THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 282, NO. 34, pp. 24990–24999, August 24, 2007

Elucidation of the Molecular Recognition of Bacterial Cell Wall by Modular Pneumococcal Phage Endolysin CPL-1*[S]

Received for publication, May 25, 2007, and in revised form, June 12, 2007 Published, JBC Papers in Press, June 19, 2007, DOI 10.1074/jbc.M704317200

Inmaculada Pérez-Dorado[‡1], Nuria E. Campillo[§], Begoña Monterroso[¶], Dusan Hesek[‖], Mijoon Lee[‖], Juan A. Páez[§], Pedro García[**], Martín Martínez-Ripoll[‡], José L. García[**], Shahriar Mobashery[‖], Margarita Menéndez[¶], and Juan A. Hermoso[‡2]

From the [‡]Grupo de Cristalografía Macromolecular y Biología Estructural, Instituto de Química-Física Rocasolano, CSIC, Serrano 119, 28006 Madrid, Spain, the [§]Departamento de Quimioterapia, Instituto de Química Médica, CSIC, Juan de la Cierva 3, 28006 Madrid, Spain, the [¶]Departamento de Química-Física de Macromoléculas Biológicas, Instituto Química-Física Rocasolano, CSIC, Serrano 119, 28006 Madrid, Spain, the [‖]Department of Chemistry and Biochemistry, University of Notre Dame, Notre Dame, Indiana 46556, and the [**]Departamento de Microbiología Molecular, Centro de Investigaciones Biológicas, CSIC, Ramiro de Maeztu 9, 28040 Madrid, Spain

Pneumococcal bacteriophage-encoded lysins are modular proteins that have been shown to act as enzymatic antimicrobial agents (enzybiotics) in treatment of streptococcal infections. The first x-ray crystal structures of the Cpl-1 lysin, encoded by the pneumococcal phage Cp-1, in complex with three bacterial cell wall peptidoglycan (PG) analogues are reported herein. The Cpl-1 structure is folded in two well defined modules, one responsible for anchoring to the pneumococcal cell wall and the other, a catalytic module, that hydrolyzes the PG. Conformational rearrangement of Tyr-127 is a critical event in molecular recognition of a stretch of five saccharide rings of the polymeric peptidoglycan (cell wall). The PG is bound at a stretch of the surface that is defined as the peptidoglycan-binding sites 1 and 2, the juncture of which catalysis takes place. The peptidoglycan-binding site 1 binds to a stretch of three saccharides of the peptidoglycan in a conformation essentially identical to that of the peptidoglycan in solution. In contrast, binding of two peptidoglycan saccharides at the peptidoglycan-binding site 2 introduces a kink into the solution structure of the peptidoglycan, en route to catalytic turnover. These findings provide the first structural evidence on recognition of the peptidoglycan and shed light on the discrete events of cell wall degradation by Cpl-1.

Streptococcus pneumoniae is one of the most common and important human pathogens, which causes serious life-threatening diseases such as acute otitis media, pneumonia, sepsis, and meningitis. Pneumococcal infections are associated with high morbidity and mortality, especially among children, the elderly, and the immune-depressed patients. The widespread emergence of antibiotic resistance and the lack of highly effective pneumococcal vaccines against all serotypes of this organism give urgency to elucidation of the molecular processes involved in its pathogenicity (1, 2).

The peptidoglycan (PG)[3] scaffold of the bacterial cell wall is a repeating GlcNAc-*N*-acetylmuramic (MurNAc) disaccharide (GlcNAc-(β-1,4)-MurNAc) unit having a pentapeptide attached to the D-lactyl moiety of each MurNAc unit. All known pneumococcal bacteriophages encode an amidase or a lysozyme, which hydrolyzes the PG at the final stage of the phage reproductive cycle, leading to bacterial cell lysis. These enzymes, known collectively as endolysins, have been shown to be highly efficient in killing pneumococci *in vitro* and can eradicate this organism from the upper respiratory tract or from the bloodstream of mice (3, 4) acting as new antimicrobial agents (*i.e.* enzybiotics). In addition, Cpl-1 lysin and Pal amidase encoded by phage Dp-1 act in a synergistic manner in a sepsis mouse model (5); this synergy has also been confirmed in *in vitro* experiments with Cpl-1 and penicillin or gentamicin (6). Very recently, the creation of a new animal model of otitis media has been reported (7). Using this new mouse model, it has been demonstrated that Cpl-1 could eliminate colonization with *S. pneumoniae* and prevent the development of otitis media (7). A paucity of information exists presently on the mechanisms of lysis of pneumococcal cell wall by the phage-encoded endolysins at the molecular level.

All known pneumococcal endolysins display a modular structure. In addition to the catalytic module, all but one possesses a choline-binding module (CBM) to facilitate their anchoring to the choline-containing teichoic acid of the pneu-

* This work was supported by Grants BFU2005-01645, BIO2003-01952, and BFU2006-10288 from Dirección General de Investigación by 08.2/0030.1/2003 from the Comunidad de Madrid, by "Factoria de Cristalización," CONSOLIDER INGENIO-2010, and by a grant from the National Institutes of Health. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The atomic coordinates and structure factors (code 2ixu, 2ixv, 2j8f, and 2j8g) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

[S] The on-line version of this article (available at http://www.jbc.org) contains supplemental Tables 2 and 3 and Figs. 1 and 2.

[1] Previously a fellow of the Consejo Superior de Investigaciones Científicas.

[2] To whom correspondence should be addressed. Tel.: 34-915619400; Fax: 34-915642431; E-mail: xjuan@iqfr.csic.es.

[3] The abbreviations used are: PG, peptidoglycan; MurNAc, *N*-acetylmuramic acid; 2S2P, GlcNAc-MurNAc-L-Ala-D-isoGln); 2S5P, GlcNAc-MurNAc-(L-Ala-D-Glu-L-Lys-D-Ala-D-Ala); $(2S5P)_2$, tetrasaccharide di-pentapeptide (GlcNAc-MurNAc-(L-Ala-D-Glu-L-Lys-D-Ala-D-Ala))$_2$; PGBS1, peptidoglycan-binding site 1; PGBS2, peptidoglycan-binding site 2; PGRPs, peptidoglycan recognition proteins; r.m.s.d., root mean square deviation; CBP, choline-binding protein.

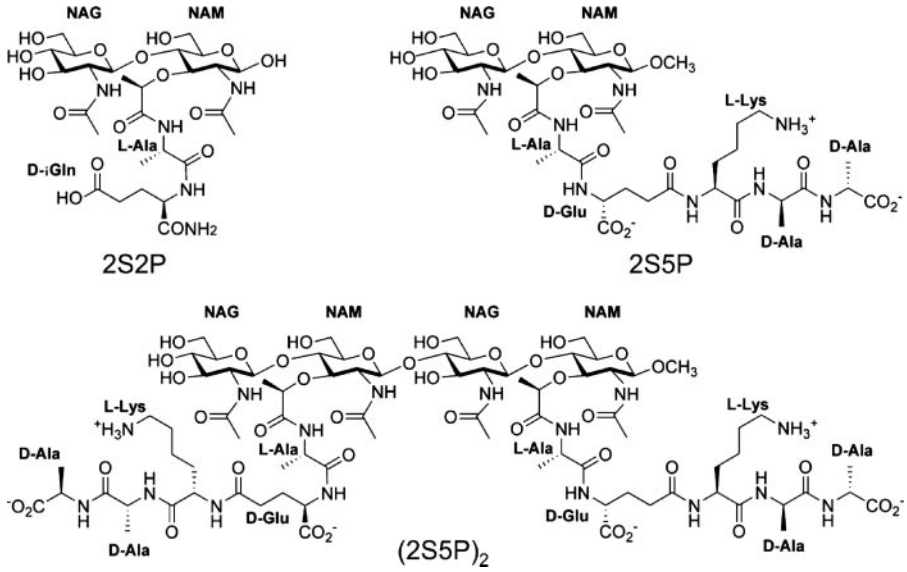

FIGURE 1. **Schematic drawing of the three peptidoglycan analogues bound to Cpl-1.** *2S2P* is disaccharide-dipeptide, 2S5P is disaccharide-pentapeptide, and $(2S5P)_2$ is tetrasaccharide-di-pentapeptide.

mococcal cell wall (8). This CBM is formed by a repeat of about 20 amino acids, found in multiple tandem copies (ranging from 4 to 18) in a large family of surface proteins (144 members identified by Pfam) from Gram-positive bacteria and from their bacteriophages. These proteins, named choline-binding proteins (CBPs) play important physiological functions in pneumococcal virulence (9). Only crystal structures of two complete CBPs, those of Cpl-1 (10) and Pce (11), have been reported.

Cpl-1 belongs to the GH25 family of glycosyl hydrolases and cleaves the $\beta 1-4$ glycosidic bond between the MurNAc and the GlcNAc residues of the pneumococcal PG. Structural knowledge of how this process takes place in Cpl-1 and in all other members of the GH25 family of glycosyl hydrolases is presently lacking.

Here we report the crystal structures of the native Cpl-1 lysozyme and of a catalytically inactive mutant variant, referred to as Cpl-1_{E94Q}, in complex with cell wall PG analogues. The combination of crystallographic and computational studies has allowed us to gain unprecedented insights into the recognition events that lead to the catalytic turnover processes. This knowledge is central in understanding how pneumococcal envelope is degraded by Cpl-1, but it also sheds light on the important question of how enzybiotics function against pneumococci and other Gram-positive bacteria.

EXPERIMENTAL PROCEDURES

PG Analogues Synthesis—The PG analogues 2S5P and $(2S5P)_2$ were synthesized as described earlier (12, 13).

Expression and Purification of Cpl-1 Endolysin—The native Cpl-1 lysozyme and its Cpl-1_{E94Q} mutant variant were expressed in *Escherichia coli* DH1 (pCIP100) and *E. coli* DH1 (pCOB7) cells, respectively, and purified from the crude extracts as described (14).

Crystallization and Data Collection—Native crystals of Cpl-1 and Cpl-1_{E94Q} were grown using the hanging drop vapor diffusion method, as reported previously (15). Complexes of the wild-type and mutant lysozymes with disaccharide dipeptide GlcNAc-MurNAc-(L-Ala-D-isoGln) (2S2P), the disaccharide pentapeptide (GlcNAc-MurNAc-(L-Ala-D-Glu-L-Lys-D-Ala-D-Ala) (2S5P), and the tetrasaccharide dipentapeptide GlcNAc-MurNAc-(L-Ala-D-Glu-L-Lys-D-Ala-D-Ala))$_2$ $(2S5P)_2$ (Fig. 1) were obtained by soaking the crystals of both proteins in solutions containing the PG analogues. Soaking time was 30 min in all cases except for the 2S5P where the time was 18 h; the concentrations for the ligands were 100 mM for the 2S2P and 25 mM for the 2S5P and the $(2S5P)_2$. GlcNAc and MurNAc monosaccharides, $(GlcNAc)_2$, $(GlcNAc)_4$, and $(GlcNAc)_6$ oligosaccharides, and the PG analogues MurNAc-(L-Ala-D-isoGln) and MurNAc-(L-Ala-D-isoGln-L-Lys-D-Ala-D-Ala) were also tested without success. The x-ray data sets of the wild-type and Cpl-1_{E94Q} mutant in complex with 2S2P were collected up to 2.3 and 2.0 Å resolution, respectively, using the graphite monochromatic CuK_α (λ = 1.5418 Å) radiation generated by an Enraf-Nonius rotating anode generator and a MAR345 image plate detector. The x-ray data sets of the Cpl-1_{E94Q} in complex with 2S5P and $(2S5P)_2$ were measured at the ID29 beamline of the ESRF up to resolutions of 1.8 and 1.7 Å, respectively. Crystals of the four complexes belong to the orthorhombic $C222_1$ space group with one molecule in the asymmetric unit. The x-ray diffraction data sets were processed and scaled using the programs MOSFLM (16) and SCALA from the CCP4 package (17).

Structure Determination and Refinement—All structures were solved by the Molecular Replacement Method. Structures of lysozyme complexes with 2S2P were solved using the program AMoRe (18), whereas those of the Cpl-1_{E94Q} in complex with 2S5P and $(2S5P)_2$ were solved using the program MOLREP (19). The crystal structure of the native Cpl-1 (10) was used as the initial model. The models were subjected to successive refinement cycles with the CNS program (20) and manual model building used the software package O (21). After this initial round of refinement, the model was further refined by iterative maximum likelihood positional and TLS refinement in REFMAC5 from the CCP4 package (17). Electron density maps of excellent quality were obtained for all ligands. In the complexes with 2S2P, the complete ligand molecule was observed. In the Cpl-1_{E94Q}-2S5P complex, two ligand molecules were seen in the crystallographic structure. Overall, the resulting electron density was of excellent quality except for two amino acid side chains of the linker region between modules (residues 191–200). Description of the atomic composition of the four complexes together with structure determination parameters and refinement statistics are summarized in Table 1.

TABLE 1

Structure determination and statistics of the Cpl-1 complexes with the three peptidoglycan analogues

Values in parentheses correspond to the highest resolution shell.

	Cpl-1–2S2P	Cpl-1_{E94Q}-2S2P	Cpl-1_{E94Q}-2S5P	Cpl-1_{E94Q}-$(2S5P)_2$
Data collection statistics				
Space group	$C222_1$	$C222_1$	$C222_1$	$C222_1$
Unit cell parameters				
a, Å	80.52	79.96	80.23	79.52
b, Å	96.50	96.44	95.56	97.39
c, Å	127.32	127.07	129.27	127.14
T (K)	120	120	100	100
Wavelength, Å	1.5418	1.5418	0.9760	0.9760
Resolution, Å	25.9 (2.41)–2.26	26.5 (2.08)–1.96	64.6 (1.97)–1.84	63.3 (1.81)–1.69
Total no. of reflections	632,248	388,319	459,898	621,365
No. of unique reflections	22,910	35,314	47,697	57,974
Redundancy	9.3 (9.1)	4.3 (4.2)	5.4 (2.9)	5.5 (3.4)
Completeness, %	99.9 (99.5)	99.4 (100)	96.3 (81.6)	92.6 (71.4)
I/σ	20.8 (4.9)	10.9 (3.2)	11.0 (3.0)	10.9 (3.2)
R_{rim} [a]	0.07 (0.45)	0.08 (0.44)	0.13 (0.61)	0.05 (0.60)
R_{pim} [b]	0.02 (0.15)	0.02 (0.14)	0.05 (0.35)	0.12 (0.32)
Wilson plot B-factor	48.5	46.7	18.3	19.6
Refinement statistics				
Resolution range, Å	25.9–2.3	26.5–2.0	64.6–1.84	63.3–1.7
Protein non-hydrogen atoms	2763	2763	2763	2763
Ligand non-hydrogen atoms	51	51	90	72
Solvent non-hydrogen atoms	172	275	333	454
R_{work}/R_{free} [c]	0.21/0.26	0.21/0.24	0.18/0.22	0.19/0.22
r.m.s.d. bond length, Å	0.007	0.006	0.017	0.006
r.m.s.d. bond angles, °	1.3	1.3	1.6	1.5
Average *B*-factor, Å^2	43.7	47.6	23.0	20.7

[a] R_{rim}, redundancy-independent (multiplicity weighted) R_{merge}.
[b] R_{pim}, precision-indicating (multiplicity weighted) R_{merge}.
[c] *R* calculated for 7% of data excluded from the refinement.

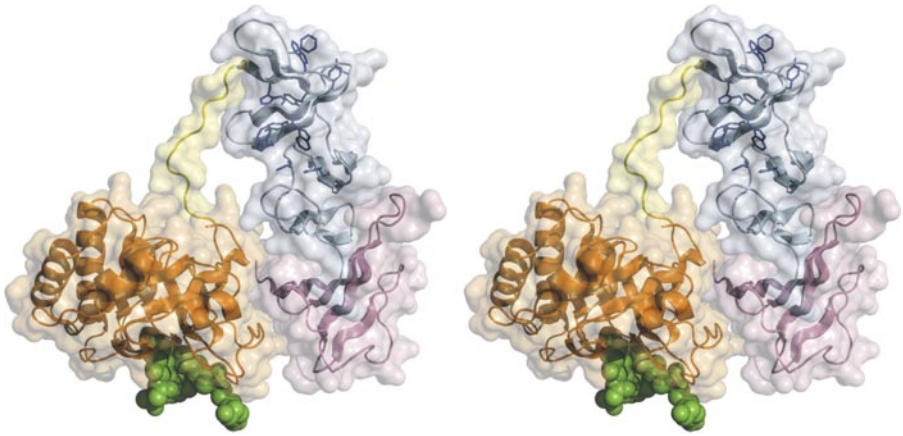

FIGURE 2. **Stereo view of the three-dimensional structure of the complex Cpl-1_{E94Q}-$(2S5P)_2$.** The catalytic module of Cpl-1 is in *orange*; the linker is in *yellow*; the CBM is in *blue* (CI domain) and in *magenta* (CII domain). The bound peptidoglycan is drawn in *green* space-filled representation.

Docking Studies—The docking studies were carried out with the FlexiDock module of the SYBYL 6.9 suite of programs (Tripos Inc., St. Louis). A PG model was constructed on the basis of the crystallographic complexes, the hexasaccharide tri-pentapeptide (GlcNAc-MurNAc-(L-Ala-D-Glu-L-Lys-D-Ala-D-Ala))$_3$ (hereafter $(2S5P)_3$).

The structure of Cpl-1 obtained from x-ray was edited and prepared. The $(2S5P)_3$ model was constructed from the crystallographic complex Cpl-1_{E94Q}-$(2S5P)_2$ by adding the GlcNAc-MurNAc-(L-Ala-D-Glu-L-Lys-D-Ala-D-Ala) fragment at positions −1 and −2. In all cases, MMFF94 force field (22, 23) together with parameters especially derived for carbohydrates (24) and charges were applied with the use of distance-dependent dielectric constants and conjugate gradient method until the gradient reached 0.005 kcal mol^{-1}·Å^{-1}. Enzyme-ligand complexes were built on the basis of the crystallographic complex and refined using FlexiDock with a genetic algorithm (25) to determine the optimum ligand geometry.

For docking procedure, the protein was considered rigid except the residues involved in the binding site (sphere of 7 Å), whereas the ligands were considered flexible. Several runs of flexidock were performed to obtain a series of model complexes. These complexes were analyzed and clustered in families based on the following: (i) score energy from flexidock results; (ii) agreement between the experimental data and the theoretical model (the interactions were examined with the LPC program (26)); and (iii) free energy of binding (ΔG_{bind}) calculated with Structural Thermodynamics Calculations version 4.3 (27). The representative conformer from each group was reoptimized.

RESULTS AND DISCUSSION

Crystal Structures of Cpl-1 in Complex with PG Analogues—The Cpl-1 structure is folded in two well defined modules connected by a linker (Fig. 2). The catalytic module consists of an irregular $(\beta/\alpha)_5\beta_3$ barrel where the PG is hydrolyzed. The CBM presents six choline-binding repeats (*p1–p6*) forming a β-hairpin each and a C-terminal tail of 16 amino acids. The first four repeats (*p1–p4*) are folded in a super-helical arrangement (CI domain, Fig. 2), whereas the other two repeats (*p5–p6*) and the C-terminal tail fold as an antiparallel-like six-stranded β-sheet

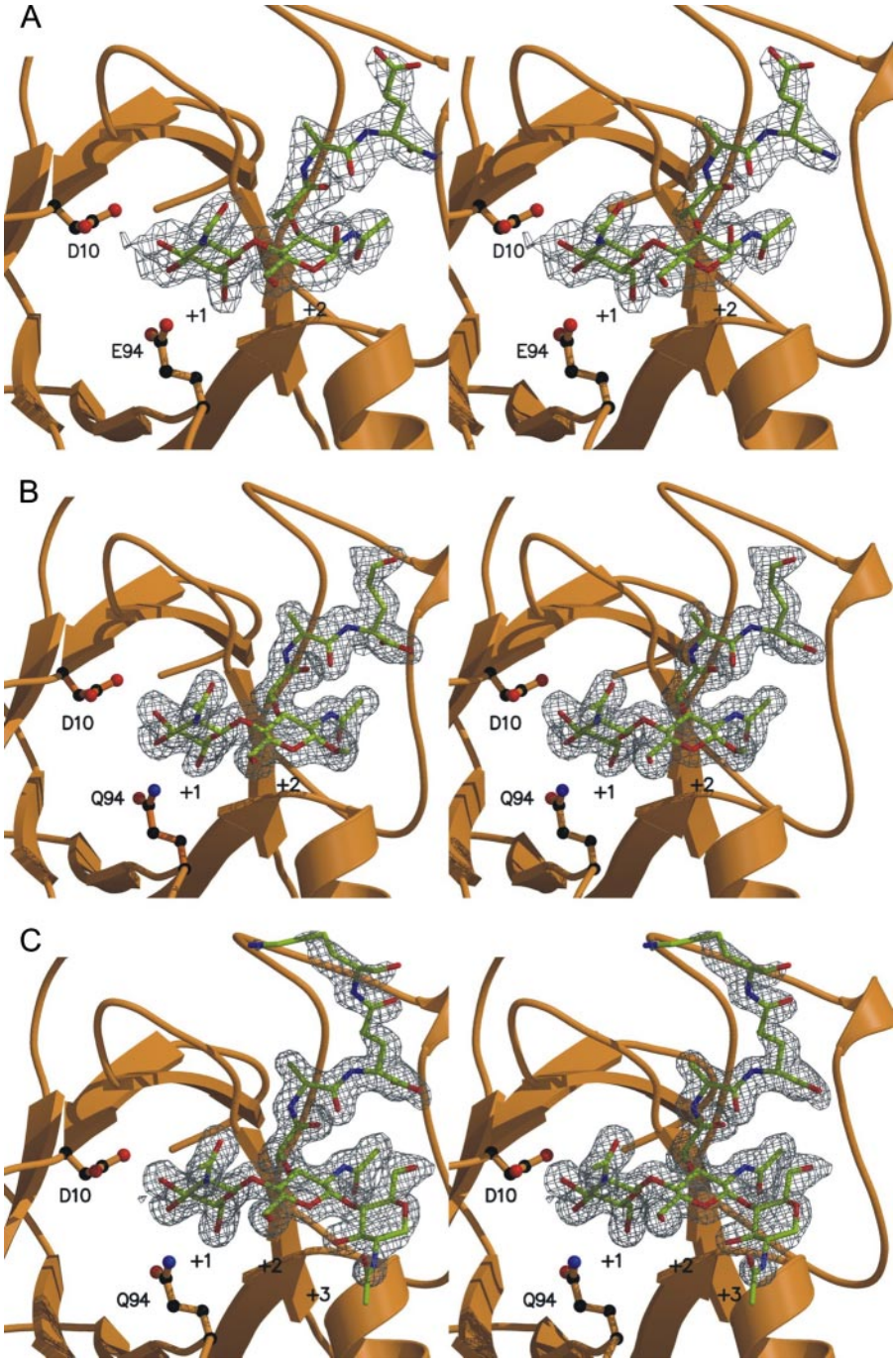

FIGURE 3. **Electron density maps observed for ligands of the crystallographic complexes.** Stereo view representations of the complexes Cpl-1-(2S2P) (*A*), Cpl-1_{E94Q}-(2S5P) (*B*), and Cpl-1_{E94Q}-$(2S5P)_2$ (*C*) at the PGBS1. Electron density maps ($2F_o - F_c$) are contoured at 1.0σ in *gray*. Carbon atoms of the ligand are in *green*, and the protein is in *orange*. The two catalytic residues are highlighted in *ball* and *sticks* with the carbon atoms in *black*.

(CII domain) that mediates interaction between modules (Fig. 2). The first four repeats provide two canonical choline-binding sites in Cpl-1, analogous to the same architecture displayed in other CBPs, such as Pce (11) and C-LytA (28). An additional choline-binding site, located at the top of the CBM, is formed by three aromatic residues (Trp-210, Phe-218, and Tyr-238).

Structural studies were performed with PG analogues shown in Fig. 1. Superimposition of the structures of the crystallized complexes with that of the native Cpl-1 gave r.m.s.d. values ranging from 0.48 to 0.64 Å.

The substrate-binding cleft of lysozymes and other glycosyl hydrolases can usually accommodate several saccharide units designated as positions −i (the nonreducing end) through +j. The saccharide units flanking the scissile glycosidic bond are assigned as positions −1 and +1. In all Cpl-1 complexes crystallized, the PG analogues were clearly identified in the electron density maps located along a groove leading to the Cpl-1 active site (Fig. 3 and supplemental Fig. 1). This groove, hereafter referred to as the peptidoglycan-binding site 1 (PGBS1), is built by a short loop (residues 125–129) after β5 and by the big loop (residues 151–173) placed between β6 and β7. The muropeptide analogues 2S2P (Fig. 3*A*) and 2S5P (Fig. 3*B*) interact through their GlcNAc and MurNAc rings with Cpl-1 at positions +1 and +2, respectively, whereas the tetrasaccharide analogue $(2S5P)_2$ binds through the three GlcNAc, MurNAc and GlcNAc rings at positions +1, +2, and +3, respectively (Fig. 3*C*). The fourth saccharide ring in $(2S5P)_2$ is not seen in the density. A smaller electron density was also found near the active site and was modeled as a formate ion from the precipitant solution (supplemental Fig. 1). The formate ion binding is made possible by interactions with Ser-13, Lys-34, and Tyr-41. Remarkably, no substrate atoms were found at positions −1 and −2, despite extensive crystallization trials with a wide variety of smaller sized analogues (data not shown). The absence of ligand atoms at positions −1 and −2 can be explained by the steric restrictions imposed by the Cpl-1 crystal packing at these positions (Trp-210 of a symmetry-related molecule is ensconced at the putative −2 position).

Considering the similarity of structural results among all three complexes, we discuss here the complex of the protein with the largest PG analogue (Cpl-1-$(2S5P)_2$). The substrate analogue binds to the enzyme via two GlcNAc and one MurNAc residues, along with its associated peptide stem. No electron density was observed for either the terminal MurNAc saccharide ring or its stem peptide, but the remainder of the structure of the $(2S5P)_2$ molecule is well defined in the electron density map sequestered in the PGBS1 (Fig. 3*C* and Fig. 4*A*).

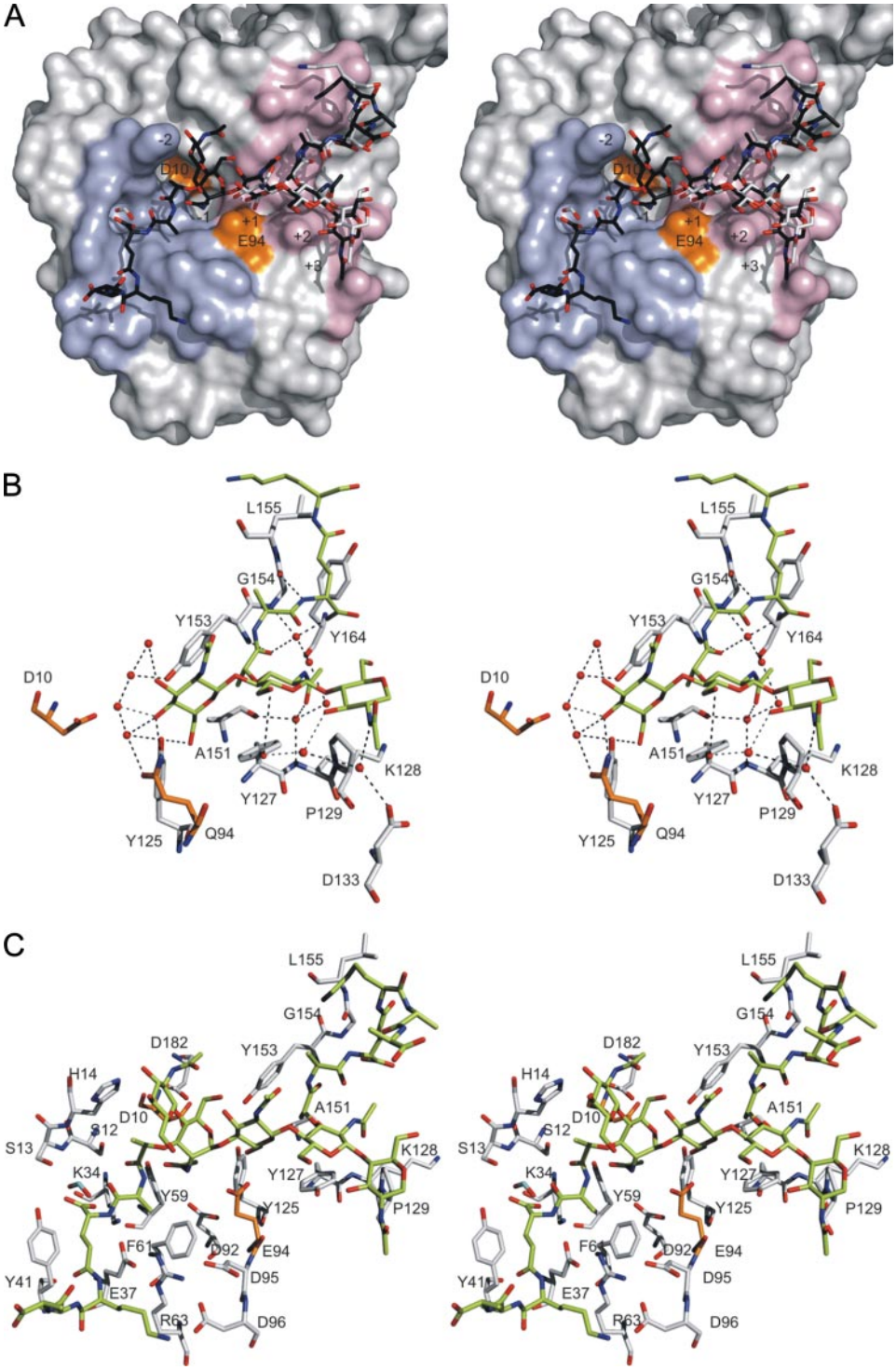

FIGURE 4. **Details of peptidoglycan recognition by Cpl-1.** *A,* superimposition of the crystallographic coordinates for Cpl-1_{E94Q}-$(2S5P)_2$ (in *white*) with the computational model for Cpl-1-$(2S5P)_3$ (in *black*). The computational model traces the crystallographically determined portions of the structure but also provides information on the portions of the structure of the ligand that did not appear in the crystallographic data. Regions involved in substrate recognition are highlighted in *pink* for PGBS1, and in *blue* for PGBS2. The catalytic residues Glu-94 and Asp-10 are in *orange*. *B,* stereo view of the crystallographic complex Cpl-1_{E94Q}-$(2S5P)_2$ showing the interactions between Cpl-1 and $(2S5P)_2$. Carbon atoms of the ligand are in *green,* and the two catalytic residues are highlighted in *orange*. Hydrogen bonds are shown as *dashed lines*. *C,* stereo view representation of the interactions in Cpl-1-$(2S5P)_3$ with the PG at positions −2 to +3. Ligand and catalytic residues are in *green* and *orange*, respectively, and the formate anion is in *cyan*.

The GlcNAc ring at position +1 is primarily stabilized by hydrogen bonds with the hydroxyl group of Tyr-125 and the carboxylic group of the catalytic Glu-94 (Gln in the inactive mutant), as well as by hydrophobic interactions with Ala-151 and Tyr-153 side chains (Fig. 4*B*). The MurNAc ring in position +2 makes a stacking interaction with Tyr-127. In addition, a polar interaction exists between the hydroxyl group of Tyr-127 and that at C-6 of the MurNAc. The peptide stem is tightly packed with a Cpl-1 loop (residues 151–154). No electron density was observed for the last two amino acids (D-Ala-D-Ala) of the ligand. We add that the solution NMR structure of $(2S5P)_2$ had revealed the L-Lys-D-Ala-D-Ala portion of the stem peptide totally devoid of structure, consistent with it being a mobile element (29). This mobility would appear to be true also for the D-Ala-D-Ala portion of the ligand in complex with Cpl-1 and hence its lack of observation in the electron density map.

A Model of Cpl-1 in Complex with a Larger Peptidoglycan Segment—Crystallization trials failed to provide information on peptidoglycan binding at positions −1 and −2. Meroueh *et al.* (29) recently reported the NMR solution structure for $(2S5P)_2$, which revealed that the saccharide backbone of the structure produced dihedral angles for the three consecutive glycosidic bonds that repeated themselves, producing a highly regular right-handed helix for the sugar backbone. Interestingly, this same right-handed helix is also observed in the three saccharide rings of the Cpl-1-$(2S5P)_2$, superimposable to the conformation determined in solution. Edified by this knowledge, we attempted to model a larger structure for the peptidoglycan into the active site of our crystal structure with occupancy at positions −2 to +4. The model revealed contacts of the MurNAc at position −1 and the GlcNAc at −2, a region (hereafter referred to as the peptidoglycan-binding site 2 (PGBS2)) formed by the active site and the β-hairpin region (residues 34–45), where the formate ion was found in the crystallographic complexes (Fig. 4*A*). A network of hydrogen bonds between the *N*-acetyl group of sugar at position −1 and residues Tyr-59, Tyr-125, and

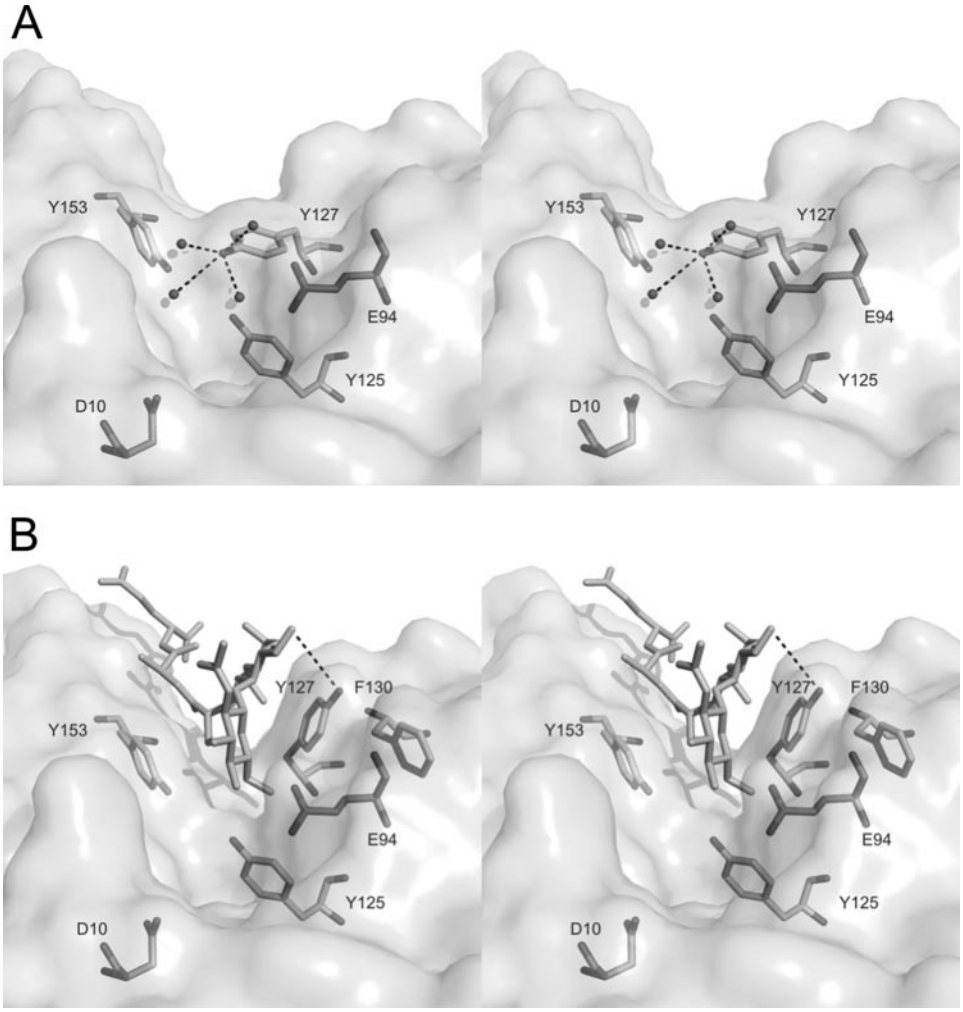

FIGURE 5. **Substrate channel before and after ligand binding.** In the absence of the PG ligand Tyr-127 blocks the entrance to the active site (*A*); upon ligand binding, Tyr-127 repositions itself to allow access to the substrate (*B*). Hydrogen bonds are drawn as *dashed lines*.

Asp-92 was also seen (Fig. 4*C*). The stem peptide of the MurNAc at position −1 experiences a network of hydrophobic and polar interactions (detailed in supplemental Tables 2 and 3). These interactions force the −1 sugar to acquire a boat conformation that introduces a kink in the PG right-handed helix. Interactions with GlcNAc at −2 are predicted to take place mainly through a stacking interaction with His-14.

Implications for the Catalytic Mechanism of Cpl-1—Hydrolytic action of lysozymes takes place via a general acid/base mechanism that requires two acidic amino acids, one behaves as the proton donor/acid and the other behaves as the nucleophile/base that promotes the hydrolytic attack by a water molecule. The previously proposed (10) electrophilic and nucleophilic residues are Glu-94 and Asp-10, respectively, and are fully conserved within the GH-25 family. As seen by the short distances, Asp-182 and Asp-92 make low barrier hydrogen bonds (30) with the catalytic residues (2.42 Å for Asp-10 to Asp-182, and 2.59 Å for Asp-92 to Glu-94), with potential implications for the enzymatic reaction. Furthermore, the distortion of the conformation of the substrate at position −1 to a higher energy species (boat *versus* chair conformers) generates the kink in the substrate, which is likely to be a high energy species en route to the transition state for the reaction, as has been also described for other glycosyl hydrolases (31).

The distance of 9.5 Å between the two catalytic residues (Asp-10 and Glu-94) in the Cpl-1 structure is consistent with that expected for an inverting enzyme (9.0 and 9.5 Å for inverting α- and β-glycosidases, respectively, and 4.8 and 5.3 Å for retaining α- and β-glycosidases, respectively) (32, 33). The distance of 7.2 Å between Asp-10 and the anomeric carbon of the −1 sugar in the model of Cpl-1-$(2S5P)_3$ complex could accommodate a water molecule between them. Asp-10 would activate the water molecule for attack at the glycosidic bond. The side chain of Glu-94 is at H-bond distance of the oxygen of the β1–4 bond; hence, it is poised for the transfer of a proton to the leaving group oxygen, making the hydrolytic process possible.

Peptidoglycan-binding Mechanism—No significant structural differences were found between the complexes with the native Cpl-1 or with the catalytically inactive mutant $Cpl\text{-}1_{E94Q}$ upon substrate binding as deduced by their low r.m.s.d. values. However, the structure of the complex with the ligands reveals a profound conformational change of the side chain of Tyr-127 (Fig. 5). This movement involves a change in the values of the two side-chain torsion angles (χ_1 and χ_2) of Tyr-127, which experience displacements of $\Delta\chi_1 = 101°$ and $\Delta\chi_2 = 25°$ from the native to the complexed state. In the absence of substrate, the side chain of Tyr-127 is placed in the center of the substrate-binding cleft, making polar interactions with Tyr-125 and Tyr-153, via their hydroxyl groups, and a few water molecules (Fig. 5*A*). In this conformation, Tyr-127 completely blocks access of the substrate to the active site at position +1. Interaction of the active site with the PG chain facilitates repositioning of the tyrosine side chain to a hydrophobic pocket created by Phe-130 and Pro-129 and places it in a suitable conformation for interacting with the MurNAc at position +2 (Fig. 5*B*). The displacement of Tyr-127 from the central position in the active site and its subsequent new interaction with MurNAc at position +2 must be pivotal to catalysis. It is not merely that the repositioning of Tyr-127 makes room for binding of PG, but it also would lower the energy barrier for the hydrolytic reaction by its interaction with the substrate, even though this electrostatic interaction is one saccharide downstream from the seat of the reaction. Other observations of note are that substrate analogues lacking the

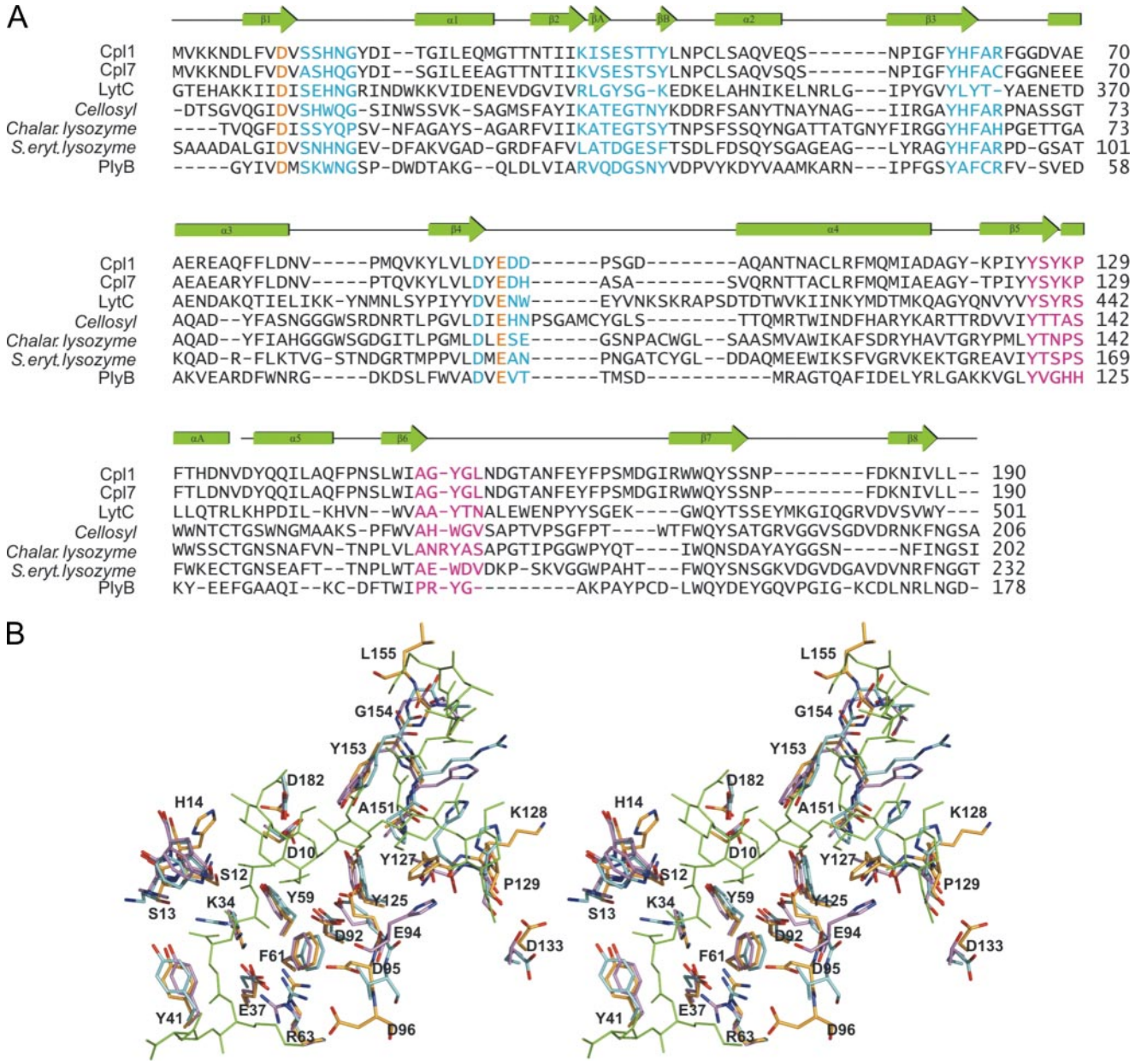

FIGURE 6. **The structure-based sequence comparison of the members of the GH25 family.** *A,* sequence alignment of muramidases of the GH25 lysozyme family: *Cpl-1*, lysozyme from the pneumococcal phage Cp-1; *Cpl-7,* lysozyme from the pneumococcal phage Cp-7; *LytC,* lysozyme from *S. pneumoniae*; *Cellosyl,* lysozyme from *S. coelicolor Muller*; lysozymes from *Chalaropsis* and *Streptomyces erythraeus* and PlyB, lysozyme from the phage BcpI. Residues belonging to the PGBS1 and PGBS2 regions are in *blue* and *pink*, respectively, and the catalytic residues are in *orange. B,* stereo view representation of the complex Cpl-1-$(2S5P)_3$ (protein residues and ligand in orange and *green,* respectively) superposed with Cellosyl (*violet*) and PlyB (*cyan*) structures. Only residues involved in PG recognition by Cpl-1 and those conserved in Cellosyl and PlyB are shown.

GlcNAc at +1 position or the peptide moiety (of at least two amino acids) on MurNAc at position +2 fail to bind to Cpl-1. These specific interactions at positions +1 and +2 might be crucial in triggering the Tyr-127 movement as a gatekeeper to the active site and to the onset of catalysis.

The crystallographic complexes showed that Cpl-1 interacts well with the first two amino acids of MurNAc at position +2 and to a lesser extent with the third (L-Lys). The expanded computational model reveals that the peptide moiety of the MurNAc at position −1 should be nicely anchored within PGBS2 region (Fig. 4, *A* and *C*). This interaction network can explain why mutation of Glu-37 (a residue within the PGBS2 site) to alanine or lysine results in a drastic loss of the catalytic activity, whereas the E37Q mutant variant still retained 67% of activity (34).

Substrate Recognition in the GH-25 Family—Sequence comparison among the GH-25 family and inspection of crystal structures of Cpl-1, Cellosyl, the muramidase encoded by *Streptomyces coelicolor* (35), and the very recently reported structure of the catalytic domain of PlyB, a lysin from the BcpI phage (36), allowed us to analyze whether the proposed peptidoglycan-Cpl-1 binding model can be applied to other family members. Besides the catalytic residues, several residues interacting with the MurNAc residue at position −1, such as Tyr-125 and Tyr-59, are also conserved in the GH-25 family (Fig. 6*A*). When the PGBS2 region of Cpl-1 is compared with the

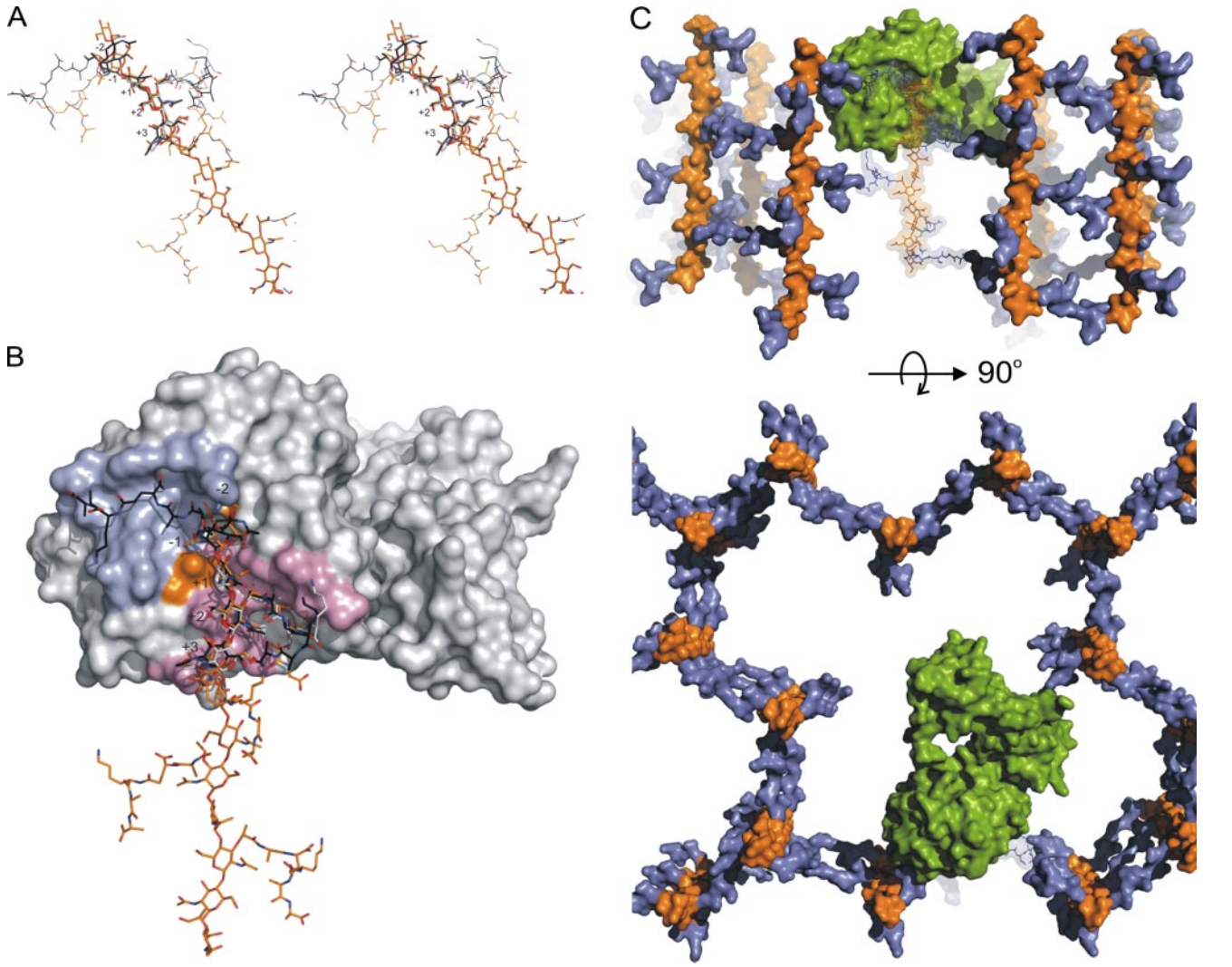

FIGURE 7. **The bacterial cell wall recognition model.** *A,* stereo view representation showing the superposition of the structures observed by x-ray crystallography (Cpl-1_{E94Q}-$(2S5P)_2$ complex; carbons in *gray*), by computation (Cpl-1-$(2S5P)_3$ complex; carbons in *black*), and the peptidoglycan solution structure determined by NMR (carbons in *orange*). *B,* protein is drawn as a Connolly solvent-accessible surface, with the same color coding given in Fig. 3*A*. The superimposition of the x-ray and computational and NMR results and their respective color coding are the same as in *A*. *C,* views of a molecule of Cpl-1 depicted as a Connolly surface in *green* docked on to the cell wall. The peptidoglycan appears in *orange* (glycan chains) and *blue* (peptide stems). The PG strand bound to Cpl-1 is represented as a transparent surface.

equivalent regions in Cellosyl and $PlyB_{cat}$, one concludes that the type of amino acids implicated in recognition at position −1 by Cpl-1 (Lys-34, Glu-37, Tyr-41, Tyr-59, Phe-61, and Arg-63; Cpl-1 numbering) is highly conserved (Fig. 6, *A* and *B*). His-14, the side chain of which is proposed to be critical in the stabilization of GlcNAc at the −2 position, is either conserved or conservatively replaced by an aromatic residue, tyrosine or tryptophan, as in the case of Cellosyl and $PlyB_{cat}$ (Trp-13 and Trp-10, respectively). This reinforces the essential role of this stacking interaction in peptidoglycan recognition at position −2 by the GH-25 family.

Concerning the PGBS1 region, the aromatic residue at position 153 (Cpl-1 numbering) stabilizing GlcNAc at +1 is systematically conserved. However, the critical Tyr-127, whose motion opens up the active site, is seen only in lysozymes encoded by pneumococcus and its bacteriophages. However, in other GH-25 family members, for which structures have been reported, the Tyr-127 position is also occupied by an aromatic residue (His-101 in Cellosyl and His-124 in PlyB) (Fig. 6*B*). Interestingly, in the case of PlyB, the side chain of His-124 is placed centrally in the binding site, implying the possibility for a similar displacement upon substrate binding seen for Tyr-127 in Cpl-1.

Molecular Recognition of Pneumococcal Cell Wall—The structure of the peptidoglycan bound to Cpl-1 at positions +1, +2, and +3 is virtually identical to the solution structure of the fragment, whereas the enzyme introduces a kink in the peptidoglycan backbone at positions −1 and −2 (Fig. 7, *A* and *B*). It would appear that Cpl-1 recognizes the solution structure of the peptidoglycan (at peptidoglycan-binding site 1) as it distorts the polysaccharide backbone (at peptidoglycan-binding site 2) en route to turnover.

It has been suggested that PG glycosyltransferases are processive enzymes, meaning that they catalyze multiple rounds of coupling without releasing the elongating product. This model for processive glycosyl chain synthesis has been documented

for the PG glycosyltransferase domain of *Aquifex aeolicus* recently (37), whose three-dimensional structure resembles that of the bacteriophage λ-lysozyme. Although no experimental information on processivity by Cpl-1 is available, the, structurally related enzyme Cellosyl from *S. coelicolor* exhibits a processive mechanism.[4] The experimental demonstration of processivity waits for Cpl-1, but the structure reported herein argues that as the hydrolytic reaction proceeds at the juncture of the peptidoglycan-binding sites 1 and 2, the enzyme might move along the backbone of the peptidoglycan in performing the catalytic process repeatedly (Fig. 7*B*).

The aforementioned right-handed helical conformation of the peptidoglycan polysaccharide backbone produced a 3-fold symmetry in the molecule (29). Each peptidoglycan strand is cross-linked to a neighboring strand to give a continuous covalent network for the cell wall that covers the entire bacterium. The 3-fold symmetry presents the opportunity for cross-linking to a maximum of three neighboring strands but requires a minimum of two, which led to the proposed honeycomb pattern for the cell wall (29). This model accommodates openings of as small as 70 Å or larger in the cell wall. For example, the absence of a single peptidoglycan strand in the perfect honeycomb generates an opening of 120 Å. The model of Cpl-1 binding to the peptidoglycan based on the x-ray structure of the enzyme and the solution NMR structure of the peptidoglycan depicted in Fig. 7*B* can be directly accommodated with this honeycomb pattern in openings of 120 Å (when one strand of peptidoglycan is missing) or larger (Fig. 7*C*).

A Common Peptidoglycan Recognition Pattern?—In the fight against infections, multicellular animals have adapted their immune system to recognize microorganisms by detecting conserved structures. This recognition constitutes the first line of defense by the host. Peptidoglycan molecules are located on the surface of all Gram-positive bacteria and therefore constitute an excellent target for recognition by the innate immune system such as CD14 and Toll-like receptors, Nod proteins, and the PGRPs (38, 39). PGRPs are proteins that are highly conserved from insects to mammals that can bind and, in some cases, hydrolyze the PG (39–42). Although the majority of the PGRPs contain a single PG-binding domain, which is structurally related to the bacteriophage T7 lysozyme, some of them have tandem domains adopting a similar fold. PGRPs act as conduits linking PG recognition to the induction of intracellular signaling or complement cascades. To fulfill this role, PGRPs present a conserved peptidoglycan-binding site and a variable locus for interacting with the host effector (39, 41, 42). Interestingly, structural comparison of Cpl-1 with the complex of human PGRPIαC bound to MurNAc-L-Ala-D-isoGln-L-Lys (43) reveals that the PG-binding site in PGRPs is similar in shape to the PGBS2 of Cpl-1 (supplemental Fig. 2, *A* and *B*). This is remarkable in light of the fact that the two proteins are unrelated to each other. A more detailed comparison shows that the nature of the interactions with the ligand is basically the same and that the conformation of the stem peptide within the binding site is practically identical in both cases. In addition, the PGRPIαC-MurNAc-L-Ala-D-isoGln-L-Lys complex revealed that PGRPs interact with both the peptide stems and the glycan backbone of the muropeptides, as we also found to be the case in Cpl-1. This binding mode, involving differential recognition of variable peptide sequences, has been postulated (43) as the structural basis for host discrimination among Gram-positive and Gram-negative bacteria. However, there is an essential difference between PGRPs and Cpl-1 proteins in that Cpl-1 uses multivalent interactions, both in the catalytic and the choline-binding modules. Furthermore, the interactions at the catalytic site are more extensive than is the case of the PGRPs, constituting two sites, which we refer to as peptidoglycan-binding sites 1 and 2.

Glycosyl hydrolases are key enzymes in reshaping the bacterial cell wall. We have described in this study the structure of the native pneumococcal phage Cpl-1 lysin in complex with three cell wall PG analogues. This structural information provides for the first time insights into how the cell wall is recognized by this enzyme and defines the incremental steps involved in the catalytic events that lead to hydrolytic fragmentation of the cell wall. In light of the fact that phage lysins such as Cpl-1 have been documented to have antibacterial effects in animal models for infections (the word "enzybiotic" has been coined) our report defines the mechanism of this antibacterial effect.

[4] M. Menéndez, unpublished results.

REFERENCES

1. Kristinsson, K. G. (1997) *Microb. Drug Resist.* **3,** 117–123
2. Pelton, S. I. (2000) *Vaccine* **19,** 96–99
3. Loeffler, J. M., Nelson, D., and Fischetti, V. A. (2001) *Science* **294,** 2170–2172
4. Loeffler, J. M., Djurkovic, S., and Fischetti, V. A. (2003) *Infect. Immun.* **71,** 6199–6204
5. Jado, I., López, R., García, E., Fenoll, A., Casal, A., and García, P. (2003) *J. Antimicrob. Chemother.* **52,** 967–973
6. Djurkovic, S., Loeffler, J. M., and Fischetti, V. A. (2005) *Antimicrob. Agents Chemother.* **49,** 1225–1228
7. McCullers, J. A., Karlström, A., Iverson, A. R., Loeffler, J. M., and Fischetti, V. A. (2007) *Plos Pathog.* **3,** e28
8. García, E., García, J. L., García, P., Arrarás, A., Sánchez-Puelles, J. M., and López, R. (1988) *Proc. Natl. Acad. Sci. U. S. A.* **85,** 914–918
9. Gosink, K. K., Mann, E., Guglielmo, C., Tuomanen, E. I., and Masure, H. R. (2000) *Infect. Immun.* **68,** 5690–5695
10. Hermoso, J. A., Monterroso, B., Albert, A., Galán, B., Ahrazem, O., García, P., Martínez-Ripoll, M., García, J. L., and Menéndez, M. (2003) *Structure (Camb.)* **11,** 1239–1249
11. Hermoso, J. A., Lagartera, L., González, L., Stelter, M., García, P., Martínez-Ripoll, M., García, J. L., and Menéndez, M. (2005) *Nat. Struct. Mol. Biol.* **12,** 533–538
12. Hesek, D., Lee, M., Morio, K., and Mobashery, S. (2004) *J. Org. Chem.* **69,** 2137–2146
13. Hesek, D., Suvorov, M., Morio, K., Lee, M., Brown, S., Vakulenko, S. B., and Mobashery, S. (2004) *J. Org. Chem.* **69,** 778–784
14. Sánchez-Puelles, J. M., Sanz, J. M., García, J. L., and García, E. (1990) *Gene (Amst.)* **89,** 69–75
15. Monterroso, B., Albert, A., Martínez-Ripoll, M., García, P., García, J. L., Menéndez, M., and Hermoso, J. A. (2002) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **58,** 1487–1489
16. Leslie, A. G. W. (1987) in *Proceedings of the CCP4 Study Weekend* (Machin, J. R., and Papiz, M. Z., eds) pp. 39–50, SERC Daresbury Laboratory, Warrington, UK
17. Collaborative Computational Project Number 4 (1994) *Acta Crystallogr.*

Sect. D Biol. Crystallogr. **50,** 760–763
18. Navaza, J. (1994) *Acta Crystallogr. Sect. A* **50,** 157–163
19. Vagin, A., and Teplyakov, A. (1997) *J. Appl. Crystallogr.* **30,** 1022–1025
20. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gross, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **54,** 905–921
21. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr. Sect. A* **47,** 110–119
22. Halgren, T. A. (1996) *J. Comput. Chem.* **17,** 553–586
23. Halgren, T. A. (1996) *J. Comput. Chem.* **17,** 490–519
24. Perez, S., Meyer, C., and Imberty, A. (1995) in *Modelling of Biomolecular Structures and Mechanisms* (Pullman, A., Jortner, J., and Pullman, B., eds) pp. 425–444, Kluwer Academic Press, Dordrecht, Netherlands
25. Judson, O. P., and Haydon, D. (1999) *J. Mol. Evol.* **49,** 539–550
26. Sobolev, V., Sorokine, A., Prilusky, J. E., Abola, E., and Edelman, M. (1999) *Bioinformatics* (*Oxf.*) **15,** 327–332
27. Lavigne, P., Bagu, J. R., Boyko, R., Willard, L., Holmes, C. F. B., and Sykes, B. D. (2000) *Protein Sci.* **9,** 252–264
28. Fernández-Tornero, C., López, R., García, E., Giménez-Gallego, G., and Romero, A. (2001) *Nat. Struct. Biol.* **8,** 1020–1024
29. Meroueh, S.-O., Bencze, K. Z., Hesek, D., Lee, M., Fisher, J. F., Stemmler, T. L., and Mobashery, S. (2006) *Proc. Natl. Acad. Sci. U. S. A.* **103,** 4404–4409
30. Cleland, W. W., Frey, P. A., and Gerlt, J. A. (1998) *J. Biol. Chem.* **273,** 25529–25532
31. David, L. Z., and Withers, S. G. (2000) *Acc. Chem. Res.* **33,** 11–18
32. Davies, G., and Henrissat, B. (1995) *Structure* (*Camb.*) **3,** 853–859
33. McCarter, J. D., and Withers, S. G. (1994) *Curr. Opin. Struct. Biol.* **4,** 885–892
34. Sanz, J. M., García, P., and García, J. L. (1992) *Biochemistry* **31,** 8495–8499
35. Rau, A., Hogg, T., Marquardt, R., and Hilgenfeld, R. (2001) *J. Biol. Chem.* **276,** 31994–31999
36. Porter, C. J., Schuch, R., Pelzek, A. J., Buckle, A. M., McGowan, S., Wilce, M. C. J., Rossjohn, J., Russell, R., Nelson, D., Fischetti, V. A., and Whisstock, J. C. (2007) *J. Mol. Biol.* **366,** 540–550
37. Yuan, Y., Barrett, D., Zhang, Y., Kahne, D., Sliz, P., and Walker, S. (2007) *Proc. Natl. Acad. Sci. U. S. A.* **104,** 5348–5353
38. Dziarski, R. (2003) *Cell. Mol. Life Sci.* **60,** 1793–1804
39. Wang, Z., Li, X., Cocklin, R. R., Wang, M., Wang, M., Fukase, K., Inamura, S., Kusumoto, S., Gupta, D., and Dziarski, R. (2003) *J. Biol. Chem.* **278,** 49044–49052
40. Dziarski, R., and Gupta, D. (2006) *Genome Biol.* **7,** 332
41. Guan, R., Malchiodi, E. L., Wang, Q., Schuck, P., and Mariuzza, A. R. (2004) *J. Biol. Chem.* **279,** 31873–31882
42. Guan, R., Wang, Q., Sundberg, E. J., and Mariuzza, A. R. (2005) *J. Mol. Biol.* **347,** 683–691
43. Guan, R., Roychowdhury, A., Ember, B., Kumar, S., Boons, G., and Mariuzza, A. R. (2004) *Proc. Natl. Acad. Sci. U. S. A.* **101,** 17168–17173