activity compared to Bc-SleL. Hence, whereas the Bc-SleL catalytic domain is sensitive to βα-loop 6 substitutions, the Bm-SleL catalytic domain can tolerate changes in this region, at least when combined with the N114A modification.

Structural analysis of SleL LysM domains - Data presented above indicate that inter-species variance in SleL enzymic activity results largely from differences in the PG-binding LysM domains. Accordingly, PG-binding assays employing both native and chimeric proteins were conducted to test this hypothesis. These assays revealed that proteins with Bm-SleL LysM domains have an apparent decreased affinity for cortex compared to proteins with Bc-SleL LysM domains (Figure S2). In an attempt to identify the structural basis for differences in PG-binding affinity, all four SleL LysM domains were superposed, revealing that they are highly conserved structurally, despite sharing relatively low sequence identity (18 - 41%) (Figure 6a). Root-mean-square deviation values for backbone atoms of the overlapping amino acids range from 0.98Å over 45 aligned residues (Bc-LysM1 versus Bc-LysM2) to 1.35Å over 43 aligned residues (Bm-LysM1 versus Bm-LysM2). Similarly, all four SleL LysM domains align well with the single LysM domain from the *B. subtilis* spore protein YkuD (PDB-ID: 1Y7M; r.m.s.d. 1.39Å for 43 aligned residues with Bc-LysM1, which shares 19% identity). The LysM2 domain from the *Arabidopsis thaliana* plant immune receptor

protein, CERK1, aligns well also (PDB-ID: 4EBZ; r.m.s.d. 1.61Å for 47 aligned residues with Bc-LysM1, sharing 17% identity), although the antiparallel β strands are longer in this protein (Figure 6a).

Superposition of the AtCERK1 LysM2-$(NAG)_5$ complex (PDB accession code 4EBZ) [35] with the four SleL LysM domains enabled structure-based analyses for potential SleL-LysM substrate binding sites (Figure 6b). The best-defined candidates bind NAG3 of the superposed substrate (NAG5 is not visible in the electron density), where the downward pointing acetamido group is accommodated in a negatively charged cavity that is located in an identical position in the putative substrate-binding groove of all four SleL LysM domains. This cavity is partially shaped by a conserved valine residue at the beginning of LysM βα loop-1, and also by a conserved leucine residue located in αβ loop-2 of the respective LysM domains (Figure 6c). The valine residue is positionally conserved in AtCERK1 LysM2, but not the leucine residue, which despite being aligned in the multiple sequence alignment is located elsewhere on the loop; instead an isoleucine residue adopts the same space. A second cavity predicted to bind the acetamido group of NAG1 of the superposed substrate is visible in Bc-SleL LysM2, for example, but in general these cavities are less well defined and appear less sterically favourable to accommodate the substrate. Since the LysM domains of SleL have been shown to have a role in recognition of PG containing MδL [23], then presumably other features of the substrate-binding groove are responsible for binding this moiety. However, the greatest structural variability in the superposed LysM domains is associated with αβ loop-2, which forms part of the binding groove, precluding at this stage putative identity of residues involved in MδL recognition.

Lastly, having identified the probable LysM substrate-binding grooves, superposition of the entire Bc-SleL and Bm-SleL molecules permitted analysis of the global positioning of the PG-binding sites. These analyses reveal that the predicted substrate binding grooves of

the LysM1 domains of both SleL orthologues are structurally aligned with respect to the respective catalytic barrels. In contrast, the LysM2 domains sites have very different topologies with respect to their catalytic domains, resulting in substrate-binding sites with different orientations. Given that LysM domains appear to function cooperatively in positioning the active sites of enzymes to their substrates [32,36-38], it seems likely that the differing topological arrangements between LysM1 and LysM2 in orthologous SleL molecules may well impact upon catalytic efficiency. However, further dissection of the affinities of the individual SleL LysM domains for cortical PG will be required to permit an assessment of the level of this impact on the function of the respective enzymes.

Discussion

The current work presents crystal structures for the *Bacillus* spore cortex lytic enzyme SleL. A major objective of the work was to identify structural features that account for previously observed differences in catalytic activity between *B. cereus* (and *B. anthracis*) SleL and the orthologous *B. megaterium* enzyme [17,21,22,24]. Structural analysis has revealed that the catalytic $(\alpha/\beta)_8$-barrel cores of both the *B. cereus* and *B. megaterium* enzymes are highly conserved, including the positioning of the three residues that participate directly in the catalytic mechanism. Similarly, structural comparison with *Serratia marcescens* ChiA bound to $(NAG)_8$ substrate has revealed that putative SleL subsites for binding spore PG substrates are highly conserved at the +2 and -1 positions, which are predicted to mediate contacts with NAG moieties. However, as might be expected, the +1 and -2 subsites, predicted to accommodate MδL and NAM respectively, are less well conserved. Accordingly, a number of residues have been identified in the putative +1 subsite that may be involved in recognition of MδL, which is essential to ensure that the enzyme degrades only cortical PG during spore germination. However, definitive identification of residues involved in shaping the MδL

binding pocket will require high resolution structural analysis of enzyme-substrate complex(es) supported by appropriate mutagenesis analyses.

Functional analyses conducted with various SleL chimeric proteins have revealed that the major structural determinant in Bm-SleL's relative inactivity against spore-PG substrates is associated with one or both of the PG-binding LysM domains. Indeed, the Bc-LysM Bm-Cat chimeric protein shows enzymatic activity at levels comparable to the native Bc-SleL protein. Despite identifying two regions near the active site that could conceivably have accounted for the relatively poor efficiency of Bm-SleL – namely the presence of N114, which protrudes at one end of the substrate binding cleft, and a short potentially disordered region of βα-loop 6 – data from the chimeric protein confer strong evidence that Bm-SleL's catalytic domain is not inherently impaired. It is compromised instead by one or more of its LysM PG-binding domains, which collectively appear to have reduced affinity for cortical PG compared to the Bc-SleL LysM domains. Whether this reduced affinity results from a localised structural defect in one or more of the LysM domains, or whether it stems from a sub-optimal topological relationship with the catalytic domain, has not been ascertained. However, structural comparison does not reveal any obvious defect in either Bm-SleL LysM domain, so it is perhaps more likely that LysM topology is the dominant structural factor in this case.

The presented functional data, considered with the high degree of structural conservation observed between Bc-SleL and Bm-SleL, would appear to discount any remaining notion that Bm-SleL is directly responsible for epimerase products that are generated during germination of *B. megaterium* (and *B. subtilis*) spores [17,25,26,39]. Instead, it seems more likely that an as yet unidentified enzyme has epimerase activity, whose function and or localisation is dependent on the presence of SleL within the spore. Similar dependencies for the localisation of PG modifying enzymes and accessory proteins are

recognised already in spores of *Bacillus* species, including the SleB-YpeB and CwlJ-GerQ pairings [4]. The possibility of a more extensive inter-dependent network of cortex-modifying enzymes that function collaboratively during spore germination is an area that requires further work.

Finally, outstanding questions remain regarding the molecular and cellular mechanisms that permit activation of mature SleL only during spore germination. Clearly if SleL is a strict CFLE, then its activity will be governed by the availability of PG fragments generated by SleB and or CwlJ during spore germination. The presented crystal structures do not, however, immediately reveal any structural basis as to why recombinant SleL should be a CFLE, since there is no obvious impediment to the binding of long (intact) glycan strands within the open-ended substrate-binding cleft. However, a comparison of the binding clefts of SleL and SleB reveals that the SleL cleft is slightly narrower at the closest point between upper and lower lobes (measuring 7.2Å) compared to the narrowest point on the SleB cleft (8.4Å) (Figure S3). The SleB cleft is also relatively straight compared to the more convoluted SleL cleft, so perhaps collectively these features impede binding of sacculus-assembled glycan chains in the latter. Alternatively, the geometric dimensions of the SleL molecule may preclude its access to the interior of the mesh-like PG sacculus, the "pore" size of which may be amenable only to smaller enzymes, such as SleB (Figure S3). Unfortunately, current knowledge of the three-dimensional structure and architecture of spore cortical PG precludes further examination of either of these hypotheses. Insights to these and other questions relating to spore CLE structure-function relationships remain the focus of future research in this area.

Acknowledgements

The crystallographic experiments were performed in the Crystallographic X-ray Facility (CXF) at the Department of Biochemistry, University of Cambridge. DYC was supported by CXF. The authors have no conflicts of interest to declare.

References

1. Nicholson WL, Munakata N, Horneck G, Melosh HJ, Setlow P. Resistance of *Bacillus* Endospores to Extreme Terrestrial and Extraterrestrial Environments. Microbiology and Molecular Biology Reviews 2000;64(3):548-572.

2. Moir A. How do spores germinate? Journal of applied microbiology 2006;101(3):526-530.

3. Setlow P. Spore germination. Curr Opin Microbiol 2003;6(6):550-556.

4. Paredes-Sabja D, Setlow P, Sarker MR. Germination of spores of *Bacillales* and *Clostridiales* species: mechanisms and proteins involved. Trends Microbiol 2011;19(2):85-94.

5. Swerdlow BM, Setlow B, Setlow P. Levels of H+ and other monovalent cations in dormant and germinating spores of *Bacillus megaterium*. Journal of bacteriology 1981;148(1):20-29.

6. Setlow P. Germination of spores of *Bacillus* species: what we know and do not know. Journal of bacteriology 2014;196(7):1297-1305.

7. Gerhardt P, Marquis RE. Spore thermoresistance mechanisms. In: Smith I, Slepecky RA, Setlow P, editors. Regulation of prokaryotic development - structural and functional analysis of bacterial sporulation and germination. Washington, DC: American Society for Microbiology; 1989. p 43-63.

8. Imae Y, Strominger JL. Relationship between cortex content and properties of *Bacillus sphaericus* spores. Journal of bacteriology 1976;126(2):907-913.

9. Warth AD, Strominger JL. Structure of the peptidoglycan of bacterial spores: occurrence of the lactam of muramic acid. Proceedings of the National Academy of Sciences of the United States of America 1969;64(2):528-535.

10. Popham DL, Helin J, Costello CE, Setlow P. Muramic lactam in peptidoglycan of *Bacillus subtilis* spores is required for spore outgrowth but not for spore dehydration or heat

resistance. Proceedings of the National Academy of Sciences of the United States of America 1996;93(26):15405-15410.

11. Heffron JD, Orsburn B, Popham DL. Roles of germination-specific lytic enzymes CwlJ and SleB in *Bacillus anthracis*. Journal of bacteriology 2009;191(7):2237-2247.

12. Ishikawa S, Yamane K, Sekiguchi J. Regulation and characterization of a newly deduced cell wall hydrolase gene (*cwlJ*) which affects germination of *Bacillus subtilis* spores. Journal of bacteriology 1998;180(6):1375-1380.

13. Setlow B, Peng L, Loshon CA, Li YQ, Christie G, Setlow P. Characterization of the germination of *Bacillus megaterium* spores lacking enzymes that degrade the spore cortex. Journal of applied microbiology 2009;107(1):318-328.

14. Jing X, Robinson HR, Heffron JD, Popham DL, Schubot FD. The catalytic domain of the germination-specific lytic transglycosylase SleB from *Bacillus anthracis* displays a unique active site topology. Proteins 2012;80(10):2469-2475.

15. Li Y, Jin K, Setlow B, Setlow P, Hao B. Crystal structure of the catalytic domain of the *Bacillus cereus* SleB protein, important in cortex peptidoglycan degradation during spore germination. Journal of bacteriology 2012;194(17):4537-4545.

16. Boland FM, Atrih A, Chirakkal H, Foster SJ, Moir A. Complete spore-cortex hydrolysis during germination of *Bacillus subtilis* 168 requires SleB and YpeB. Microbiology (Reading, England) 2000;146 (Pt 1):57-64.

17. Christie G, Ustok FI, Lu Q, Packman LC, Lowe CR. Mutational analysis of *Bacillus megaterium* QM B1551 cortex-lytic enzymes. Journal of bacteriology 2010;192(20):5378-5389.

18. Heffron JD, Sherry N, Popham DL. In vitro studies of peptidoglycan binding and hydrolysis by the *Bacillus anthracis* germination-specific lytic enzyme SleB. Journal of bacteriology 2011;193(1):125-131.

19. Li Y, Butzin XY, Davis A, Setlow B, Korza G, Ustok FI, Christie G, Setlow P, Hao B. Activity and regulation of various forms of CwlJ, SleB, and YpeB proteins in degrading

cortex peptidoglycan of spores of *Bacillus* species in vitro and during spore germination. Journal of bacteriology 2013;195(11):2530-2540.

20. Paidhungat M, Ragkousi K, Setlow P. Genetic requirements for induction of germination of spores of *Bacillus subtilis* by Ca(2+)-dipicolinate. Journal of bacteriology 2001;183(16):4886-4893.

21. Chen Y, Fukuoka S, Makino S. A novel spore peptidoglycan hydrolase of *Bacillus cereus*: biochemical characterization and nucleotide sequence of the corresponding gene, *sleL*. Journal of bacteriology 2000;182(6):1499-1506.

22. Lambert EA, Popham DL. The *Bacillus anthracis* SleL (YaaH) protein is an N-acetylglucosaminidase involved in spore cortex depolymerization. Journal of bacteriology 2008;190(23):7601-7607.

23. Lambert EA, Sherry N, Popham DL. In vitro and in vivo analyses of the *Bacillus anthracis* spore cortex lytic protein SleL. Microbiology (Reading, England) 2012;158(Pt 5):1359-1368.

24. Ustok FI, Packman LC, Lowe CR, Christie G. Spore germination mediated by *Bacillus megaterium* QM B1551 SleL and YpeB. Journal of bacteriology 2014;196(5):1045-1054.

25. Atrih A, Zollner P, Allmaier G, Williamson MP, Foster SJ. Peptidoglycan structural dynamics during germination of *Bacillus subtilis* 168 endospores. Journal of bacteriology 1998;180(17):4603-4612.

26. Chirakkal H, O'Rourke M, Atrih A, Foster SJ, Moir A. Analysis of spore cortex lytic enzymes and related proteins in *Bacillus subtilis* endospore germination. Microbiology (Reading, England) 2002;148(Pt 8):2383-2392.

27. Geertsma ER, Poolman B. High-throughput cloning and expression in recalcitrant bacteria. Nature methods 2007;4(9):705-707.

28. Kabsch W. Integration, scaling, space-group assignment and post-refinement. Acta crystallographica Section D, Biological crystallography 2010;66(Pt 2):133-144.

29. Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta crystallographica Section D, Biological crystallography 2010;66(Pt 2):213-221.

30. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. Acta crystallographica Section D, Biological crystallography 2010;66(Pt 4):486-501.

31. Murshudov GN, Vagin AA, Dodson EJ. Refinement of macromolecular structures by the maximum-likelihood method. Acta crystallographica Section D, Biological crystallography 1997;53(Pt 3):240-255.

32. Buist G, Steen A, Kok J, Kuipers OP. LysM, a widely distributed protein motif for binding to (peptido)glycans. Molecular microbiology 2008;68(4):838-847.

33. Wierenga RK. The TIM-barrel fold: a versatile framework for efficient enzymes. FEBS letters 2001;492(3):193-198.

34. Papanikolau Y, Prag G, Tavlas G, Vorgias CE, Oppenheim AB, Petratos K. High resolution structural analyses of mutant chitinase A complexes with substrates provide new insight into the mechanism of catalysis. Biochemistry 2001;40(38):11338-11343.

35. Liu T, Liu Z, Song C, Hu Y, Han Z, She J, Fan F, Wang J, Jin C, Chang J, Zhou JM, Chai J. Chitin-induced dimerization activates a plant immune receptor. Science (New York, NY) 2012;336(6085):1160-1164.

36. Steen A, Buist G, Horsburgh GJ, Venema G, Kuipers OP, Foster SJ, Kok J. AcmA of *Lactococcus lactis* is an N-acetylglucosaminidase with an optimal number of LysM domains for proper functioning. The FEBS journal 2005;272(11):2854-2868.

37. Wong JE, Alsarraf HM, Kaspersen JD, Pedersen JS, Stougaard J, Thirup S, Blaise M. Cooperative binding of LysM domains determines the carbohydrate affinity of a bacterial endopeptidase protein. The FEBS journal 2013.

38. Mesnage S, Dellarole M, Baxter NJ, Rouget JB, Dimitrov JD, Wang N, Fujimoto Y, Hounslow AM, Lacroix-Desmazes S, Fukase K, Foster SJ, Williamson MP. Molecular basis

for bacterial peptidoglycan recognition by LysM domains. Nature communications 2014;5:4269.

39. Atrih A, Bacher G, Korner R, Allmaier G, Foster SJ. Structural analysis of *Bacillus megaterium* KM spore peptidoglycan and its dynamics during germination. Microbiology (Reading, England) 1999;145 (Pt 5):1033-1041.

Figure legends

Figure 1. Crystal structures of *B. cereus* and *B. megaterium* SleL. Cartoon representations of *B. cereus* SleL, viewed from the (a) C-terminal face of the $[\alpha/\beta]_8$ barrel, and (b) rotated clockwise by 90°. The SleL PG-binding modules, LysM1 (I2-K50) and LysM2 (G51-P97), in blue and cyan respectively, are positioned on top of the catalytic barrel domain. (c) Molecular surface representation of Bc-SleL, showing the entrance to the putative substrate-binding groove. (d) and (e) Cartoon representations of *B. megaterium* SleL. Whereas the LysM1 domain (Q2–I46) occupies a similar position to the Bc-SleL -LysM1 domain, LysM2 (P47–P96) is evidently displaced with respect to the TIM-barrel compared to the analogous domain in Bc-SleL. (f) Superposition of Bc-SleL (orange) and Bm-SleL (green), viewed from the N-terminal face of the $[\alpha/\beta]_8$ barrels, showing the different global position of the LysM2 domains.

Figure 2. SleL active site. (a) Electrostatic surface potential of the active site of *B. cereus* SleL (generated using PyMOL version 1.6.0). A patch of negative charge localised around the central cavity marks the centre of the putative substrate-binding groove; the predicted catalytic residues, shown in stick form, are positioned nearby. (b) Superposition of the $(\alpha/\beta)_8$-barrel domains of Bc-SleL (orange), Bm-SleL (green) and *S. marcescens* ChiA (PDB 1EDQ; light blue), showing close structural alignment of catalytic residues.

Figure 3. Site-directed mutagenesis of *B. cereus* SleL active site. (a) Location of Bc-SleL residues that were subject to mutagenesis in this study. (b) Effects of defined mutations on Bc-SleL activity against purified spore PG fragments. In all cases examined, lytic activity is compromised severely compared to native Bc-SleL activity, which was set as 100%. Data represent the means ± standard deviations of three independent experiments.

Figure 4. Conservation of residues in the Bc-SleL and ChiA -1 NAG binding site. Superposition of Bc-SleL (orange) with ChiA-$(NAG)_8$ (PDB 1EHN; light blue) reveals a high degree of structural conservation of residues in the -1 substrate-binding subsite. Only the non-catalytic residues are shown for clarity (all three catalytic residues participate in NAG binding at the -1 position in ChiA also). Y418 acts to close the binding cleft in ChiA; there is no equivalent residue in SleL, hence the binding cleft is open at both sides of the molecule.

Figure 5. Front and side-view cartoon representations of Bm-SleL, illustrating major structural differences to Bc-SleL. These include potential structural disorder in the βα-loop 6 region of the $(\alpha/\beta)_8$-barrel domain (residues flanking the region that lacks electron density are shown), residue N114 (shown as sticks), which may restrict substrate access to the binding cleft, and intrinsic mobility associated with LysM2, as indicated by higher than average values of temperature factor in this region (the colouring of the cartoon is according to the temperature factor values, blue – low, red – high values).

Figure 6. Sequence and structural comparison of SleL LysM domains. (a) Superposition of Bc-SleL LysM1 (green) and LysM2 (blue) and Bm-SleL LysM1 (yellow) and LysM2 (magenta), together with the *B. subtilis* YkuD LysM domain (cyan) and the *A. thaliana* CERK1 LysM2 domain (orange). (b) Structural superposition of the NAG_5 moiety from the AtCERK1 LysM2–NAG_5 complex with Bc-SleL LysM1. All four SleL LysM domains have a putative substrate-binding groove that is marked by a small cavity with electronegative surface potential. This feature, by analogy with the AtCERK1 LysM2 structure, may accommodate the acetamido group from a bound NAG moiety, as illustrated in the figure. (c)

Multiple sequence alignment of LysM domains from Bc-SleL, Bm-SleL, YkuD and AtCERK1. Residues within the black box, which are largely conserved, form the cavity described in (b). AtCERK1 residues within red boxes interact with NAG sugars in the bound chitin pentamer. The cartoon on top indicates secondary structures as observed in the Bc-SleL LysM1 domain.

Figure S1. Clustal Omega sequence alignment of SleL, YdhD and ChiA. Residues directly involved in chitin hydrolysis in ChiA, which are conserved in the *Bacillales* spore enzymes SleL and YdhD, are boxed.

Figure S2. SleL peptidoglycan pull-down assay. Purified *B. subtilis* spore sacculi were incubated with 0.05 – 0.1 mg of recombinant SleL proteins, incubated on ice, and then subject to centrifugation to separate supernatant (unbound) and pellet (bound) fractions before analysis by SDS-PAGE. Key: B, bound; U, unbound; T, total (suspension of sacculi and protein).

Figure S3. Structural comparison of SleL and SleB. (a) Surface representations of *B. cereus* SleL (pale green) and the C-terminal catalytic domain of *B. cereus* SleB (light orange), illustrating the difference in size between the two molecules. (b) Substrate-binding clefts for both molecules comprise relatively open channels that traverse the face of the molecules, although the SleL channel is narrower at the closest point between upper and lower lobes. Either feature – relative molecule size or differences in substrate binding-cleft geometry – could potentially account for differences in preferred substrates between SleL and SleB.

Table I Crystallographic data collection and refinement statistics

	B. cereus SleL	*B. megaterium* SleL
Data collection		
Radiation Source	Diamond (UK), I04-1	ESRF (France), ID23-1
Wavelength (Å)	0.9200	0.9795
Space group	P1	$P2_12_12_1$
Cell dimensions:		
a, *b*, *c* (Å)	51.33 62.73 119.88	63.69 81.02 90.75
α, β, γ (°)	92.70 100.99 108.40	90.0 90.0 90.0
Resolution (Å)	50.0 – 1.60 (1.64 – 1.60)[1]	43.63 – 1.70 (1.79 – 1.70)[1]
R_{meas} [2] (%)	6.8 (56.2)	9.8 (80.8)
$\langle I / \sigma(I) \rangle$	18.7 (3.4)	11.6 (2.1)
Completeness (%)	92.7 (78.3)	99.4 (99.3)
Redundancy	7.3 (7.1)	5.0 (5.1)
Number of unique reflections	169,576	51,055
Refinement		
Resolution (Å)	47.81 – 1.60	40.25 – 1.70
Number of reflections used:		
Total	164,510	50,990
R_{free} set	1,942	2,000
R_{cryst} / R_{free} (%)	15.7 / 17.7	16.7 / 18.4
Solvent content, %	48.8	47.6
Number of SleL molecules in asymmetric unit	3	1
Number of non-hydrogen of atoms in asymmetric unit:		
Protein atoms	10,143	3402
Ligand/ion	12	10
Water atoms	1,336	441
B-factor, (Å^2):		
Average	32.1	30.7
Wilson	17.2	17.7
Ramachandran plot analysis, number of residues in:		
Favoured regions, %	97.72	97.88
Allowed regions, %	2.12	2.12
Disallowed regions, %	0.16	0
R.m.s. deviations:		
Bond lengths (Å)	0.009	0.004
Bond angles (°)	1.180	0.864

[1] The statistics shown in parentheses are for the highest-resolution shell.
[2] $R_{meas} = (\Sigma_{hkl} [N/(N-1)]^{1/2} \Sigma_i | I_i(hkl) - I_{mean}(hkl)|) / \Sigma_{hkl} \Sigma_i I_i(hkl)$, where N is redundancy.
[3] $R_{cryst} = \Sigma_{hkl} ||F_{obs}(hkl)| - |F_{calc}(hkl)|| / \Sigma_{hkl} |F_{obs}(hkl)|$
[4] R_{free} is the same as R_{cryst} for a random subset not included in the refinement of 10% of total reflection.

Table II Enzymatic activities of chimeric and variant SleL proteins against partially digested spore PG sacculi[1]

Enzyme	OD_{600} loss (%)	Activity versus Bc-SleL (%)	Standard deviation (%)
Buffer	<1	<1	<1
Bc-SleL	43	100	4
Bm-SleL	<1	<1	<1
Bm-SleL N114A	<1	<1	<1
Bm-SleL Bc-loop 6	<1	<1	<1
Bc-LysM Bm-Cat	43	100	2
Bc-LysM Bm-Cat N114A Bc-loop 6	42	98	4
Bm-LysM Bc-Cat	17	40	10
Bc-SleL Bm-loop 6	29	67	6

[1] SleL proteins (0.5 μM) were incubated with purified *B. subtilis* spore PG sacculi (OD_{600} 0.5) that had been pre-digested with *B. megaterium* SleB (0.5 μM) for 40 min. PG hydrolysis was quantified as the total loss in optical density over a period of 150 min, with the native Bc-SleL value being set to 100 %. Presented data are the mean ±SD of three independent assays.

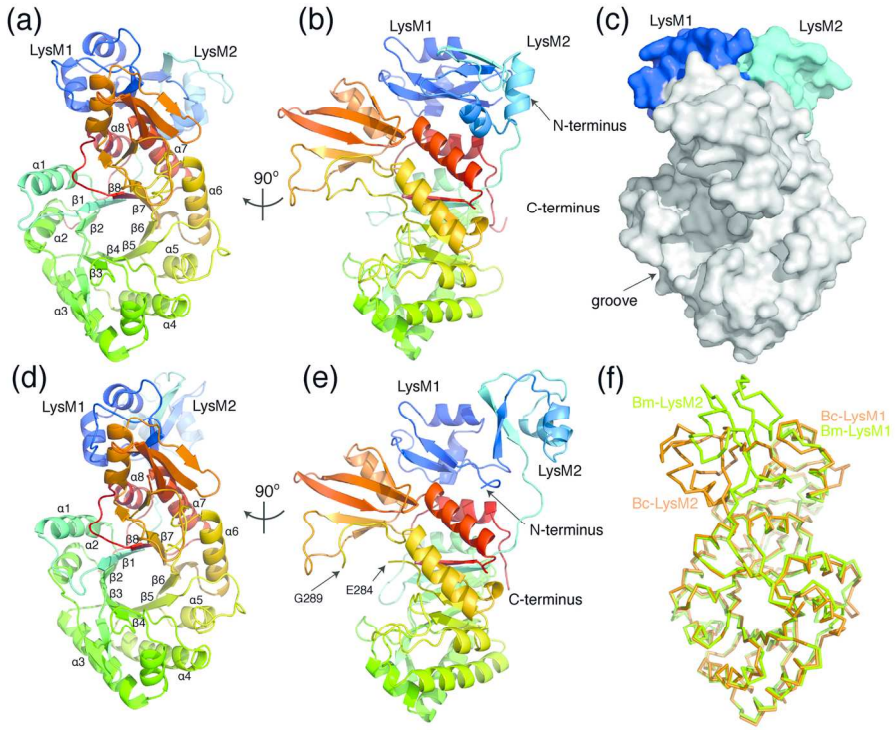

Figure 1. Crystal structures of B. cereus and B. megaterium SleL. Cartoon representations of B. cereus SleL, viewed from the (a) C-terminal face of the [α/β]8 barrel, and (b) rotated clockwise by 90o. The SleL PG-binding modules, LysM1 (I2-K50) and LysM2 (G51-P97), in blue and cyan respectively, are positioned on top of the catalytic barrel domain. (c) Molecular surface representation of Bc-SleL, showing the entrance to the putative substrate-binding groove. (d) and (e) Cartoon representations of B. megaterium SleL. Whereas the LysM1 domain (Q2–I46) occupies a similar position to the Bc-SleL -LysM1 domain, LysM2 (P47–P96) is evidently displaced with respect to the TIM-barrel compared to the analogous domain in Bc-SleL. (f) Superposition of Bc-SleL (orange) and Bm-SleL (green), viewed from the N-terminal face of the [α/β]8 barrels, showing the different global position of the LysM2 domains.
153x111mm (300 x 300 DPI)

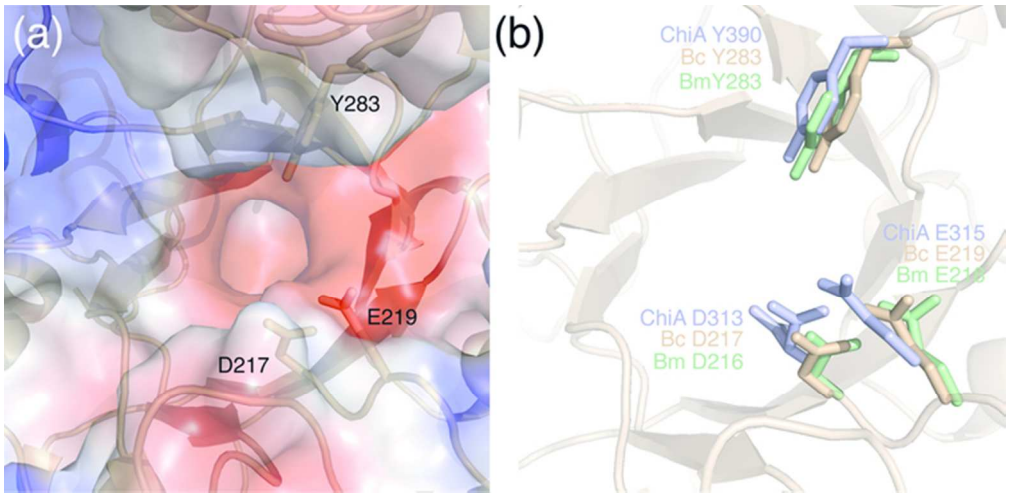

Figure 2. SleL active site. (a) Electrostatic surface potential of the active site of B. cereus SleL (generated using PyMOL version 1.6.0). A patch of negative charge localised around the central cavity marks the centre of the putative substrate-binding groove; the predicted catalytic residues, shown in stick form, are positioned nearby. (b) Superposition of the (α/β)8-barrel domains of Bc-SleL (orange), Bm-SleL (green) and S. marcescens ChiA (PDB 1EDQ; light blue), showing close structural alignment of catalytic residues.
63x30mm (300 x 300 DPI)

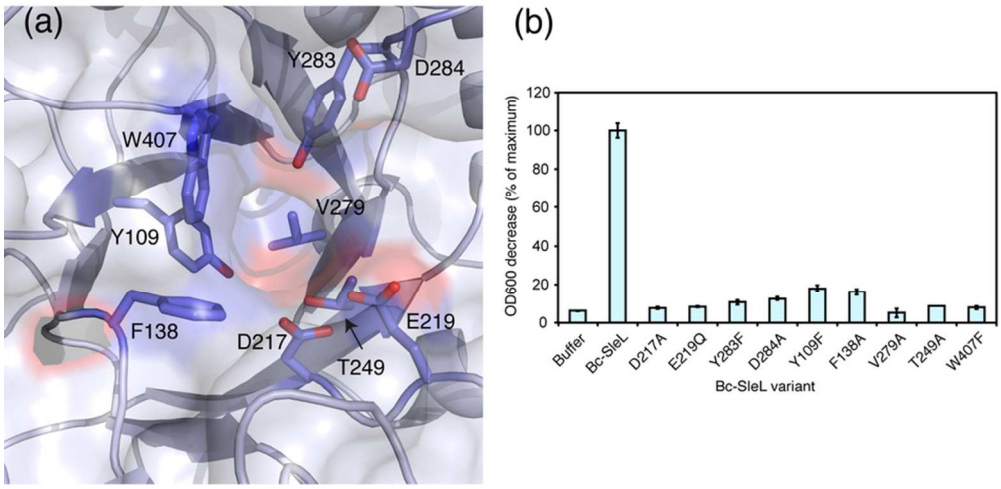

Figure 3. Site-directed mutagenesis of B. cereus SleL active site. (a) Location of Bc-SleL residues that were subject to mutagenesis in this study. (b) Effects of defined mutations on Bc-SleL activity against purified spore PG fragments. In all cases examined, lytic activity is compromised severely compared to native Bc-SleL activity, which was set as 100%. Data represent the means ± standard deviations of three independent experiments.
80x38mm (300 x 300 DPI)

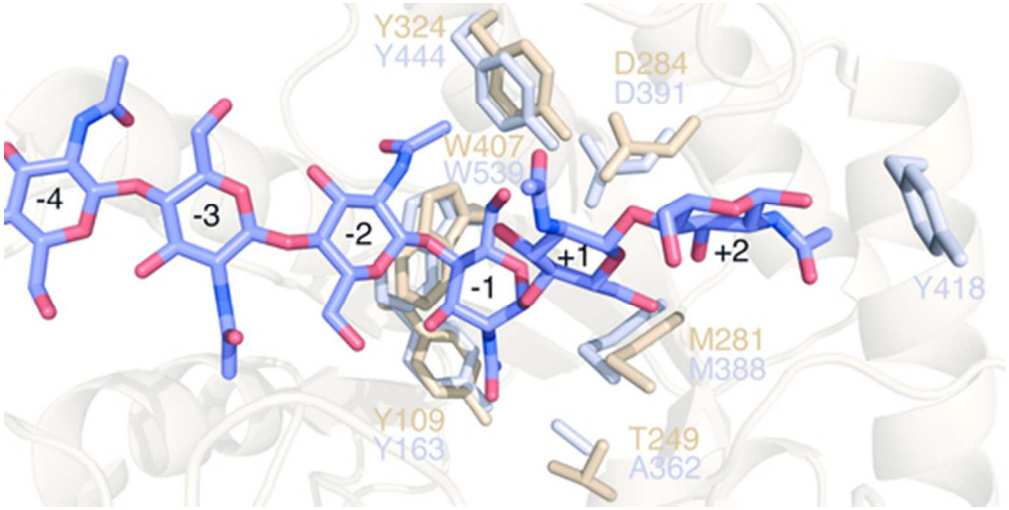

Figure 4. Conservation of residues in the Bc-SleL and ChiA -1 NAG binding site. Superposition of Bc-SleL (orange) with ChiA-(NAG)8 (PDB 1EHN; light blue) reveals a high degree of structural conservation of residues in the -1 substrate-binding subsite. Only the non-catalytic residues are shown for clarity (all three catalytic residues participate in NAG binding at the -1 position in ChiA also). Y418 acts to close the binding cleft in ChiA; there is no equivalent residue in SleL, hence the binding cleft is open at both sides of the molecule.
45x22mm (300 x 300 DPI)

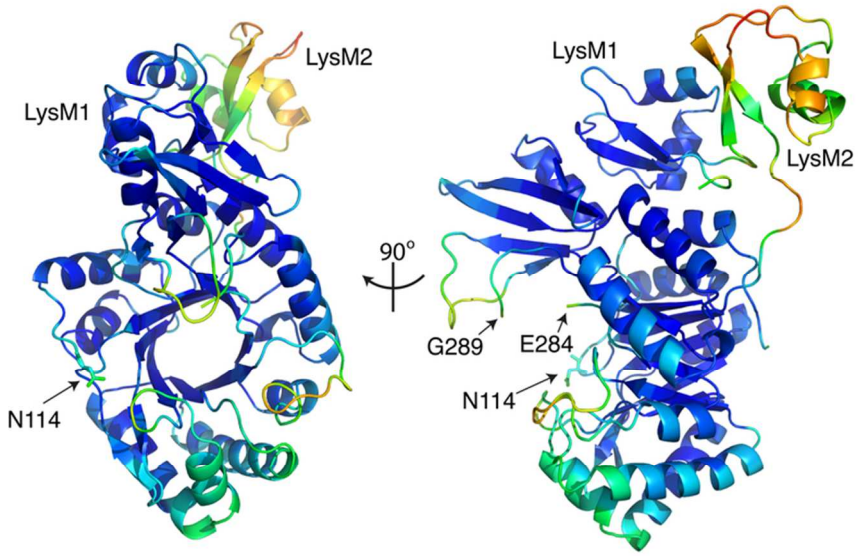

Figure 5. Front and side-view cartoon representations of Bm-SleL, illustrating major structural differences to Bc-SleL. These include potential structural disorder in the βα-loop 6 region of the (α/β)8-barrel domain (residues flanking the region that lacks electron density are shown), residue N114 (shown as sticks), which may restrict substrate access to the binding cleft, and intrinsic mobility associated with LysM2, as indicated by higher than average values of temperature factor in this region (the colouring of the cartoon is according to the temperature factor values, blue – low, red – high values).
77x44mm (300 x 300 DPI)

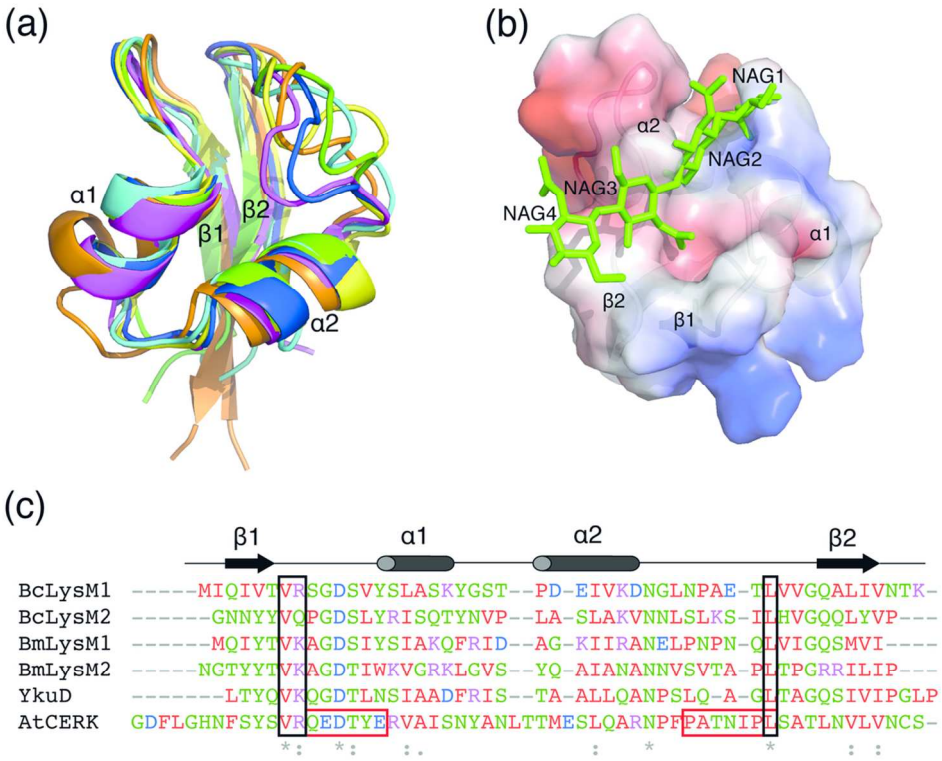

Figure 6. Sequence and structural comparison of SleL LysM domains. (a) Superposition of Bc-SleL LysM1 (green) and LysM2 (blue) and Bm-SleL LysM1 (yellow) and LysM2 (magenta), together with the B. subtilis YkuD LysM domain (cyan) and the A. thaliana CERK1 LysM2 domain (orange). (b) Structural superposition of the NAG5 moiety from the AtCERK1 LysM2–NAG5 complex with Bc-SleL LysM1. All four SleL LysM domains have a putative substrate-binding groove that is marked by a small cavity with electronegative surface potential. This feature, by analogy with the AtCERK1 LysM2 structure, may accommodate the acetamido group from a bound NAG moiety, as illustrated in the figure. (c) Multiple sequence alignment of LysM domains from Bc-SleL, Bm-SleL, YkuD and AtCERK1. Residues within the black box, which are largely conserved, form the cavity described in (b). AtCERK1 residues within red boxes interact with NAG sugars in the bound chitin pentamer. The cartoon on top indicates secondary structures as observed in the Bc-SleL LysM1 domain.
107x81mm (300 x 300 DPI)