X-ray crystal structure of the streptococcal specific phage lysin PlyC

Sheena McGowan[a,1], Ashley M. Buckle[a,1], Michael S. Mitchell[b,1], James T. Hoopes[b], D. Travis Gallagher[c], Ryan D. Heselpoth[b], Yang Shen[b], Cyril F. Reboul[a], Ruby H. P. Law[a], Vincent A. Fischetti[d,2,3], James C. Whisstock[a,e,2,3], and Daniel C. Nelson[b,d,f,2,3]

[a]Department of Biochemistry and Molecular Biology, Monash University, Clayton 3800, Australia; [b]Institute for Bioscience and Biotechnology Research, University of Maryland, Rockville, MD 20850; [c]National Institute of Standards and Technology, Rockville, MD 20850; [d]Laboratory of Bacterial Pathogenesis and Immunology, The Rockefeller University, New York, NY 10021; [e]Australian Research Council Centre of Excellence in Structural and Functional Microbial Genomics, Monash University, Clayton 3800, Australia; and [f]Department of Veterinary Medicine, University of Maryland, College Park, MD 20742

Edited by Bonnie L Bassler, Howard Hughes Medical Institute, Princeton University, Princeton, NJ, and approved June 22, 2012 (received for review May 18, 2012)

Bacteriophages deploy lysins that degrade the bacterial cell wall and facilitate virus egress from the host. When applied exogenously, these enzymes destroy susceptible microbes and, accordingly, have potential as therapeutic agents. The most potent lysin identified to date is PlyC, an enzyme assembled from two components (PlyCA and PlyCB) that is specific for streptococcal species. Here the structure of the PlyC holoenzyme reveals that a single PlyCA moiety is tethered to a ring-shaped assembly of eight PlyCB molecules. Structure-guided mutagenesis reveals that the bacterial cell wall binding is achieved through a cleft on PlyCB. Unexpectedly, our structural data reveal that PlyCA contains a glycoside hydrolase domain in addition to the previously recognized cysteine, histidine-dependent amidohydrolases/peptidases catalytic domain. The presence of eight cell wall-binding domains together with two catalytic domains may explain the extraordinary potency of the PlyC holoenyzme toward target bacteria.

enzybiotic | protein crystallography | peptidoglycan | multimeric protein

Bacteriophage (or phage) typically encode holins and lysins as part of their lytic system to achieve virus exit from the host bacterial cell (1). Holins are responsible for forming pores in the cytoplasmic membrane, following which lysins, having accumulated in the cytoplasm, are responsible for degradation of the peptidoglycan (i.e., cell wall) layer. Damage to this layer results in rapid cell rupture and concomitant virus release through loss of osmotic integrity. These enzymes are also able to rapidly and specifically destroy Gram-positive bacteria when applied exogenously, making the purified lysins functional "inside-out" enzymes (2, 3). There is thus great interest in the development of phage lysins as enzybiotic agents to treat antibiotic-resistant bacterial infections (4).

Lysins derived from phage that infect Gram-positive bacteria are generally composed of a single polypeptide consisting of an N-terminal catalytic domain and a C-terminal cell wall binding domain (CBD) held together by a short flexible linker (5). In some rare cases, two to three catalytic domains may be linked to a single binding domain (6, 7). With few exceptions, the catalytic domain is usually represented by one of four families of peptidoglycan hydrolases: *N*-acetylglucosaminidases, *N*-acetylmuramidases (lysozymes), *N*-acetylmuramoyl-L-alanine amidases, and endopeptidases (8). In contrast, the CBDs are markedly divergent and can distinguish discrete epitopes present within the cell wall, typically carbohydrates or teichoic acids, giving rise to the species- or strain- specific activity of a particular lysin. Consequently, it is possible to combine the catalytic domain of one lysin and the CBD of a second lysin to make a chimeric protein with altered specificity or activity (9–11).

The streptococcal C1 phage lysin, PlyC, is the most potent lysin described to date, with a specific activity ~100 fold that of the next most catalytic lysin. Previous research spanning over 50 y has shown that PlyC can rapidly lyse cultures of groups A, C, and E streptococci in addition to *Streptococcus uberis* and *Streptococcus equi* and has been shown to protect mice from streptococcal challenge (12, 13). PlyC is furthermore unique among the Gram-positive lysins in that it consists of two separate proteins—a single 50-kDa PlyCA subunit that is suggested to form a complex with at least eight 8-kDa PlyCB subunits (14). The two proteins are transcribed from two genes located in a single operon (14). PlyCA is known to contain an active cysteine, histidine-dependent amidohydrolases/peptidase (CHAP) domain, a fold distantly related to the papain-like cysteine-protease family, with Cys_{333} and His_{420} shown to be essential for amidase catalytic activity (14). The PlyCB octamer is suggested to represent the CBD because purified material lacking the PlyCA subunit was able to specifically bind *Streptococcus pyogenes*, *S. uberis*, *S. equi*, and groups C and E streptococci, but not other bacterial species (i.e., *Streptococcus agalactiae*, *Streptococcus mutans*, and *Staphylococcus aureus*) (14).

Although the potent lytic activity of PlyC has been extensively characterized, its structural architecture and mechanism of action has remained unclear. In this study, we have made a pivotal step forward in our understanding of this enzyme by determining the structure and mechanism of PlyC action, including cell wall binding and catalytic activity.

Results

Crystallography and Structure Determination. Initial crystal trials on PlyCB demonstrated that the cell wall-binding subunit crystallized readily, and these crystals diffracted to high resolution. The structure of PlyCB was determined through sulfur single anomalous dispersion phasing (Tables S1 and S2). In contrast, the protein crystallography and structure determination of the PlyC holoenzyme was particularly challenging. More than 500 crystals were screened, and typical diffraction data obtained were limited to ~10 Å. A single native dataset to 3.3 Å was collected from

Author contributions: S.M., A.M.B., V.A.F., J.C.W., and D.C.N. designed research; S.M., A.M.B., M.S.M., J.T.H., D.T.G., R.D.H., Y.S., C.F.R., R.H.P.L., J.C.W., and D.C.N. performed research; D.C.N. contributed new reagents/analytic tools; S.M., A.M.B., M.S.M., J.T.H., D.T.G., R.D.H., Y.S., V.A.F., J.C.W., and D.C.N. analyzed data; and S.M., V.A.F., J.C.W., and D.C.N. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

Data deposition: The atomic coordinates and structure factors have been deposited in the Protein Data Bank, www.pdb.org (PDB ID codes 4F87 and 4F88).

[1]S.M., A.M.B., and M.S.M. contributed equally to this work.

[2]V.A.F., J.C.W., and D.C.N. contributed equally to this work.

[3]To whom correspondence may be addressed. E-mail: james.whisstock@monash.edu, vaf@rockefeller.edu, or nelsond@umd.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1208424109/-/DCSupplemental.

a single crystal by using synchrotron radiation (Tables S1 and S2). Attempts to obtain heavy atom derivatives and/or crystals using selenomethionine labeling to acquire experimental phases for PlyC were all unsuccessful. Ultimately, the structure of the octameric 1.4 Å PlyCB ring was used as a molecular replacement probe for the native PlyC dataset. Molecular replacement attempts using CHAP domains failed to identify the location of the CHAP domain in the PlyCA molecules. Preliminary refinement of the molecular replacement model, however, revealed significant regions of positive connective density that clearly did not form part of the PlyCB probe. Accordingly, the structure of PlyCA was determined through iterative rounds of refinement and manual building (as detailed in *Methods*).

Overall Architecture of PlyC Holoenzyme. The 3.3-Å X-ray crystal structure of the PlyC holoenzyme consists of a nine-protein assembly that comprises a single PlyCA catalytic subunit in complex with eight PlyCB molecules (Fig. 1*A*, Fig. S1, and Movie S1). The PlyCB CBD is arranged in a planar octameric ring that is 80 Å in diameter and 20 Å high (Fig. 1*B*). One surface of the PlyCB octomer is quite planar, whereas the other is more convex in shape. A similar octameric arrangement of PlyCB monomers was also observed in the 1.4-Å structure of PlyCB alone. The single PlyCA molecule is oriented and positioned entirely on the flat (planar) surface of the PlyCB octamer (Fig. 1*A* and Fig. S1).

In the PlyCB octamer, each monomer is comprised of a four-stranded β-sheet capped on each side by a short α-helix (Fig. S2). Oligomerization is mediated through strand/helix hydrogen bonding interactions at each interface (Fig. S2). Sequence-based searches reveal that PlyCB contains no significant sequence similarity to any other protein and therefore represents a rare example of a protein family that is apparently defined by a unique family member. Searches using DALI and VAST, however, reveal distant structural similarity between the PlyCB monomer with a functionally uncharacterized *S. mutans* protein (3L9A; rmsd 2.2 Å over 54 Cα atoms). We suggest that PlyCB may represent a circular permutation of this fold (Fig. S2), although no pattern of conserved residues common to PlyCB and 3L9A is apparent.

The catalytic PlyCA subunit contains three distinct domains—an N-terminal domain (residues 1–205) that comprises a loose bundle of α-helices (Fig. 1*A*), a short central helical domain that we term the docking domain (residues 228–286; Fig. 1*A*), and a C-terminal CHAP domain (residues 309–465; Fig. 1*A*). Linker regions between the central helical bundle and the two terminal domains showed partial disorder in electron density, consistent with limited proteolysis data that demonstrated that both domains could be readily dissociated from the holoenzyme. Outside the docking domain, a single contact is made with the body of the PlyCB ring. Mutation at this site (PlyCB-Gln_{46Ala}) had no effect on lytic activity (Table S3 and Fig. S3).

The CHAP domain is most closely related to an enzyme from *Staphylococcus saprophyticus* (2K3A) (15) and comprises a small half β-barrel packed against a cluster of two helices. As previously described for related enzymes, the active site (Cys_{333}/His_{420}) is located in the cleft between these two subdomains (16). Interestingly in the structure, the region linking the N-terminal domain to the central helical structure (linker 1) blocks the CHAP active site (Fig. 2). Rearrangement of this loop would be anticipated to be necessary to expose the CHAP active site.

PlyCB Assembly Localizes Eight Binding Sites for Streptococcal Peptidoglycan. To identify the region of the PlyCB octamer that was involved in cell wall binding, we undertook a mutagenesis study of both surfaces of the ring. Mutation of residues on the flat surface retained WT PlyC assembly, cell-wall binding, and lytic activity (Table S3 and Fig. S3). On the convex surface, we identified a prominent groove in each monomer that was lined by residues Tyr_{28}, Lys_{59}, and Arg_{66} (Fig. 3, Table S3, and Fig. S3). Mutation of Arg_{66} completely abolished cell wall binding and PlyC activity (Table S3). Mutation of Tyr_{28} and Lys_{59} also had a major impact on the ability of the enzyme to bind the cell wall as well as PlyC activity (Table S3 and Fig. S3). However, none of the mutations interfered with formation of the PlyCB octameric ring or formation of the holoenzyme (Table S3 and Fig. S3). Thus, we suggest that the interface between the cavity formed by the PlyCB C-terminal helix and central sheet forms a binding site for the bacterial cell wall. Presently, it is not clear if the binding

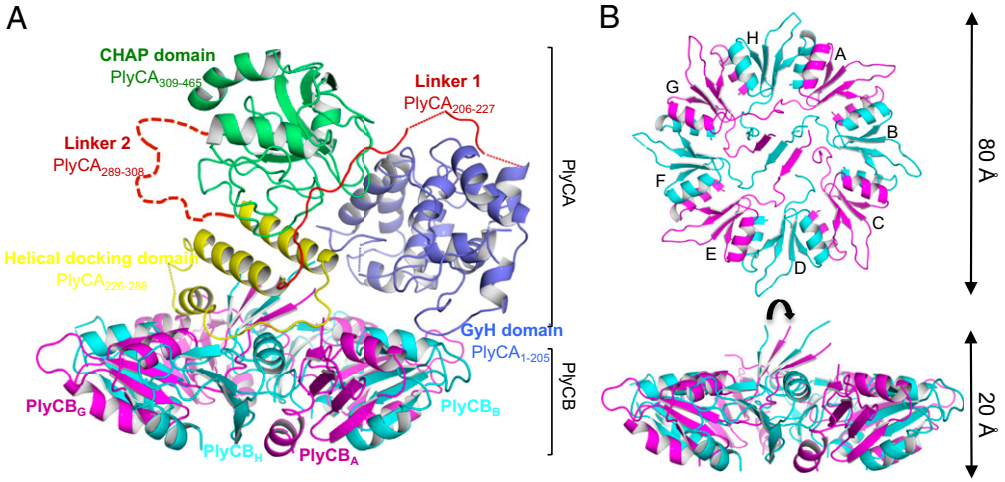


Fig. 1. Structure of PlyC. Two complete PlyC molecules were located within the asymmetric unit with 16 PlyCB molecules (chains A–P) and two PlyCA molecules (chains 1 and 2). Regions of disorder are shown as broken lines. Monomers A–P are essentially identical (average rmsd, 0.14 Å over 61 Cα atoms), as are monomers 1 and 2 (average rmsd, 0.86 Å over 460 Cα atoms); however, poor electron density was observed for monomers J, K, and L. The cartoons depicted represent PlyCB (A–H) and PlyCA (1). (*A*) The 3.3-Å X-ray crystal structure of PlyC whereby PlyCB monomers are colored alternately in magenta/cyan and labeled monomers A–H. The PlyCA molecule is colored by domain as indicated. The Cα atoms of the model show the N-terminal residues 1–205 in light blue, the disordered linker 1 (residues 206–227) in red, the helical structure (residues 226–288) that docks PlyCA to PlyCB in yellow, the second disordered linker 2 (residues 289–308) in a dashed red line, and the CHAP domain (residues 309–465) in green. Regions of disordered/absent density are depicted by dashed lines. (*B*) The PlyCB CBD alone colored alternately in magenta/cyan and labeled A–H.

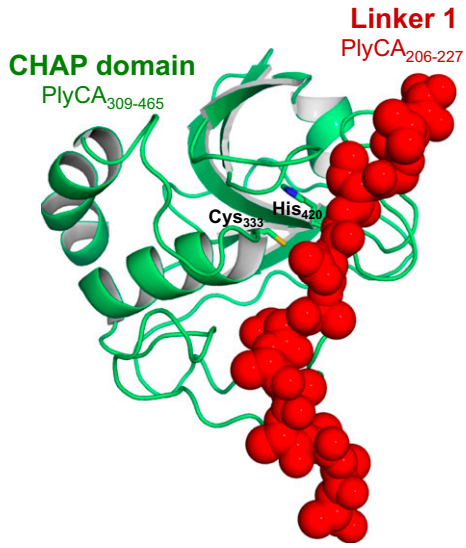


Fig. 2. The active site of the CHAP is occluded by linker region 1 (between the GyH domain and CHAP). Carbon atoms of CHAP catalytic residues Cys333 and His420 are shown in green, and the linker region is shown as red spheres and clearly blocks access to the CHAP active site. Mutation of a linker residue ($PlyCA_{S224Q}$) had no effect on the activity of the PlyC holoenzymes (Table S3 and Fig. S3).

groove of a single monomer is sufficient to mediate binding the holoenzyme or whether all eight such sites function to coordinate the PlyC lysin onto target bacterial cell walls (Fig. 3).

Catalytic Protein PlyCA and Cell Wall Binding PlyCB Interact Through Unique Protein–Protein Interaction. Despite the high resolution of these data, in the PlyCB octamer alone, the N-terminal region (residues 1–8) of each PlyCB monomer on the planar side is disordered in electron density. In contrast, in the PlyC holoenzyme structure, four of the PlyCB N-termini assemble to form a four-stranded parallel β-sheet (Fig. 1*B*). The docking domain of PlyCA, which comprises an antiparallel bundle of three α-helices, sits on top of this platform (Fig. 1*A*). Together, these structures form an exclusively hydrophobic interface (707.5 $Å^2$ of buried accessible surface area) and include the majority of interactions between PlyCA and the PlyCB octomer. Sequential deletion of the PlyCB N-terminal residues (Table S3 and Fig. S3) confirms the importance of this interaction, as deletion or mutation of PlyCB residues 3 to 8 results in loss of holoenzyme assembly (Table S3 and Fig. S3). Mutagenesis of the PlyCB-Ile_3 to Ala or Asp also resulted in the loss of holoenzyme assembly, whereas PlyCB-Ile_3Lys resulted in the formation of inclusion bodies (Table S3 and Fig. S3).

PlyCA Has Two Distinct Catalytic Domains Contributing Amidase and Glycosidase Activity to Achieve Cell Lysis. Although the presence of a CHAP domain in PlyCA has been previously reported (14), limited proteolysis of PlyCA indicated that a second globular domain was present in the protein. The holoenzyme structure revealed this domain was present in the N terminus of the protein. VAST and DALI searches show the PlyCA N-terminal domain (residues 1–205) is homologous to class IV family 19 chitinases—a branch of the glycoside hydrolase enzyme superfamily. We therefore have referred to the N-terminal domain of PlyCA as a putative glycosyl hydrolase (GyH) domain. Structural superposition between the PlyCA GyH domain and the top scoring chitinase structure (3CJL) reveals that only a central core of five helices is conserved (Fig. 4). Importantly, however, these data, together with sequence alignments of similar GyH domains (identified by using PSI-BLAST; Fig. S4), revealed that PlyCA-Glu_{78}, a position conserved in other GyH domains (Fig. S4), corresponds to a key catalytic residues of the family 19 chitinases (Glu_{68} in 2CJL). Two other absolutely conserved residues cluster in this region—PlyCA-Tyr_{74} and PlyCA-Asn_{87} (Fig. S4). Together, we suggest these residues represent a catalytic center.

Family 19 chitinases function to cleave β(1→4) glycosidic bonds of unbranched chains of *N*-acetylglucosamine polymers. As the bacterial peptidoglycan varies in structure from chitin [peptidoglycan is characterized by alternating β(1→4) linked sugars of *N*-acetylglucosamine (GlcNAc) and *N*-acetylmuramic acid (MurNAc) with each MurNAc attached to a short (4- to 5-residue) stem peptide of L- and D-form amino acids], we wanted to ascertain if this domain did indeed act as a glucosaminidase to release free reduced sugars. Notwithstanding any potential glucosaminidase activity associated with the GyH domain, PlyC has been previously shown to have an amidase activity that cleaves between MurNAc and L-alanine of the stem peptide, producing a free N terminus (and releasing amines) (17). Later studies demonstrated that inactivation of the CHAP domain, presumed to be responsible for the amidase activity, significantly reduced the lytic activity of PlyC on streptococcal cells (14). In light of the distinct CHAP and GyH domains revealed by the crystal structure and the potential for multiple catalytic activities, we used two separate biochemical assays to analyze the free sugar (18) and amine (19) release that occurs after PlyC hydrolysis of purified streptococcal peptidoglycan. Purified WT PlyC, PlyCA, and the PlyCB octamer were used as controls. Streptococcal peptidoglycan was the substrate for both assays. For analysis of the contribution of the CHAP domain, we used a catalytic residue mutant PlyCA-Cys_{333}Ser (Fig. S5) (14). To analyze the contribution of the GyH domain (and furthermore to test the functional importance of the putative catalytic center) we used four mutations: Glu_{78}Ala, Asn_{87}Ala, His_{88}Ala, and Asn_{104}Ala (Fig. S5). Biophysical and gel filtration studies suggested all of the mutations formed PlyC holoenzyme and were properly folded (as detailed in *Methods*).

Our experiments revealed that PlyC indeed possesses significant glycosidase activity (Fig. 5*A*). We further demonstrated that mutation of putative catalytic residues or surrounding residues in the GyH domain reduces this activity by ∼85%. Interestingly, however, even the CHAP active-site KO (Cys333Ser) significantly decreased glycosidase activity suggesting that the CHAP domain itself has intrinsic glycosidase activity, or that it cooperatively interacts with the GyH domain. Finally, we showed that double mutations in the CHAP and GyH domains resulted in total loss of glycosidase activity (Fig. 5*A*). Next, we performed a similar series of experiments by using an amidase assay (Fig. 5*B*). Consistent with previous studies (14), we showed that the Cys333Ser mutation resulted in an almost complete loss of amidase function and corresponds to ∼99% loss in streptococcal lytic activity. Interestingly, catalytic mutations in the GyH domain also significantly reduced overall amidase activity, but not to the extent of the Cys333Ser mutation. Similar to the glycosidase assay, double mutations to both the CHAP and GyH domains were found to be completely devoid of any measurable amidase activity.

Taken together, our biochemical data reveal that glycosidase and amidase activities exist in the PlyC holoenzyme. Based on our analysis of active-site mutants, homology to characterized domains, and previous studies (14, 17), we suggest the CHAP domain contains the amidase activity, specifically an *N*-acetylmuramoyl-L-alanine amidase activity, and the GyH domain contains the glycosidase activity. As expected, PlyCB was found to be absent of any type of catalytic activity (Fig. 5 *A* and *B*). However, we note that neither recombinant purified CHAP domain nor full-length PlyCA, which possesses both CHAP and GyH catalytic domains, displayed significant amidase,

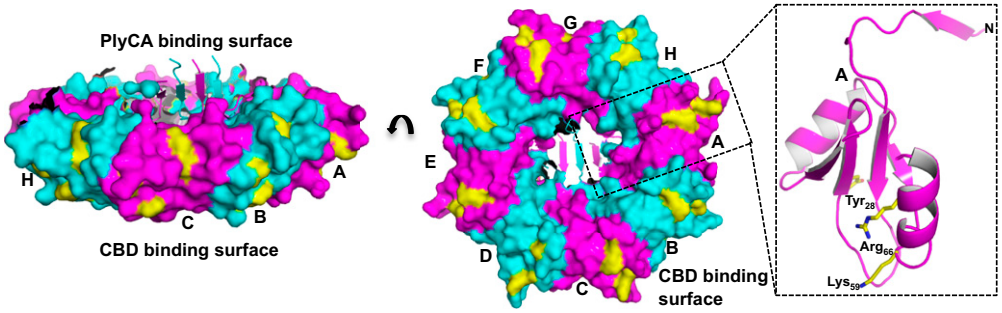


Fig. 3. PlyCB has eight cell wall-binding grooves. The cell wall-binding surface of PlyCB alone (colored as in Fig. 1) shows the residues involved in cell wall binding in yellow. *Inset*: Cartoon depiction of PlyCB monomer A shows cell wall binding residues in yellow sticks as indicated.

glycosidase, or lytic activities (Fig. 5). Any minor activity observed for these constructs is presumably a consequence of random collisions of the catalytic domain(s) with the cell wall in the absence of the PlyCB binding domain. These data serve to further highlight the critical importance of the cell wall binding function of PlyCB to the overall activity of PlyC.

To investigate if catalytic efficiency was a result of cooperative or synergistic activity, we examined the activity of each domain in isolation. First, we constructed two new mutants, PlyCΔCHAP and PlyCΔGyH, each of which contain WT PlyCB and PlyCA with the specified catalytic domain deletion. Next, we examined the lytic activity of equimolar amounts (44 nM) of PlyC, PlyCA alone, CHAP alone, PlyCΔCHAP, and PlyCΔGyH (Fig. 5*C*). The two deletion mutants had <1% of WT activity, and complementation with purified CHAP did not restore WT activity. Even when we diluted WT PlyC to 1 nM, it still displayed more than twice the lytic activity of 44 nM PlyCΔCHAP complemented with 44 nM CHAP (Fig. 5*C*). We therefore suggest that the unusually high processivity of the PlyC holoenzyme can be explained by synergistic effects resulting from the positioning of both catalytic domains in PlyCA, and in the context of the PlyCB binding domain. The synergy/cooperativity between catalytic domains is further supported by biochemical assays showing a significant decrease in glycosidase activity of the CHAP (i.e., amidase domain) active-site KO and a decrease in amidase activity of the GyH (i.e., glycosidase domain) active-site KOs (Fig. 5 *A* and *B*).

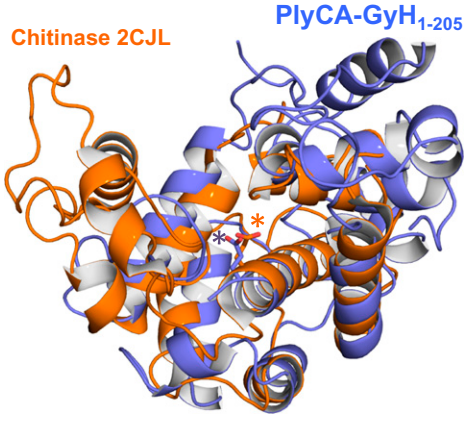


Fig. 4. The PlyCA GyH domain is homologous to class IV chitinases. Superposition of the GyH domain (light blue) with a representative chitinase (orange; 2CJL: rmsd of 3.3 Å over 107 Cα atoms). The position of PlyCA GyH (Glu_{78}) (blue asterisk) and the catalytic residue 2CJL(Glu_{68}) (orange asterisk) is shown as a stick.

Discussion

The PlyC lysin is by far the most active lysin described to date, yet it is larger than most lysins by nearly a factor of four and remains the only known lysin composed of distinct subunits. Accordingly, the molecular basis for this potency is interesting from the perspective of the development of novel antibacterial therapeutic agents and molecular dynamics. Here we show through structural and biochemical studies that PlyC contains two functional catalytic domains. This finding was unexpected, as the relationship between the GyH domain and the bacterial chitinases is not detectable through sequence-based searches alone. We further show that PlyC displays amidase and glycosidase activity, and thus can cleave both classes of bond present in the bacterial peptidoglycan. It is notable that several traditional lysins have been identified that likewise contain tandem catalytic subunits. However, when studied at the biochemical level, most of these "extra" domains were found to be silent. For example, the staphylococcal Φ11 lysin has both D-alanyl-glycyl endopeptidase and *N*-acetylmuramoyl-L-alanine amidase catalytic domains (6). Although the endopeptidase domain was active by itself (20, 21), the amidase domain was shown to be silent by deletion analysis (21). Likewise, the LysK lysin has an active endopeptidase domain and an inactive amidase domain (22, 23), and the streptococcal λSA2 lysin contains an active D-glutaminyl-L-lysine endopeptidase and an inactive *N*-acetylglucosaminidase domain (24, 25). Thus, PlyC is one of the first characterized lysins with two active and distinct catalytic domains.

As much as the structural and biological data reveal about the unique properties of PlyC, the question remains, however, why the combination of PlyCB and PlyCA forms such an effective enzyme. Synergy or cooperation between the CHAP and GyH domains may, in part, begin to explain the enhanced lytic activity displayed by PlyC compared with traditional lysins. Significantly, despite possessing amidase and glycosidase activities, a single point mutation to the catalytic center of the CHAP or the GyH domain reduces streptococcal lytic activity by 90% to 99% (Fig. 3 and Table S3). At present, however, it is not known if the two domains act synergistically or cooperatively. Under a synergistic scenario, the two domains could simultaneously cleave the peptidoglycan between GlcNAc and MurNAc, as well as MurNAc and L-alanine, thereby destabilizing both the vertical and horizontal scaffolds of the peptidoglycan. These large defects in the superstructure of the cell wall would lead to rapid bacterial lysis, which would be observed as enhanced activity over a typical lysin. For the cooperative scenario, one domain may be largely silent catalytically, but may function to position the peptidoglycan for optimal hydrolysis by the other catalytic domain. In this way, cooperativity between the two domains may significantly enhance the turnover rate.

Another unique feature of PlyC that stands out in comparison with other lysins is the presence of the octameric cell wall

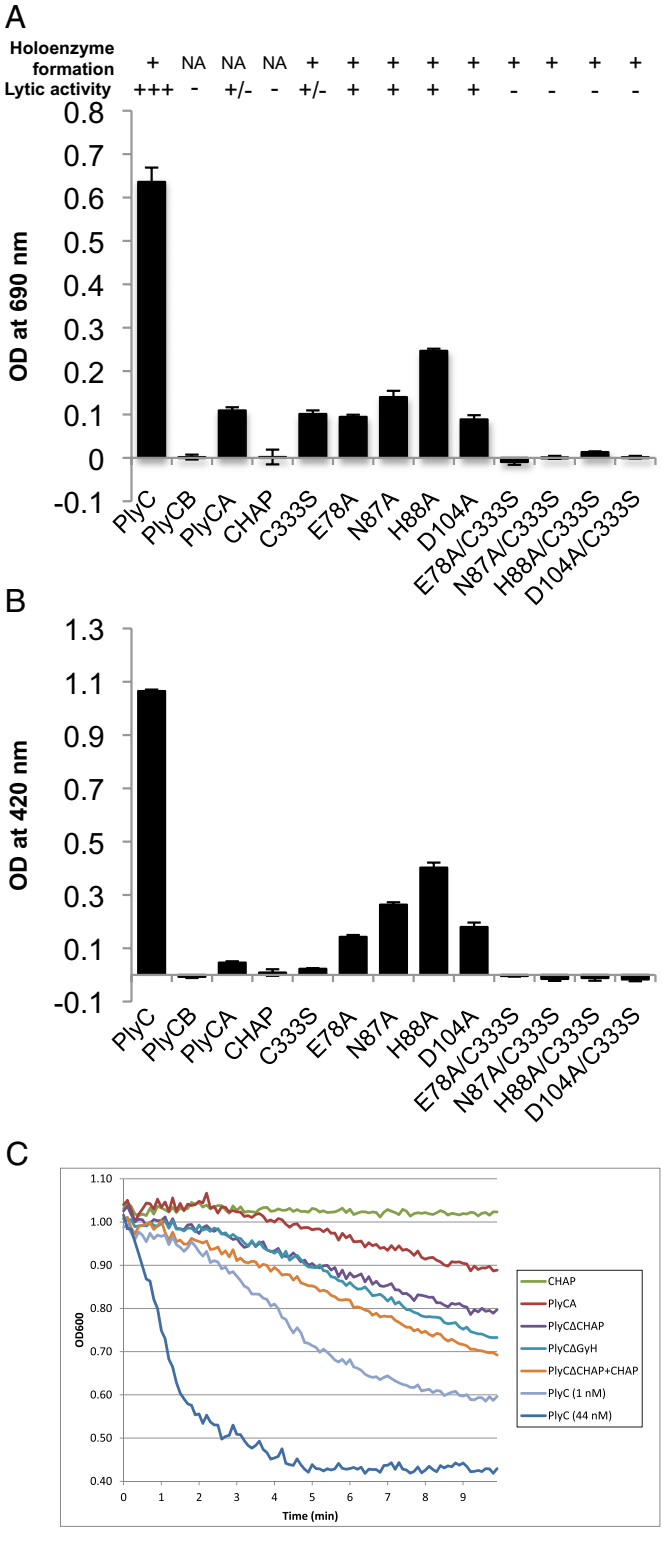


Fig. 5. Biochemical activity of PlyC domains. PlyC, PlyCB alone, PlyCA alone, CHAP, the PlyCA CHAP domain active-site KO (C333S), the four PlyCA GyH mutants with diminished lytic activity (E78A, N87A, H88A, D104A), and the four CHAP/GyH double mutants were tested on purified streptococcal peptidoglycan for (*A*) glycosidase activity or (*B*) amidase activity (*Methods*). Additionally, holoenzyme formation and lytic activity on live streptococcal cells is noted at the top of the figure. +++, WT lytic activity; +, significantly diminished lytic activity (≤10% WT); +/–, negligible lytic activity (≤1% WT), and –, no detectable lytic activity. (*C*) Additional lytic activity on selected construct. PlyC, CHAP, PlyCA, PlyCΔCHAP, PlyCΔGyH, and PlyCΔCHAP + CHAP were all at a final concentration of 44 nM. In addition, PlyC is also shown at 1 nM final concentration to highlight the diminished activity of all non-WT constructs. Plots shown are the average OD reading at each data point from independent experiments, each in triplicate.

binding subunit, PlyCB. Our data further highlight the crucial functional importance of the PlyCB subunit to direct binding to the streptococcal surface. Although PlyCB itself has no integral enzymatic function, PlyCA is almost devoid of any lytic activity (<1% of WT) in the absence of PlyCB. Furthermore, our mutational data reveal that PlyCB contains eight potential binding sites for cell wall components. In contrast, most characterized lysins contain only a single CBD. Our structural data do not, however, provide evidence whether all eight binding sites participate simultaneously to coordinate holoenzyme binding to the streptococcal cell wall. Concurrent binding may indicate a very tight, stable interaction between the holoenzyme and cell wall, whereas a consecutive interaction may lead to a more transient interaction. It is worthy to note that similar classes of enzymes have used tight binding domains in combination with flexible catalytic domains to hydrolyze multiple substrate bonds within a range between domains allowed by the linker flexibility. Interestingly, the cellulase family of enzymes, responsible for degrading the insoluble, polysaccharide cellulose, display a similar architecture and mechanism for substrate recognition and cleavage to that which we propose for PlyC. Most cellulases have a modular structure with a catalytic module and a cellulose-binding module. Long, flexible linkers are pivotal for the ability of the enzyme to access the complex surface of cellulose and allow a caterpillar-like displacement to move through the substrate (26, 27). For example, small-angle X-ray scattering studies of the fungal cellulase, Cel45, suggest as many as four cellobiose units on a cellulose chain are within range of the catalytic domain when binding domain is attached to a specific site and the linker is fully extended (26). Appreciably, both catalytic PlyCA domains are connected to the central helical domain (the primary point of contact with PlyCB) by extensive flexible linker sequences, and neither the GyH domain nor the CHAP domain makes substantial contact with the PlyCB ring. Together, these data suggest that the position of the catalytic domains is not spatially restricted by strong interactions with the PlyCB ring. Whether PlyC shares a similar mechanism as put forth for certain cellulases remains to be elucidated. Nonetheless, PlyC is distinctive among known lysin family members for its two functionally active catalytic domains, extreme flexibility between subunits, octameric binding domain, and high processivity. Identification of additional multisubuint lysins may help further define the molecular mechanisms of this new lysin subfamily.

Methods

Biological Methods. ***Cloning, generation of mutants, and purification techniques.*** The constructs for pBAD24::*plyC*, pBAD24::*plyCA*, and pBAD24::*plyCB* were previously described (14). Construction of the CHAP, PlyCΔCHAP and PlyCΔGyH domain are described in detail in *SI Methods*. Expression of constructs was as previously described (13, 14) with the exception of the CHAP domain (*SI Methods*). Details of the mutagenesis and analysis of mutants are given in *SI Methods*.

Lytic activity and cell wall binding. Deletions, mutants, and subunits were evaluated for lytic activity, proper folding, formation of the holoenzyme (or PlyCB octamer where appropriate), and binding to the streptococcal peptidoglycan. Spectrophotometric lytic assays were performed with an overnight culture of *S. pyogenes* D471 grown at 37 °C in Todd–Hewitt broth supplemented with 1% wt/vol yeast extract. Streptococci were washed in PBS solution (pH 7.4) and resuspended at an OD_{600} of 1.5. In a 96-well plate, 100 μL of bacteria was mixed with 100 μL of enzyme (final concentration of 44 nM), and the OD_{600} was measured kinetically on a SpectraMax 190 (Molecular Devices) every 6 s for 10 min. All assays were performed in triplicate.

To assess cell wall binding, the fluorescent dye Alexa Fluor 488 (Molecular Probes) was conjugated to primary amines of purified (1 mg each) proteins through a tetrafluorophenyl ester. The reaction was quenched by the addition of 100 mM Tris and immediately applied to a PD10 desalting column to separate the labeled protein from unreacted dye. A total of 10 μg of labeled protein was then incubated with a fresh overnight culture of *S. pyogenes* D471 resuspended in PBS solution. The cells were washed in PBS solution and viewed on a Nikon Eclipse 80i epifluorescent microscope equipped with a Retiga 2000R CCD camera. The degree of fluorescent staining of the cell wall (i.e., binding of the protein) was quantified with the Q Capture Pro or NIS-Elements software packages.

Biochemical assays. The *S. pyogenes* D471 peptidoglycan was purified by a French press method as previously described (7), lyophilized, and used for all biochemical assays. For all assays, purified peptidoglycan was resuspended in PBS solution, pH 7.2, to an OD_{550} of 1.0. WT PlyC or mutants were added at a concentration of 3.5 μg/mL (30 nM) and allowed to react for with peptidoglycan for 5, 15, 30, or 60 min at 37 °C. All experiments consisted of at least two independent experiments assayed in triplicate. For analysis of reducing sugars released from the peptidoglycan, we used the method of Park and Johnson (28) as modified by Spiro (18). This method is based on the reduction of ferricyanide by sugars in an alkaline solution followed by the formation of Prussian blue (ferric ferrocyanide) upon addition of ferric ions. To determine an increase in free amine groups, we used the trinitrophenylation reaction originally described by Satake et al. (29) and modified by Mokrasch (30).

Structural Biology Methods. ***Crystallization, X-ray data collection, structure determination, and refinement.*** Crystals were grown in all cases by using the hanging drop vapor diffusion method, with 1:1 (vol/vol) ratio of protein to mother liquor (0.5 mL well volume). Data were collected by using in-house and synchrotron radiation (Industrial Macromolecular Crystallography Association Collaborative Access Team, Advanced Photon Source, Argonne, IL). The structure of PlyCB was solved by using phases determined using sulfur single anomalous dispersion phasing using the program AutoSHARP (31). Density modification was applied by using the program SOLOMON followed by automated building by using the program ARP/wARP (32). The refined structure of the octameric PlyCB ring was used as a molecular replacement search probe for the native PlyC dataset using PHASER (33), resulting in two clear solutions. Electron density maps were then calculated using the AutoBUSTER (34) program (*SI Methods*). The initial Fo-Fc map revealed significant unbiased regions of positive connective density that clearly did not form part of PlyCB probe. This strongly indicated a correct molecular replacement solution, and these initial phases were used to manually build the two PlyCA molecules. We noted four short regions of the PlyCA molecule that contained residues identified as Ramachandran outliers (*SI Methods*). Data collection and refinement statistics are detailed in Tables S1 and S2. Further details of structural determination are provided in *SI Methods*.

Figures and structural coordinates. Superpositions were conducted using PDBeFold (35) (http://www.ebi.ac.uk/msd-srv/ssm/). Pymol was used to produce all structural representations (http://www.pymol.org). The coordinates and structure factors are available from the Protein Data Bank under ID codes 4F87 and 4F88. Raw data and images are available from TARDIS (36) (www.tardis.edu.au).

ACKNOWLEDGMENTS. The authors thank the Industrial Macromolecular Crystallography Association Collaborative Access Team (Advanced Photon Source) for beam time and for technical assistance, Jamie Rossjohn for technical assistance with data collection, Joseph Kotarek for FTIR assistance, and Amanda Altieri for CD assistance. This work was supported in part by US Public Health Service Grant AI11822 (to V.A.F.) and US Department of Defense Grant DM102823 (to D.C.N.). S.M. is an Australian Research Council (ARC) Future Fellow. A.M.B. is a National Health and Medical Research Council (NHMRC) Senior Research Fellow. J.C.W. is an ARC Federation Fellow and an NHMRC Honorary Principal Research Fellow.

1. Young R (1992) Bacteriophage lysis: Mechanism and regulation. *Microbiol Rev* 56:430–481.
2. Fischetti VA (2011) Exploiting what phage have evolved to control gram-positive pathogens. *Bacteriophage* 1:188–194.
3. Hermoso JA, García JL, García P (2007) Taking aim on bacterial pathogens: From phage therapy to enzybiotics. *Curr Opin Microbiol* 10:461–472.
4. Fischetti VA, Nelson D, Schuch R (2006) Reinventing phage therapy: Are the parts greater than the sum? *Nat Biotechnol* 24:1508–1511.
5. Hermoso JA, et al. (2003) Structural basis for selective recognition of pneumococcal cell wall by modular endolysin from phage Cp-1. *Structure* 11:1239–1249.
6. Navarre WW, Ton-That H, Faull KF, Schneewind O (1999) Multiple enzymatic activities of the murein hydrolase from staphylococcal phage phi11. Identification of a D-alanyl-glycine endopeptidase activity. *J Biol Chem* 274:15847–15856.
7. Pritchard DG, Dong S, Baker JR, Engler JA (2004) The bifunctional peptidoglycan lysin of Streptococcus agalactiae bacteriophage B30. *Microbiology* 150:2079–2087.
8. Stark CJ, Hoopes JT, Bonocora RP, Nelson DC (2010) Bacteriophage lytic enzymes as antimicrobials. *Bacteriophage in the Detection and Control of Foodborne Pathogens*, eds Sabour PV, Griffith MW (ASM Press, Washington DC), pp 137–156.
9. Croux C, Ronda C, López R, García JL (1993) Interchange of functional domains switches enzyme specificity: Construction of a chimeric pneumococcal-clostridial cell wall lytic enzyme. *Mol Microbiol* 9:1019–1025.
10. Daniel A, et al. (2010) Synergism between a novel chimeric lysin and oxacillin protects against infection by methicillin-resistant Staphylococcus aureus. *Antimicrob Agents Chemother* 54:1603–1612.
11. Schmelcher M, Tchang VS, Loessner MJ (2011) Domain shuffling and module engineering of Listeria phage endolysins for enhanced lytic activity and binding affinity. *Microb Biotechnol* 4:651–662.
12. Krause RM (1957) Studies on bacteriophages of hemolytic streptococci. I. Factors influencing the interaction of phage and susceptible host cell. *J Exp Med* 106:365–384.
13. Nelson D, Loomis L, Fischetti VA (2001) Prevention and elimination of upper respiratory colonization of mice by group A streptococci by using a bacteriophage lytic enzyme. *Proc Natl Acad Sci USA* 98:4107–4112.
14. Nelson D, Schuch R, Chahales P, Zhu S, Fischetti VA (2006) PlyC: A multimeric bacteriophage lysin. *Proc Natl Acad Sci USA* 103:10765–10770.
15. Rossi P, et al. (2009) Structural elucidation of the Cys-His-Glu-Asn proteolytic relay in the secreted CHAP domain enzyme from the human pathogen Staphylococcus saprophyticus. *Proteins* 74:515–519.
16. Pai CH, et al. (2006) Dual binding sites for translocation catalysis by Escherichia coli glutathionylspermidine synthetase. *EMBO J* 25:5970–5982.
17. Fischetti VA, Zabriskie JB, Gotschlich EC (1972) Physical, chemical and biological properties of Type 6 M-protein extracted with purified streptococcal phage-associated lysin. *Fifth International Symposium on Streptococcus pyogenes*, ed Haverkorn MJ (Excerpta Medica, Amsterdam), pp 26–36.
18. Spiro RG (1966) Analysis of sugars found in glycoproteins. *Methods Enzymol* 8:3–26.
19. Snyder SL, Sobocinski PZ (1975) An improved 2,4,6-trinitrobenzenesulfonic acid method for the determination of amines. *Anal Biochem* 64:284–288.
20. Donovan DM, Lardeo M, Foster-Frey J (2006) Lysis of staphylococcal mastitis pathogens by bacteriophage phi11 endolysin. *FEMS Microbiol Lett* 265:133–139.
21. Sass P, Bierbaum G (2007) Lytic activity of recombinant bacteriophage phi11 and phi12 endolysins on whole cells and biofilms of Staphylococcus aureus. *Appl Environ Microbiol* 73:347–352.
22. Becker SC, et al. (2009) LysK CHAP endopeptidase domain is required for lysis of live staphylococcal cells. *FEMS Microbiol Lett* 294:52–60.
23. Horgan M, et al. (2009) Phage lysin LysK can be truncated to its CHAP domain and retain lytic activity against live antibiotic-resistant staphylococci. *Appl Environ Microbiol* 75:872–874.
24. Donovan DM, Foster-Frey J (2008) LambdaSa2 prophage endolysin requires Cpl-7-binding domains and amidase-5 domain for antimicrobial lysis of streptococci. *FEMS Microbiol Lett* 287:22–33.
25. Pritchard DG, Dong S, Kirk MC, Cartee RT, Baker JR (2007) LambdaSa1 and LambdaSa2 prophage lysins of Streptococcus agalactiae. *Appl Environ Microbiol* 73:7150–7154.
26. Receveur V, Czjzek M, Schülein M, Panine P, Henrissat B (2002) Dimension, shape, and conformational flexibility of a two domain fungal cellulase in solution probed by small angle X-ray scattering. *J Biol Chem* 277:40887–40892.
27. von Ossowski I, et al. (2005) Protein disorder: Conformational distribution of the flexible linker in a chimeric double cellulase. *Biophys J* 88:2823–2832.
28. Park JT, Johnson MJ (1949) A submicrodetermination of glucose. *J Biol Chem* 181: 149–151.
29. Satake K, Okuyama T, Ohashi M, Shinoda T (1960) The spectrophotometric determination of amine, amino acid and peptide with 2,4,6,-trinitrobenzene 1-sulfonic acid. *J Biochem* 47:654–660.
30. Mokrasch LC (1967) Use of 2,4,6-trinitrobenzenesulfonic acid for the coestimation of amines, amino acids and proteins in mixtures. *Anal Biochem* 18:64–71.
31. Vonrhein C, Blanc E, Roversi P, Bricogne G (2007) Automated structure solution with autoSHARP. *Methods Mol Biol* 364:215–230.
32. Cohen SX, et al. (2008) ARP/wARP and molecular replacement: The next generation. *Acta Crystallogr D Biol Crystallogr* 64:49–60.
33. McCoy AJ, Grosse-Kunstleve RW, Storoni LC, Read RJ (2005) Likelihood-enhanced fast translation functions. *Acta Crystallogr D Biol Crystallogr* 61:458–464.
34. Bricogne G, et al. (2011) BUSTER, version 2.8 (Global Phasing, Cambridge, UK).
35. Krissinel E, Henrick K (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol* 372:774–797.
36. Androulakis S, et al. (2008) Federated repositories of X-ray diffraction images. *Acta Crystallogr D Biol Crystallogr* D64:810–814.

Supporting Information

McGowan et al. 10.1073/pnas.1208424109

SI Methods

Biological Methods. ***Cloning, generation of mutants, and purification techniques.*** The constructs for pBAD24::*plyC*, pBAD24::*plyCA*, and pBAD24::*plyCB* were previously described (1). A cysteine, histidine-dependent amidohydrolases/peptidase (CHAP) domain construct was made that comprised of the C-terminal 166 aa residues of PlyCA. The CHAP domain sequence was amplified by PCR from pBAD24::*plyC* using the forward primer (5′-GTCCCCGGGATGTCACCAGGTCAAAACGATTTAG-3′) containing a SmaI restriction site (underlined) and reverse primer 5′-GTCAAGCTTTTATTTAAATGTTATCA AACCAGT-3′ containing a HindIII restriction site (underlined). The product was ligated to a SmaI/HindIII-digested pBAD24 (Amp^r) vector and used to transform *Escherichia coli* DH5α or BL21 (DE3) cells. Plasmid was rescued from *E. coli* and sequenced to confirm insert.

All point mutants and/or deletion mutants were created using the QuikChange Site-Directed Mutagenesis kit (Stratagene) according to the manufacturer's instructions. Briefly, plasmid DNA was purified from DH5α/pBAD24::*plyC* or DH5α/pBAD24::*plyCB* and *Pfu*Turbo was used to amplify the genes via PCR in conjunction with overlapping primers that incorporated each point mutation. The parental, methylated DNA was digested by *DpnI*, and subsequent plasmids were transformed into DH5α and BL21(DE3) cells according to standard techniques. For PlyCΔCHAP, the C-terminal 166 aa (i.e., all of the CHAP domain and half of linker 2) were deleted from PlyCA by introduction of a premature stop codon at position 894 of the *plyCA* gene in the DH5α/pBAD24::*plyC* background. Similarly, the PlyCΔGyH truncation mutant was created by using overlapping primers in conjunction with the QuikChange kit to delete the first 225 aa of PlyCA (i.e., GyH domain and linker 1 region).

Expression of constructs was induced with 0.5% arabinose for 4 h. PlyC, PlyCA, and PlyCB were purified as previously described (1, 2). All point and deletion mutants likewise followed the same protocols, taking advantage of affinity purification on hydroxyapatite via interactions with PlyCB. Only the CHAP domain required a de novo purification strategy. For this protein, harvested cells (3 L) were pelleted, resuspended in 40 mL of PBS solution (pH 7.4), and sonicated on ice. Following centrifugation of the lysate, ammonium sulfate was slowly added to the supernatant to 40% saturation and allowed to stir at 4 °C overnight. The sample was then centrifuged, and the pellet was resuspended in 20 mM Tris, pH 8.0, and dialyzed against the same. Next, the sample was applied to a Mono Q column (GE Healthcare) and eluted with 20 mM Tris, 1 M NaCl (pH 8.0) over 30 column volumes. Final purification was achieved on an S-200 26/60 gel filtration column (GE Healthcare) equilibrated in PBS solution (pH 7.4)

For selenomethionine labeling, *E. coli* B834 cells (Novagen) were transformed with pBAD24::*plyC* or pBAD24::*plyCB* and methionine auxotrophy was confirmed by observing growth in M9 minimal media supplemented with 50 mg/L methionine, whereas no growth was noted in M9 media alone. Fresh 20-mL overnight cultures of B834/pBAD24::*plyC* and B834/pBAD24::*plyCB* were washed twice with M9 media and used to inoculate a 2-L flask of M9 minimal media supplemented with 50 mg/L selenomethionine. Expression was induced with 0.5% arabinose, and the proteins were purified as described earlier.

Characterization of domains, deletions, and mutants. Deletions, mutants, and subunits were evaluated as described in *Methods*. All point mutants or deletions were analyzed by a Chirascan CD spectrometer (Applied Photophysics) with the Global 3 software suite to determine if they were properly folded. A total of 400 μL of a 1 mg/mL solution in 20 mM phosphate buffer (pH 7.4) was used for each protein. The change in ellipticity at 222 nm, which monitors unfolding of α-helix structures, was used to monitor unfolding over a range of 25 °C to 90 °C at a rate of 0.2 °C/min. Additionally, several of the samples were also analyzed by FTIR. In this case, samples were placed in a 7-μm Biotools Biocell and transmission IR spectra were collected with an MCT detector in a Hyperion microscope (Nikon) with a Vertex 80 FTIR spectrometer (Bruker). A total of 480 scans were taken for both background and sample spectra. Contributions from water vapor were subtracted and secondary structure was evaluated using Bruker OPUS 6.5 software.

Next, purified proteins were subjected analytical gel filtration to determine if mutants affected formation of the PlyC holoenzyme (~114 kDa) or the PlyCB octamer (~64 kDa). A total of 100 μL of each sample (100 μg) was applied to a preequilibrated Superose12 gel filtration column (GE Healthcare) and run under isocratic conditions in PBS solution for 1.5 column volumes on an AKTA FPLC system (GE Healthcare). Molecular mass was estimated from a standard curve [linear regression of log (molecular mass) against retention volume] generated by using gel filtration standards (Bio-Rad), as well as superimposition on WT PlyC and PlyCB control runs.

By definition, all constructs that had lytic activity also bound the streptococcal peptidoglycan. However, as several mutants were made to probe potential cell wall-binding pockets on the surface of PlyCB, we fluorescently labeled select proteins to visually observe direct binding through fluorescent microscopy. This assay is described in *Methods*.

Structural Biology Methods. ***Crystallization, X-ray data collection, structure determination, and refinement of PlyCB.*** The purified PlyCB enzyme was dialyzed against 25 mM Hepes, pH 6.0, 100 mM NaCl, 2 mM EDTA and concentrated to 17 mg/mL The crystals were grown by using the hanging drop vapor diffusion method, with 1:1 (vol/vol) ratio of protein to mother liquor (0.5 mL well volume). Cubic crystals appeared overnight in 25% (vol/vol) methylpentanediol, 0.1 M Hepes (pH 6.0), and 0.2 M sodium citrate and reached full size in 5 d.

A dataset to 1.7-Å resolution was collected at 100 K by using an in-house CuKa source, and processed and scaled using MOSFLM and SCALA (3, 4). Phases were determined by using sulfur single anomalous dispersion phasing using the program AutoSHARP (5). Density modification was applied using the program SOLOMON followed by automated building using the program ARP/wARP (6).

A further dataset to 1.4-Å resolution was collected at 100 K using synchrotron radiation [Industrial Macromolecular Crystallography Association Collaborative Access Team, Advanced Photon Source (APS), Argonne, IL]. Data were processed and scaled at APS by using the HKL2000 suite (3). For all processed data, 5% of each dataset was flagged for calculation of R_{Free} (7), with neither a sigma nor a low-resolution cutoff applied to the data. A summary of statistics is provided in Tables S1 and S2.

Starting from the auto-built 1.7-Å model and then the 1.4-Å data, maximum-likelihood refinement using REFMAC (3), incorporating translation, liberation and screw-rotation displacement refinement, was carried out, using a bulk solvent correc-

tion (Babinet model with mask). Building was guided by manual inspection of the model and R_{Free}. All model building and structural validation was done by using COOT (8). Water molecules were added to the model by using ARP/wARP (6) when the R_{Free} reached 25%. Solvent molecules were retained only if they had acceptable hydrogen-bonding geometry contacts of 2.5 to 3.5 Å with protein atoms or with existing solvent and were in good $2F_o$-F_c and F_o-F_c electron density. Subsequent crystallographic and structural analysis was performed by using the CCP4i interface (9) to the CCP4 suite (3) unless stated otherwise.

Crystallization and X-ray data collection, structure determination, and refinement of PlyC. Purified PlyC enzyme was concentrated to ~13 mg/mL and crystals grown using the hanging drop vapor diffusion method, with 1:1 (vol/vol) ratio of protein to mother liquor (0.5 mL well volume). The crystals appeared in 0.1 M Hepes, pH 7.4, 12% (vol/vol) polyethylene glycol 4000, and 10% (vol/vol) isopropanol. Crystals took 5 d to grow and reached full size by 7 d. Crystal morphology were typically highly stacked plates.

The PlyC crystals were flash-frozen before data collection with 10% (vol/vol) glycerol as the cryoprotectant. More than 500 crystals were screened and typical diffraction data obtained were limited to ~10 Å. A single dataset to 3.3 Å was collected from one single crystal at 100 K by using synchrotron radiation (at APS). Diffraction images were processed using MOSFLM (4), pointless (10), and SCALA (10) from the CCP4 suite (3). For all processed data, 5% of each was flagged for calculation of R_{Free} (7), with neither a sigma nor a low-resolution cutoff applied to the data. To avoid bias in the R_{Free} test set resulting from NCS-related reflections, we selected a test set with thin resolution shells by using SFTools from the CCP4 suite. A summary of statistics is provided in Tables S1 and S2.

The Matthews coefficient for the primitive orthorhombic cell ($P2_12_12_1$; a = 93.9; b = 117.4; c = 222.2; β= 90.0°) indicated that there could be two complete PlyC molecules present in the asymmetric unit with a solvent content of 53.8%. The structure of the octameric PlyCB ring was used as a molecular replacement probe for the native PlyC dataset in PHASER (11). Two clear solutions in both the rotation and translation functions were evident. Molecular replacement attempts by using CHAP domains (2IO7, 2K3A, 2VOB, or a core assembly of all) failed to identify the location of the CHAP domain in the PlyCA molecules. These molecular replacement experiments were attempted with no initial phasing information or the initial phases obtained from the placement of the PlyCB probe. No further improvement of the phases was obtained from the initial PHASER solution of two complete PlyCB rings.

This initial PHASER solution model was used as input into the AutoBUSTER (12) program, running *–autoncs* and *–TLSbasic* with no B factor refinement (i.e., *-B* none) and a target bond angle of 0.007 (i.e., -r 0.007) (13). The expected percentage solvent was held at 50% (i.e., *-s* 0.50), and the number of missing atoms was estimated to be 5,000 (i.e., *-M* Missingatoms nmiss = 5,000). The initial Fo-Fc map (after five BIG cycles) revealed significant unbiased regions of positive connective density that clearly did not form part of PlyCB probe. This strongly indicated a correct molecular replacement solution, and these initial phases were used to manually build the two PlyCA molecules.

The program COOT was used to place polyalanine helices into regions of electron density that had "helix-like" character. Further, polyalanine β-strands were also placed by COOT into electron density that clearly showed a four-stranded sheet within the gap in the PlyCB ring. These secondary structure elements were refined by rigid-body minimization using CNS defining each secondary structure element as a rigid body. The polyalanine model was further extended to the central β-sheet of the PlyCA CHAP domain and other regions of apparent helical density. This manual building occurred iteratively for several rounds of rigid-body minimization and autoBUSTER (running the same initial parameters). Medium-temperature (~1,500 K) torsion-angle simulated annealing was also used after several cycles of manual building and rebuilding. Alanine residues were replaced by serine if unbiased side-chain density was apparent. When the autoBUSTER R_{Free} value had decreased to less than 0.35, rigid body refinement was stopped and replaced by maximum-likelihood refinement using REFMAC and incorporating NCS restraints imposed only on each of the 16 copies of the PlyCB monomers (residues 15–60). Subsequent rounds of refinement (REFMAC and autoBUSTER) and manual building extended the secondary structure elements into loop connectivity throughout a majority of the PlyCA molecule. The final PlyCA models had 433 (chain 1) and 419 (chain 2) residues of the 465-aa protein.

Building side chains of the CHAP domain of PlyCA was guided by sequence similarity to other CHAP domain structures and known active site residues (1). Placement of the catalytic cysteine (Cys_{333}) and histidine (His_{420}) provided the initial starting point for PlyCA sequencing, and this was supported by the clear density for Trp residues 336, 339, and 340. Tryptophan residues in the GyH domain also aided the successful sequence assignment of the entire PlyCA molecule. The program PHENIX (14) was used to aid the correction of stereochemistry and perform final refinements. Iterative refinement with PHENIX and autoBUSTER produced the final model with an R/R_{Free} of 26.7/29.7. The final asymmetric unit consisted of two complete PlyC holoenzyme structures. We noted that PlyC copy 2 had weaker/less-connected electron density than copy 1. All final analysis and figures were thus based on PlyC copy 1.

Investigation of the Ramachandran plot for the PlyC structure revealed that most of the outliers were located in the PlyCA molecule. The two copies of the PlyCA molecule had similar Ramachandran outliers (5.83% and 6.5% for chain 1 and chain 2, respectively). We identified four small regions of the PlyCA structure that included the majority of the outliers. One was in the GyH domain (three outliers between residues 120 and 126) and three in the CHAP domain (seven outliers between residues 354 and 389, three outliers between residues 411 and 417, and three outliers between residues 438 and 443). Despite extensive rebuilding and refinement, it was not possible to improve the backbone stereochemistry for these regions, and we reason that these portions of the molecule are most likely intrinsically flexible. The presence of these outliers do not interfere with the findings and interpretations presented in this manuscript. A breakdown of the components of the PlyC structure and their associated Ramachandran outliers and MolProbity scores are shown below.

Molecule	Ramachandran outliers, %	MolProbity Score*
PlyC asymmetric unit	3.07 (53/1725 residues)	2.42 (98th percentile)
PlyC copy 1†	2.70 (25/927 residues)	2.38 (98th percentile)
PlyC copy 2	3.51 (28/798 residues)	2.46 (98th percentile)
PlyCA copy 1	5.83 (24/412 residues)	2.93 (89th percentile)
PlyCA copy 2	6.50 (26/400 residues)	2.88 (91st percentile)
PlyCB copy 1	0.19 (1/515 residues)	1.50 (100th percentile)
PlyCB copy 2	0.50 (2/398 residues)	1.65 (100th percentile)

*n = 892, 3.30 ± 0.25 Å.
†From which graphical figures were produced in the present paper.

Superpositions were conducted by using PDBeFold (15) (http://www.ebi.ac.uk/msd-srv/ssm/). Pymol was used to produce

all structural representations (http://www.pymol.org). The coordinates and structure factors are available from the Protein Data Bank under ID codes 4F87 and 4F88. Raw data and images are available from TARDIS (16) (www.tardis.edu.au).

1. Nelson D, Schuch R, Chahales P, Zhu S, Fischetti VA (2006) PlyC: A multimeric bacteriophage lysin. *Proc Natl Acad Sci USA* 103:10765–10770.
2. Nelson D, Loomis L, Fischetti VA (2001) Prevention and elimination of upper respiratory colonization of mice by group A streptococci by using a bacteriophage lytic enzyme. *Proc Natl Acad Sci USA* 98:4107–4112.
3. Collaborative Computational Project, Number 4 (1994) The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* 50:760–763.
4. Leslie AGW (1992) *Joint CCP4 + ESF-EAMCB Newsletter on Protein Crystallography*, No. 26.
5. Vonrhein C, Blanc E, Roversi P, Bricogne G (2007) Automated structure solution with autoSHARP. *Methods Mol Biol* 364:215–230.
6. Cohen SX, et al. (2008) ARP/wARP and molecular replacement: The next generation. *Acta Crystallogr D Biol Crystallogr* 64:49–60.
7. Brünger AT (1993) Assessment of phase accuracy by cross validation: The free R value. Methods and applications. *Acta Crystallogr D Biol Crystallogr* 49:24–36.
8. Emsley P, Cowtan K (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60:2126–2132.
9. Potterton E, Briggs P, Turkenburg M, Dodson E (2003) A graphical user interface to the CCP4 program suite. *Acta Crystallogr D Biol Crystallogr* 59:1131–1137.
10. Evans P (2006) Scaling and assessment of data quality. *Acta Crystallogr D Biol Crystallogr* 62:72–82.
11. McCoy AJ, Grosse-Kunstleve RW, Storoni LC, Read RJ (2005) Likelihood-enhanced fast translation functions. *Acta Crystallogr D Biol Crystallogr* 61:458–464.
12. Bricogne G, et al. (2011) BUSTER, version 2.8 (Global Phasing, Cambridge, UK).
13. Smart OS, et al. (2008) Refinement with local structure similarity restraints (LSSR) enables exploitation of information from related structures and facilitates use of NCS. *Annual Meeting of the American Crystallography Association,* p 117.
14. Adams PD, et al. (2010) PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66:213–221.
15. Krissinel E, Henrick K (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol* 372:774–797.
16. Androulakis S, et al. (2008) Federated repositories of X-ray diffraction images. *Acta Crystallogr D Biol Crystallogr* D64:810–814.

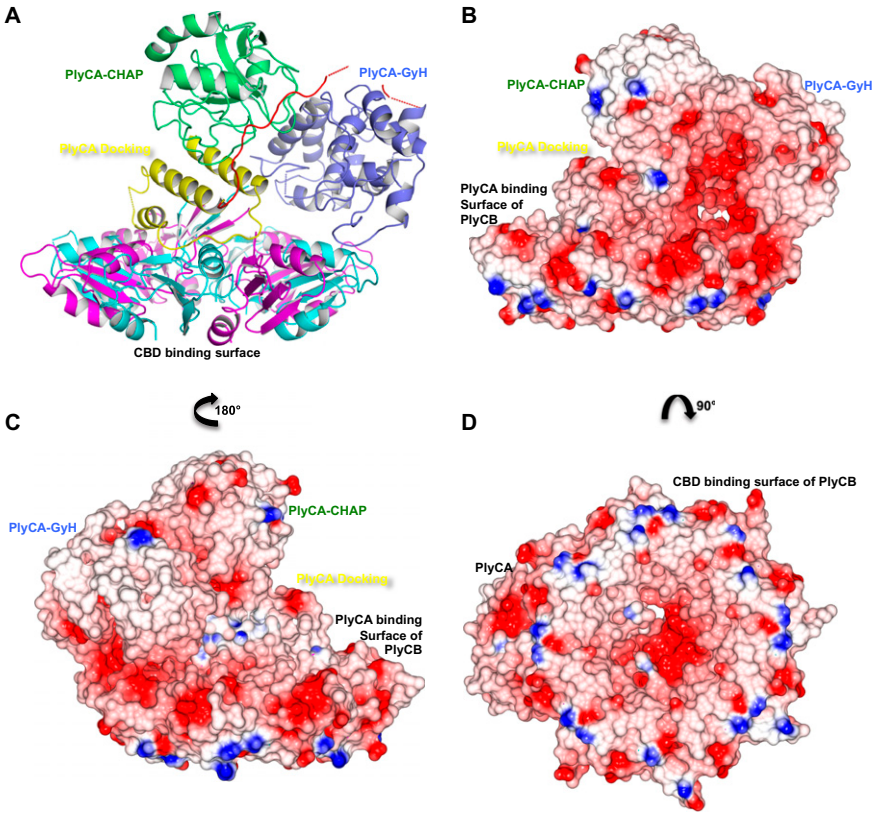


Fig. S1. Electrostatic surface potential of the 3.3-Å X-ray crystal structure of PlyC. (*A*) Cartoon model of PlyC as shown in Fig. 1*A* for orientation and domains. (*B*) Electrostatic surface potential of PlyC as oriented in *A*. CCP4MG was used to produce *B*–*D*. Surfaces are color-coded according to electrostatic potential (calculated by the Poisson–Boltzmann solver within CCP4MG). Lys and Arg residues were assigned a single positive charge, and Asp and Glu residues were assigned a single negative charge; all other residues were considered neutral. The calculation was done assuming a uniform dielectric constant of 80 for the solvent and two for the protein interior. The ionic strength was set to zero. The color of the surface represents the electrostatic potential at the protein surface, going from blue (potential of +10 kT/e) to red (potential of −10 kT/e), where T is temperature, e is the charge of an electron, and k is the Boltzmann constant. The probe radius used was 1.4 Å. *C* and *D* are the same surface with altered orientation (as indicated).

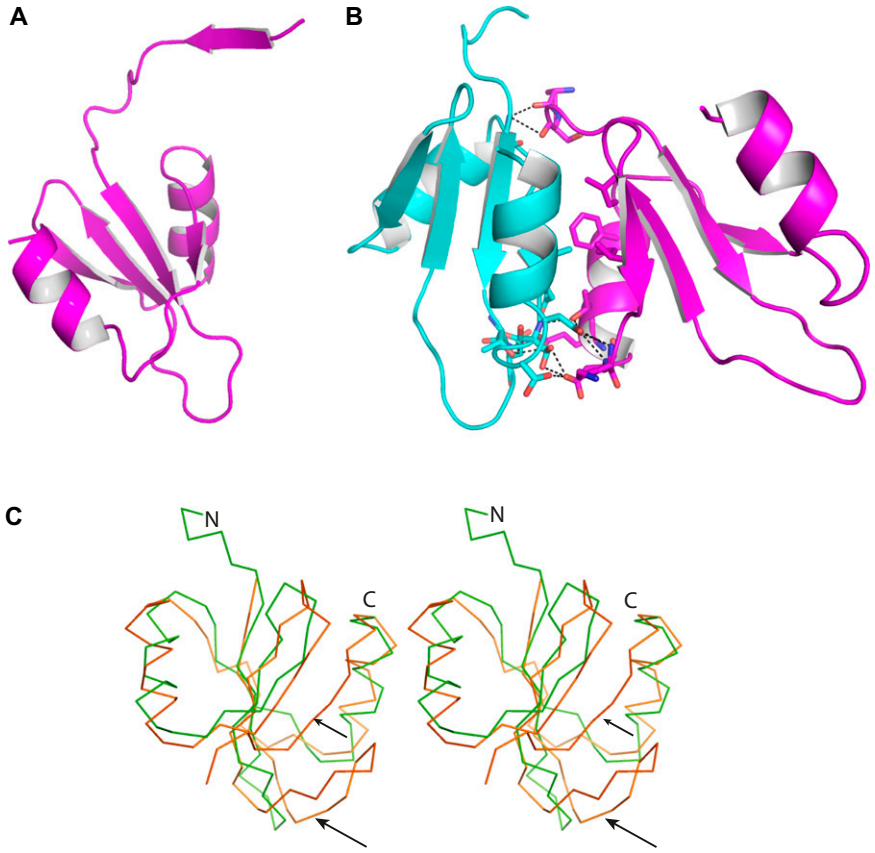


Fig. S2. Cartoon model of the X-ray crystal structure of PlyCB. (*A*) The 3.3-Å X-ray crystal structure of a PlyCB monomer, containing ordered N-terminal residues, as found in the PlyC model. (*B*) The 1.4-Å X-ray crystal structure of the PlyCB monomer (magenta)/monomer (cyan) interface. The interaction is mediated by 12 H-bonds and three vdw packing interactions and buries an average surface area of 770 $Å^2$ (excluding residues 1–9). This structure shows no order in N-terminal PlyCB residues 1 to 10. (*C*) Structural comparisons between PlyCB (green) and the C-terminal domain from *Streptococcus mutans* hypothetical protein (Protein Data Bank ID code 3L9A, orange). Stereo representation of structural alignment, showing Cα trace. N- and C-termini are labeled. Arrows indicate locations of insertions in 3L9A.

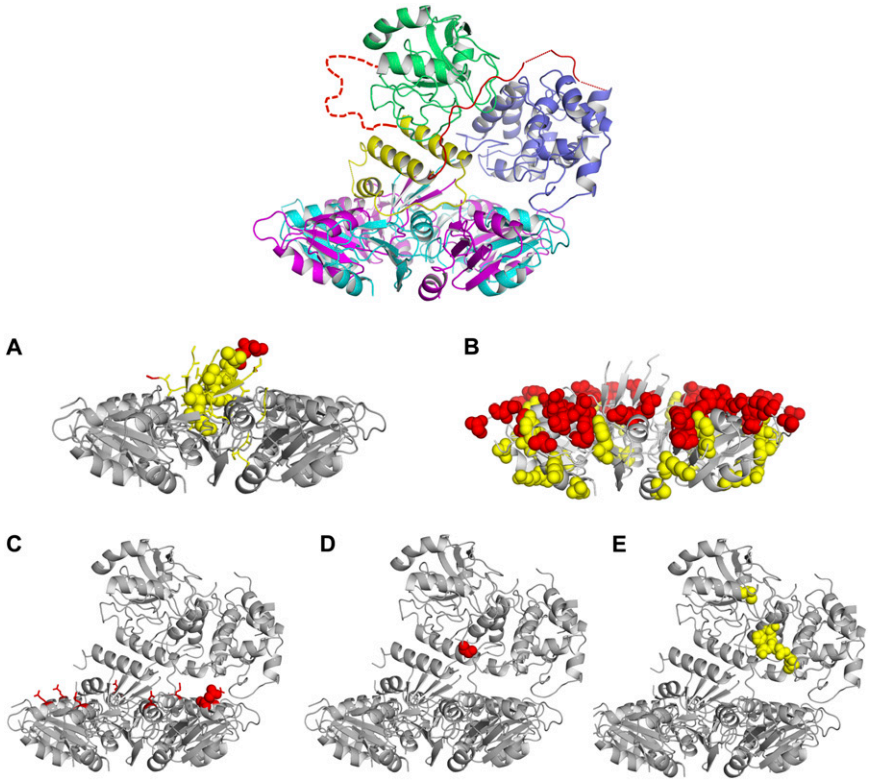


Fig. S3. Location of residues mutated in PlyC. A cartoon model of PlyC as shown in Fig. 1*A* for orientation and domains. For *A–E*, the PlyC cartoon is depicted in gray. The residues that have been mutated as per Table S3 are shown as spheres. The residues colored as red sticks and/or spheres had no effect on PlyC activity. The residues colored as yellow spheres affected PlyC activity. (*A*) PlyCA/PlyCB docking platform mutants. The PlyCB octamer has eight copies of each particular N-terminal mutant. Four of these will be located in the central β-sheet. Seven of the eight PlyCB residues are shown as sticks, with one representative monomer shown in spheres. (*B*) PlyCB cell wall binding mutants. (*C*) GyH domain/PlyCB interaction mutant. The PlyCB octamer has eight copies of the mutant. Seven of the eight residues are shown as sticks, and the residue proposed to interact with GyH (as per model) is shown as a sphere. (*D*) Linker occluding CHAP active site mutant. (*E*) GyH domain mutants with CHAP-Cys$_{333}$ are also indicated.

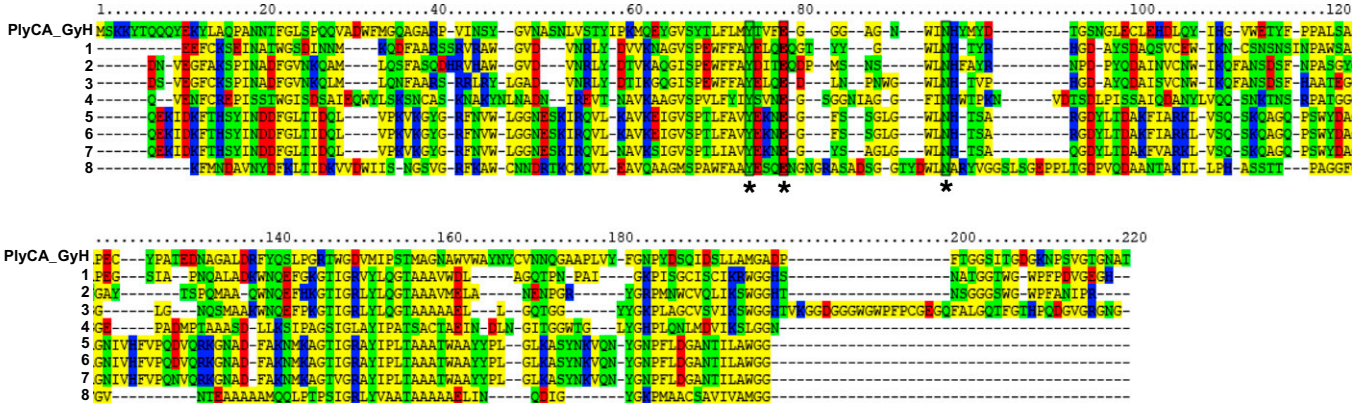


Fig. S4. Alignment of GyH with similar domains identified by PSI-Blast. Hydrophobic residues are boxed in yellow, hydrophilic in green, basic in blue, and acidic in red. Conserved residues are boxed and indicated by asterisks. Domain identification is as follows: 1, peptidase M23 (*Lactobacillus reuteri* 100–23) ZP_03072736.1; 2, phage lysin (*Lactobacillusantri* DSM16041) ZP_05746373.1; 3, peptidase M23 (*Lactobacillusoris* PB013-T2-3) ZP_07730732.1; 4, peptidoglycan-binding lysin domain (*Acetivibrio cellulolyticus* CD2) ZP_07327691.1; 5, ORF004 (*Staphylococcus* phage 66) YP_239474.1; 6, hypothetical protein 44AHJD_11 (*Staphylococcus* phage 44AHJD) NP_817306.1; 7, hypothetical protein SAP2_gp10 (*Staphylococcus* phage SAP-2) YP_001491535.1; and 9, gp076 (*Lactococcus* phage KSY1) YP_001469075.1.

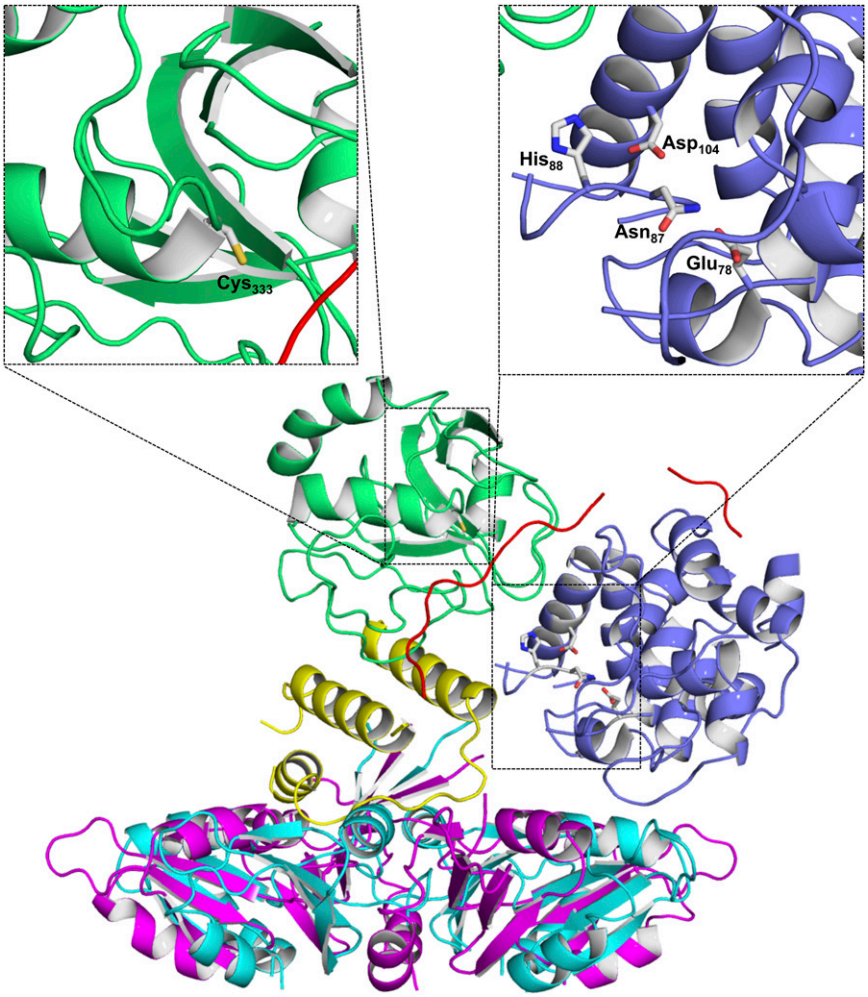


Fig. S5. Cartoon model of PlyC showing PlyCA residues mutated to test the biochemical activity of the two catalytic domains. Cartoon model of PlyC as shown in Fig. 1*A* for orientation and domains. *Insets*: Magnifications of regions indicated with residues shown as sticks and labeled.

Table S1. Crystallography data

Data collection	PlyCB (native)	PlyCB (S-SAD)	PlyC
Space group	*I*23	*I*23	$P2_12_12_1$
Cell dimensions			
a/b/c, Å	130.968	130.7	93.9/117.4/222.2
β, °	90.0	90.0	90.0
Resolution, Å	30–1.40 (1.45–1.40)	29.2–1.68 (1.77–1.68)	103.7–3.3 (3.48–3.30)
Total no. of observations	71,861	654,427	151,249
No. of unique observations	68,128	39,364	37,425
Multiplicity	7.1 (6.8)	16.6 (2.9)	4.0 (3.8)
Data completeness, %	98.4 (100.0)	93.4 (56.2)	99.2 (96.7)
$\langle I/\sigma_I \rangle$	48.9 (2.8)	36.4 (1.9)	8.6 (2.0)
R_{merge}, %[†]	5.6 (58.0)	5.0 (41.4)	10.0 (65.8)
$CC_{1/2}$	N/A	1.0 (0.81)	0.928 (0.63)
CC* (1)	N/A	1.0 (0.95)	0.98 (0.88)

S-SAD,sulfur single anomalous dispersion phasing; N/A, not applicable.
[†]Two complete PlyC molecules in asymmetric unit (1xPlyCA and 8xPlyCB = PlyC; x2).

1. Karplus PA, Diederichs K (2012) Linking crystallographic model and data quality. *Science* 336:1030–1033.

Table S2. Refinement statistics

Molecule	PlyCB AU*	PlyC				
		AU†	PlyC_1	PlyC_2	PlyCA(1/2)	PlyCB(1/2)
Nonhydrogen atoms						
Protein	1,894	12,296	N/A	N/A	N/A	N/A
Solvent	391	N/A	N/A	N/A	N/A	N/A
R_{free}, %	19.7	29.7	N/A	N/A	N/A	N/A
R_{cryst}, %	18.7	26.7	N/A	N/A	N/A	N/A
Bond lengths, Å	0.01	0.008	N/A	N/A	N/A	N/A
Bond angles, °	1.10	1.63	N/A	N/A	N/A	N/A
Ramachandran plot						
Unfavored, %	0.0	10.4	9.0	12.0	18.8	2.8
Outliers, %	0.0	3.07	2.70	3.51	6.16	0.33
B factors, Å^2						
Mean main chain	21.1	97.7	91.0	105.1	99.2	96.3
Mean side chain	24.7	99.6	95.4	105.3	99.4	99.7
Mean water molecule	25.6	N/A	N/A	N/A	N/A	N/A
rmsd bonded Bs						
Main chain	1.61	2.09	2.20	2.47	2.33	2.32
Side chain	3.46	4.67	4.97	4.38	4.11	5.19
MolProbity score (1)	1.21	2.42	2.38	2.46	2.91	1.6
MolProbity percentile	96th‡	98th§	98th§	98th§	90th§	100th§

N/A, not applicable.
*Four molecules of PlyCB in asymmetric unit (AU).
†Two complete PlyC molecules in asymmetric unit (1xPlyCA and 8xPlyCB = PlyC; x2).
‡n = 3363, 1.40Å ± 0.25Å.
§n = 892, 3.30 Å ± 0.25 Å.

1. Chen VB, et al. (2010) MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* D66:12–21.

Table S3. Properties of engineered PlyC mutants

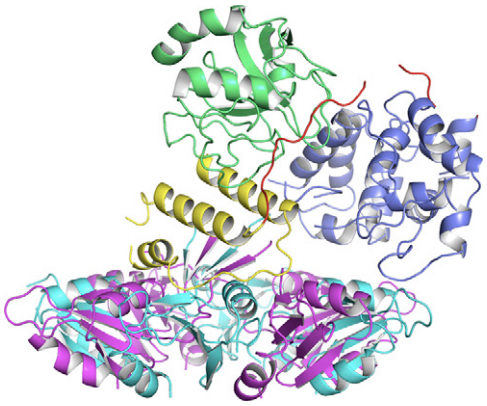
Mutation	Lytic activity	Holoenzyme formation	Cell wall binding
PlyCA/PlyCB docking platform mutants (Fig. S3*A*)			
PlyCBΔ1	**+++**	*	*
PlyCBΔ2	**+++**	*	*
PlyCBΔ3	–	No	NT
PlyCBΔ4	–	No	NT
PlyCBΔ5	–	No	NT
PlyCBΔ6	–	No	NT
PlyCBΔ7	–	No	NT
PlyCBΔ8	–	No	+++
PlyCB(I3A)	–	No	NT
PlyCB(I3D)	–	No	NT
PlyCB(I3K)	†	N/A	N/A
PlyCB(N4A)	**+++**	*	*
PlyCB streptococcal peptidoglycan binding mutants (Fig. S3*B*)			
PlyCB(D21A)	**+++**	*	*
PlyCB(D21N)	**+++**	*	*
PlyCB(K23A)	**+++**	*	*
PlyCB(K23E)	**+++**	*	*
PlyCB(Y28A)	++ (50%)	Yes	*
PlyCB(Y28F)	**+++**	*	*
PlyCB(R29E)	†	N/A	N/A
PlyCB(Q49A)	**+++**	*	*
PlyCB(Q50A)	**+++**	*	*
PlyCB(K59E)	+ (10%)	Yes	+
PlyCB(R66E)	+/– (<1%)	Yes	–
GyH domain/PlyCB interaction mutant (Fig. S3*C*)			
PlyCB(Q46A)†	+++	*	*
Linker occluding CHAP active site (Fig. S3*D*)			
PlyCA(S224Q)	**+++**	*	*
GyH domain mutants (Fig. S3*E*)			
PlyCA(E78A)	+ (10%)	Yes	‡
PlyCA(N87A)	+ (10%)	Yes	‡
PlyCA(H88A)	+ (10%)	Yes	‡
PlyCA(D104A)	+ (<3%)	Yes	‡
GyH/CHAP domain double mutants (Fig. S3*E*)			
PlyCA(E78A/C333S)	–	Yes	‡
PlyCA(N87A/C333S)	–	Yes	‡
PlyCA(H88A/C333S)	–	Yes	‡
PlyCA(D104A/C333S)	–	Yes	‡
Catalytic domain deletion mutants			
PlyCAΔCHAP	+/– (<1%)	Yes	‡
PlyCAΔGyH	+/– (<1%)	Yes	‡

+++ WT activity or binding; ++, diminished activity or binding; +, significantly diminished activity or binding (≤10% WT); +/–, negligible activity or binding (≤1% WT); –, no activity or binding observed; N/A, not applicable; NT, not tested.

*All constructs that exhibited WT lytic activity (+++) were not tested further as it was inferred that they were able to form the holoenzyme and bind the streptococcal surface. Boldface indicates PlyC mutants possessing WT activity.

†Protein formed inclusion bodies/misfolded. Attempts to refold were unsuccessful.

‡These mutants all possessed a WT PlyCB molecule so were able to bind to the streptococcal surface.



Movie S1. The cartoon model of the PlyC holoenzyme, colored as in Fig. 1.

Movie S1