online data

Unravelling the structure of the pneumococcal autolytic lysozyme

Begoña MONTERROSO*[1], Consuelo LÓPEZ-ZUMEL*, José L. GARCÍA†, José L. SÁIZ*, Pedro GARCÍA†, Nuria E. CAMPILLO‡[2] and Margarita MENÉNDEZ*[2]

*Instituto de Química-Física "Rocasolano" (CSIC), Serrano 119, 28006 Madrid, Spain, †Centro de Investigaciones Biológicas (CSIC), Ramiro de Maeztu 9, 28040 Madrid, Spain, and ‡Instituto de Química Médica (CSIC), Juan de la Cierva 3, 28006 Madrid, Spain

The LytC lysozyme of *Streptococcus pneumoniae* forms part of the autolytic system of this important pathogen. This enzyme is composed of a C-terminal CM (catalytic module), belonging to the GH25 family of glycosyl hydrolases, and an N-terminal CBM (choline-binding module), made of eleven homologous repeats, that specifically recognizes the choline residues that are present in pneumococcal teichoic and lipoteichoic acids. This arrangement inverts the general assembly pattern of the major pneumococcal autolysin, LytA, and the lytic enzymes encoded by pneumococcal bacteriophages that place the CBM (made of six repeats) at the C-terminus. In the present paper, a three-dimensional model of LytC built by homology modelling of each module and consistent with spectroscopic and hydrodynamic studies is shown. In addition, the putative catalytic-pair residues are identified. Despite the inversion in the modular arrangement, LytC and the bacteriophage-encoded Cpl-1 lysozyme most probably adopt a similar global fold. However, the distinct choline-binding ability and their substrate-binding surfaces may reflect a divergent evolution directed by the different roles played by them in the host (LytC) or in the bacteriophage (Cpl-1). The tight binding of LytC to the pneumococcal envelope, mediated by the acquisition of additional choline-binding repeats, could facilitate the regulation of the potentially suicidal activity of this autolysin. In contrast, a looser attachment of Cpl-1 to the cell wall and the establishment of more favourable interactions between its highly negatively charged catalytic surface and the positively charged chains of pneumococcal murein could enhance the lytic activity of the parasite-encoded enzyme and therefore liberation of the phage progeny.

Key words: choline-binding protein, hydrodynamic particle, modular organization, pneumococcal lysin, secondary structure, structural modelling.

INTRODUCTION

GHs (glycosyl hydrolases) are a widespread group of enzymes that break down glycosidic bonds between carbohydrates. They are classified into over 100 sequence-unrelated families [1], whose members share a common general fold, active-site topology and stereochemical course [2]. The lysozymes encoded by *Streptococcus pneumoniae* [3] and its bacteriophages [4] cleave the β1-4-glycosydic bond between the *N*-acetylmuramoyl-*N*-glucosaminyl residues that form the polysaccharide chain of the bacterial cell wall. They are modular enzymes that harbour a CM (catalytic module) and a CBM (choline-binding module), which generally binds to the choline residues of the pneumococcal teichoic and lipoteichoic acid chains [5]. The CBM is made up of homologous repeats of approx. 20 amino acids (Pfam accession code PF01473). This modular organization is shared between all the pneumococcal murein hydrolases and enhances the enzymatic activity by several orders of magnitude [6]. The catalytic module of the pneumococcal cell wall lysozymes belongs to the GH25 family of GHs (chalaropsis-type lysozymes) [3] whose sequences and structures are unrelated to enzymes from the other three lysozyme families (GH22, GH23 and GH24) [7]. Only two members of the GH25 group have known 3D (three-dimensional) structure, i.e. Cpl-1 from pneumococcal bacteriophage Cp-1 [8] and cellosyl from *Streptomyces coelicolor* [9]. The catalytic residues that are involved in the cleavage of the scissile bond have been identified in the Cpl-1 lysozyme (Asp^{10} and Glu^{94}) [8,10] and appear to be conserved within the GH25 family [8,9]. Their position is compatible with an inverting mechanism in which regeneration of the protonation states of the acid donor (Glu^{94}) and the general base (Asp^{10}) would be facilitated by low-energy hydrogen bonding to residues Asp^{92} and Asp^{182} respectively.

On the other hand, the crystal structures of the Cpl-1 lysozyme and C-LytA, the isolated CBM of the major pneumococcal LytA autolysin, showed that each sequence repeat forms a β-hairpin, followed by a loop and a coiled region [8,11]. The repeats tend to adopt a left-handed β-solenoid, although they can also arrange in an extended β-sheet [8,11]. Choline residues bind at the interface of two consecutive repeats of the β-solenoid in a cavity lined, primarily, by aromatic side chains [8,11].

Cpl-1 and LytC are representative examples of module shuffling. As in most pneumococcal murein hydrolases [5], the CM of Cpl-1 is located at the N-terminus, while, in LytC, it comprises the C-terminal region. Furthermore, the CBM of the former is made up of six homologous repeats (p1–p6) plus a C-terminal tail of 16 residues, while, in LytC, it comprises 11 repeating units (p1–p11), with a lower conservation of the consensus sequence [3]. The sequence differences between LytC and Cpl-1, together with their inverted modular arrangement, yield two lysozymes with very different catalytic efficiencies and temperature-activity profiles [3]. In the present paper, as a first approach towards the elucidation of the LytC structure, we propose a 3D model for the pneumococcal-encoded lysozyme based on homology-modelling strategies. This approach can give very reliable

Abbreviations used: 3D, three-dimensional; CBM, choline-binding module; C-Cpl-1, choline-binding domain of Cpl-1; CM, catalytic module; GH, glycosyl hydrolase(s); TIM, triose phosphate isomerase.

[1] Present address: Section on Physical Biochemistry, Laboratory of Biochemistry and Genetics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892, U.S.A.

[2] Correspondence may be addressed to either of these authors (email nuria@suricata.iqm.csic.es or mmenendez@iqfr.csic.es).

results, even when the sequence identity is lower than 40 %, if the target and the template proteins belong to the same structural family and the homologous proteins of known structures show a high conservation of functional and structural features [12,13]. The proposed model is compatible with the hydrodynamic behaviour and the secondary structure of LytC, which have been characterized, respectively, by analytical ultracentrifugation and by CD and IR spectroscopy. The soundness of the validation tests was checked on the Cpl-1 lysozyme. Despite having an inverted modular arrangement, LytC and Cpl-1 can adopt similar overall structures whose specific features are discussed in relation to their functional roles.

EXPERIMENTAL

Protein purification and chemicals

LytC (55 210 Da) and Cpl-1 (39 249 Da) lysozymes were purified from *Escherichia coli* BL21(DE3)[pLCC14] and DH5α[pCIP100] cells respectively by affinity chromatography on DEAE-cellulose [14], followed by size-exclusion chromatography on a dextran–agarose column (Superdex 200; Amersham Biosciences) in the case of LytC. The purity of samples was analysed routinely by SDS/PAGE. Before use, proteins were dialysed extensively at 4 °C against the appropriate buffer and centrifuged for 5 min at 11 600 ***g***. Protein concentration was determined spectrophotometrically using molar absorption coefficients of 186 107 $M^{-1} \cdot cm^{-1}$ (LytC) and 97 115 $M^{-1} \cdot cm^{-1}$ (Cpl-1) at 280 nm. Choline concentration was measured by differential refractometry [15]. All reagents were from Sigma Chemical Co. and were of analytical grade.

Sedimentation velocity

Sedimentation velocity experiments were performed by centrifugation of 400 μl samples of LytC (2.4 μM) and Cpl-1 (5.2 μM) at 45 000 rev./min and 4 °C in an Optima-XLA analytical ultracentrifuge (Beckman Instruments) using an An50Ti rotor. Radial concentration profiles were taken every 8 min at 280 nm. Sedimentation coefficients (s) were calculated from the ratio of the velocity of the solute boundary (Svedberg program) [16] and corrected to the standard values in water at 20 °C, $s_{20,w}$ [17]. The translational frictional coefficient (f) was determined from the protein molecular mass and the $s_{20,w}$ value. The frictional coefficient of the equivalent hydrated sphere (f_0) was calculated assuming hydration coefficients (∂_w) of 0.41 $g_{water}/g_{protein}$ (LytC) and 0.39 $g_{water}/g_{protein}$ (Cpl-1) estimated from the amino acid sequences [18]. A gross estimation of the hydrodynamic shape of the protein was obtained by generating a family of revolution ellipsoids compatible with the hydrodynamic properties from the f/f_0 ratio [19].

CD

Far-UV CD spectra (average of four scans) were recorded at 4 °C (LytC) or 20 °C (Cpl-1) in a JASCO-810 spectropolarimeter (Jasco), equipped with a Peltier system for temperature control, using 0.02 cm path-length quartz cells and a scan rate of 20 $nm \cdot min^{-1}$. Addition of choline impaired data acquisition below 195 nm. The buffer spectrum was subtracted from the experimental data, and the corrected ellipticities were converted into mean residue ellipticities using average molecular masses per residue of 118.0 (LytC) and 115.8 (Cpl-1). Secondary-structural content was estimated by deconvolution of the experimental curves using the CONTIN (16 protein spectra data set) and CDNN (33 protein spectra data set) programs [20,21] that evaluate four conformations (α-helix, β-sheet, β-turn and remainder), although CDNN discriminates between parallel and antiparallel β-sheet. Structure assignment in CONTIN was based on the Kabsch and Sander method [22].

Fourier-transform IR spectroscopy

Infrared spectra were recorded at 20 °C in a Bruker IFS66 spectrophotometer equipped with a DTGS detector at 4 cm^{-1} resolution using the OPUS program. Protein samples were prepared as follows: 600 μg of protein equilibrated in 10 mM phosphate buffer, pH 8.0, in the absence or in the presence of 20 mM choline, were freeze-dried under a nitrogen stream, then were rehydrated by adding 30 μl of ^2H_2O, and maintained at 20 °C under vacuum for 2 h to ensure isotopic substitution of protein-exchangeable NH groups. Samples were assembled in a thermostatically controlled cell between two calcium fluoride windows with 50 μm spacers. For baselines, the process was repeated using the dialysis buffer. Data analysis (average of four spectra) was carried out following the previously described procedure [23]. Percentages of secondary structure were determined from the relative areas of the amide I band components, assuming that all they have equivalent molar absorbances [23].

Template searching and secondary-structure predictions

The identification of homologues for each module of LytC (Swiss-Prot code Q9Z4J8) in the protein databases was carried out using the Psi-BLAST server [24]. The secondary-structure prediction was performed at the PHD [25], PSIPRED [26] and Jpred [27] servers. Only those structural elements predicted with a reliability above 85.7 % (score $\geqslant 6$; range 0–9) by at least two of them were considered for the sequence-to-structure alignment.

Sequence-to-structure alignment and homology modelling of LytC

Sequence alignment based on the correlation between the secondary structure of the template and the predicted secondary structure of the target was performed with ClustalW [28], followed by a manual adjustment with SeaView [29] guided by the secondary-structure prediction. The alignment was formatted with JOY [30]. In the case of the CBM we used the LALIGN program [31] to compare the CBM of LytC and the template in order to find the best local alignments. Two overlapping fragments of the CBM of Cpl-1, which showed a significant similarity with the CBM of LytC, were thus identified.

3D models of each module containing all non-hydrogen atoms were obtained automatically from the best alignment using the method implemented in MODELLER [32]. The program deduces spatial restraints from the input template(s), subsequently applied to the model of the query protein. 3D models are obtained by optimization of the molecular pdf (probability density function), with simultaneous minimization of input-restraint violations. Initially, 15 models were generated for each input, but only that with the fewest restraint violations and the lowest energy was evaluated. Several cycles of realignment, modelling and validation were repeated until the model showed no further improvement.

The SVRs (structurally variable regions) were built with the standard procedure in MODELLER and they were, afterwards, refined. For short loops, MODELLER procedures yielded good results using the loop parent structures. Long loops were re-modelled using the Search-Loop option in the SYBYL program software suite coupled database (Tripos, St. Louis, MO, U.S.A.). This database contains a large set of high-resolution protein structural fragments. A total of 25 loops clustered in families were identified, and the fragment that showed a lower number of

van der Waals contacts and a higher sequence identity ($\geqslant 25\,\%$) with each loop was selected for remodelling. The final model was subjected to an additional minimization procedure using the Tripos forcefield in order to avoid the unfavourable possible contacts. All models were validated using PROCHECK [34] and VERIFY 3D [35]. JOY and COMPARER [36] were used to generate structure-based alignments between the target and the template in every cycle of the modelling process.

RESULTS AND DISCUSSION

Homology modelling of LytC 3D structure

Template search and multiple sequence alignment

There are only two homologous 3D structures of the CM of LytC available in the PDB (Protein Data Bank), i.e. the catalytic module of Cpl-1 (Swiss-Prot accession number P15057; PDB code 1H09) [8] and cellosyl (Swiss-Prot accession number P25310; PDB code 1JFX) [9], that showed a sequence identity of 21.5 % and 20 % respectively. Figure 1(a) shows the sequential alignment of the target (CM of LytC) with the templates. The irregular $(\alpha/\beta)_5\beta_3$ fold characteristic of cellosyl and the CM of Cpl-1, and the positioning of the potential catalytic residues (Asp^{306} and Glu^{398}) at the end of β_1 and in the loop connecting β_4 to α_4, seem to be conserved in LytC. The reliability of the predicted target structure is supported by the good correlation found between the structural elements present in the 3D structure of Cpl-1 and those predicted by the theoretical methods employed throughout the present work (see Figure 1a).

The search of close homologues for the whole CBM of LytC in several databases was unsuccessful, and only the structures of the CBM of Cpl-1 (C-Cpl-1) and C-LytA (PDB code 1HCX) [11] were considered to be significant hits. The identity was approx. 27 % with both homologues, but the former was used as template, since its 3D structure contains a higher number of repeats. Using the LALIGN program, we identified two fragments of the CBM of Cpl-1, comprising repeats p1–p4 and p1–p6, which showed the best alignment with the N- and the C-terminal regions of the target (Figure 1b). Thus both fragments were used as templates for modelling the CBM of LytC. We found a good correlation between the predicted and the experimental distribution of secondary-structure elements in the CBM of Cpl-1 (Figure 1b). The aromatic residues that are characteristic of choline-binding repeats are conserved in the CBM of LytC, and the predicted secondary structure suggests a conservative folding of the module. Nevertheless, the higher variability in sequence and size of the LytC repeats yield a more irregular distribution of the structural elements. The conservation of the glycine residue in the segment connecting the two β-strands of each repeat will aid them to adopt the antiparallel disposal into the β-hairpin observed in the crystallographic structures [8,11].

3D modelling of LytC catalytic module

Three 3D models of the CM of LytC were built using cellosyl, C-Cpl-1 and both structures as templates. We selected, through the Ramachandran structural validation in the PROCHECK program and VERIFY 3D, the model using the structure of cellosyl as a template for the best model in terms of average stereochemical properties and energy profile. The 3D model showed that 97.9 % of the residues were in allowed regions and only 2.1 % were in the disallowed region. Moreover, the analysis of the template and the model with VERIFY 3D showed that their Z-scores were comparable (0.68 and 0.77 for the model and for cellosyl

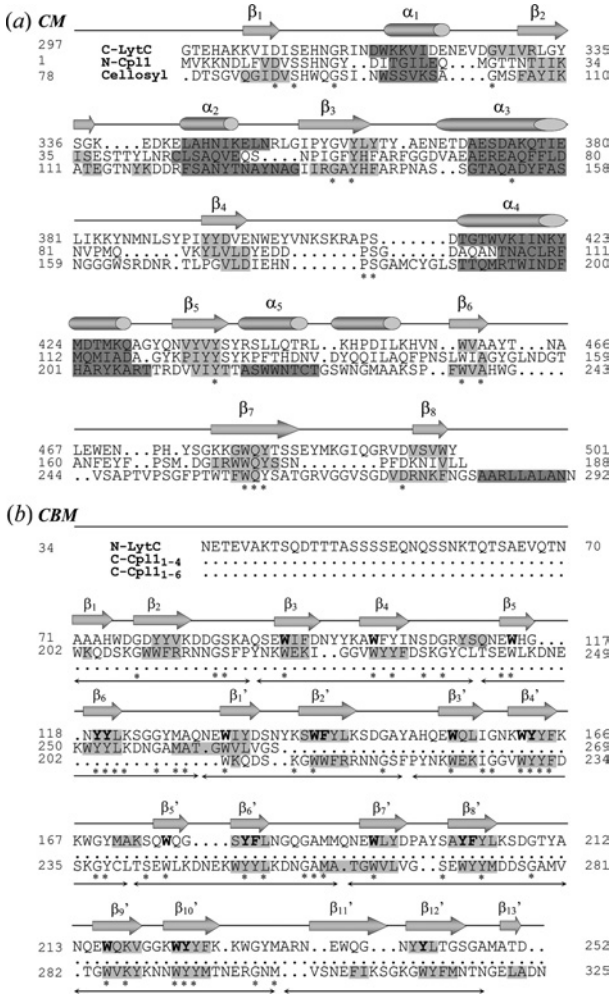

Figure 1 Sequence alignments of LytC modules with their structural templates

(**a**) Sequence alignment of the CMs of LytC (C-LytC), Cpl-1 (N-Cpl1) and the cellosyl muramidase. Numbers at both sides correspond to residue numbers including the signal peptide. Strictly conserved residues are marked with an asterisk (*). Predicted elements of secondary structure in Cpl-1 and LytC are presented as dark (α-helix) and light (β-strands) grey-shaded regions. The same colour code is used for the secondary-structure elements present in the cellosyl 3D structure. Structural elements in the Cpl-1 crystal structure are depicted as cylinders (α-helix) and arrows (β-strands) over the sequence. (**b**) Alignment of the first nine repeats (p1–p9) of the CBM of LytC (N-LytC) with two fragments of the Cpl-1 CBM (C-Cpl1$_{1-4}$ and C-Cpl1$_{1-6}$). Symbols, colour coding and numbers are as in (**a**). Horizontal arrows under the sequence alignment indicate the regions that comprise each repeating unit. Conserved residues potentially involved in choline binding to LytC are in bold letters. β-Strands in the 3D structure of the Cpl-1 CBM are named β_i (fragment p1–p4) and β_i' (fragment p1–p6).

respectively). Remodelling of the loops, focused on the regions structurally different to the template [loops 1 (α_1–β_2), 3 (α_3–β_4) and 5 (α_5–β_6) or potentially involved in functionality loop 4 (β_4–α_5); see below], was performed using the Loop-Search option of SYBYL. No severely disallowed atomic contacts were detected upon re-evaluation of the final model. These results suggest that the proposed model is reasonable and represents an acceptable approximation of the real structure of the CM of LytC. The 3D superimposition of the model and template structures made by COMPARER and formatted by JOY is shown in Supplementary Figure 1(a) (available at http://www.BiochemJ.org/bj/391/bj3910041add.htm). The structural superimposition is made by aligning the pairs of residues selected automatically

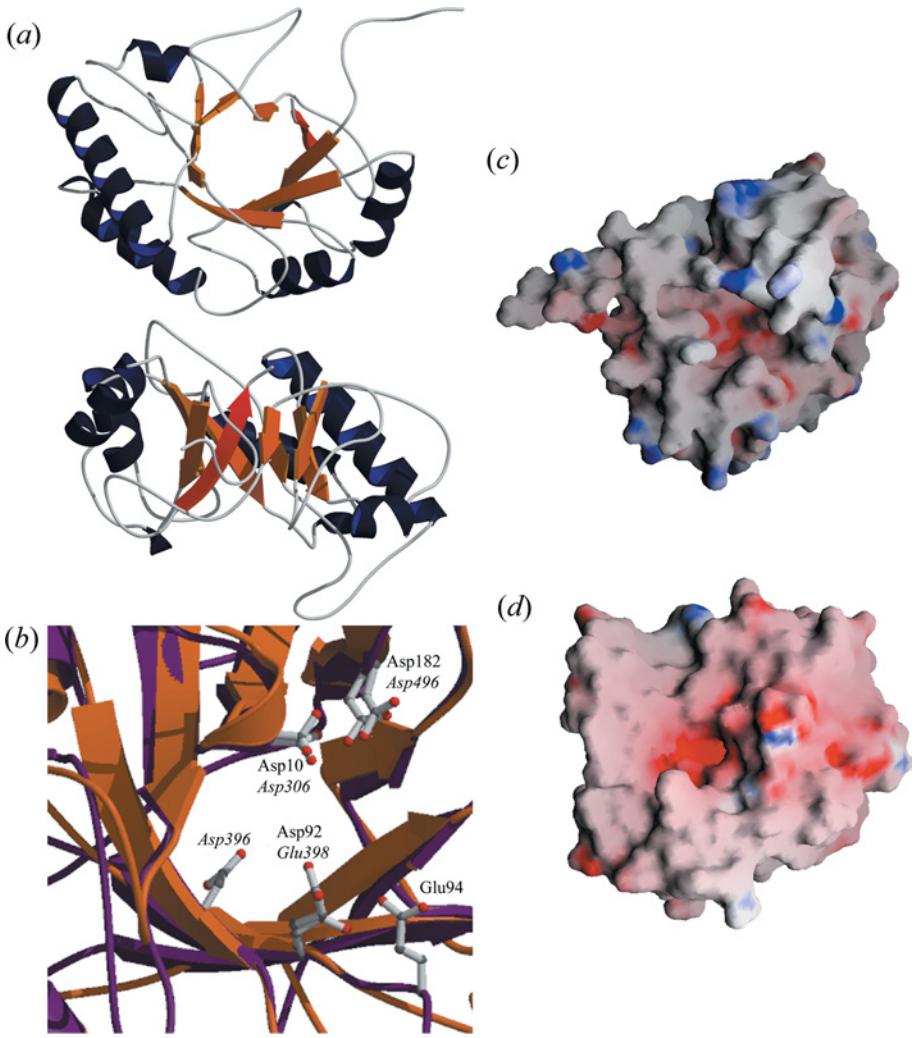

Figure 2 Structural model of the catalytic module of LytC

(**a**) Views from the N-terminal side of the barrel (top) and the region lacking α_6–α_8 helices (bottom). Helices and strands are coloured dark blue and orange respectively; antiparallel β_8 is highlighted in red. (**b**) Superposition of the catalytic cavities of LytC (orange) and Cpl-1 (purple). Catalytic residues of Cpl-1 and putative catalytic residues in LytC are in ball-and-stick representation (C atoms, grey; N atoms, blue; O atoms, red). Labels indicate their position in the sequence of Cpl-1 (roman) and LytC (italic). (**c** and **d**) Surface electrostatic potentials of LytC and Cpl-1 respectively. Negatively charged regions are coloured red, and positive regions are blue.

from the highest scoring local sequence alignment made by COMPARER, which identifies, as equivalent positions in the 3D structures of the model and the template, pairs of residues whose Cα atoms are located less than 3.0 Å (1 Å = 0.1 nm) apart. The structurally corrected alignment generated in this way and the sequence-based one can show some differences in highly variable regions, as actually occurs in the loops of the CM connecting β_4 to α_4 and β_6 to β_7 (see Figure 1a, and Supplementary Figure 1a at http://www.BiochemJ.org/bj/391/bj3910041add.htm).

As shown in Figure 2(a), the model of the CM of LytC exhibits a TIM (triose phosphate isomerase) barrel-like $(\alpha/\beta)_5\beta_3$ fold with the antiparallel orientation of the last β-strand being characteristic of the GH25 family [8,9]. The active site is at the C-terminal side of the β-barrel in a highly negatively charged cavity (Figure 2c). In agreement with sequence alignment (Figure 1a), the pair Asp^{306} and Asp^{496} superimpose with the Cpl-1 pair formed by Asp^{10} and Asp^{182} (Figure 2b), supporting the assignment of Asp^{306} as the general base that activates the nucleophilic water molecule. Although Glu^{398} (the putative donor) is slightly shifted in relation to the position of Glu^{94} in Cpl-1 (Figure 2b), they are both located (like Glu^{177} in cellosyl) at the end of the β_4 strand and can be considered to be structurally conserved. Moreover, the distance of 9 Å between Glu^{398} and Asp^{306} is compatible with an inverting mechanism. The observation that the proton donor in all GHs with a TIM barrel structure characterized so far is a glutamic acid, usually located at the end of the β_4 strand [37,38], gives additional support to the assignment of Glu^{398} as the proton donor. In addition, the position of Asp^{396} is consistent with its potential role as regulator of Glu^{398} protonation state. The shift of the Asp^{396} and Glu^{398} catalytic pair might result from a misalignment of sequences. However, the model derived from the alignment in which Asp^{396} and Glu^{398}, and Asp^{92} and Glu^{94} were superimposed was significantly worse than the proposed model in relation to stereochemistry (7.3 % residues in the disallowed region) and energetic profile.

As observed in cellosyl and Cpl-1, the catalytic cavity is layered by the side chains of conserved aromatic residues that can contribute, as in other carbohydrate-binding proteins, to the substrate specificity [39,40], and the shape of the active site pocket suggests that they could act as exoglycosidases.

3D modelling of the CBM

The 3D modelling of the CBM of LytC (p1–p9 repeats) was based on the alignment shown in Figure 1(b). Repeats p10 and p11 were not modelled because of their poor sequence conservation and the failure of protein-fold recognition methods to find a template for this region. The fact that prediction methods indicate low conservation of secondary structure in both repeats (results not shown) and the high content of polar residues, characteristic of linker regions [41], suggests that these repeats probably act as a linker holding the CM and the CBM of LytC in place.

Figure 3 and Supplementary Figure 1(b) (available at http://www.BiochemJ.org/bj/391/bj3910041add.htm) show respectively the 3D model of the CBM of LytC and the JOY output with the final structural alignment of the model and the templates. The PROCHECK summary of the model showed that 97.4 % residues were in allowed regions and only 1.6 % were in the disallowed region. The average score of 0.67 obtained with VERIFY 3D is comparable with the value of 0.72 of the CBM of Cpl-1. According to the model, the CBM of LytC folds into a β-solenoid fold (Figure 3a), similar to those found in Cpl-1 and C-LytA structures [8,11]. Each repeat comprises a β-hairpin followed by a loop and a coiled region. Up to p8, the hairpin β_n is oriented parallel to $\beta_{(n+3)}$ and is turned approx. 120° around the central axis with respect to the previous hairpin. Beyond p6, the CBM of LytC shows a small curvature, enhanced by a short helix at the end of the hairpin in p8, that modifies the orientation of p9. The choline-binding sites would be located at the interface of two consecutive repeats in the cavity lined by a tryptophan residue from the first strand of the upper β-hairpin and two aromatic side chains provided respectively by the second strand of both repeats [8,11]. The aromatic residues forming the choline-binding sites are structurally conserved, and the hydrophobic pocket is properly formed in the sites from p2 to p9 (Figure 3b). The solvent accessibility of the aromatic residues involved in choline recognition in the model of LytC and in the Cpl-1 structure are comparable, except for Phe^{181} at the interface of p6 and p7 (Figure 1b). Owing to the slight curvature of the CBM of LytC, the β-hairpin of p8 partially occludes the site. Nevertheless, a reliable characterization of the choline-binding sites available in LytC will require direct stoichiometric measurements. The presence of shorter repeats (17 amino acids) in LytC affects the loops connecting the β-hairpins and reduces the angle of 120° formed between two consecutive β-hairpins in the more regular modules, such as those of Cpl-1 and LytA. Thus the alternation of shorter (p3, p6 and p9) and longer repeats leads to an uneven distribution of choline-binding sites along the module surface (Figure 3b).

Overall fold of LytC

Inspection of the LytC structure suggests that the linker connecting the CM and the CBM could have arisen from the degeneration of canonical choline-binding repeats and would be composed of residues from the poorly conserved repeats p10 and p11. A tentative model for the entire molecule of LytC was obtained by structural superimposition of the 3D models of the CM and the CBM on the basis of their spatial orientation in Cpl-1 lysozyme (Figure 3c). The interacting surfaces of catalytic barrel (β_6–β_8) and the CBM (p7–p8 repeats) have a high content of aromatic and aliphatic residues, also observed in the equivalent regions of Cpl-1, but absent in the unimodular cellosyl muramidase, that would contribute to the intermodular interactions stabilizing the overall structure of LytC. Thermal denaturation experiments indicate that LytC is composed of two strongly interacting modules (B. Monterroso, J. L. Saiz, P. García, J. L. García and M. Menéndez, unpublished work). In order to test the proposed 3D model, the secondary structure and the hydrodynamic behaviour of LytC were characterized by CD and IR spectroscopies and analytical centrifugation respectively. The reliability of the analysis was checked by characterizing in parallel the Cpl-1 lysozyme.

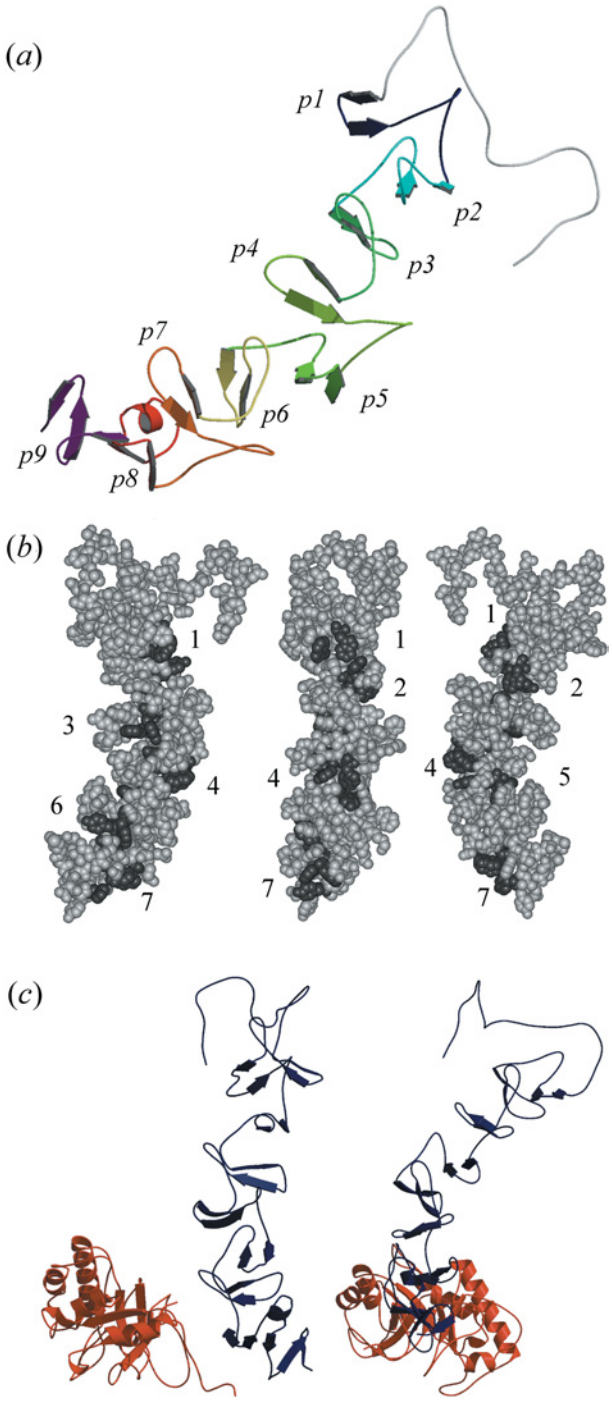

Figure 3 Structural models of the CBM of LytC and the whole enzyme

(**a**) The CBM folds in a β-solenoid structure. β-Hairpins of repeating units appear differently coloured, from dark blue (p1) to purple (p9). (**b**) Disposition of the potential choline-binding sites. Numbers indicate the order of the hydrophobic cavities (dark grey) from the N- to the C-terminal end starting at the p2 repeat. Sites 2 and 5 are located between a short and a large repeat. (**c**) Different views of the model generated for the whole structure of LytC. The region comprising strands β_6–β_8 of the CM (red) interacts with repeats p7–p8 of the CBM (dark blue). The C-terminal side of the barrel, where the catalytic cavity is located, points towards the bottom of the Figure.

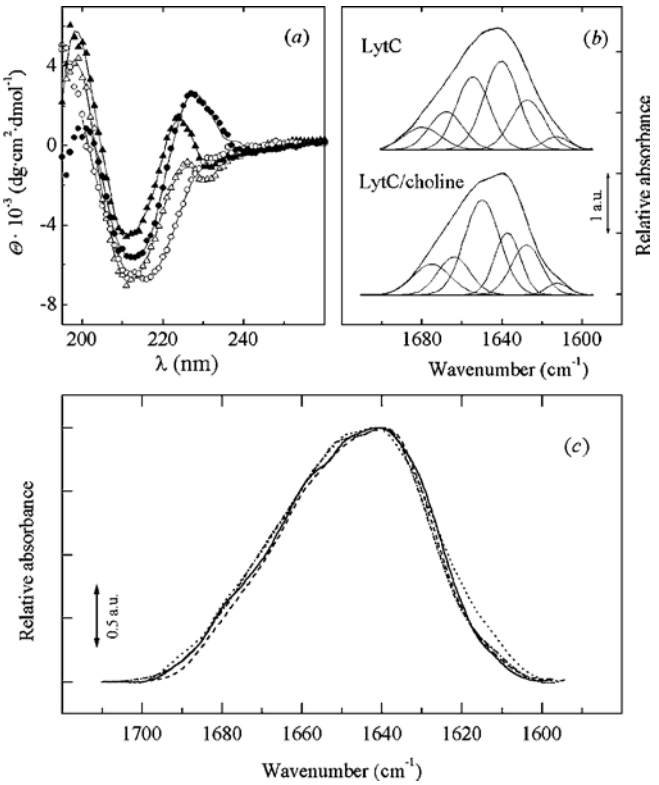

Figure 4 Spectroscopic data of LytC and Cpl-1

(**a**) Far-UV CD spectra of LytC (13.4 μM; circles) and Cpl-1 (21.5 μM; triangles) lysozymes in the absence (open symbols) and in the presence (filled symbols) of 140 mM choline. Data were collected at 4 °C (LytC) and 20 °C (Cpl-1) in 20 mM phosphate buffer, pH 8.0. Solid lines are the theoretical curves calculated by CONTIN. (**b**) Decomposition of LytC amide I band in ^2H_2O. Parameters derived for the band components are shown in Table 2. (**c**) Amide I bands (^2H_2O) of LytC and Cpl-1 in arbitrary units. · · · · ·, LytC;·–·–·–, LytC–choline; ——, Cpl-1; – – –, Cpl-1–choline.

Spectroscopic determination of secondary structure

The far-UV region of the CD spectrum of pneumococcal murein hydrolases is characterized by the presence of a band in the region of 224–230 nm, which is sensitive to the interaction of choline with the tryptophan residues present at the binding site [15,42–44]. Interestingly, the presence of this band in the LytC spectrum is linked to choline binding (Figure 4a). This indicates that, in LytC, choline mediates a reorganization of the polypeptide chain that strongly affects the environment of the aromatic residues of the CBM. Table 1 summarizes the percentages of secondary structure estimated with CONTIN and CDNN, and solid lines in Figure 4(a) are the theoretical fits obtained by CONTIN. In general, there is a good correlation between the results obtained by both procedures. However, CONTIN slightly overestimates the β-turns in Cpl-1 considering the 3D structure [8], and the β-sheet percentage of LytC without choline is also higher than the CDNN estimate. Considering the aggregated state of LytC in the absence of choline (see below), this difference could reflect the specific capability of each method to overcome chiral effects unrelated to protein secondary structure [45]. The mean values from CDNN and CONTIN estimates are 10 % α-helix, 46 % β-strand and 18 % β-turn for the monomer of LytC stabilized in the presence of choline, and 8 % α-helix, 43 % β-strand and 23 % β-turn for Cpl-1 (average values of the free and choline-bound forms).

Table 1 Secondary structure estimated by analysis of far-UV CD spectra of LytC and Cpl-1

Secondary-structure content was estimated using CDNN and CONTIN (values in parentheses) methods from spectra recorded in the absence (–) and presence (+) of choline. Errors in the estimated values are 1–6 % for CDNN and 1–4 % for CONTIN. Values are rounded off to the nearest integer.

	Cpl-1 (%)				LytC (%)			
Structural element	–		+		–		+	
α-Helix	11	(7)	8	(4)	12	(8)	7	(12)
β-Sheet	40	(42)	48	(42)	37	(52)	45	(47)
Antiparallel	35		42		31		40	
Parallel	5		6		6		5	
β-Turn	17	(30)	15	(31)	18	(17)	17	(19)
Random structure	32	(21)	29	(23)	33	(23)	31	(22)

The unambiguous analysis of secondary-structure content may be hampered when the experimentally available CD spectra do not include data down to 185 nm. For this reason, the secondary structure of LytC and Cpl-1 was also characterized by IR spectroscopy. The amide I bands of LytC and Cpl-1 in ^2H_2O present two maxima at 1640 cm^{-1} and 1650 cm^{-1} (Figure 4c). In addition, Cpl-1 also presents a shoulder around 1660 cm^{-1}. Choline binding induces a narrowing of the amide I band of LytC in the region around 1615 cm^{-1}, and a slight decrease in the intensity of the spectra of both enzymes around 1650 cm^{-1}. Figure 4(b) shows the resolved components derived from the analysis of the amide I band of LytC, and the corresponding parameters, i.e. band position and area percentage, for both proteins are shown in Table 2. The assignment of these bands has been carried out previously [23,46]. The bands around 1629 and 1637 cm^{-1} indicate the presence of extended structures [23,46], while those between 1660 cm^{-1} and 1680 cm^{-1} can be attributed primarily to β-turns [46], since the contribution of the high-frequency component of the β-sheet component (1672–1694 cm^{-1}) would be at most 10 % of the low-frequency component around 1630 cm^{-1} [23]. On the other hand, bands around 1642–1660 cm^{-1} are usually assigned to α-helix [23,46], although contributions from long loops [47] or structures with dihedral angles close to those found in the α-helix, such as the 3_{10} helix, have been described around this frequency. Finally, the component centred around 1640 cm^{-1} is due to unordered segments and loops [23]. The presence of a component centred at 1659 cm^{-1} in the spectrum of Cpl-1 without choline could indicate some degree of overlap of the components from α-helix and long loops with those from the β-turns.

The simultaneous presence of bands at 1637 cm^{-1} and 1628 cm^{-1} in the spectra of LytC and Cpl-1 is consistent with the structure of Cpl-1 and the model of LytC. The contribution of the TIM-barrel β-strands would be centred around the values expected for the low-frequency component of large β-sheets (approx. 1629 cm^{-1}), while the CBM double strands would account for the component at 1637 cm^{-1} [46]. Indeed, the appearance of the latter in the amide I band of the LytC–choline complex, with the simultaneous loss of the unordered segments signal at 1640 cm^{-1}, suggests that formation of β-hairpins in the CBM of LytC could be to some degree mediated by the interaction with choline. The small differences found in the components of Cpl-1 with and without choline agree with the rather small variations found between the free and the choline-bound structures of Cpl-1 [8].

Taking together the results with and without choline, and assuming that the contribution of the high-frequency component of β-strands in the region of around 1672–1694 cm^{-1} can increase

Table 2 Band frequencies, fractional areas and proposed structure assignments of the components derived from IR measurements of LytC and Cpl-1

Spectra were recorded in ^2H_2O in the absence (−) or presence (+) of choline (10 mM phosphate, p^{2}H 8.0). Values are rounded off to the nearest integer. Odd-numbered columns are values in cm^{-1}, even-numbered columns are percentages. $\uparrow\upsilon$ and $\downarrow\upsilon$ indicate high- and low-frequency components respectively.

LytC				Cpl-1				
−		+		−		+		Assignment
1680	9	–	–	–	–	1683	1	β-Sheet ($\uparrow\upsilon$) and β-turns
–	–	1675	13	1677	11	1673	12	β-Sheet ($\uparrow\upsilon$) and β-turns
1668	13	1664	13	1659*	33	1660	18	β-Turns
1655	25	1650†	37	–	–	1650†	22	α-Helix, 3_{10} helix and long loops
1640	32	–	–	1646	14	–	–	Unordered structure
–	–	1637	17	1637	24	1635	45	β-Sheet ($\downarrow\upsilon$)
1627	17	1628	17	1629	17	–	–	β-Sheet ($\downarrow\upsilon$)
1613‡	4	1612‡	3	1612‡	1	1614‡	2	

* Maximum position could indicate some overlap with the α-helix component.
† Maximum position could indicate some overlap with the component from unordered structures.
‡ Bands at around 1612 cm^{-1} in native proteins arise from amino acid side chains.

to 10 % of the intensity of the low-frequency absorption, the secondary structure of Cpl-1 appears to be 15 % α-helix (the difference between the intensity of the component around 1660 cm^{-1} in the Cpl-1–choline complex and the free enzyme), 47 % β-strand and 27 % β-turn. Analogously, IR data for the monomer of LytC stabilized by choline-binding are consistent with the following estimations for the secondary-structure composition: 25 % α-helix (intensity of the component at 1655 cm^{-1} in the absence of choline), 37 % β-strand and 23 % β-turn.

It is worth mentioning the agreement between the IR and CD estimations of secondary structure for each protein, and the reasonable correlation found with the percentages derived from either the Cpl-1 3D structure (17 % α-helix, 3 % 3_{10} helix, 12 % parallel β-sheet, 23 % antiparallel β-sheet and 22 % β-turns) or the proposed 3D model of LytC (14 % α-helix, 25 % β-sheet and 25 % β-turns), particularly if we consider that the length of strands and helices in the model could be underestimated.

Hydrodynamic properties of LytC and Cpl-1

The sedimentation velocity profiles of Cpl-1 showed an apparently single boundary both in the absence and in the presence of choline that can be described in terms of a single sedimenting species (results not shown) with $s_{20,w}$ of 3.28 ± 0.01 S (free enzyme) and 4.46 ± 0.01 S (Cpl-1–choline complex). These values are compatible with the molecular masses of the monomer and the dimer of Cpl-1 respectively. Thus choline regulates, as in the amidases with CBM composed of six repeats [42–44], the self-association of the phage-encoded lysozyme, inducing the dimerization of Cpl-1. In contrast, the complete sedimentation of LytC at very short times and sedimentation rates as low as 3000 rev./min revealed the formation of high-order molecular-mass aggregates that dissociate upon choline binding into a single species whose sedimentation coefficient ($s_{20,w}$ of 3.73 ± 0.01 S) is compatible with the monomer molecular mass. Dissociation of LytC aggregates is parallel to the appearance of the band at 224 nm in the CD spectra. The model generated for LytC was tested further by calculating the expected $s_{20,w}$ value from the model co-ordinates using the bead modelling approach implemented in HYDROPRO [48] that yielded an excellent estimate of $s_{20,w}$ for Cpl-1 using the crystallographic co-ordinates (Table 3). The good agreement between the calculated (3.67 S) and the experimental (3.73 S) values of $s_{20,w}$ supports further the proposed model of LytC.

The translational frictional ratios (f/f_0) derived from the sedimentation coefficients (Table 3) clearly deviate from the values

Table 3 Hydrodynamic parameters of LytC and Cpl-1

Measurements were made in 20 mM phosphate buffer, pH 8.0, in the absence (Cpl-1 monomer) and in the presence of 60 mM (Cpl-1 dimer) and 140 mM choline (LytC monomer).

	LytC	Cpl-1	
	Monomer	Monomer	Dimer
$s_{20,w}$ (S)	3.73 ± 0.01	3.28 ± 0.01	4.46 ± 0.01
$s^T_{20,w}$ (S)*	3.67 ± 0.01	3.14 ± 0.01	–
f ($g \cdot s^{-1}$)	6.83×10^{-8}	5.47×10^{-8}	7.82×10^{-8}
f/f_0	1.41	1.27	1.44
a/b (prolate)†	4.6	2.9	5.0
a/b (oblate)†	5.0	3.1	5.5

* Calculated with HYDROPRO from the Cpl-1 structure and the 3D model of LytC.
† Axial ratio between length (a) and diameter (b) of the theoretical ellipsoids.

expected for rigid spherical particles and are compatible with the prolate and oblate ellipsoids with the axial ratios reported in Table 3. Interestingly, the f/f_0 ratio of the choline-bound forms of LytC (monomer) and Cpl-1 (dimer) are very close and suggest that both complexes can have similar hydrodynamic shapes. The hydrodynamic particle model that best agrees with the dimensions of Cpl-1 in the crystal structure [8] is an oblate ellipsoid of 25 Å × 76 Å. Analogously, the dimensions (21 Å × 88 Å) derived from the 3D model of LytC reasonably agree with the oblate ellipsoid model (20 Å × 101 Å).

Structural comparison of Cpl-1 and LytC lysozymes

In spite of having an inverted disposition of their constituting modules, LytC and Cpl-1 seem to adopt a similar global fold. However, there are substantial differences between the 3D model generated for LytC and the crystal structure of Cpl-1, including the electrostatic potential of the catalytic surface (Figures 2c and 2d), the length of the loops at the C-terminal side of the β-strands (related in α/β barrels to substrate recognition [49]), the number of potential choline-binding sites (seven in LytC compared with four in the Cpl-1 dimer) and the shape of the active particle. Moreover, the structural changes promoted by choline in the host-encoded enzyme suggest that the conformation adopted by the CBM of LytC is strongly dependent on choline binding, as revealed CD and ultracentrifugation studies.

In contrast with LytC and cellosyl [9], Cpl-1 exhibits a broad negatively charged catalytic surface (Figures 3c and 3d) that

would favour its interaction with a positively charged substrate (80 % of glucosamine residues of pneumococcal peptidoglycan are deacetylated) [50]. Besides, its looser attachment to the bacterial cell wall (the choline IC_{50} is 15 times higher for LytC than for Cpl-1) will facilitate the diffusion of the protein through the bacterial envelope. These features might account for the higher specific activity of the parasite-encoded enzyme [3], especially if the active species is the Cpl-1 dimer bearing two catalytic centres.

Choline-induced self-association has been observed in all the pneumococcal lysins containing a CBM formed by six repeats located at the C-terminal extreme characterized to date [42–44]. Therefore the acquisition of additional choline-binding motifs by LytC may constitute an alternative adaptation mechanism leading to the enhancement of cell-wall-binding capability that reduces, at the same time, the curvature associated with the dimerization of CBM displayed by the crystal structure of C-LytA [11]. The lack of stabilizing contacts between sequence-distant repeats, a characteristic of the solenoid fold, would enhance the flexibility of the CBM of LytC and therefore its ability to selectively interact with those regions of the bacterial envelope where the complementarity of the target is maximal, providing a tight binding [51]. Indeed, shallow grooves in solenoids with small curvature are frequently found in structures that can bind extended and flexible structures (like those of teichoic and lipoteichoic acids) through multiple interactions [52]. High-affinity binding of peptidoglycan hydrolases has been proposed to be related to the bacterial regulation of the suicidal activity of the autolytic enzymes. Thus the acquisition of additional repeats by LytC and the uneven distribution of choline-binding sites in the CBM could reflect a divergent evolution toward an improved regulation of the autolytic activity of the pneumococcal lysozyme.

Concluding remarks

The family of pneumococcal choline-binding proteins, known as CBPs, represents one of the most paradigmatic examples that illustrate the theory of the modular evolution of proteins. Until very recently, it was assumed that the murein hydrolases encoded by *Pneumococcus* and its phages evolved following a general assembling pattern that placed the functional/catalytic domain at the N-terminus and the CBM at the C-terminus. However, the reversed order of these modules in the LytC lysozyme encoded by *Pneumococcus* confirmed further the versatility of this modular design and raised the question of why the bacteria had not taken advantage of this already existing gene to evolve its own autolytic enzyme. Lysozymes are extremely aggressive enzymes whose expression and activity can be tailored at the genetic level by natural selection to suit the prevailing physiological requirements. For the lytic bacteriophages, it is essential that a general and rapid destruction of the cell wall occur once the virions are formed, and Cpl-1 has probably evolved to fulfil this requirement for high efficiency, making it difficult for the host to adapt such an active enzyme. Thus *Pneumococcus* would have acquired a different, less active, version of this lytic gene to remodel the cell wall at its convenience and to carry out a slow autolysis at 30 °C. Our data provide evidence that, by increasing the number and variability of the repeats, LytC has acquired a more unstable, irregular and extended CBM whose conformation is extremely dependent on its binding to the choline-containing teichoic acids and provides tight binding to the cell wall. This fact may confine the activity of the host enzyme to specific areas of cell envelope, since complete dissociation from teichoic acids will be highly unfavourable. Furthermore, the catalytic activity can be also limited by steric boundaries, since the distinct binding patterns of LytC and Cpl-1 can result in a different orientation of the glycanic chain within the active site. Moreover, the dissimilarities found in the catalytic cavity can also contribute to reduce the activity of LytC relative to Cpl-1.

This work was supported by DGICYT (Dirección General de Investigación Científica y Técnica) Grants BIO2000-1307, BMC2000-1002, BIO2003-01952, BMC2003-00074 and by CAM Grant of the Program for Strategic Groups. We are grateful to Dr Germán Rivas for many helpful discussions, to Dr Concepción Domingo for assistance in IR experiments, to Dr Douglas Laurents for linguistic revision, and to Victoria López for technical support. B. M. was supported by a fellowship from the Spanish Ministerio de Ciencia y Tecnología.

REFERENCES

1 Coutinho, P. M. and Henrissat, B. (1999) Carbohydrate-active enzymes: an integrated database approach. In Recent Advances in Carbohydrate Bioengineering (Gilbert, H. J., Davies, G., Henrissat, B. and Svensson, B., eds.), pp. 3–12, The Royal Society of Chemistry, Cambridge
2 Gebler, J., Gilkes, N. R., Claeyssens, M., Wilson, D. B., Beguin, P., Wakarchuk, W. W., Kilburn, D. G., Miller, R. C., Warren, R. A. J. and Withers, S. G. (1992) Stereoselective hydrolysis catalyzed by related β-1,4-Glucanases and β-1,4-xylanases. J. Biol. Chem. **267**, 12559–12561
3 García, P., González, M. P., García, E., García, J. L. and López, R. (1999) The molecular characterization of the first autolytic lysozyme of *Streptococcus pneumoniae* reveals evolutionary mobile domains. Mol. Microbiol. **33**, 128–138
4 García, P., García, J. L., García, E., Sánchez-Puelles, J. M. and López, R. (1990) Modular organization of the lytic enzymes of *Streptococcus pneumoniae* and its bacteriophages. Gene **86**, 81–88
5 López, R., García, E., García, P. and García, J. L. (1997) The pneumococcal cell wall degrading enzymes: a modular design to create new lysins? Microb. Drug Resist. **3**, 199–211
6 Sanz, J. M., Díaz, E. and García, J. L. (1992) Studies on the structure and function of the N-terminal domain of the pneumococcal murein hydrolases. Mol. Microbiol. **6**, 921–931
7 Henrissat, B. and Bairoch, A. (1993) New families in the classification of glycosylhydrolases based on amino acid sequence similarities. Biochem. J. **293**, 781–788
8 Hermoso, J. A., Monterroso, B., Albert, A., Galán, B., Ahrazem, O., García, P., Martínez-Ripoll, M., García, J. L. and Menéndez, M. (2003) Structural basis for selective recognition of pneumococcal cell wall by modular endolysin from phage Cp-1. Structure **11**, 1239–1249
9 Rau, A., Hogg, T., Marquardt, R. and Hilgenfeld, R. (2001) A new lysozyme fold. J. Biol. Chem. **276**, 31994–31999
10 Sanz, J. M., García, P. and García, J. L. (1992) Role of Asp-9 and Glu-37 in the active site of the pneumococcal Cpl-1 lysozyme: an evolutionary perspective of lysozyme mechanism. Biochemistry **31**, 8495–8499
11 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
12 Tramontano, A. (1998) Homology modelling with low sequence identity. Methods **14**, 293–300
13 Facchiano, A. M., Struso, P., Chiusano, M. L., Caraglia, M., Giuberti, G., Marra, M., Abbruzzese, A. and Colonna, G. (2001) Homology modelling of the human eukaryotic initiation factor 5A (eIF-5A). Protein Eng. **14**, 881–890
14 Sanz, J. M., López, R. and García, J. L. (1988) Structural requirements of choline derivatives for "conversion" of pneumococcal amidase. FEBS Lett. **232**, 308–312
15 Medrano, F. J., Gasset, M., López-Zumel, C., Usobiaga, P., García, J. L. and Menéndez, M. (1996) Structural characterization of the unligated and choline-bound forms of the major pneumococcal autolysin LytA amidase. J. Biol. Chem. **271**, 29152–29161
16 Philo, J. (1997) An improved function for fitting sedimentation velocity data for low-molecular weight solutes. Biophys. J. **72**, 435–444
17 van Holde, K. E. (1985) Sedimentation. In Physical Biochemistry, pp. 110–136, Prentice Hall, Englewood Cliffs
18 Kuntz, I. D. (1971) Hydration of macromolecules IV. Polypeptide conformation in frozen solutions. J. Am. Chem. Soc. **93**, 516–518
19 Waxman, E., Laws, W. R., Laue, T. M., Nemerson, Y. and Ross, J. B. A. (1993) Human factor VIIa and its complex with soluble tissue factor: evaluation of asymmetry and conformational dynamics by ultracentrifugation and fluorescence anisotropy decay methods. Biochemistry **32**, 3005–3012
20 Provencher, S. W. and Glöckner, J. (1981) Estimation of globular protein secondary structure from circular dichroism. Biochemistry **20**, 33–37

21 Böhm, G., Muhr, R. and Jaenicke, R. (1992) Quantitative analysis of protein far UV circular dichroism spectra by neural networks. Protein Eng. **5**, 191–195

22 Kabsch, W. and Sander, C. (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. Biopolymers **22**, 2577–2637

23 Arrondo, J. L. R., Muga, A., Castresana, J. and Goñi, F. M. (1993) Quantitative studies of the structure of proteins in solution by Fourier-transform infrared spectroscopy. Prog. Biophys. Mol. Biol. **59**, 23–56

24 Altschul, S. F., Madden, T. L., Schaffer, 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

25 Rost, B. (1996) PHD: predicting one-dimensional protein structure by profile based neural networks. Methods Enzymol. **266**, 525–539

26 McGuffin, L., Bryson, K. and Jones, D. T. (2000) The PSIPRED protein structure prediction server. Bioinformatics **16**, 404–405

27 Cuff, J. A. and Barton, C. J. (2000) Application of multiple sequence alignment profiles to improve protein secondary structure prediction. Proteins Struct. Funct. Genet. **40**, 502–511

28 Thompson, J. D., Higgins, D. G. and Gibson, T. J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. **22**, 4673–4680

29 Galtier, N., Gouy, M. and Gautier, C. (1996) SeaView and Phylo_win, two graphic tools for sequence alignment and molecular phylogeny. CABIOS Comput. Appl. Biosci. **12**, 543–548

30 Mizuguchi, K., Deane, C. M., Blundell, T. M., Johnson, M. S. and Overington, J. P. (1998) JOY: protein sequence structure representation and analysis. Bioinformatics **14**, 617–623

31 Pearson, W. R., Wood, T., Zhang, Z. and Miller, W. (1997) Comparison of DNA sequences with protein sequences. Genomics **46**, 24–36

32 Sali, A. and Blundell, T. (1993) Comparative protein modelling by satisfaction of spatial restraints. J. Mol. Biol. **234**, 779–815

33 Reference deleted

34 Laskowski, R. A., McArthur, W., Moss, D. and Thornton, J. M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. **26**, 283–291

35 Luthy, R., Bowie, J. U. and Eisenberg, D. (1992) Assessment of protein models with three-dimensional profiles. Nature (London) **356**, 83–85

36 Sali, A. and Blundell, T. L. (1990) Definition of general topological equivalence in protein structures: a procedure involving comparison of properties and relationships thought simulated annealing and dynamic programming. J. Mol. Biol. **212**, 403–428

37 van Roey, P., Rao, V., Plummer, Jr, T. H. and Tarentino, A. L. (1994) Crystal structure of endo-β-N-acetylglucosaminidase F_1, an α/β-Barrel enzyme adapted for a complex substrate. Biochemistry **33**, 13989–13996

38 Mark, B. L., Vocadlo, D. J., Knapp, S., Triggs-Raine, B. L., Withers, S. G. and James, M. N. (2001) Crystallographic evidence for a substrate-assisted catalysis in a bacterial β-hexosaminidase. J. Biol. Chem. **276**, 10330–10337

39 Simpson, P. J., Bolam, D. N., Cooper, A., Ciruela, A., Hazlewood, G. P., Gilbert, H. J. and Williamson, M. P. (1999) A family IIb xylan-binding domain has a similar secondary structure to a homologous family IIa cellulose-binding domain but different ligand specificity. Structure **7**, 853–864

40 Nagy, T., Simpson, P. J., Williamson, M. P., Hazlewood, G. P., Orosz, L. and Gilbert, H. J. (1998) All three surface tryptophans in Type IIa cellulose binding domains play a pivotal role in binding both soluble and insoluble ligands. FEBS Lett. **429**, 312–316

41 Udwary, D. W., Merski, M. and Townsend, C. A. (2002) A method for prediction of the locations of linker regions within large multifunctional proteins, and application to a Type I polyketide synthase. J. Mol. Biol. **323**, 585–598

42 Sáiz, J. L., López-Zumel, C., Monterroso, B., Varea, J., Arrondo, J. L. R., Iloro, I., García, J. L., Laynez, J. and Menéndez, M. (2002) Characterization of EjI, the cell-wall amidase coded by the pneumococcal bacteriophage Ej-1. Protein Sci. **11**, 1788–1799

43 Varea, J., Sáiz, J. L., López-Zumel, C., Monterroso, B., Medrano, F. J., Arrondo, J. L. R., Iloro, I., Laynez, J., García, J. L. and Menéndez, M. (2000) Do sequence repeats play an equivalent role in the choline-binding module of pneumococcal LytA amidase? J. Biol. Chem. **275**, 26842–26855

44 Varea, J., Monterroso, B., Sáiz, J. L., López-Zumel, C., García, J. L., Laynez, J., García, P. and Menéndez, M. (2004) Structural and thermodynamic characterization of PaI, a phage natural chimeric lysin active against Pneumococci. J. Biol. Chem. **279**, 43697–43707

45 Johnson, Jr, W. C. (1990) Protein secondary structure and circular dichroism: a practical guide. Proteins Struct. Funct. Genet. **7**, 205–214

46 Barth, A. and Zscherp, B. (2002) What vibrations tell us about proteins. Q. Rev. Biophys. **35**, 369–430

47 Wilder, C. L., Friedrich, A. D., Potts, R. O., Daumy, G. O. and Francoeur, M. L. (1992) Secondary structural analysis of two recombinant murine proteins, interleukins 1α and 1β: is infrared spectroscopy sufficient to assign structure? Biochemistry **31**, 27–31

48 García de la Torre, J., Huertas, M. L. and Carrasco, B. (2000) Calculation of hydrodynamic properties of globular proteins from their atomic-level structure. Biophys. J. **78**, 719–730

49 Branden, C. and Tooze, J. (1991) Alpha/beta structures. In Introduction to Protein Structure, pp. 43–56, Garland Publishing, New York and London

50 Vollmer, W. and Tomasz, A. (2000) The *pgdA* gene encodes for a peptidoglycan N-acetylglucosamine deacetylase in *Streptococcus pneumoniae*. J. Biol. Chem. **275**, 20496–20501

51 Dyson, H. J. and Wright, P. E. (2002) Coupling of folding and binding for unstructured proteins. Curr. Opin. Struct. Biol. **12**, 54–60

52 Kobe, B. and Kajava, A. (2000) When protein folding is simplified to protein coiling: the continuum of solenoid protein structures. Trends Biochem. Sci. **25**, 509–515

Received 14 April 2005/7 June 2005; accepted 8 June 2005
Published as BJ Immediate Publication 8 June 2005, doi:10.1042/BJ20050612